

# Qualitative and Quantitative Radiological Measures of Fracture Healing

John R. Field<sup>1</sup> Greg R. Ruthenbeck<sup>1</sup>

<sup>1</sup> Pattagansett Consulting Pty Ltd, Upper Sturt, South Australia, Australia

Address for correspondence John R. Field, DVSc, PhD, Pattagansett Consulting Pty Ltd, 92 Upper Sturt Road, Upper Sturt, SA 5150, Australia (e-mail: lantbruks@bigpond.com).

Vet Comp Orthop Traumatol 2018;31:1–9.

## Abstract

The formulation of appropriate postoperative strategies, following fracture repair, currently involves an understanding of radiological and clinical outcome measures. This study has evaluated several modalities used to assess the progression of bone healing in a sheep tibial segmental defect model. Measures of defect optical density and volumetric data including bone density (BD), bone volume (BV) and bone mass (BM) were compared with qualitative data involving visual appraisal of radiographs [% bridging callus and modified radiographic union score tibia (mRUST)] and a clinical outcome measure (locomotory function). Percent bridging callus and mRUST measures displayed strong correlation ( $r = 0.999$ ), while locomotory function was weakly correlated with bridging callus ( $r = 0.029$ ) and mRUST ( $r = 0.046$ ). There was moderate to strong correlation between the qualitative and quantitative data. Bone density, BV and BM showed strong correlations within this dataset (BD–BV,  $r = 0.814$ ; BD–BM,  $r = 0.818$ ; BV–BM,  $r = 1.000$ ). Likewise, optical density measures were strongly correlated with BD ( $r = 0.824$ ), BV ( $r = 0.957$ ) and BM ( $r = 0.959$ ). The utilization of both qualitative and quantitative data, in assessment of the progression of fracture healing, has provided valuable insight. Measures of optical density have been shown to make a substantial contribution to this assessment and which should be considered for use in studies evaluating fracture healing.

## Keywords

- advanced imaging techniques
- bone defect healing
- animal models of human disease
- comparative research
- fracture repair

## Introduction

In clinical practice, defining the progression of fracture healing involves, most commonly, either serial radiography or serial computed tomography (CT), combined with an assessment of clinical wellness. These measures rely on qualitative and quantitative interpretations as to the state of repair and response in healing.

Measures of wellness provide an important aspect in human clinical medicine and are commonly based on patient questionnaire responses or clinician impressions.<sup>1,2</sup> In veterinary medicine, a key indicator of patient wellness is their willingness to ambulate. Objective measurement and analysis of locomotory function in animals<sup>3</sup> involves either kinematic (temporal

and geometric motion characteristics) or kinetics (forces producing or modifying motion). A locomotion scoring scale has been developed specifically for the postoperative assessment of sheep based primarily on observable characteristics of their gait and posture.<sup>4</sup> As in human clinical medicine, the qualitative data obtained from measures of wellness are considered in conjunction with quantitative measures.

Diagnostic radiography provides insight as to whether a fracture has healed. It should also underscore the progression of healing, enabling inference as to whether non-union might occur allowing for early reintervention.<sup>5</sup> Several options are available where standardized, plain radiographs can provide both qualitative and quantitative measures of fracture healing progress. The percentage of bridging callus has been used

received

March 16, 2017

accepted after revision

October 7, 2017

Copyright © 2018 Schattauer

DOI <https://doi.org/10.3415/VCOT-17-03-0042>  
ISSN 0932-0814.

to track the performance of various treatments in animal models of long bone segmental defects.<sup>6–9</sup> Four circumferential quadrants were established requiring craniocaudal and lateromedial radiographs allowing the measurement of callus formation over time and the comparison of treatments. Scoring the degree of fusion apparent is another semi-quantitative method. As with bridging callus quantitation, the radiographic union score for tibial fractures (RUST) applies a score to each of four quadrants with the minimum score indicative of failed union and a maximum score indicating the fracture has healed.<sup>10</sup> With both % bridging callus and RUST determinations, there is considerable inter-operator variability (20–25%), impacting both the accuracy and precision of the measures obtained which detracts from their use in clinical cases.<sup>11,12</sup>

Quantitation of optical density (OD) has been utilized in both clinical and research environments.<sup>13–18</sup> The methods employed ultimately provide a value for OD in a given region of interest (ROI). This has been achieved using two primary methodologies, which have allowed comparison of OD across multiple radiographic images while controlling for factors that may influence the results, such as exposure factors.

Some have used a reference penetrometer (aluminium step wedge) to obtain baseline values with which to compare the progression of bone healing over the period of evaluation.<sup>13,15,19–21</sup> Evidence of strong linearity between steps (penetrometer) is essential to confer high accuracy and precision when comparing different images. Other studies have based their determination of OD on the manipulation and standardization of intensity (greyscale) magnitudes.<sup>5,22,23</sup> A high level of association was found when comparing computer algorithm measures and the visual interpretation made by clinicians ( $R^2 = 0.94$ ). The average inter-observer variation in computer algorithm measures was 4% while clinician-based observations had a variation of 22%. By limiting any subjective contribution, the computer algorithm has displayed a substantially lower variability when compared with visual clinical observations.<sup>5,22,24</sup>

The hypothesis of the current study was that change in OD, during bone regeneration, could be adequately quantified, using computer-based algorithms in the greyscale environment and which would reflect the progression of bone regeneration. This study presents data derived from plain radiographs, manipulated using different approaches and compares these with volumetric data obtained by CT from the same sample set over a 16-week in vivo period. The degree of linear association (correlation) between the measures has been used to determine if the quantitation of OD may have a role in defining regeneration of bone in fracture repair.

## Materials and Methods

Data obtained from an in vivo study were utilized to evaluate the utility of using qualitative or quantitative data as a means of tracking bone regeneration following surgical reconstruction of a mid-diaphyseal, tibial segmental defect. The study involved the use of skeletally mature, 3-year castrated male Merino sheep. The animal utilization protocol was approved

by the Animal Care and Use Review Committee, Department of the Army (US – DARPA-FY08–019) and the undertaking of Institutional Animal Care Committees (SA Path: 111.12).

