Title
High-precision method for cyclic loading of small-animal vertebrae to assess bone quality.

Permalink
https://escholarship.org/uc/item/2n45968x

Authors
Pendleton, Megan M
Sadoughi, Saghi
Li, Alfred
et al.

Publication Date
2018-12-01

DOI
10.1016/j.bonr.2018.10.002

Peer reviewed
High-precision method for cyclic loading of small-animal vertebrae to assess bone quality

Megan M. Pendletona, Saghi Sadoughib, Alfred Li, Grace D. O'Connell, Joshua S. Alwood, Tony M. Keavenc,e,⁎

a Department of Mechanical Engineering, University of California, Berkeley, CA, USA
b Endocrine Research Unit, University of California and Veteran Affairs Medical Center, San Francisco, CA, USA
c Department of Orthopaedic Surgery, University of California, San Francisco, CA, USA
d Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA, USA
e Department of Bioengineering, University of California, Berkeley, CA, USA

ARTICLE INFO

Keywords:
Fatigue
Bone mechanics
Mouse
Vertebrae
Bone quality

ABSTRACT

One potentially important bone quality characteristic is the response of bone to cyclic (repetitive) mechanical loading. In small animals, such as in rats and mice, cyclic loading experiments are particularly challenging to perform in a precise manner due to the small size of the bones and difficult-to-eliminate machine compliance. Addressing this issue, we developed a precise method for ex vivo cyclic compressive loading of isolated mouse vertebrae. The method has three key characteristics: 3D-printed support jigs for machining plano-parallel surfaces of the tiny vertebrae; pivotable loading platens to ensure uniform contact and loading of specimen surfaces; and specimen-specific micro-CT-based finite element analysis to measure stiffness to prescribe force levels that produce the same specified level of strain for all test specimens. To demonstrate utility, we measured fatigue life for three groups (n = 5–6 per group) of L5 vertebrae of C57BL/6J male mice, comparing our new method against two methods commonly used in the literature. We found reduced scatter of the mechanical behavior for this new method compared to the literature methods. In particular, for a controlled level of strain, the standard deviation of the measured fatigue life was up to 5-fold lower for the new method (F-ratio = 4.9; p < 0.01). The improved precision for this new method for biomechanical testing of small-animal vertebrae may help elucidate aspects of bone quality.

1. Introduction

A number of diseases, including osteoporosis (Bouxsein, 2003a; Saito and Marumo, 2010) and type II diabetes (Janghorbani et al., 2007; Vestergaard, 2007; Farr and Khosin, 2016; Rubin and Patsch, 2016), are thought to increase fracture risk due in part to deteriorated bone quality. That is, these diseases may result in detrimental changes to mechanical behavior that are not directly associated with reductions in bone mass (Hernandez and Keaveny, 2006). Concerns over possible changes in bone quality have also caused the FDA to require extensive pre-clinical testing on any new treatments (Moskalkide, 1995; Ominsky et al., 2011). In assessing bone quality, various mechanical properties can be measured (Hernandez and Keaveny, 2006; Turner and Burr, 1993; Fyhrie and Christiansen, 2015), typically stiffness, strength, toughness, and ductility. All these properties are obtained from monotonic loading. In addition, the response of bone to cyclic (repetitive) loading may be important for bone quality assessment if such loading produces unique failure mechanisms (e.g. slow crack growth, microdamage accumulation) that are not manifested in monotonic loading (Kruzić and Ritchie, 2008; Nalla et al., 2003). Biomechanical testing of any type is challenging to perform in a precise manner in small animals, such as in rats and mice, due to the small size of the bones and the inevitable machine-compliance effects that are difficult to eliminate. Furthermore, for cyclic loading, loss of test specimens during the experiment and the substantial variation in cyclic-loading responses both serve to reduce statistical power (Brouwers et al., 2009), which in turn can lead to inconclusive findings regarding the role of bone quality.

