

# MICROSTRUCTURE AND HOMOGENEITY OF DISTRIBUTION OF MINERALISED STRUTS DETERMINE CALLUS STRENGTH

Manav Mehta<sup>1,2,3,§,\*</sup>, Mark Heyland<sup>1,§</sup>, Daniel Toben<sup>4</sup> and Georg N. Duda<sup>1,5</sup>

<sup>1</sup>Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité – Universitätsmedizin Berlin, Campus – Virchow Klinikum, Berlin, Germany

<sup>2</sup>School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

<sup>3</sup>Wyss Institute for Biologically Inspired Engineering at Harvard University, Cambridge, MA, USA

<sup>4</sup>Department of Trauma-, Hand- and Reconstructive Surgery, University Medical Centre Hamburg-Eppendorf (UKE), Hamburg, Germany

<sup>5</sup>Berlin-Brandenburg Center for Regenerative Therapies, Berlin, Germany

§ Manav Mehta and Mark Heyland have contributed equally to this work.

## Abstract

Non-invasive assessment of fracture healing, both in clinical and animal studies, has gained favour as surrogate measure to estimate regain of mechanical function. Micro-computed tomography ( $\mu$ CT) parameters such as fracture callus volume and mineralisation have been used to estimate callus mechanical competence. However, no in-depth information has been reported on microstructural parameters in estimating callus mechanical competence. The goal of this study is to use differently conditioned mice exhibiting good and impaired fracture healing outcomes and investigate the relationship between  $\mu$ CT imaging parameters (volume, mineralisation, and microstructure) that best estimate the callus strength and stiffness as it develops over time. A total of 99 mice with femoral fracture and intramedullary stabilisation were divided into four groups according to conditioning: wild type, NF1 knock-out, RAG1 knock-out and macrophage depleted. Animals were sacrificed at 14, 21, 28 or 35 days and  $\mu$ CT parameters and torsional stiffness and strength were assessed post-sacrifice. Using linear regression for all groups and time points together, torsional stiffness could be estimated with strut thickness, strut number and strut homogeneity ( $R^2 = 0.546$ ,  $p < 0.0001$ ); torsional strength could be estimated using bone mineral density, strut thickness and strut homogeneity ( $R^2 = 0.568$ ,  $p < 0.0001$ ). Differently conditioned mice that result in different fracture healing outcomes have been shown to result in varying structural, material and volumetric  $\mu$ CT parameters which can be used to estimate regain of bone strength. This study is the first to demonstrate that microstructure and strut homogeneity influence callus stiffness and strength.

**Keywords:** Fracture healing; callus; microstructure; homogeneity; heterogeneity.

## \*Address for correspondence:

Manav Mehta  
Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration,  
Charité – Universitätsmedizin Berlin, Campus – Virchow Klinikum,  
Augustenburger Platz 1, Institutsgebäude Süd, 13353 Berlin, Germany

Telephone Number: +49 (0)30-450 55 90 48

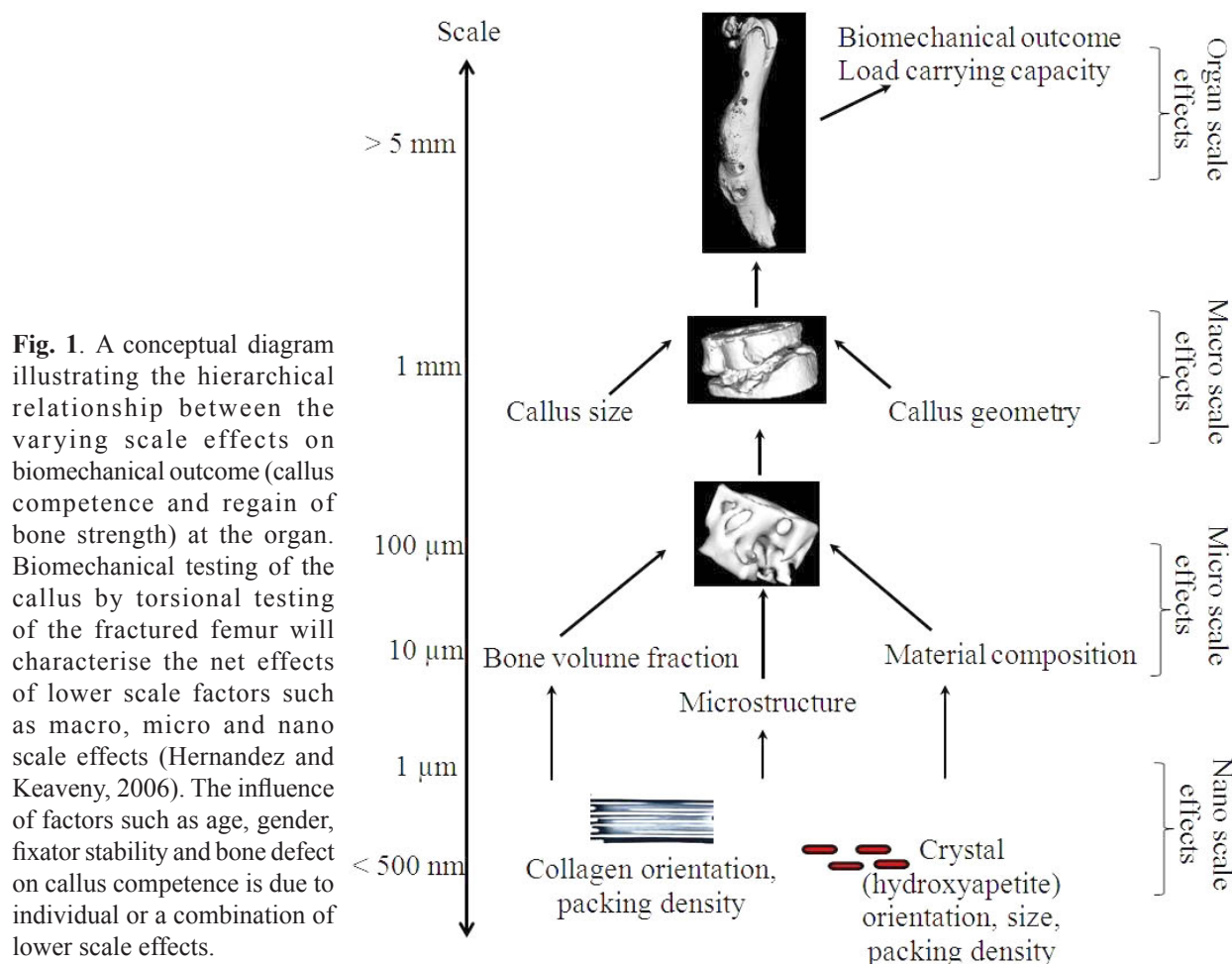
FAX Number: +49 (0)30-450 55 99 69

E-mail: manavben@gmail.com

## Introduction

Fracture stiffness has the potential to provide both a measure of the rate of healing and an objective definition of union. Since healing is the return of function and the chief function of bone is the resistance to stress, it is logical to measure the elastic modulus. Fortunately, the end-point is at a relatively low stiffness compared with the intact bone when measurements are relatively easy and stiffness correlates with strength (Chehade *et al.*, 1997). Strength is defined by failure and as such directly measuring strength is only possible in the laboratory or by using fracture failure as an outcome measure (Wade and Richardson, 2001). The monitoring of stiffness is useful primarily in assessing progress towards union but is inherently limited as an assessment of strength at the time of clinical union (Chehade *et al.*, 1997). Torsional or bending strength, and stiffness have been used as gold standard parameters in assessing fracture healing outcome in small and large animal models (Schell *et al.*, 2005; Claes *et al.*, 2009; Goodship *et al.*, 2009; Mehta *et al.*, 2011a; O'Neill *et al.*, 2012). However, biomechanical assessments require destructive testing of bone, making *post mortem* analysis using other techniques on a structurally intact sample difficult. Furthermore, current biomechanical assessments involve testing of the fracture callus tissue at an organ level, giving no information on the influence of lower scale effects (Fig. 1), which might give earlier and causal insights into progression or impairment of fracture healing. Recent bone healing studies have shown the influence of age, gender, fixation stability, and bone defect size on callus microstructure (Mehta *et al.*, 2010; Mehta *et al.*, 2011b). However, an in-depth analysis of the effects of microstructure and its relationship to the strength and stiffness of callus have not been reported. Furthermore, the evolution of the callus structure and its relationship to callus strength and stiffness is also unknown. With the advancing super high resolution X-ray imaging technologies that will see clinical installations in the future, the knowledge on structure-function relationships over the time course of healing will provide as a valuable tool for diagnosis of fracture healing outcome in animal and clinical studies.

Micro-Computed tomography ( $\mu$ CT) imaging, a non-destructive technique for assessment of healing outcome, has been gaining favour (Kalpakcioglu *et al.*, 2008; Brandi, 2009; Bouxsein *et al.*, 2010). It has been shown that volumetric, geometrical, and compositional



parameters derived from  $\mu$ CT imaging provide as a surrogate measure of mechanical competence of a fracture callus (Reynolds *et al.*, 2007; Morgan *et al.*, 2009). However, it is unclear how volumetric, geometrical, compositional parameters – specifically mineral density, callus volume and polar moment of inertia, influence the predictability of mechanical competence of a fracture callus. Furthermore, to what degree microstructural parameters (strut number, thickness, and separation and microstructural homogeneity) relate to callus strength and stiffness remains unknown. The goal of this study is to use differently conditioned mice exhibiting good and impaired fracture callus at different time points in the time course of healing, and relate the  $\mu$ CT imaging parameters (material, volume, and microstructure) that best estimate the torsional strength and stiffness of a developing mineralised callus tissue. Due to limitations in resolution of the  $\mu$ CT, nano-scale effects are not reported in this study.