### Study Protocol

A unilateral, 3.5-cm mid-diaphyseal tibial segmental defect ( $n = 8$ ), was created and reconstructed using a locking intramedullary nail in combination with autograft bone using methodologies developed elsewhere.<sup>6,7,9–11</sup> It has been shown that a 3.0-cm defect in the tibial diaphysis is a 'critical defect' resulting in 100% non-union when void-fillers are not used.<sup>9</sup>

Following pre-medication (xylazine: 0.1 mg/kg intramuscularly), anaesthesia was induced using thiopentone [50–70 mg/kg body weight intravenously (i.v.)]. A cuffed endotracheal tube allowed the maintenance of gaseous anaesthesia with isoflurane (2%) in oxygen (flow rate: 15 L/min). All sheep received pre-emptive cephalosporin (Kefzol: 22 mg/kg body weight, i.v. daily; Sdpen Pharmacare Australia Pty Ltd, St. Leonards, NSW, Australia) and meloxicam (0.2 mg/kg i.v. daily; Boehringer Ingelheim Pty Ltd, North Ryde, NSW, Australia) during surgical preparation and analgesia, via a constant infusion backpack pump (xylazine: 0.5 mg/h i.v.) initiated at the time of extubation; all drugs were administered to all sheep for an initial 3-day period postoperatively and continued if deemed necessary.

In all sheep, a 24 cm × 8.5 mm titanium Trigen Metanail (Advanced Surgical Devices, Smith and Nephew Surgical, Memphis, Tennessee, United States) was used in reconstructing the 3.5-cm mid-shaft defect. A standardized mid-diaphyseal defect, using known anatomical landmarks (proximal medial tibia and distal medial malleolus) was created prior to reconstruction; a 3.5-mm cutting guide provided precise and reproducible defect dimensions.

Autograft bone was obtained by morselizing the segment of cortical bone removed into 3 to 4 mm chips (Aesculap Bone Mill, B. Braun, Bella Vista, NSW, Australia). Bone chips were combined with medullary content (fat) forming a gruel, which was packed into the defect following the application of four 4.0-mm titanium locking screws (Trigen internal hex capture screw, Advanced Surgical Devices, Smith and Nephew Surgical), two proximal and two distal. Wound closure was performed using 3.5 metric polyglactin 910 suture (Vicryl: Ethicon Inc., Cornelia, Georgia, United States). Metal staples were used for skin closure (Covidien, Medtronic, Minneapolis, Minnesota, United States).

All sheep were recovered in a weight-bearing sling in which they remained for 3 weeks; thereafter, they were maintained in soil-floored pens allowing unrestricted movement. Serial in vivo examinations (radiology and CT) occurred immediately postoperatively and at 4, 8, 12 and 16 weeks while the sheep were under general anaesthesia.

Data from serial radiographic and CT examinations were incorporated into the analysis performed. Comparisons were made between qualitative clinical measures (% bridging callus, RUST and locomotory function) and quantitative measures (volumetric analysis, OD) obtained from each ROI or volume of interest (VOI).

Using standardized radiographic technique (Villa Visitor mobile X-ray, Buccinasco, Italy: 54–56 kVp, 6–8 mA, focal

length 120 cm), both craniocaudal and lateromedial DICOM (Digital Imaging and Communications in Medicine) images were obtained (Carestream Vita LA processor, Carestream, Rochester, New York, United States). Computed tomography was performed (Phillips Brilliance, North Ryde, NSW, Australia) using the acquisition parameters 120 kVp, 200 mA/slice, 1-mm slice thickness and 200 mm field of view, producing a matrix of  $512 \times 512$ . Primary manipulation of image files was performed using the OSIRIX MD software (v2.9, 64-bit, Pixmeo, Switzerland).

For accurate and reproducible determination of both the radiographic ROI and the computer tomographic VOI, the immediate postoperative images provided a template for all subsequent ROI and VOI delineation; this relied on the known dimensions of the nail and the distance (mm or pixels) from the locking screw proximal and distal to the defect (►Fig. 1). For both radiographic and CT evaluations, the ROI and VOI related specifically to the defect created and did not involve either proximal or distal cortical bone.

Following collection of raw data, image sets were evaluated to build a picture of the progression of defect healing and to evaluate the degree of association between the methods employed using Pearson product-moment correlation.