As a result of these challenges, and because of the continuing interest in bone quality (Bouxsein, 2003b; Seeman and Delmas, 2006), there remains a need to refine the experimental protocols for biomechanical testing of small-animal bone specimens, particularly for...
cyclic loading. Here, we report on a method for cyclic biomechanical testing of the mouse lumbar vertebra ex vivo that has improved precision. This technique should help optimize statistical power in animal experiments, thus improving the ability of researchers to detect mechanically relevant changes in bone quality due to aging, disease, or treatment, particularly for any situation in which samples sizes are limited in number.

2. Materials and methods

We first present our new method in detail. Then, to demonstrate utility, we describe results from cyclic testing experiments for three groups \((n = 5–6 \text{ per group})\) of L5 mouse vertebrae, comparing results for our new method (the referred to as “KFEA method”) against two existing methods (Brouwers et al., 2009; Caler and Carter, 1989; Bowman et al., 1998; Haddock et al., 2004; Lambers et al., 2013; McCubbrey et al., 1995).

2.1. Animals

Our experiments used L5 vertebrae obtained post-mortem from male C57BL/6J mice (Jackson Labs, Sacramento, CA) that were in the control groups of various previous studies (unpublished). All original studies, and this study, were approved by NASA Ames Research Center (ARC) Animal Care & Use Committee, and all experimental testing was conducted at NASA ARC. Ages of the mice ranged from 15 to 18 weeks. After sacrifice, the L1-L5 spinal segment was excised from the mouse with a scalpel via an anterior approach. The L5 vertebrae was then separated by slicing through the adjacent intervertebral discs, gently cleaned of soft tissue, wrapped in saline-soaked gauze (Gibco PBS 1X, pH7.4), and stored at \(-20^\circ\text{C}\) until specimen preparation.

2.2. Specimen preparation

Obtaining true plano-parallel surfaces, which is difficult for such small specimens (Brouwers et al., 2009), is important to ensure full contact — and thus uniform compression — between the specimen surface and loading platens. If plano-parallel surfaces are not achieved, or if the loading platens are not perfectly parallel, local contact can occur over just part of the specimen surface. This introduces unwanted scatter in the mechanical response and perhaps also premature failure. Thus, we developed a detailed method for this aspect of the specimen preparation (Fig. 1). Each vertebrae was aligned inside a 3D-printed (Stratasys Dimension SST 1200es, Eden Prairie, MN), disposable jig with a metal pin (diameter range 0.91 to 0.99 mm; length of 38 mm) (Fig. 1A) and secured with polymethyl methacrylate (PMMA). The combined jig-bone construct was then positioned in a diamond-saw microtome (Leica SP1600 Saw Microtome, Wetzlar, Germany). Two parallel cuts were made on the cranial and caudal surfaces under constant irrigation to remove only the endplates (Fig. 1B), preserving the maximum amount of bone tissue. A final cut through the posterior pedicles released the vertebral body from the jig (Fig. 1C).

2.3. Quantitative micro-CT imaging

After machining, specimens were imaged using quantitative micro-CT (μCT 50, Scanco Medical AG, Bruttisellen, Switzerland) at a 10-μm voxel size (55 kV, 109 μA, 1000 projections per 180°, 500 ms integration time) (Fig. 2). Specimens were returned to the freezer after imaging. The micro-CT images of the whole specimen (cortical and trabecular compartments) were then analyzed for height, minimum tissue cross-sectional area (minimum value of tissue area at any cross-section, over all cross-sections), and total bone volume fraction (BV/TV) (ImageJ 1.51h, Java 1.6.0). Each specimen with cortices removed was then evaluated for trabecular micro-architecture (Scanco Evaluation Software v6.0): trabecular bone volume fraction (Tb.BV/TV), number (Tb.N), thickness (Tb.Th), and separation (Tb.Sp).

2.4. Calculating specimen-specific cyclic forces

We used a force-control loading protocol, which allows both creep and damage to occur during cyclic loading. However, since the cyclic loading response of bone appears to be determined by the magnitude of the applied strain (as opposed to the magnitude of stress) (Nalla et al., 2003; Keaveny et al., 1994a; Keaveny et al., 1994b), our protocol required us to calculate a specimen-specific level of applied force that would result in the same level of initial apparent strain (calculated over the entire specimen) for all specimens. Cyclic loading in force control in this way should minimize scatter in the resulting fatigue life despite the heterogeneity in size, shape, and microarchitecture of the mouse vertebrae.