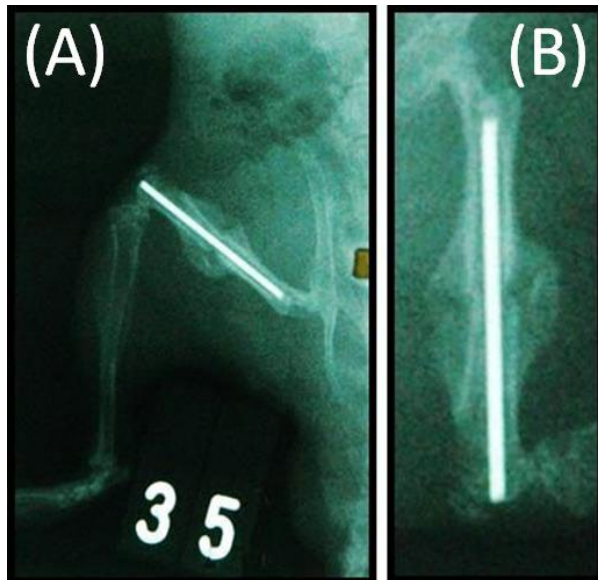
### Materials and Methods

Animal experiments were carried out in accordance with the institution guidelines and the national welfare guidelines. Animal testing was approved by local legal authorities (Landesamt für Gesundheit und Soziales, Berlin, Germany). A total of  $n = 99$  mice were available and all were analysed, grouped according to wild type,

C57Bl6/N (WT) with a total animal number of  $n = 40$ , a conditional NF1 gene knock-out,  $Nf1^{\text{flox/flox}} \text{Prx1}^{\text{Cre}}$  (NF1) with  $n = 14$ , recombination activating gene 1 knockout (RAG1 $^{-/-}$ ) mice lacking mature B and T lymphocytes (RAG) with  $n = 24$  and a macrophage depletion (MAC) with  $n = 21$ .

A standard transverse femoral fracture was created in 8-10 week old mice with a special fracture device (Bonnarens and Einhorn, 1984; Toben *et al.*, 2011). The murine femora were stabilised with an intramedullary nail (Fig. 2). Animals were sacrificed at 14, 21, 28 or 35 days and femora were harvested (Table 1). Fixation was removed and samples were scanned in tubes with isotonic solution aligned along the diaphyseal axis *ex vivo* at an isotropic voxel resolution of  $10.5 \mu\text{m}$  (70 kVp, 114 mA; vivaCT 40, Scanco Medical, Brüttisellen, Switzerland). The scan included the fracture callus in all dimensions. Cortical bone was manually excluded from the volume of interest (VOI, Fig. 3) (Morgan *et al.*, 2009; Mehta *et al.*, 2010). Exclusion of the cortical remnants avoids assessment of influences of cortical bone among the different mice, including transgenic. Therefore, biomechanical testing could be clearly attributed to the newly formed callus area for all groups. A fixed global threshold corresponding to  $190 \text{ mg hydroxyapatite (HA)/cm}^3$  was selected through manual inspection of tomographic slices isolating the mineralised tissue and preserving its morphology while excluding unmineralised tissues. Frozen storing of the

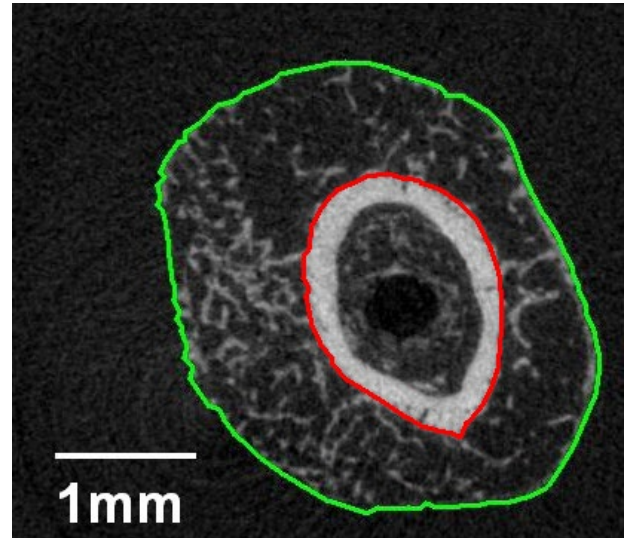
femora wrapped in saline-soaked gauze was followed by embedding of both proximal and distal ends of the femora centrally into device pots using polymethyl methacrylate leaving the diaphysis untouched for testing (Fig. 4). The free diaphysis length was set using a temporally applied metal fixator attached to the device pots to preserve bone integrity during embedding and handling. The different bone sizes required two groups of temporal external stabilisation length of 6 mm for NF1 and MAC and of 8 mm for NF1 and RAG. Biomechanical testing of the healing callus without the temporal stabilisation fixator was performed through whole bone torsional testing by application of a constant axial preload force of 0.3 N and a following torsional deflection of 0.5 %/s until failure (BOSE, ElectroForce 3200 TestBench Instruments, Eden Prairie, MN, USA). The following parameters were computed from  $\mu$ CT scans using scanner software (Scanco Medical microCT systems software suite standard Evaluation Software v6.0, Scanco Medical): total callus volume (TV), mineralised callus volume (BV), callus mineralised volume fraction (BV/TV), bone mineral density (BMD), tissue mineral density (TMD), degree of anisotropy (DA), connectivity density (Conn.D), structure modelling index (SMI), strut number (Tb.N), strut separation (Tb.Sp), strut thickness (Tb.Th), standard deviations of strut parameters (Tb.Sp.SD, Tb.Th.SD). BV and TV were normalised to callus length. The callus length was determined by the length between the most distal and proximal apparent accretions to the cortical bone. The following parameters were calculated from mechanical testing: torsional stiffness (rigidity) and strength (MTF, maximum torque at failure). Torsional stiffness was normalised to the length of the tested diaphysis.



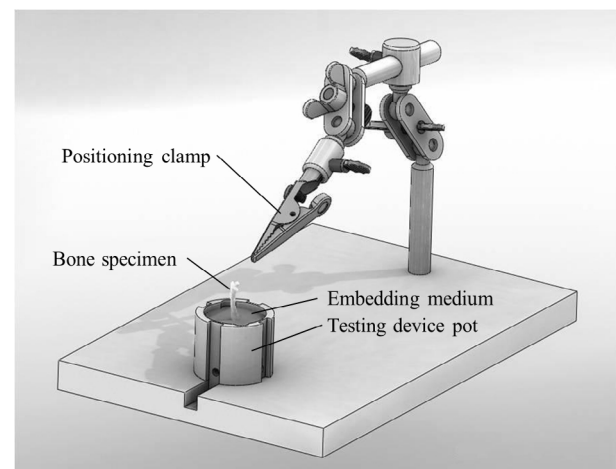
**Fig. 2.** Exemplary x-ray image of the stabilisation of an animal from the wild-type group at time point 35 days post fracture. The image on the left (A) shows the overview of the mouse's hind limbs with the metal intramedullary stabilisation rod within the femur. The image on the right (B) shows the same stabilisation and the healing callus tissue in more detail, turned and enlarged.

**Table 1:** Investigated groups and sample sizes for each time point of harvest.

Animal groups	Day 14	Day 21	Day 28	Day 35
Wildtype (WT)	8	16	8	8
Nf1 <sup>fllox/fllox</sup> Prx1 <sup>Cre</sup> (NF1)	5	9	-	-
RAG1 <sup>-/-</sup> (RAG)	8	8	8	-
Macrophage depletion (MAC)	8	8	5	-

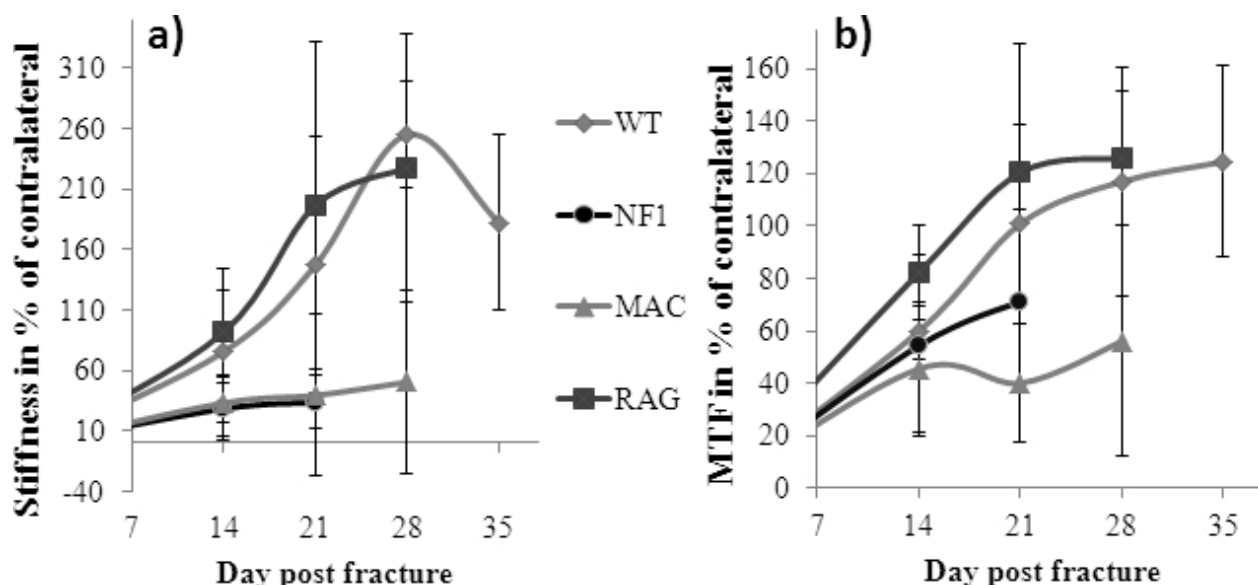


**Fig. 3.** Image segmentation was done in a standard fashion (Morgan *et al.*, 2009; Mehta *et al.*, 2010) semi-automated using scanner software (Scanco Medical microCT systems software suite standard Evaluation Software v6.0, Scanco Medical) by defining the outer boundary of the callus (green) and the periosteal surface of the cortex (red) on each 2D tomogram. The volume of interest (VOI) is the region enclosed by these two boundaries. The total volume (TV) includes all voxels that are within the included VOI, bone volume (BV) would be all of these voxels that are above the threshold.



**Fig. 4.** Illustration of a potting device to precisely place the mural femur bone in the device pots with temporary specimen mount for potting. The testing device pots on each end of the bone (only one shown here) were used as attachments to the BOSE mechanical test setup. The diaphysis of the bone remained untouched.





**Fig. 5.** Biomechanical result of fracture healing for the different groups (abbreviations in text) relative to the intact contralateral side displaying (a) Stiffness on the left and (b) maximum torque at failure (MTF, torsional strength) on the right (mean  $\pm$  standard deviation).