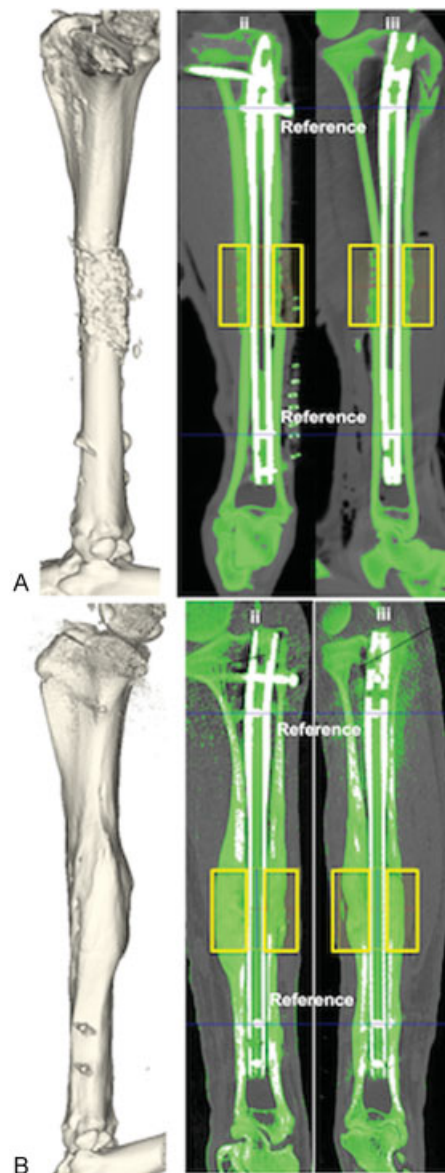
## Qualitative Assessments

### Bridging Callus

The craniocaudal radiographic view allowed quantitation of % bridging callus along the lateral and medial aspects of the defect.<sup>6,7,25,26</sup> Similarly, the latero-medial view allowed quantitation of the bridging callus along the cranial and caudal aspects. Because of possible differences in image magnification, pixel values were used to enable quantitation. A value of 0% indicated no bridging callus, while 100% indicated complete bridging. For each individual animal, at each time point (postoperative, 4, 8, 12 and 16 weeks), and for each anatomical location (cranial, caudal, lateral and medial), the mean and standard deviation of data were compiled allowing comparison with the other methodologies under consideration.

### Modified Radiographic Union Score for Tibial Fractures

Using the same radiographic projections and anatomical locations as above, visual scoring assigned a numerical value for each anatomical location. A modified version of RUST (mRUST) was used, providing more appropriate scoring for callus formation in the current segmental defect study, as opposed to that which is used in clinical cases. A score of 1 reflected no evidence of callus; a score of 2, minimal callus formation; a score of 3, moderate callus formation; a score of 4, advanced callus formation; and a score of 5, complete bridging. Individual quadrant scores were summed to give a total for that particular examination (4 being the minimum score indicating poor healing or non-union and 20 being the maximum score with all quadrants indicating complete bridging). The mean circumferential value was compiled by pooling data obtained from each quadrant.



**Fig. 1** A schematic depiction of immediate postoperative (A) and 16-week (B) 3-D reconstructions (i) and mutually orthogonal views used in the analysis showing craniocaudal (ii) and lateromedial (iii) projections. The yellow-lined rectangles highlight the volumes of interest used in the analysis. Reference locations (proximal and distal locking screws) are depicted.

## Locomotor Function

A validated 7-point numeric scale (0 to –6) was used to grade the postoperative locomotory performance of sheep enrolled in the study.<sup>4</sup> The scale ranged from 'normal' (zero) to 'unable' to stand or move with descriptions of locomotion for each increase in score severity. Scores were recorded preoperatively and at 4, 8, 12 and 16 weeks postoperatively for all individuals.

## Quantitative Assessments

### Volumetric Analysis

Computed tomography was used to generate DICOM files of the VOI (defect). This was performed immediately post-

surgery and at 4, 8, 12 and 16 weeks postoperatively for all individuals. Pixel resolution of the images obtained was 0.408 mm with a slice interval of 0.5 mm. The images were processed using custom software (GPU Volumetric re-sampler; GRIT, Adelaide, Australia) to remove variations in the VOI, resolution and voxel shape. Raw binary files containing the voxel data were subject to a 3-D Gaussian blur, with sub-voxel radius applied to the datasets to remove speckle and artefact contributed by the nail. Re-sampling was then undertaken using 3-D bilinear sampling to maximize the data quality and minimize information loss. The final output was a raw 16-bit greyscale image stack with a voxel size of 0.123 mm<sup>3</sup> (0.5 mm × 0.5 mm × 0.5 mm). Automated thresholding, based on a histogram of each single bone volume (BV), was performed. This removed subjectivity that might have been present had the threshold been manually selected; a value of 108% was chosen ensuring the lowest possible threshold was selected so as to include bone from which the soft tissue was excluded.

Two, mutually orthogonal axial views (►Fig. 1), within the intensity threshold range, formed the basis for voxel counts and the quantitation of bone regeneration. Scaling to one cubic centimetre of the VOI allowed for direct comparison of treatments using equivalent data for change in bone density (BD;  $\Delta BD$  – gr/cm<sup>3</sup>), BV ( $\Delta BV$  – cm<sup>3</sup>) and bone mass (BM;  $\Delta BM$  – gr). Voxel intensity (grey-level) is proportional to density; this value was multiplied by voxel count (volume) to provide an estimate of BM.

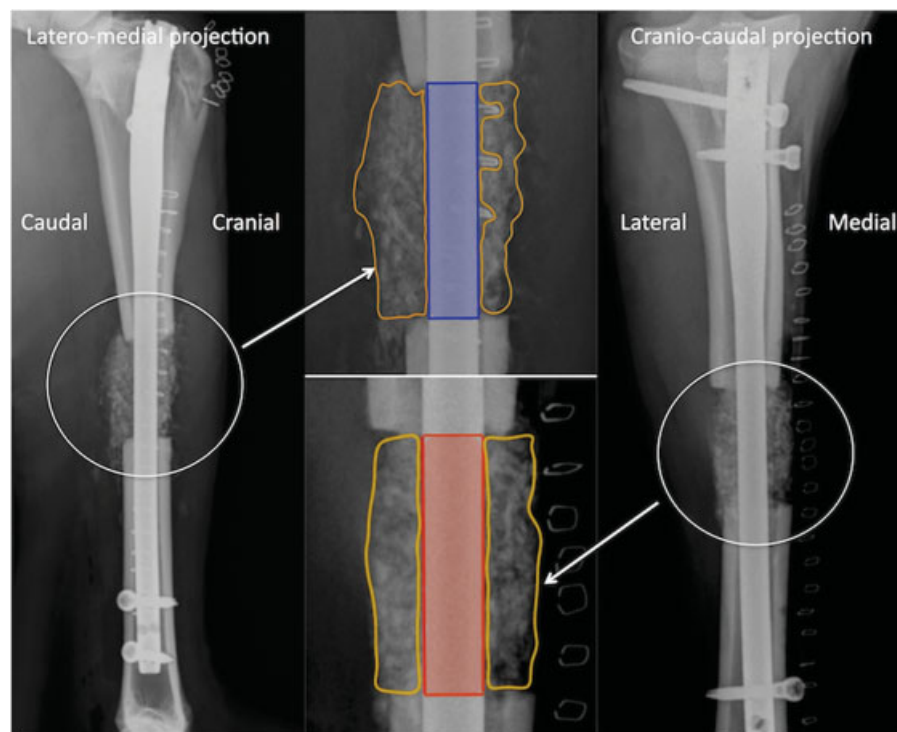
All data were 'normalized'; each of the data counts obtained were divided by the first data output (postopera-

tive); as all VOIs were the same, it was reasonable to make comparisons of the data between the different sheep.