Pearson coefficients of correlation for evaluation of the influence of the  $\mu$ CT parameters on the mechanical outcome have been calculated for each group and time point and altogether for the whole population. Further stepwise linear regression analyses allowing models with up to three parameters for all time points and groups together, for each single group and for each single time point (14/21/28/35 days) with all groups were performed to find dependencies of mechanical and  $\mu$ CT measures. Group offset was then incorporated as a fixed factor for the regression of all groups and time points together in order to account for unconsidered parameters. Calculations were performed using SPSS (Chicago, IL, USA) 17 statistics software.

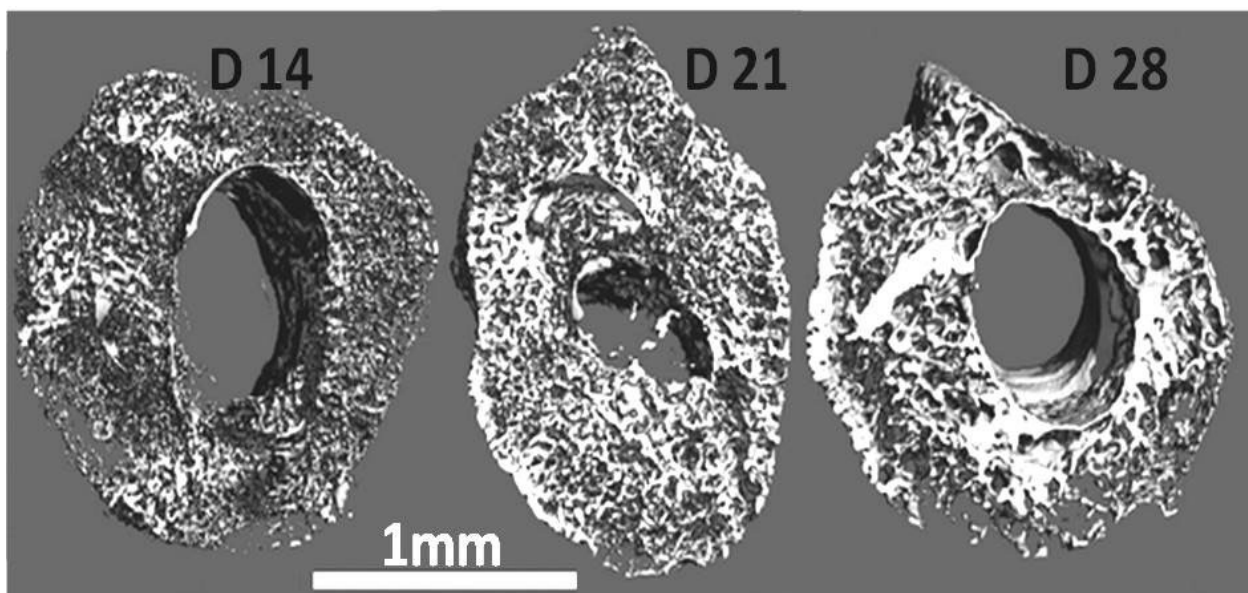
## Results

The study groups exhibit diverse healing outcomes reflected by varying mean stiffness (Fig. 5a) and strength (Fig. 5b) over the time course of healing and changing  $\mu$ CT imaging parameters (Table 4a-c). Determined by Mann-Whitney-U-test ( $p < 0.05$ ), at day 14, WT mice have a significantly higher strength (mean  $\pm$  standard deviation:  $14.33 \pm 3.25$  Nmm) and stiffness ( $8.25 \pm 5.66$  Nmm mm/deg) than MAC ( $8.33 \pm 2.31$  Nmm and  $3.64 \pm 2.17$  Nmm mm/deg) and a significantly lower strength and stiffness than RAG ( $18.71 \pm 3.45$  Nmm and  $30.84 \pm 16.96$  Nmm mm/deg). At day 14, RAG also has a higher MTF and stiffness than NF1 ( $7.94 \pm 6.60$  Nmm and  $1.39 \pm 1.39$  Nmm mm/deg). Stiffness of WT at day 14 is also higher than in the NF1 group. At day 21 all groups except MAC and NF1 differ in stiffness (RAG:  $52.37 \pm 25.79$  Nmm mm/deg, WT:  $23.41 \pm 14.28$  Nmm mm/deg, NF1:  $1.81 \pm 0.87$  Nmm mm/deg, MAC:  $3.69 \pm 2.47$  Nmm mm/deg) and MTF (RAG:  $30.76 \pm 5.80$  Nmm, WT:  $22.22 \pm 6.44$  Nmm, NF1:  $9.32 \pm 3.79$  Nmm, MAC:  $7.44 \pm 1.82$  Nmm).

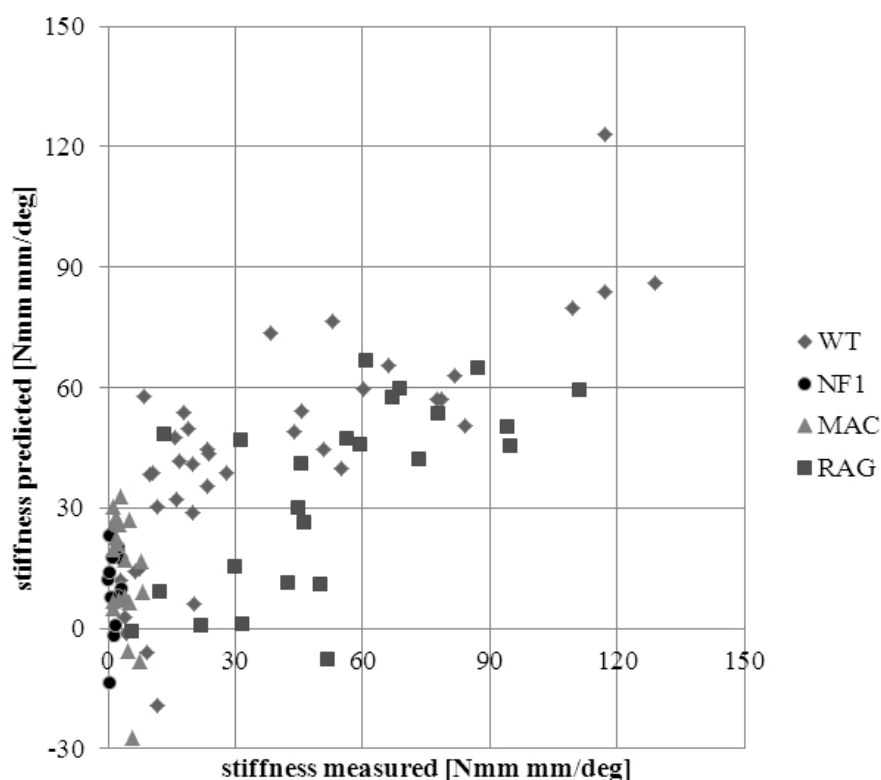
At day 28, MAC shows significantly lower strength ( $9.38 \pm 4.52$  Nmm) and stiffness ( $3.22 \pm 2.45$  Nmm mm/deg) than WT ( $30.98 \pm 8.81$  Nmm and  $79.01 \pm 48.22$  Nmm mm/deg) and RAG ( $33.29 \pm 5.93$  Nmm and  $76.81 \pm 19.51$  Nmm mm/deg).

Significant Pearson correlations ( $p < 0.05$ ) showing the relationship of  $\mu$ CT measures and mechanical outcome were observed (Table 2a, b). Specifically, the WT group showed correlations between volumetric parameters (TV, BV) and mechanical outcome. The RAG group showed strong correlations of BV/TV and BMD with the mechanical outcome. Additionally, WT, RAG and NF1 show correlations of Tb.Th and mechanical outcome.

Correlations of individual  $\mu$ CT parameters with biomechanical outcome show an evolution over time (Table 2a, b). Interestingly, it is observed that different parameters dominate at different time points of the healing course. There is an increase in the number of struts (Tb.N) between day 14-21, followed by an increase in their thickness (Tb.Th) between day 21-28, thereafter an increase in mineralisation of the struts is observed at day 28 and onwards (Figure 6). The change in callus microstructure over different time points in healing impacts mechanical competence, as illustrated by linear regressions of  $\mu$ CT parameters with stiffness and MTF (Table 3). Good predictors for stiffness and MTF determined by stepwise linear regression vary for parameters over time. Early on, callus stiffness was influenced by volume (BV) between day 14 and 21, and subsequently by microstructure (Tb.Sp, SD, and others) between day 21 and 28 (Table 3). For MTF, which is an indicator of strength, microstructure had an impact from early on. Strut number (Tb.N) between day 14 and 21, Tb.Sp, SD between day 21 and 28 were observed to be good predictors. Both, MTF and stiffness at later time points can be predicted from  $\mu$ CT parameters with higher coefficients of determination, i.e. explaining more of the variability (Table 3).



**Fig. 6.** Evolution of microstructure over time in WT (different animals) from day 14 (D 14) over day 21 (D 21) to day 28 (D 28).



**Fig. 7.** Result of linear regression (all time points) for the dependent variable torsional stiffness (using Tb.Sp, SD, Tb.Th, Tb.N, all predictors  $p < 0.0001$ ,  $R^2 = 0.546$ ,  $p < 0.0001$ ).