### Optical Density

All standard radiographic images (DICOM) were converted to greyscale (GIMP, GNU Image Manipulating Program v2.8.20, <https://www.gimp.org>). The intensity of each pixel was expressed within a given range between a minimum (0 = black) and a maximum (255 = white) with fractional values of grey in between. Following conversion to greyscale, each image was re-introduced to the analytical software (OSIRIX MD v2.9, 64-bit, Pixmeo) for delineation of ROIs and generation of OD values (►Fig. 2). One observer (J.R.F.) performed the analysis; intra-observer kappa values were generated.

All data were normalized with regards to the range of greyscale values (0–255). Assuming a linear relationship between 0 and 255, the raw OD value of the nail was normalized to OD 255 for both radiographic views (cranial, caudal, lateral and medial), at each time point and in each individual. Regions of interest from raw images were defined using the draw function of the OSIRIX MD software. The number of outlined pixels was converted to a metric area and the mean OD of that area automatically quantitated. The resultant raw OD values were normalized using the factor defined in normalizing the OD of the nail. The normalized OD value of the nail in the craniocaudal projection was used to normalize the ROI data in the lateral and medial anatomical aspects of the reconstruction, while normalized OD values of the nail in the lateromedial projection were used to



**Fig. 2** A schematic depiction of the radiographic projection (lateromedial or craniocaudal) and anatomic location (cranial, caudal, lateral or medial) of measures obtained. Regions of interest (ROIs) are defined for each projection; the solid rectangles reflect the ROI for quantitation of greyscale value for the nail. The ROIs for autograft, within the defect, are also highlighted (yellow line).



normalize the ROI data in the cranial and caudal anatomical aspects of the reconstruction. Normalizing OD data from each quadrant allowed the development of a circumferential impression of normalized OD for each treatment at each time point considered (postoperatively, 4, 8 and 12 weeks) and for each individual, on which to make comparisons. The normalized, raw data were also modified defining the immediate postoperative values to time zero; this assisted in removing possible artefact associated with the nail.

All data were subject to multivariate analysis of variance with Bonferroni post hoc assessment; a probability of  $p < 0.05$  was deemed significant. An assessment of association was undertaken by submitting the data to a Pearson product-moment correlation (Data Desk v6.3, Data Description, Ithaca, New York, United States). Quantitative data measures (volumetric analysis and OD) were repeated in a blinded manner and evaluated for intra-observer variation (kappa statistic).

## Results

Qualitative and quantitative data were collected from eight animals that underwent creation and reconstruction of a tibial segmental defect using a locking intramedullary nail combined with autograft bone.

There were no surgical complications, nor withdrawals from the study.

### Qualitative Assessments

Visual appraisal of plain radiographs, documenting bridging callus formation (►Fig. 3A) and compilation of mRUST scores (►Fig. 3B), combined with assessment of locomotory function (►Fig. 3C), allowed description of qualitative data as follows.

#### Bridging Callus

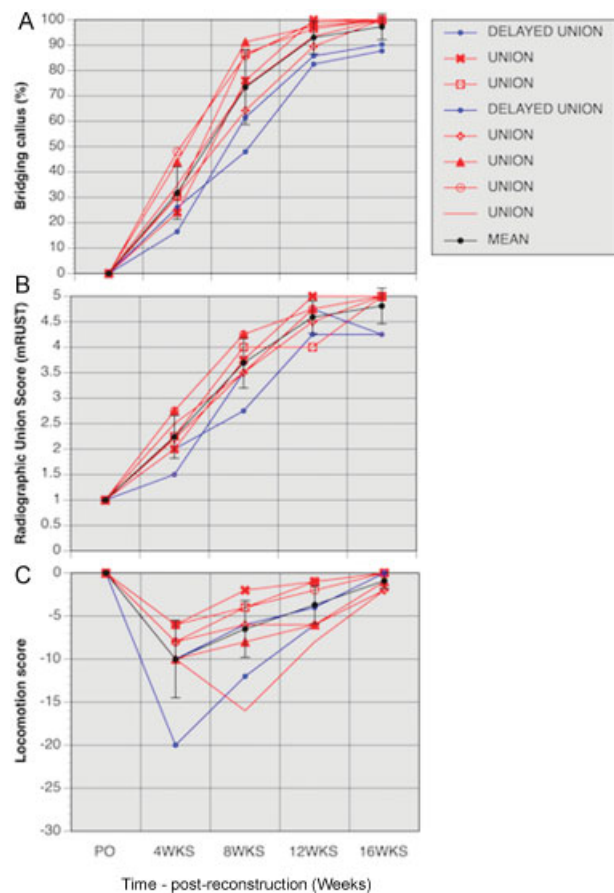
There was a steady increase in the amount of bridging callus formation over 8 to 12 weeks post-reconstruction (►Fig. 3A). In all but two animals, bridging was complete by 16 weeks. The performance of two sheep, with delayed union, is highlighted graphically. Initially, there was large variation around the means as might be expected in rapidly forming new bone. By 12 weeks, this variation was largely eliminated. The passage of time proved a significant factor in the formation of bridging callus ( $p = 0.0001$ ). Pearson product-moment correlation was  $r = 0.999$  in consideration of bridging callus (%) and mRUST, indicating a strong association between these two qualitative measures.