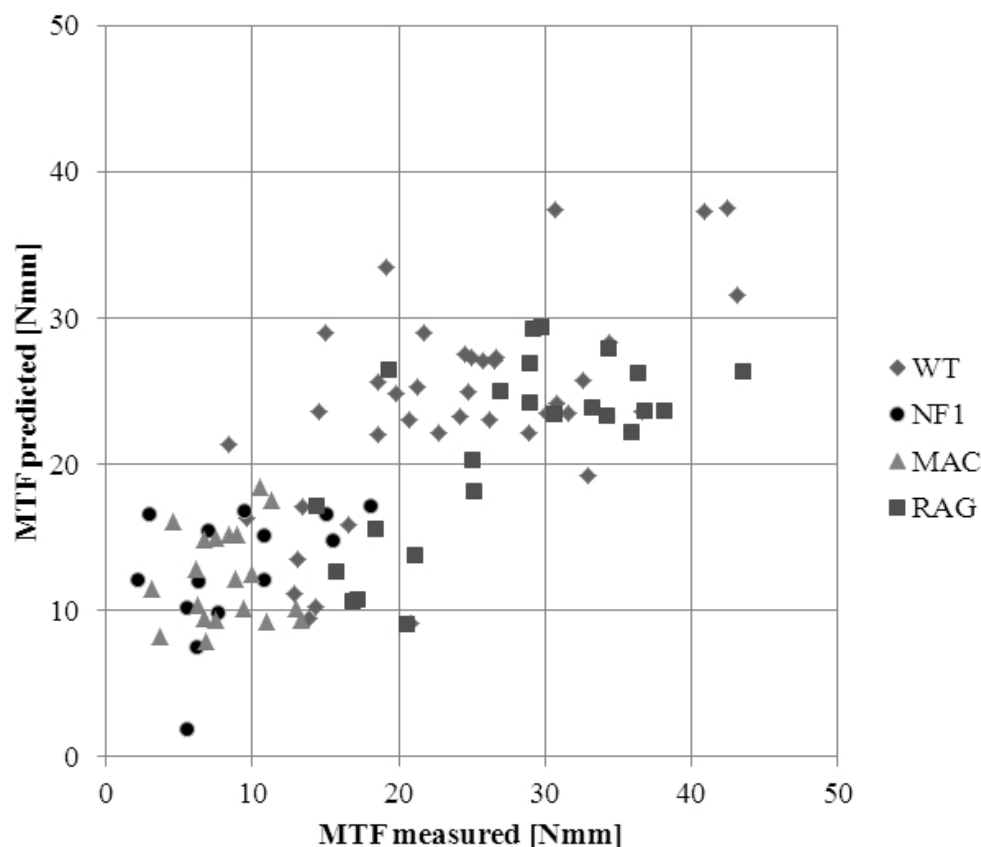
Using a linear regression for all time points together, stiffness can be predicted for this basic population with  $\mu$ CT parameters with a good coefficient of determination ( $R^2$ ) value (Fig. 7). Torsional stiffness could be predicted ( $R^2 = 0.665$ ,  $p < 0.0001$ ) using: Tb.Sp, SD, Tb.Th, Tb.N and group (all predictors  $p < 0.05$ ); same prediction model without group ( $R^2 = 0.546$ ,  $p < 0.0001$ ). A linear model with all  $\mu$ CT measures (without group) lead to  $R^2 = 0.623$ ,  $p < 0.0001$ , only TV and Tb.N with  $p < 0.05$ .

The prediction of strength leads to superior  $R^2$  values (Fig. 8). Torsional strength could be predicted ( $R^2 = 0.793$  or  $0.794$ ,  $p < 0.0001$ ) using: BMD or BV/TV, Tb.Th, Tb.Sp, SD and group (for Tb.Th and group  $p < 0.05$ ), without

group ( $R^2 = 0.568$  or  $0.538$ ,  $p < 0.0001$ ). A linear model with all  $\mu$ CT measures (without group) lead to  $R^2 = 0.687$ ,  $p < 0.0001$ , only BV/TV and BMD, Tb.N and Tb.Sp, SD with  $p < 0.05$ . Results of linear regression for individual time points differ from the whole set and from each other (Table 3) except for day 28 and MTF.

## Discussion

This is the first study to highlight changes of microstructure over time course of healing and its relationship to mechanical competence of a fracture callus. The study



**Fig. 8.** Result of linear regression (all time points) for the dependent variable torsional strength (MTF) (using BMD, Tb.Th, Tb.Sp, SD, all predictors  $p < 0.0001$ ,  $R^2 = 0.568$ ,  $p < 0.0001$ ).

results have demonstrated (Table 4a-b, see development of Tb.N, Tb.Th, and TMD over time) that newly mineralised struts appear to populate a callus region by day 14, followed by thickening of the structures by day 21, and subsequently an increase in its mineralisation density. Only one previous study has highlighted the mineralisation wave front kinetics to occur in a 2-step process (Liu *et al.*, 2010). The landmark study focused on spatial and temporal mineralisation of a 2D callus, and not callus strut microstructures. In coherence, this study furthermore illustrates the microstructural changes and its relationship to mechanical competence of a healing fracture callus.

This study could show that although there is no single universal predictor of bone stiffness and strength for all time points of healing, stiffness and strength could be estimated in the samples analysed in the present study with the value of microstructural parameters (Tb.N, Tb.Th) and a measure for the heterogeneity of the distribution of struts (Tb.Sp, SD). This held true for the population groups of mice analysed within the present study that exhibit a broad spectrum of fracture healing outcomes. A coefficient of determination of  $R^2 = 0.546$  (Fig. 7) indicates that up to 54.6 % of the variability of stiffness and with  $R^2 = 0.568$  (Fig. 8) up to 56.8 % of the variability of strength are explained by the predictive parameters of these models, callus microstructure and its heterogeneity.

Variations in callus microstructure (Tb.Sp, SD, Tb.N, Tb.Th) estimated most of the variability in the mechanical outcome. Further differences between the groups used in this study could only explain up to 11.9 % more of the variability in stiffness and up to 25.6 % more of the variability in strength as seen in the coefficients

of determination with the group factor included in the otherwise same models for stiffness  $R^2 = 0.665$  and strength  $R^2 = 0.794$ . The inter-group variability could only be partially explained with the further  $\mu$ CT measures: 7.7 % of group differences out of 11.9 % variability in stiffness could be explained with a model including all  $\mu$ CT measures ( $R^2 = 0.623$ ) and up to 14.9 % of group differences out of 25.6 % of the variability in strength could be explained with a model including all  $\mu$ CT measures ( $R^2 = 0.687$ ). Torsional testing is very sensitive and especially within this magnitude of bone sizes inherently prone to comparably large variability. This is undermined by the fact that further  $\mu$ CT parameters as well as a group offset could only explain a small fraction of the variability that could not be explained by microstructure or its heterogeneity. Group was incorporated into the linear regression as a fixed factor in order to account for unknown and thus unconsidered parameters that vary among groups. Notably, there is only a small variance between groups that could not be explained by the  $\mu$ CT parameters used in this study. This small group dependent offset might be caused by differences in the detailed local quality of microstructure (Green *et al.*, 2011) or nanostructure such as a higher flaw rate associated with group. Although there is an effect of nanoscale variability (Burket *et al.*, 2011; Ishimoto *et al.*, 2013), this could not be considered in the  $\mu$ CT imaging approach. Another cause of unexplained variability in strength and stiffness may originate from an inadequate degree of callus bridging or non-uniform location of bridges, especially at early time points. In this study, bony bridging was regarded as a prerequisite, evaluating only later time points of healing after day 14 already showing bony bridging. Although

**Table 2a:** Pearson correlations of different  $\mu$ CT measures.

$\mu$ CT measure	time point	WT		RAG		NF1		MAC		ALL	
		MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.
TV	14	<b>0.92</b>	0.48	-0.23	-0.48	-0.38	-0.53	0.70	-0.11	<b>0.67</b>	<b>0.45</b>
	21	0.11	-0.25	0.15	0.38	0.43	0.26	0.39	0.29	<b>0.52</b>	<b>0.36</b>
	28	0.66	<b>0.86</b>	0.25	0.13			0.79	0.39	0.35	0.42
	35	0.38	<b>0.88</b>								
	ALL	0.06	0.18	<b>-0.60</b>	<b>-0.52</b>	-0.06	-0.21	<b>0.66</b>	0.20	<b>0.28</b>	<b>0.26</b>
BV	14	0.34	<b>0.75</b>	0.29	-0.35	0.71	0.58	0.55	-0.10	<b>0.48</b>	0.28
	21	0.42	-0.28	0.70	0.62	<b>0.69</b>	0.66	0.32	0.41	<b>0.58</b>	<b>0.38</b>
	28	<b>0.76</b>	<b>0.75</b>	0.63	0.45			0.83	0.44	0.19	0.19
	35	0.50	<b>0.81</b>								
	ALL	<b>0.33</b>	0.23	0.20	0.07	0.34	0.27	<b>0.67</b>	0.24	<b>0.38</b>	<b>0.29</b>
BV/TV	14	-0.23	0.44	<b>0.91</b>	0.34	0.83	<b>0.90</b>	-0.46	0.17	<b>-0.39</b>	-0.36
	21	<b>0.57</b>	-0.23	<b>0.80</b>	0.41	0.01	0.30	-0.53	0.02	-0.17	-0.16
	28	0.36	-0.25	0.33	0.32			-0.85	-0.52	<b>-0.49</b>	<b>-0.50</b>
	35	0.06	-0.58								
	ALL	<b>0.46</b>	0.08	<b>0.85</b>	<b>0.66</b>	0.46	<b>0.58</b>	-0.40	-0.08	-0.10	-0.12
Tb.N	14	-0.51	0.28	0.63	0.60	0.80	<b>0.87</b>	-0.63	0.58	-0.08	-0.07
	21	0.18	0.12	0.58	0.17	-0.01	0.33	-0.51	0.16	0.07	0.12
	28	-0.54	<b>-0.91</b>	0.40	0.30			-0.43	0.08	-0.23	-0.33
	35	-0.31	-0.64								
	ALL	-0.12	<b>-0.33</b>	<b>0.65</b>	<b>0.45</b>	0.39	<b>0.56</b>	<b>-0.43</b>	0.24	-0.03	-0.10
Tb.Th	14	0.15	0.66	0.25	-0.47	0.71	0.81	0.51	0.06	-0.13	-0.34
	21	<b>0.58</b>	-0.23	<b>0.78</b>	0.60	0.57	0.52	-0.49	-0.33	<b>0.53</b>	0.31
	28	<b>0.73</b>	<b>0.77</b>	<b>0.78</b>	<b>0.77</b>			0.38	-0.18	0.35	<b>0.46</b>
	35	0.31	0.65								
	ALL	<b>0.72</b>	<b>0.68</b>	<b>0.90</b>	<b>0.72</b>	<b>0.57</b>	<b>0.57</b>	0.27	-0.11	<b>0.57</b>	<b>0.58</b>
Tb.Sp	14	<b>0.74</b>	0.01	-0.42	-0.62	-0.67	-0.77	<b>0.71</b>	-0.42	0.13	-0.14
	21	-0.11	-0.21	-0.58	-0.12	-0.06	-0.36	0.29	-0.31	<b>-0.38</b>	<b>-0.36</b>
	28	0.61	<b>0.90</b>	-0.35	-0.25			0.31	-0.29	-0.21	-0.05
	35	0.15	0.49								
	ALL	0.02	0.14	<b>-0.66</b>	<b>-0.53</b>	-0.30	-0.48	0.37	-0.31	<b>-0.22</b>	-0.14
Tb.Th.SD	14	0.30	<b>0.84</b>	0.16	-0.42	0.57	0.66	0.41	0.17	0.17	-0.09
	21	<b>0.51</b>	-0.25	0.64	0.49	0.47	0.40	0.04	0.10	<b>0.58</b>	<b>0.35</b>
	28	0.53	0.64	<b>0.74</b>	<b>0.72</b>			0.58	0.07	0.22	0.33
	35	0.40	<b>0.75</b>								
	ALL	<b>0.62</b>	<b>0.62</b>	<b>0.84</b>	<b>0.62</b>	0.43	0.42	0.40	0.01	<b>0.56</b>	<b>0.56</b>