#### Modified Radiographic Union Score for Tibial Fractures

Similarly, there was a steady increase in mRUST scores through 12 weeks post-reconstruction (►Fig. 3B). The performance of two individuals with delayed union, is also highlighted graphically. The passage of time proved a significant factor in the scores obtained ( $p = 0.0001$ ).

#### Locomotory Function

Conversely, locomotory function (►Fig. 3C) showed poor association with bridging callus ( $r = 0.029$ ) and mRUST ( $r = 0.046$ ). As expected postoperatively, there was a rapid



**Fig. 3** A graphical depiction of the trends observed in qualitative data (% bridging callus, A; mRUST, B; and locomotory score, C). Individual measures have been labelled either union or delayed union; individuals defined as delayed union are shown in blue. The mean and standard deviation of the pooled individual data are presented.

decrease in locomotory function over 4 to 8 weeks post-reconstruction. Thereafter, there was a gradual return to preoperative levels by 16 weeks. The passage of time was a significant factor in the scores obtained ( $p = 0.0001$ ).

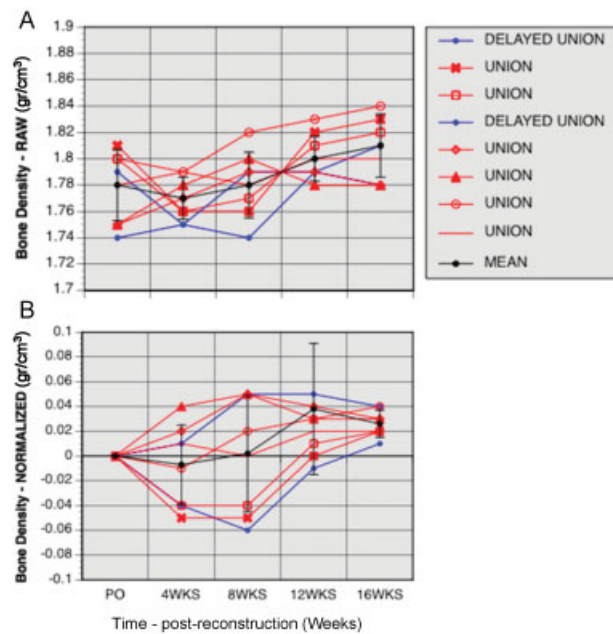
### Quantitative Assessments

A graphical depiction of volumetric parameters ( $\Delta BD$ ,  $\Delta BV$ ,  $\Delta BM$ ) is shown in ►Figs. 4–6. Optical density measures are depicted graphically in ►Fig. 7. For each measure, the trend was similar, with a decrease apparent through 4 weeks post-reconstruction followed by a gradual and consistent increase to the time of retrieval.

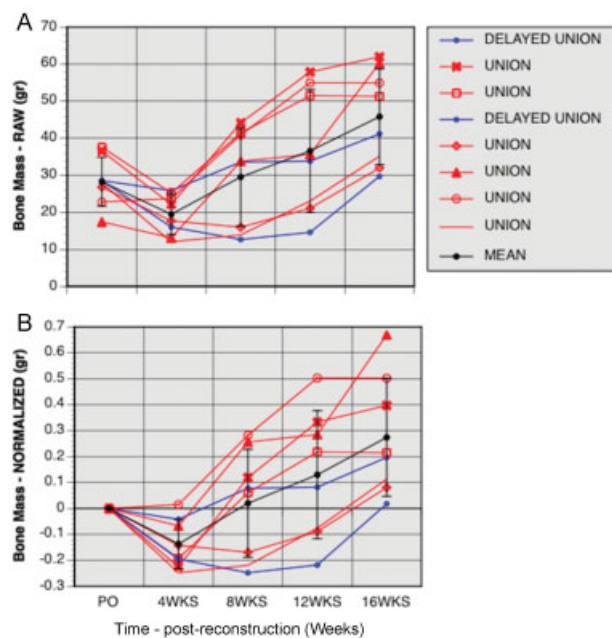
#### Volumetric Analysis

Changes in BD (►Fig. 4A, B) were not uniform with the passage of time and were not significant ( $p = 0.089$ ). Bone density had strong correlation with both BV ( $r = 0.814$ ) and BM ( $r = 0.818$ ). Bone volume (►Fig. 5A, B) and BM (►Fig. 6A, B) displayed significant differences with the passage of time (BV:  $p = 0.001$ , BM:  $p = 0.002$ ) and were strongly associated ( $r = 1.000$ ).

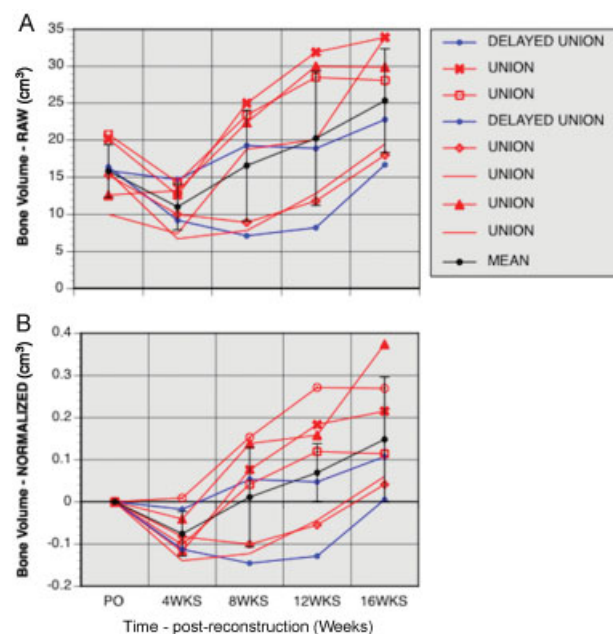
For each parameter, those specimens adjudged to be delayed union have been highlighted graphically.



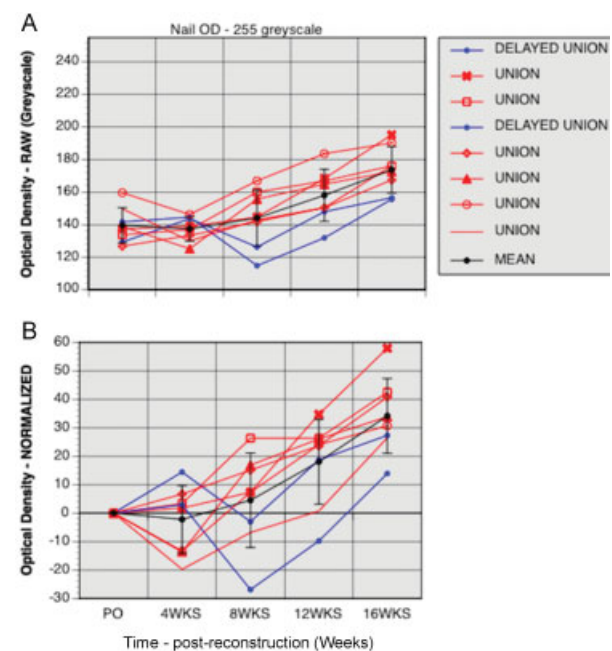
**Fig. 4** A graphical depiction of changes in bone density ( $\Delta BD$ ), over the in vivo period post-reconstruction. The raw data (A) and normalized data (B) are shown. Individual measures have been labelled either union or delayed union; individuals defined as delayed union are shown in blue. The mean and standard deviation of the pooled individual data are presented.



**Fig. 6** A graphical depiction of changes in bone mass ( $\Delta BM$ ), over the in vivo period post-reconstruction. The raw data (A) and normalized data (B) are shown. Individual measures have been labelled either union or delayed union; individuals defined as delayed union are shown in blue. The mean and standard deviation of the pooled individual data are presented.



**Fig. 5** A graphical depiction of changes in bone volume ( $\Delta BV$ ), over the in vivo period post-reconstruction. The raw data (A) and normalized data (B) are shown. Individual measures have been labelled either union or delayed union; individuals defined as delayed union are shown in blue. The mean and standard deviation of the pooled individual data are presented.



**Fig. 7** A graphical depiction of the trend in optical density measures for individuals over the in vivo period post-reconstruction. The raw data (A) and normalized data (B) are shown. Individual measures have been labelled either union or delayed union; individuals defined as delayed union are shown in blue. The mean and standard deviation of the pooled individual data are presented.

**Table 1** Pearson's product-moment correlation ( $r$ ) values within and between qualitative and quantitative datasets

Qualitative data				Quantitative data				
	Callus	mRUST	LOCOMOTION		BD	BV	BM	OD
Callus	1.000				0.775	0.728	0.731	0.802
mRUST	0.999	1.000			0.791	0.743	0.746	0.823
LOCOMOTION	<i>0.029</i>	<i>0.046</i>	1.000		<i>0.449</i>	0.684	0.681	0.501
				BD	1.000			
				BV	0.814	1.000		
				BM	0.818	1.000	1.000	
				OD	0.824	0.957	0.959	1.000

Abbreviations: BD, bone density; BM, bone mass; BV, bone volume; mRUST, modified radiographic union score for tibial fracture; OD, optical density. Note: Italicized text reflects poor to moderate correlation.

### Optical Density

The passage of time was shown to have a significant impact on the measured OD ( $p = 0.0001$ ) reflecting the regenerative process. ► **Table 1** presents Pearson product-moment correlation values for all interactions within and between the qualitative and quantitative data.

Measures of OD have shown strong correlation with the qualitative measures of bridging callus ( $r = 0.802$ ) and mRUST ( $r = 0.823$ ). Likewise, OD was strongly correlated with BD ( $r = 0.824$ ), BV ( $r = 0.957$ ) and BM ( $r = 0.959$ ). A graphical depiction of the changes in OD show a similar trend, as seen in other quantitative values, with an initial decrease, during the resorptive phase, followed by a steady increase in value between 8 and 16 weeks post-reconstruction (► **Fig. 7A, B**). Of note are the OD changes observed for both delayed union individuals, although they could not be shown to be significantly different from 'union' individuals ( $p > 0.05$ ).

► **Fig. 8** depicts the plot of qualitative data (► **Fig. 8A**) and quantitative data (► **Fig. 8B**) versus OD; correlation values are highlighted and are strong.

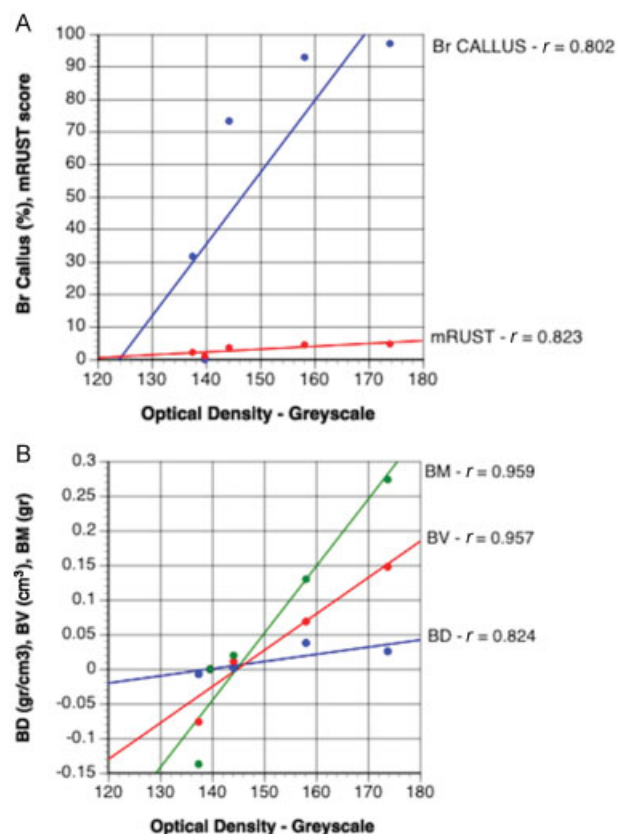
### Discussion

The establishment of standardized clinical measures of fracture healing remains elusive. The key aspects of fracture repair are the following: when is the repair complete and when can normal ambulation be undertaken?