Abbreviations are given in the text, Stiffness (Stiffn.) and MTF for different time points and groups and altogether. Significant correlations are marked bold ( $p < 0.05$ ), but are given only as orientation, because the test is only fully valid for normally distributed data.

there are semi-quantitative methods to characterise callus bridging, they remain controversial. Therefore, callus bridging as means of estimating mechanical competence was not investigated in depth here. Measurements with  $\mu$ CT and their assessment also produce errors. Both image resolution, as well as the threshold used to segment the bone from soft tissue, influence the accuracy of parameter identification (Burghardt *et al.*, 2007). The minimum ratio

of voxels to object size should be 2, but this is associated with substantial local errors, which become smaller when averaged over the entire structure (Bouxsein *et al.*, 2010). This was met in our approach, although the resolution of 10.5  $\mu$ m is a general limitation of the presented method, this resolution is the current image resolution of *in vivo* scanners and has been used before to assess murine bone microstructure (Schulte *et al.*, 2011). The standard



**Table 2b:** Pearson correlations of different  $\mu$ CT measures.

$\mu$ CT measure	time point	WT		RAG		NF1		MAC		ALL	
		MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.	MTF	Stiffn.
Tb.Sp.SD	14	0.67	-0.03	-0.47	-0.54	-0.72	-0.80	0.70	-0.48	0.18	-0.06
	21	0.08	-0.27	-0.45	0.09	-0.09	-0.43	0.34	-0.33	<b>-0.52</b>	<b>-0.52</b>
	28	0.66	<b>0.90</b>	-0.34	-0.24			0.46	-0.19	<b>-0.48</b>	-0.38
	35	0.26	0.48								
	ALL	-0.28	-0.22	<b>-0.72</b>	<b>-0.57</b>	-0.31	-0.49	0.43	-0.29	<b>-0.43</b>	<b>-0.38</b>
BMD	14	-0.47	0.24	<b>0.96</b>	0.39	0.79	<b>0.88</b>	-0.46	0.38	<b>-0.44</b>	<b>-0.38</b>
	21	<b>0.53</b>	-0.29	<b>0.82</b>	0.42	0.00	0.22	-0.46	0.13	-0.26	-0.25
	28	0.43	-0.15	0.41	0.39			<b>-0.92</b>	-0.60	<b>-0.55</b>	<b>-0.52</b>
	35	-0.04	-0.63								
	ALL	<b>0.39</b>	0.02	<b>0.84</b>	<b>0.67</b>	0.46	<b>0.56</b>	-0.40	-0.04	-0.19	-0.19
TMD	14	<b>-0.78</b>	-0.15	0.22	0.20	0.20	0.35	0.27	0.67	-0.31	-0.16
	21	0.24	0.27	-0.02	-0.10	0.21	0.14	-0.05	0.00	<b>0.43</b>	<b>0.38</b>
	28	-0.49	-0.23	0.07	0.19			-0.59	-0.48	-0.15	-0.02
	35	-0.30	0.13								
	ALL	<b>0.64</b>	<b>0.65</b>	<b>0.77</b>	<b>0.66</b>	0.18	0.20	-0.04	-0.06	<b>0.53</b>	<b>0.56</b>
Conn.D.	14	0.07	0.57	<b>0.80</b>	0.36	-0.15	-0.29	<b>-0.85</b>	-0.04	0.07	0.12
	21	0.26	-0.24	0.58	0.22	-0.15	-0.04	0.21	0.46	0.21	0.09
	28	-0.56	<b>-0.91</b>	0.54	0.42			-0.41	0.31	<b>-0.55</b>	<b>-0.60</b>
	35	0.03	-0.57								
	ALL	<b>-0.55</b>	<b>-0.66</b>	-0.14	-0.28	-0.16	-0.21	-0.36	0.22	<b>-0.39</b>	<b>-0.50</b>
DA	14	<b>-0.84</b>	-0.57	0.18	0.55	0.16	-0.03	-0.22	-0.46	-0.22	0.00
	21	0.27	0.24	-0.42	-0.65	-0.62	<b>-0.72</b>	-0.66	<b>-0.79</b>	0.30	0.17
	28	-0.48	-0.48	-0.41	-0.19			-0.67	-0.45	-0.03	-0.11
	35	0.14	-0.61								
	ALL	0.08	0.03	<b>0.51</b>	<b>0.52</b>	-0.25	-0.38	<b>-0.45</b>	<b>-0.58</b>	<b>0.20</b>	0.15
SMI	14	-0.40	-0.54	-0.53	0.17	0.05	0.21	-0.04	0.70	<b>0.59</b>	<b>0.50</b>
	21	-0.32	0.43	-0.68	-0.33	0.42	0.12	0.49	0.18	<b>0.45</b>	<b>0.38</b>
	28	0.12	0.50	0.08	0.07			-0.45	0.05	<b>0.53</b>	<b>0.61</b>
	35	-0.31	0.44								
	ALL	<b>-0.34</b>	-0.01	<b>-0.72</b>	<b>-0.54</b>	0.23	0.10	0.01	0.28	<b>0.34</b>	<b>0.32</b>

Abbreviations are given in the text, Stiffness (Stiffn.) and MTF for different time points and groups and altogether. Significant correlations are marked bold ( $p < 0.05$ ), but are given only as orientation, because the test is only fully valid for normally distributed data.

deviation trabecular dimension parameters, which quantify the degree of structural heterogeneity and can be measured with a clinical *in vivo* protocol, have previously been shown to be potential predictors of fracture risk (Laib *et al.*, 2002) and have been shown to be correlated to direct high-resolution  $\mu$ CT values (Burghardt *et al.*, 2007).

Pearson correlations of different  $\mu$ CT measures with stiffness and strength for different time points and groups indicate multifaceted associations (Table 2a,b). As observed in Table 2a, at early time points in WT subjects, higher TV is strongly associated with higher strength, and higher BV with higher stiffness. At later time points in WT, both higher TV and BV are associated with higher stiffness, and higher

BV is associated with higher strength. In our accelerated healing model (RAG), higher TV was associated with lower strength and stiffness while in the MAC group, TV had the opposite effect on strength. Depending on group, TV does not always correspond to better healing. Also, a higher strut number is only associated with a higher strength and stiffness, if the struts are aligned, connected and thick as seen for instance in WT day 28 where strut number and thickness compete in correlations. Strut number has to be evaluated together with strut thickness, connectivity density and anisotropy. The dependencies of the different parameters on each other, the high degree of multi-collinearity, make it difficult to find a small set



**Table 3:** Linear regression analyses for MTF and stiffness with  $\mu$ CT parameters for different time points.

Time point	Predictive variables		R <sup>2</sup>	
	MTF	Stiffness	MTF	Stiffness
Day 14	TV, Tb.N	SMI, BV, DA	.621	.424
	TV, BV, BMD, TMD, BV/TV	TV, BV, BMD, TMD, BV/TV	.572	.291
	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	.381	.312
Good	volume	volume		
Day 21	Tb.Th.SD, Tb.Sp.SD, Tb.N	Tb.Sp.SD, BV, Tb.N	.629	.460
	TV, BV, BMD, TMD, BV/TV	TV, BV, BMD, TMD, BV/TV	.690	.476
	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	.656	.423
Good	volume, microstructure	volume, microstructure		
Day 28	BMD, Tb.Sp.SD, Tb.Th	SMI, Tb.Th, Tb.Sp.SD	.729	.702
	TV, BV, BMD, TMD, BV/TV	TV, BV, BMD, TMD, BV/TV	.614	.567
	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	.805	.805
Good	microstructure, density	microstructure		
Day 35	BV, SMI, BV/TV	TV, TMD	.945	.876
	TV, BV, BMD, TMD, BV/TV	TV, BV, BMD, TMD, BV/TV	.864	.934
	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	.584	.837
Good	volume, density	volume, density		
ALL	BMD or BV/TV, Tb.Th, Tb.Sp.SD	Tb.Sp.SD, Tb.Th, Tb.N	0.568 or 0.538	.546
	TV, BV, BMD, TMD, BV/TV	TV, BV, BMD, TMD, BV/TV	.619	.544
	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	Tb.N, Tb.Th, Tb.Sp, Tb.Th.SD, Tb.Sp.SD	.545	.549
Good	density, volume, microstructure	density, volume, microstructure		

**Table 4a:** Average value  $\pm$  standard deviations of  $\mu$ CT parameters for different time points and groups. Explanation of abbreviations are given in the text.