No currently available qualitative or quantitative measure is capable of accurately defining these requirements. We continue, therefore, to place reliance on a compilation of qualitative and quantitative measures to provide insight, but these are not definitive. Recent studies<sup>27–29</sup> have used differentially loaded radiostereometric analysis as a means of providing a non-invasive quantitation of construct stiffness for healing tibial fractures. From this, reconstruction strength can be extrapolated, but it does not provide a measure of tolerance to loading, which is critical in defining an appropriate end to convalescence.

The current study has investigated what, if any, contribution measures of OD could make in the assessment of fracture healing. The potential utility of this method was

its reliance on serial plain radiographs, simply analysed using computer-based algorithms in the greyscale environment. A key factor in developing this method was based on the normalization of OD data using the metallic fixation device to represent the maximum OD value (255) while assuming a linear relationship between 0 and 255 greyscale units. The OD data obtained were then correlated with both qualitative and quantitative measures from the same specimens. In



**Fig. 8** A graphical depiction of the data points for qualitative data (A: % bridging callus, mRUST) and quantitative data (B: bone density, BD; bone volume, BV; and bone mass, BM) when plotted against optical density (greyscale). The lines represent best fit to the data with the resultant Pearson product-moment correlation value for each of the correlations depicted.

this regard, OD was strongly correlated with the other measures (bridging callus,  $r = 0.802$ ; mRUST,  $r = 0.823$ ;  $\Delta BD$ ,  $r = 0.824$ ;  $\Delta BV$ ,  $r = 0.957$ ; and  $\Delta BM$ ,  $r = 0.959$ ).

Qualitative measures (bridging callus and mRUST) suffer from poor inter- and intra-observer agreement.<sup>4,12–14</sup> This variability is diminished with the removal of operator effects through the use of computer-based algorithms. To evaluate this aspect, both the volumetric and OD data measures were repeated to determine the degree of intra-observer agreement,<sup>30</sup> which was substantial ( $k = 0.82$ ). Any future study should address inter-observer agreement to confirm the utility of these measures.

At the time of retrieval, two of eight specimens were adjudged as delayed union; this was based on their radiographic appearance, response to torsional loading and histology (loading and histological data presented elsewhere). These specimens have been highlighted in each of the graphical depictions of the data (►Figs. 3–7). For bridging callus, mRUST scores, volumetric measures and OD, these two individuals have displayed variable responses in all the parameters measured and which could not be utilized to confirm delayed union. This highlights the fact that, even in the face of strong radiographic and CT evidence for union, it is ultimately the tolerance to loading that provides the conclusive assessment.

There are several issues that require further investigation to obtain a more complete picture as to the state of fracture healing:

1. The determination of bone union is fraught; current clinical practice relies on subjective assessments of radiographs and CT. A key element is load-bearing capacity; serial measures of torsional stiffness in human patients is currently being evaluated using differentially loaded radiostereometric analysis<sup>27–29</sup> and shows great promise; however, it also is used in combination with radiographic and CT imaging to make a final judgement as to the state of healing.
2. It can be asserted that the presence of autograft bone in the defect and the presence of the locking nail may interfere with the accuracy of the measures obtained (bridging callus and mRUST). In evaluating radiographs for bridging callus formation, the new bone growth along the outer aspect of the defect was evaluated and compared with the post-operative images. New bone immediately adjacent to the nail could not be evaluated in these measures.

In performing the volumetric analysis, the potential 'artefact' provided by either the nail or the autograft bone has been taken into account by normalizing all later time point data to the data outputs obtained immediately postoperatively. The initial decrease in parameter value over the first 4 weeks likely reflects the resorptive process commonly associated with the use of bone autograft. Bone density varied with the passage of time and was most likely a consequence of substantial variations in BD between individuals with resultant large variations about the means. This may have been impacted through variability inherent in the autograft bone although a standardized amount of bone was used in each animal.

This study has enabled the evaluation of fracture healing in a sheep tibial critical defect model. Both qualitative (% bridging callus, mRUST and locomotion score) and quantitative data (CT volumetric analysis and OD) have been utilized to build a picture of the progression of healing over a 16-week in vivo period post-reconstruction. None of the measures provide a comprehensive assessment as to the state of healing in isolation but do allow trends to be delineated. The degree of association between the parameters measured has been shown to be variable ranging from poor to strong. Ultimately, however, it is the response to load bearing that provides the ultimate test of fracture healing in combination with the measures undertaken in this study. As discussed, differentially loaded radiostereometric analysis is one possible means of determining fracture stiffness, in a serial manner, in the clinical setting; this is the subject of an on-going evaluation.

Changes in OD have been evaluated as a clinical tool,<sup>15–20</sup> but its utilization is not widespread. This study lends support to the concept that measures of OD can contribute to our understanding of bone regeneration in fracture healing and warrants consideration for use in the clinical setting.

#### Conflicts of interest

The authors declare that they have no conflicts of interest with regard to this manuscript.

#### Author contributions

J. Field contributed to the conception of the study and the study design. Both authors contributed to acquisition of data, data analysis and interpretation, as well as drafting and revising of the manuscript. Both authors approved the submitted manuscript.