Group	Day	TV/I [mm <sup>2</sup> ]	BV/I [mm <sup>2</sup> ]	BV/TV	BMD [mgHA/cm <sup>3</sup> ]	TMD [mgHA/cm <sup>3</sup> ]
WT	14	5.89 $\pm$ 1.23	1.31 $\pm$ 0.37	0.23 $\pm$ 0.06	278.3 $\pm$ 44.9	762.8 $\pm$ 24.1
	21	4.56 $\pm$ 1.89	1.38 $\pm$ 0.68	0.31 $\pm$ 0.07	327.3 $\pm$ 61.7	886.9 $\pm$ 22.2
	28	4.22 $\pm$ 1.35	1.30 $\pm$ 0.44	0.31 $\pm$ 0.04	331.6 $\pm$ 27.7	953.5 $\pm$ 17.1
	35	4.84 $\pm$ 3.23	1.38 $\pm$ 0.76	0.32 $\pm$ 0.09	338.6 $\pm$ 78.0	984.9 $\pm$ 34.7
NF1	14	3.07 $\pm$ 0.99	1.22 $\pm$ 0.14	0.42 $\pm$ 0.11	429.0 $\pm$ 90.8	813.9 $\pm$ 59.3
	21	2.27 $\pm$ 0.77	0.85 $\pm$ 0.21	0.39 $\pm$ 0.07	408.5 $\pm$ 48.5	801.3 $\pm$ 58.1
MAC	14	3.70 $\pm$ 0.93	1.25 $\pm$ 0.24	0.35 $\pm$ 0.05	379.2 $\pm$ 45.9	799.0 $\pm$ 22.1
	21	2.64 $\pm$ 1.26	1.07 $\pm$ 0.38	0.44 $\pm$ 0.12	457.2 $\pm$ 97.0	880.7 $\pm$ 27.2
	28	3.36 $\pm$ 2.12	1.48 $\pm$ 0.65	0.48 $\pm$ 0.08	510.8 $\pm$ 83.3	946.7 $\pm$ 50.9
RAG	14	6.77 $\pm$ 1.53	1.51 $\pm$ 0.33	0.23 $\pm$ 0.04	275.3 $\pm$ 26.2	776.8 $\pm$ 19.4
	21	4.39 $\pm$ 1.17	1.56 $\pm$ 0.56	0.36 $\pm$ 0.09	361.9 $\pm$ 79.3	890.3 $\pm$ 24.4
	28	3.29 $\pm$ 0.90	1.26 $\pm$ 0.28	0.39 $\pm$ 0.06	400.8 $\pm$ 55.2	944.2 $\pm$ 12.9
ALL	14	5.04 $\pm$ 1.90	1.33 $\pm$ 0.30	0.29 $\pm$ 0.10	331.3 $\pm$ 80.3	785.5 $\pm$ 35.0
	21	3.65 $\pm$ 1.75	1.24 $\pm$ 0.58	0.36 $\pm$ 0.10	377.3 $\pm$ 84.8	867.6 $\pm$ 48.7
	28	3.66 $\pm$ 1.42	1.33 $\pm$ 0.43	0.38 $\pm$ 0.09	400.6 $\pm$ 87.5	948.4 $\pm$ 26.4
	35	4.84 $\pm$ 3.23	1.38 $\pm$ 0.76	0.32 $\pm$ 0.09	338.6 $\pm$ 78.0	984.9 $\pm$ 34.7
	ALL	4.15 $\pm$ 1.96	1.30 $\pm$ 0.49	0.34 $\pm$ 0.10	365.6 $\pm$ 86.6	870.1 $\pm$ 77.8

of parameters that completely describe callus stiffness and strength. A high heterogeneity of strut thickness represents an intact bone mineralisation and remodelling process adapted to load (Glatt *et al.*, 2007). One would expect that a low heterogeneity of strut thickness is best for stiffness and strength. However, we found a positive correlation of Tb.Th.SD with MTF and stiffness, indicating that a broad range of thicknesses is associated with better mechanical outcome. Generally, the strut heterogeneity

along with other microstructural parameters appear to be good surrogate measures for the prediction of strength and stiffness, especially when common measures such as volume and degree of mineralisation fail (Stenström *et al.*, 2000). BMD is strongly correlated with BV/TV, practically containing the same information in the regression model found in this study. TMD is a strong predictor of strength for WT and RAG groups with physiologically high mineralisation and remodelling rate, compare (Toben

**Table 4b:** Average value  $\pm$  standard deviations of  $\mu$ CT parameters for different time points and groups. Explanation of abbreviations are given in the text.

Group	Day	Tb.N [1/mm <sup>3</sup> ]	Tb.Th [mm]	Tb.Sp [mm]	Tb.Th.SD [mm]	Tb.Sp.SD [mm]
WT	14	3.9 $\pm$ 1.3	0.074 $\pm$ 0.005	0.325 $\pm$ 0.140	0.051 $\pm$ 0.006	0.304 $\pm$ 0.113
	21	5.7 $\pm$ 0.8	0.103 $\pm$ 0.019	0.188 $\pm$ 0.045	0.072 $\pm$ 0.018	0.104 $\pm$ 0.066
	28	4.7 $\pm$ 0.7	0.117 $\pm$ 0.024	0.228 $\pm$ 0.042	0.082 $\pm$ 0.025	0.113 $\pm$ 0.026
	35	3.9 $\pm$ 0.8	0.136 $\pm$ 0.037	0.281 $\pm$ 0.077	0.092 $\pm$ 0.038	0.131 $\pm$ 0.048
NF1	14	6.1 $\pm$ 1.9	0.093 $\pm$ 0.022	0.219 $\pm$ 0.080	0.055 $\pm$ 0.015	0.168 $\pm$ 0.068
	21	5.6 $\pm$ 1.4	0.084 $\pm$ 0.019	0.241 $\pm$ 0.083	0.048 $\pm$ 0.018	0.200 $\pm$ 0.075
MAC	14	4.2 $\pm$ 1.7	0.086 $\pm$ 0.011	0.344 $\pm$ 0.113	0.051 $\pm$ 0.008	0.287 $\pm$ 0.086
	21	6.2 $\pm$ 2.8	0.102 $\pm$ 0.005	0.243 $\pm$ 0.130	0.065 $\pm$ 0.006	0.187 $\pm$ 0.091
	28	5.6 $\pm$ 1.9	0.118 $\pm$ 0.016	0.260 $\pm$ 0.101	0.081 $\pm$ 0.022	0.206 $\pm$ 0.098
RAG	14	4.1 $\pm$ 1.7	0.075 $\pm$ 0.008	0.315 $\pm$ 0.117	0.050 $\pm$ 0.005	0.290 $\pm$ 0.126
	21	6.3 $\pm$ 1.0	0.107 $\pm$ 0.014	0.162 $\pm$ 0.032	0.075 $\pm$ 0.011	0.073 $\pm$ 0.013
	28	5.5 $\pm$ 0.7	0.116 $\pm$ 0.010	0.189 $\pm$ 0.029	0.073 $\pm$ 0.010	0.087 $\pm$ 0.019
ALL	14	4.4 $\pm$ 1.7	0.081 $\pm$ 0.013	0.309 $\pm$ 0.120	0.052 $\pm$ 0.008	0.272 $\pm$ 0.110
	21	5.9 $\pm$ 1.5	0.100 $\pm$ 0.018	0.205 $\pm$ 0.079	0.066 $\pm$ 0.018	0.135 $\pm$ 0.083
	28	5.2 $\pm$ 1.1	0.117 $\pm$ 0.017	0.221 $\pm$ 0.061	0.078 $\pm$ 0.019	0.125 $\pm$ 0.067
	35	3.9 $\pm$ 0.8	0.136 $\pm$ 0.037	0.281 $\pm$ 0.077	0.092 $\pm$ 0.038	0.131 $\pm$ 0.048
	ALL	5.2 $\pm$ 1.6	0.101 $\pm$ 0.025	0.245 $\pm$ 0.100	0.066 $\pm$ 0.022	0.173 $\pm$ 0.107

**Table 4c:** Average value  $\pm$  standard deviations of  $\mu$ CT parameters for different time points and groups. Explanation of abbreviations are given in the text.

Group	Day	Conn.D. [1/mm <sup>3</sup> ]	DA	SMI	Callus length l [mm]
WT	14	400.2 $\pm$ 141.1	1.13 $\pm$ 0.04	2.99 $\pm$ 0.75	6.87 $\pm$ 0.41
	21	295.2 $\pm$ 71.8	1.36 $\pm$ 0.11	2.81 $\pm$ 0.75	7.17 $\pm$ 0.41
	28	165.0 $\pm$ 35.6	2.13 $\pm$ 1.89	2.20 $\pm$ 0.52	5.54 $\pm$ 0.92
	35	123.5 $\pm$ 39.4	1.66 $\pm$ 0.25	2.53 $\pm$ 1.54	5.67 $\pm$ 1.73
NF1	14	307.5 $\pm$ 114.5	1.18 $\pm$ 0.10	0.69 $\pm$ 0.50	5.11 $\pm$ 0.09
	21	297.9 $\pm$ 50.7	1.21 $\pm$ 0.12	0.48 $\pm$ 0.84	5.26 $\pm$ 0.08
MAC	14	367.6 $\pm$ 74.4	1.18 $\pm$ 0.04	0.76 $\pm$ 0.74	6.17 $\pm$ 0.38
	21	302.1 $\pm$ 59.3	1.22 $\pm$ 0.12	0.83 $\pm$ 1.58	5.36 $\pm$ 0.81
	28	279.3 $\pm$ 78.7	1.19 $\pm$ 0.05	0.43 $\pm$ 0.84	5.12 $\pm$ 0.83
RAG	14	347.3 $\pm$ 76.4	1.15 $\pm$ 0.07	2.98 $\pm$ 0.41	6.00 $\pm$ 0.78
	21	320.8 $\pm$ 98.3	1.29 $\pm$ 0.09	1.97 $\pm$ 0.68	5.31 $\pm$ 1.15
	28	205.7 $\pm$ 40.7	1.54 $\pm$ 0.13	1.55 $\pm$ 0.65	6.28 $\pm$ 0.73
ALL	14	360.6 $\pm$ 103.5	1.16 $\pm$ 0.06	1.97 $\pm$ 1.29	6.13 $\pm$ 0.76
	21	302.1 $\pm$ 69.6	1.29 $\pm$ 0.13	1.75 $\pm$ 1.36	6.03 $\pm$ 1.12
	28	207.7 $\pm$ 65.4	1.68 $\pm$ 1.19	1.53 $\pm$ 0.93	5.72 $\pm$ 0.92
	35	123.5 $\pm$ 39.4	1.66 $\pm$ 0.25	2.53 $\pm$ 1.54	5.67 $\pm$ 1.73
	ALL	284.8 $\pm$ 106.0	1.36 $\pm$ 0.59	1.83 $\pm$ 1.28	5.97 $\pm$ 1.05

*et al.*, 2011). In NF1 and MAC groups, this parameter failed to correlate with stiffness or MTF. Interestingly, a higher connectivity density is correlated with lower strength and stiffness. Connectivity does not assess the mineral density or further structural quality of struts, but only their connectivity within the whole volume. A high connectivity does not necessarily result in high strength or stiffness. In our approach, the degree of anisotropy influences the mechanical outcome, but positively as in

RAG or negatively as in MAC, reasonably according to the orientation of the anisotropy.

Gross bone volume and material density have been analytically and empirically shown to predict strength and stiffness (Chakkalakal *et al.*, 1990; Martin, 1991; Van der Meulen *et al.*, 2001; Reynolds *et al.*, 2007; Morgan *et al.*, 2009). The WT and RAG groups also showed this pronounced dependency. The analytical concept of (weighted) polar moment of inertia could only achieve

success in predicting stiffness and strength for cortical bone (Morgan *et al.*, 2009; Entezari *et al.*, 2011). This is due to a number of limitations of this parameter, for instance no warping of cross sections and a Hookean material. Little is known about microstructural influences on callus mechanical competence. This study could show empirically that evenly distributed struts seem to be essential, expressed as low standard deviation of strut separation (Tb.Sp.SD). It is reasonable, that the found regression does not display a heterogeneity correlation, because there is a mechanical dependency of the homogeneity of the distribution of bars of a stochastic, many times over-determined truss and its ultimate strength and stiffness. A callus can be considered as a framework (truss) that is statically undetermined. It is evident that the stiffness (rigidity) is a stochastic distribution whose attributes depend on the number of bars, the homogeneity of their distribution and connections and their thickness. In engineered truss structures this effect is used and most such structures display homogeneous distances between the bars.

In osteoporosis and osteoarthritis studies, it has been shown before that patients with verifiably low bone strength exhibit low bone density, trabecular number and thickness and a heterogeneity of microstructure shown as high Tb.Sp.SD (Laib *et al.*, 2002; Sornay-Rendu *et al.*, 2009; Blaizot *et al.*, 2012) or regional variations in the trabecular microstructure (Nazarian *et al.*, 2007; Tassani *et al.*, 2010; Wegrzyn *et al.*, 2010; Hussein and Morgan, 2012). Improved bone microarchitecture has been shown to enhance bone strength by Boyd *et al.* (2011). Given that some studies (Lochmüller *et al.*, 2008a; Lochmüller *et al.*, 2008b) could not find an improvement of failure load prediction including bone microstructure additionally to BMD, it has to be recognised that in these studies, heterogeneity of microstructure was not considered at all, although microstructural variability has a significant effect on mechanical properties (Yeni *et al.*, 2011). However, to the authors' knowledge this is the first study to empirically show that fracture callus microstructure and strut homogeneity correlate with strength and stiffness of the mineralised callus tissue. Study results have shown that torsional stiffness is not only dependent on bone material (TMD, BMD), callus volume (TV, BV), but especially on microstructure (negatively with Tb.Sp.SD, positively with Tb.Th), compare Table 3. In contrast to our expectations and the emphasis on callus volume in previous studies, compared to microstructure, callus size (girth) and geometry have reduced influence on strength and stiffness in our approach.

The regression analyses of each time point separately (Table 3) emphasise the change and evolution of variables over time. This shows the stages of the fracture healing process and what parameters mainly influence the stiffness and strength at what time point. First, a high number of uniformly distributed struts, then increase in their thickness, followed by further mineralisation of these structures contributes most to ultimate strength and stiffness of the callus tissue (compare Fig. 6 and Table 4), but only if a defect healing does not greatly alter the

other parameters. Tissue density and bone volume gain high importance in later, equilibrium stages. However, only microstructure and its homogeneity seem to remain a robust predictor over a long course of fracture healing. It may be very tempting to build a scoring system with these  $\mu$ CT parameters. However, the evolution, along with conflictive and synergetic parameters, for instance many thin struts may have a similar effect as few thick struts, but many thick struts do not strongly change the result; make it difficult to implement a universal model to apply onto all healing cases. The results from this study motivate further investigations to observe individual and combined parameters in  $\mu$ FE models for a precise cause and effect relationship with strength and stiffness (Boyd *et al.*, 2011), also considering the physiological load application (Isaksson *et al.*, 2009).

The practical value of our approach is to show that microstructural parameters are important in addition to parameters that describe volume and mineralisation, and their influence varies temporally and spatially (Table 3). At different time points in fracture healing, different parameters show varying power in estimating strength and stiffness of the healing callus. Furthermore, this study shows that torsional stiffness and strength can be estimated using  $\mu$ CT assessment of mineralised fracture callus. Assessment of the various healing groups has shown that  $\mu$ CT analysis of mechanical competence of callus tissue can be quite robust when considering microstructure and its heterogeneity. While tissue density and volumetric data is widely used to characterise fracture healing, structural parameters and especially their (standard) deviations, which represent the homogeneity, will have to be investigated more thoroughly and should be reported in future bone healing studies. Additionally to the parameters listed by Bouxsein *et al.* (2010), homogeneity of the structural parameters should be reported, but the clinical evaluation possibilities of callus microstructure remain difficult (Cheung *et al.*, 2013).

### Acknowledgements

This research was supported in part by grants from the Berlin-Brandenburg Centre for Regenerative Therapies (BCRT), the German Research Foundation (SFB 760), and Nachwuchsakademie Medizintechnik (DFG- ME 4083/1-1). The authors would like to thank Mr. Jan-Erik Ode and Mr. Mario Thiele for their excellent technical assistance. The authors would like to thank Mr. Thaqif el-khassawna for assisting with animal surgeries.

### Conflict of interest statement

The authors declare that there exists no conflict of interest. No author has any financial and personal relationships with other people or organisations that could inappropriately influence their work.

## References

- Blaizot S, Boutroy S, Vilayphiou N, Boonen S, Chapurlat R, Szulc P (2012) Poor bone microarchitecture in older men with impaired physical performance – the STRAMBO study. *Osteoporos Int* **23**: 2785-2796.
- Bonnarens F, Einhorn TA (1984) Production of a standard closed fracture in laboratory animal bone. *J Orthop Res* **2**: 97-101.
- Bouxsein ML, Boyd SK, Christiansen BA, Guldberg RE, Jepsen KJ, Müller R (2010) Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J Bone Miner Res* **25**: 1468-1486.
- Boyd SK, Szabo E, Armann P (2011) Increased bone strength is associated with improved bone microarchitecture in intact female rats treated with strontium ranelate: A finite element analysis study. *Bone* **48**: 1109-1116.
- Brandi ML (2009) Microarchitecture, the key to bone quality. *Rheumatology (Oxford)* **48 Suppl 4**: iv3-8.
- Burghardt AJ, Kazakia GJ, Majumdar S (2007) A local adaptive threshold strategy for high resolution peripheral quantitative computed tomography of trabecular bone. *Ann Biomed Eng* **35**: 1678-1686.
- Burket J, Gourion-Arsiquaud S, Havill LM, Baker SP, Boskey AL, van der Meulen MC (2011) Microstructure and nanomechanical properties in osteons relate to tissue and animal age. *J Biomech* **44**: 277-284.
- Chakkalakal DA, Lippiello L, Wilson RF, Shindell R, Connolly JF (1990) Mineral and matrix contributions to rigidity in fracture-healing. *J Biomech* **23**: 425-434.
- Chehade MJ, Pohl AP, Percy MJ, Nawana N (1997) Clinical implications of stiffness and strength changes in fracture healing. *J Bone Joint Surg Br* **79**: 9-12.
- Cheung AM, Adachi JD, Hanley DA, Kendler DL, Davison KS, Josse R, Brown JP, Ste-Marie L-G, Kremer R, Erlandson MC (2013) High-resolution peripheral quantitative computed tomography for the assessment of bone strength and structure: A review by the Canadian Bone Strength Working Group. *Curr Osteoporos Rep* **11**: 136-146.
- Claes L, Blakytyn R, Göckelmann M, Schoen M, Ignatius A, Willie B (2009) Early dynamization by reduced fixation stiffness does not improve fracture healing in a rat femoral osteotomy model. *J Orthopaed Res* **27**: 22-27.
- Entezari V, Basto PA, Vartanians V, Zurakowski D, Snyder BD, Nazarian A (2011) Non-invasive assessment of failure torque in rat bones with simulated lytic lesions using computed tomography based structural rigidity analysis. *J Biomech* **44**: 552-556.
- Glatt V, Canalis E, Stadmeier L, Bouxsein ML (2007) Age-related changes in trabecular architecture differ in female and male C57BL/6J mice. *J Bone Miner Res* **22**: 1197-1207.
- Goodship AE, Lawes TJ, Rubin CT (2009) Low-magnitude high-frequency mechanical signals accelerate and augment endochondral bone repair: Preliminary evidence of efficacy. *J Orthopaed Res* **27**: 922-930.
- Green JO, Nagaraja S, Diab T, Vidakovic B, Guldberg RE (2011) Age-related changes in human trabecular bone: Relationship between microstructural stress and strain and damage morphology. *J Biomech* **44**: 2279-2285.
- Hernandez C, Keaveny T (2006) A biomechanical perspective on bone quality. *Bone* **39**: 1173-1181.
- Hussein AI, Morgan EF (2013) The effect of intravertebral heterogeneity in microstructure on vertebral strength and failure patterns. *Osteoporos Int* **24**: 979-989.
- Isaksson H, Gröngröft I, Wilson W, van Donkelaar CC, van Rietbergen B, Tami A, Huiskes R, Ito K (2009) Remodeling of fracture callus in mice is consistent with mechanical loading and bone remodeling theory. *J Orthopaed Res* **27**: 664-672.
- Ishimoto T, Nakano T, Umakoshi Y, Yamamoto M, Tabata Y (2013) Degree of biological apatite c-axis orientation rather than bone mineral density controls mechanical function in bone regenerated using rBMP-2. *J Bone Miner Res* **28**: 1170-1179.
- Kalpakioglu BB, Morshed S, Engelke K, Genant HK (2008) Advanced imaging of bone macrostructure and microstructure in bone fragility and fracture repair. *J Bone Joint Surg Am* **90A**: 68-78.
- Laib A, Newitt DC, Lu Y, Majumdar S (2002) New model-independent measures of trabecular bone structure applied to *in vivo* high-resolution MR images. *Osteoporos Int* **13**: 130-136.
- Liu Y, Manjubala I, Schell H, Epari DR, Roschger P, Duda GN, Fratzl P (2010) Size and habit of mineral particles in bone and mineralized callus during bone healing in sheep. *J Bone Miner Res* **25**: 2029-2038.
- Lochmüller EM, Kristin J, Matsuura M, Kuhn V, Hudelmaier M, Link T, Eckstein F (2008a) Measurement of trabecular bone microstructure does not improve prediction of mechanical failure loads at the distal radius compared with bone mass alone. *Calcif Tissue Int* **83**: 293-299.
- Lochmüller EM, Pöschl K, Würstlin L, Matsuura M, Müller R, Link TM, Eckstein F (2008b) Does thoracic or lumbar spine bone architecture predict vertebral failure strength more accurately than density? *Osteoporos Int* **19**: 537-545.
- Martin RB (1991) Determinants of the mechanical properties of bones. *J Biomech* **24 Suppl 1**: 79-88.
- Mehta M, Duda GN, Perka C, Strube P (2011a) Influence of gender and fixation stability on bone defect healing in middle-aged rats: A pilot study. *Clin Orthopaed Rel Res* **469**: 3102-3110.
- Mehta M, Schell H, Schwarz C, Peters A, Schmidt-Bleek K, Ellinghaus A, Bail HJ, Duda GN, Lienau J (2011b) A 5 mm femoral defect in female but not in male rats leads to a reproducible atrophic non-union. *Arch Orthopaed Trauma Surg* **131**: 121-129.
- Mehta M, Strube P, Peters A, Perka C, Hutmacher D, Fratzl P, Duda G (2010) Influences of age and mechanical stability on volume, microstructure, and mineralization of the fracture callus during bone healing: Is osteoclast activity the key to age-related impaired healing? *Bone* **47**: 219-228.
- Morgan EF, Mason ZD, Chien KB, Pfeiffer AJ, Barnes GL, Einhorn TA, Gerstenfeld LC (2009) Micro-computed tomography assessment of fracture healing: relationships among callus structure, composition, and mechanical function. *Bone* **44**: 335-344.
- Nazarian A, Muller J, Zurakowski D, Muller R, Snyder BD (2007) Densitometric, morphometric and mechanical