#### References

- 1 McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis Rheum* 2001; 45(05):453–461
- 2 Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res (Hoboken)* 2011;63(Suppl 11): S200–S207
- 3 Field JR, Schnurr D, Horney FD. Evaluation of an instrumented horseshoe system. *Vet Comp Orthop Traumatol* 1992;19:58–64
- 4 Kaler J, Wassink GJ, Green LE. The inter- and intra-observer reliability of a locomotion scoring scale for sheep. *Vet J* 2009; 180(02):189–194
- 5 Augat P, Morgan EF, Lujan TJ, MacGillivray TJ, Cheung WH. Imaging techniques for the assessment of fracture repair. *Injury* 2014;45 (Suppl 2):S16–S22
- 6 Field JR, McGee M, Wildenauer C, Kurmis A, Margerrison E. The utilization of a synthetic bone void filler (JAX) in the repair of a femoral segmental defect. *Vet Comp Orthop Traumatol* 2009;22 (02):87–95
- 7 Field JR, McGee M, Stanley R, et al. The efficacy of allogeneic mesenchymal precursor cells for the repair of an ovine tibial segmental defect. *Vet Comp Orthop Traumatol* 2011;24(02):113–121



- 8 den Boer FC, Patka P, Bakker FC, et al. New segmental long bone defect model in sheep: quantitative analysis of healing with dual energy x-ray absorptiometry. *J Orthop Res* 1999;17(05):654–660
- 9 Reichert JC, Epari DR, Wulschleger ME, et al. Establishment of a preclinical ovine model for tibial segmental bone defect repair by applying bone tissue engineering strategies. *Tissue Eng Part B Rev* 2010;16(01):93–104
- 10 Gao C, Seuntjens J, Kaufman GN, et al. Mesenchymal stem cell transplantation to promote bone healing. *J Orthop Res* 2012;30(08):1183–1189
- 11 Christou C, Oliver RA, Pelletier MH, Walsh WR. Ovine model for critical-size tibial segmental defects. *Comp Med* 2014;64(05):377–385
- 12 Whelan DB, Bhandari M, Stephen D, et al. Development of the radiographic union score for tibial fractures for the assessment of tibial fracture healing after intramedullary fixation. *J Trauma* 2010;68(03):629–632
- 13 Bhandari M, Guyatt GH, Swiontkowski MF, Tornetta P III, Sprague S, Schemitsch EH. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. *J Orthop Trauma* 2002;16(08):562–566
- 14 Whelan DB, Bhandari M, McKee MD, et al. Interobserver and intraobserver variation in the assessment of the healing of tibial fractures after intramedullary fixation. *J Bone Joint Surg Br* 2002;84(01):15–18
- 15 Barnes K, Lanz O, Were K, et al. Comparison of autogenous cancellous bone grafting and extra-corporeal shock wave therapy on osteotomy healing in the tibial tuberosity advancement procedure in dogs. Radiographic densitometric evaluation. *Vet Comp Orthop Traumatol* 2015;28(03):207–214
- 16 Filho LA, Primo BT, Gassen HT, et al. Evaluation of optical density of bone defects filled with calcium phosphate cement and bioactive glass in rats. *Acta Cir Bras* 2011;26(11):1–6
- 17 Souza TF, Andrade AL, Ferreira GT, et al. Healing and expression of growth factors (TGF- $\beta$  and PDGF) in canine radial osteotomy gap containing platelet-rich plasma. *Vet Comp Orthop Traumatol* 2012;25(06):445–452
- 18 Farinazzo CS, Iwaki LC, Filho LI, et al. Optical density evaluation around the zirconia granules used as graft in rat calvaria. *Acta Scientiarum* 2014;36(02):265–272
- 19 Janic J, Mijovic Z, Mihailovic D, et al. Optical density of cortical bone matrix is diminished in experimentally induced osteoporosis. *Acta Medica Medianae* 2016;55(02):35–39
- 20 Pekkan G, Aktas A, Pekkan K. Comparative radiopacity of bone graft materials. *J Craniomaxillofac Surg* 2012;40(01):e1–e4
- 21 Grassl U, Schulze RK. In vitro perception of low-contrast features in digital, film, and digitized dental radiographs: a receiver operating characteristic analysis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2007;103(05):694–701
- 22 Ostlere SJ, Gold RH. Osteoporosis and bone density measurement methods. *Clin Orthop Relat Res* 1991;(271):149–163
- 23 Watts DC, McCabe JF. Aluminium radiopacity standards for dentistry: an international survey. *J Dent* 1999;27(01):73–78
- 24 Lujan TJ, Madey SM, Fitzpatrick DC, Byrd GD, Sanderson JM, Bottlang M. A computational technique to measure fracture callus in radiographs. *J Biomech* 2010;43(04):792–795
- 25 Lujan TJ, Henderson CE, Madey SM, Fitzpatrick DC, Marsh JL, Bottlang M. Locked plating of distal femur fractures leads to inconsistent and asymmetric callus formation. *J Orthop Trauma* 2010;24(03):156–162
- 26 Henderson CE, Lujan TJ, Kuhl LL, Bottlang M, Fitzpatrick DC, Marsh JL. 2010 mid-America Orthopaedic Association Physician in Training Award: healing complications are common after locked plating for distal femur fractures. *Clin Orthop Relat Res* 2011;469(06):1757–1765
- 27 Chehade MJ, Solomon LB, Callary SA, Benveniste SH, Pohl AP, Howie DW. Differentially loaded radiostereometric analysis to monitor fracture stiffness: a feasibility study. *Clin Orthop Relat Res* 2009;467(07):1839–1847
- 28 Solomon LB, Stevenson AW, Callary SA, Sullivan TR, Howie DW, Chehade MJ. The accuracy and precision of radiostereometric analysis in monitoring tibial plateau fractures. *Acta Orthop* 2010;81(04):487–494
- 29 Chehade MJ, Vakaci IA, Callary SA, et al. Differentially loaded radiostereometric analysis (DLRSA) in torsion adds essential information in diaphyseal bone healing: the example of a tibial osteotomy. *J Bioeng Biomed Sci* 2011;1:1–5
- 30 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20(01):37–46