distributions in the human proximal femur. *J Biomech* **40**: 2573-2579.

O'Neill KR, Stutz CM, Mignemi NA, Burns MC, Murry MR, Nyman JS, Schoenecker JG (2012) Micro-computed tomography assessment of the progression of fracture healing in mice. *Bone* **50**: 1357-1367.

Reynolds DG, Hock C, Shaikh S, Jacobson J, Zhang X, Rubery PT, Beck CA, O'Keefe RJ, Lerner AL, Schwarz EM, Awad HA (2007) Micro-computed tomography prediction of biomechanical strength in murine structural bone grafts. *J Biomech* **40**: 3178-3186.

Schell H, Epari D, Kassi J, Bragulla H, Bail H, Duda G (2005) The course of bone healing is influenced by the initial shear fixation stability. *J Orthopaed Res* **23**: 1022-1028.

Schulte FA, Lambers FM, Kuhn G, Müller R (2011) *In vivo* micro-computed tomography allows direct three-dimensional quantification of both bone formation and bone resorption parameters using time-lapsed imaging. *Bone* **48**: 433-442.

Sornay-Rendu E, Cabrera-Bravo J, Boutroy S, Munoz F, Delmas P (2009) Severity of vertebral fractures is associated with alterations of cortical architecture in postmenopausal women. *J Bone Miner Res* **24**: 737-743.

Stenström M, Olander B, Lehto-Axtelius D, Erik Madsen J, Nordsletten L, Alm Carlsson G (2000) Bone mineral density and bone structure parameters as predictors of bone strength: an analysis using computerized microtomography and gastrectomy-induced osteopenia in the rat. *J Biomech* **33**: 289-297.

Tassani S, Öhman C, Baleani M, Baruffaldi F, Viceconti M (2010) Anisotropy and inhomogeneity of the trabecular structure can describe the mechanical strength of osteoarthritic cancellous bone. *J Biomech* **43**: 1160-1166.

Toben D, Schroeder I, El Khassawna T, Mehta M, Hoffmann JE, Frisch JT, Schell H, Lienau J, Serra A, Radbruch A (2011) Fracture healing is accelerated in the absence of the adaptive immune system. *J Bone Miner Res* **26**: 113-124.

Van der Meulen M, Jepsen K, Mikić B (2001) Understanding bone strength: size isn't everything. *Bone* **29**: 101-104.

Wade R, Richardson J (2001) Outcome in fracture healing: a review. *Injury* **32**: 109-114.

Wegrzyn J, Roux JP, Arlot ME, Boutroy S, Vilayphiou N, Guyen O, Delmas PD, Chapurlat R, Bouxsein ML (2010) Role of trabecular microarchitecture and its heterogeneity parameters in the mechanical behavior of *ex vivo* human L3 vertebrae. *J Bone Miner Res* **25**: 2324-2331.

Yeni YN, Zinno MJ, Yerramshetty JS, Zauel R, Fyhrie DP (2011) Variability of trabecular microstructure is age-, gender-, race- and anatomic site-dependent and affects stiffness and stress distribution properties of human vertebral cancellous bone. *Bone* **49**: 886-894.

## Discussion with Reviewers

**R. Schnettler:** Can the (absence of) correlations noted in this study suggest the underlying mechanical nature

(Hookean *versus* non-Hookean) of the individual mice healing models?

**Authors:** This is a very interesting question since it underlines the importance of mechanical assumptions, for instance, when using finite element models instead of morphological parameters to evaluate bone quality as assessed in this study. As an example, it has been shown that the apparent-level strength of trabecular bone varies appreciably depending on the assumption of fully brittle *versus* fully ductile failure (Nawathe *et al.*, 2013, additional reference). In finite element assessments, the mechanical assumptions have to be explicitly made while in a regression model they are automatically implicitly incorporated. The mechanical nature (bone material model at the tissue level) was not investigated here, but the model is implicitly incorporated into the regression model. We cannot derive a certain stress-strain response curve from the imaging parameters, because we have only investigated characteristic points of the curve (indeed implicitly assuming a certain part of the curve with Hookean behaviour). The stress-strain curves did not lead us not to assume Hookean behaviour for the first slope of the curve. A finite element assessment should be able to derive the underlying definitive stress-strain curve from the imaging parameters and would be the method of choice to determine the mechanical nature at tissue level (Hookean *versus* non-Hookean). As we mentioned in the manuscript: "The results from this study motivate further investigations to observe individual and combined parameters in  $\mu$ FE models for a precise cause and effect relationship with strength and stiffness (Boyd *et al.*, 2011, text reference), also considering the physiological load application (Isaksson *et al.*, 2009) (text reference).

**R. Schnettler:** Although the authors justifiably argue against developing a diagnostic scoring system, could they elaborate on the explicit limitations for future investigators?

**Authors:** Table 3 now includes qualitative descriptors to clarify the quality of the evolution of different imaging parameters groups and achieved more consistency in the evolution going from volume, over microstructure and density to the state when all factors are equally important. However, we could not find a small, definitive set of parameters that can predict fracture healing outcome for all different groups reliably.

In this manuscript, we use the different healing models to test for varying healing outcomes. This approach is used as a test platform for robust prediction analysis. Therefore, the results from the study based on healing outcomes can be used in human clinical situations as well, upon availability of high resolution clinical imaging technology.

Possibly, in order to identify the limitations of certain predictive parameters for certain healing groups, finite element models display a better option, because for this method, the assumptions have to be explicitly made and can be varied and the results can then be compared to biomechanical test data.

We also highlight the need for further studies: "The results from this study motivate further investigations to observe individual and combined parameters in  $\mu$ FE

models for a precise cause and effect relationship with strength and stiffness ...”.

**P. Arany:** Is torsional testing relevant for callus quality and fracture healing, since torsion is not similar to the *in vivo* loading situation. Doesn't torsional properties give the wrong idea about fracture healing success?

**Authors:** Torsional stiffness and strength measurements are routine biomechanical test parameters in fracture healing studies, and have been shown to correlate with other measures of fracture healing outcome, e.g., with parameters from histology, radiographs, computed tomography and vibrational analysis (Den Boer *et al.*, 1998; Luppen *et al.*, 2002; Morgan *et al.*, 2009; Morshed *et al.*, 2008, additional references). During biomechanical testing, failure (fracture) could be clearly attributed to the newly formed callus area for all groups. Therefore, the most relevant area for mechanical bone function (least stability) after fracture can be reduced to the fracture callus healing area, consequently, this has been the area of focus and torsional testing is used to assess restoration of biomechanical outcome. One reason we chose torsional testing was that the cortical remnants have smaller influence than for instance in compression testing.

**L. Mulder:** Can we omit biomechanical testing in the future only by  $\mu$ CT analysis? This would reduce animal group sizes and could be an alternative method. This would be of great value in animal research.

**Authors:** It would clearly provide great advantages, such as reduced animal numbers, less time and material effort if the number of biomechanical tests could be reduced by  $\mu$ CT analysis predictions. Additionally,  $\mu$ FE analysis can provide subject-specific predictions even with small

deviations from the established characteristics. Indeed, both methods, finite element calculation and regression from morphological parameters, will have to be compared to assess their appropriateness or more probable individual strengths and weaknesses of each method. A minimal biomechanical testing to assure the validity of  $\mu$ CT or even  $\mu$ FE analyses is compulsory. In the near future, biomechanical testing cannot be omitted, but the amount of biomechanical testing might be reduced to a minimum.

### Additional References

Den Boer F, Bramer J, Patka P, Bakker F, Barentsen R, Feilzer A, De Lange E, Haarman H (1998) Quantification of fracture healing with three-dimensional computed tomography. *Arch Orthopaed Trauma Surg* **117**: 345-350.

Luppen C, Blake C, Ammirati K, Stevens M, Seeherman H, Wozney J, Bouxsein M (2002) Recombinant human bone morphogenetic protein-2 enhances osteotomy healing in glucocorticoid-treated rabbits. *J Bone Miner Res* **17**: 301-310.

Morgan EF, Mason ZD, Chien KB, Pfeiffer AJ, Barnes GL, Einhorn TA, Gerstenfeld LC (2009) Micro-computed tomography assessment of fracture healing: relationships among callus structure, composition, and mechanical function. *Bone* **44**: 335-344.

Morshed S, Corrales L, Genant H, Miclau III T (2008) Outcome assessment in clinical trials of fracture-healing. *J Bone Joint Surg* **90**: 62-67.

Nawathe S, Juillard F, Keaveny TM (2013) Theoretical bounds for the influence of tissue-level ductility on the apparent-level strength of human trabecular bone. *J Biomech* **46**: 1293-1299.