This protocol required measuring the overall vertebral stiffness before starting the cyclic loading so that the prescribed apparent strains \((\varepsilon_{\text{min}} \text{ and } \varepsilon_{\text{max}})\) could be converted into specimen-specific levels of applied force \((F_{\text{min}} \text{ and } F_{\text{max}})\). However, experimentally measuring stiffness for any bone is challenging due to different types of machine compliance errors, which can introduce both fixed and random errors (Odgaard et al., 1989; Odgaard and Linde, 1991; Keaveny et al., 1993; Keaveny et al., 1997). The fixed errors have implications for external validity, but otherwise do not compromise comparisons within any given experiment and are therefore oftentimes acceptable. However, large random errors compromise statistical power, particularly when sample sizes are < 10 per group. Thus, to circumvent any errors associated with machine compliance, we used a computationally derived stiffness \((K_{\text{FEA}})\) that was obtained from specimen-specific, micro-CT-based, linearly elastic, voxel-based finite element analysis. The models used 10-μm sized cube voxels, and assumed a nominal value of 10 GPa for the tissue-level (voxel) elastic modulus for all specimens (see Appendix A for details). The specimen-specific cyclic forces \((F_{\text{min}}, \text{ Eq. 1 and } F_{\text{max}}, \text{ Eq. 2})\) were then calculated for each specimen using \(K_{\text{FEA}}\), the specimen height \((H)\) as measured from micro-CT, and two assumed nominal values of initial apparent elastic strain \((\varepsilon_{\text{min}} = 0.05\% \text{ and } \varepsilon_{\text{max}} = 0.5\%):\)

\[
F_{\text{min}} = K_{\text{FEA}} H \varepsilon_{\text{min}} \\
F_{\text{max}} = K_{\text{FEA}} H \varepsilon_{\text{max}}
\]  

The assumed value of \(\varepsilon_{\text{max}} = 0.5\%\) was chosen empirically by trial and error based on earlier preliminary experiments on 6 mouse vertebral specimens. There, we found that the resulting \(F_{\text{max}}\) value was approximately half of the directly measured ultimate force, which generally placed a specimen in the middle region of its fatigue life. The minimum initial elastic apparent strain value was chosen such that the resulting \(F_{\text{min}}\) values were near 1 N.

2.5. Mechanical testing

Within one week after imaging, specimens were thawed again to room temperature and prepared for cyclic loading in uniaxial compression to failure. The vertebral body was positioned onto the center of the lower platen of the material testing device (TA ElectroForce 3200, Eden Prairie, MN). Because obtaining perfectly parallel loading platens can be challenging, the upper platen comprised a spherically-seated platen on a lubricated ball-bearing which allows the platen surface to rotate and mate flush with the sample surface to achieve uniform contact (Fig. 3); with the pre-load in place, the set screws on the upper platen are locked. We used a compressive pre-load of 1 N, and after the platen was locked room-temperature saline-solution (Gibco PBS 1X, pH7.4) was added to the bath until the specimen was fully submerged. The pre-load was then adjusted to the predetermined \(F_{\text{min}}\), and then sinusoidal cycles of compression between \(F_{\text{min}}\) and \(F_{\text{max}}\) were applied at

\[\varepsilon_{\text{min}} = 0.05\% \text{ and } \varepsilon_{\text{max}} = 0.5\%\]:

\[
F_{\text{min}} = K_{\text{FEA}} H \varepsilon_{\text{min}} \\
F_{\text{max}} = K_{\text{FEA}} H \varepsilon_{\text{max}}
\]
a frequency of 8 Hz until the specimen reached 10% strain, measured via the actuation of the testing device platens.

2.6. Outcome measurements

Three plots were produced from each specimen fatigue test to define outcome measurements. First, the force-displacement data were plotted for each loading-unloading cycle, from which stiffness was defined as the secant stiffness, or range of force divided by range of displacement per cycle (Fig. 4A). Second, the apparent strain per cycle was plotted, and was separated into three curves representing total strain, creep strain, and damage strain (Fig. 4B). The total-strain curve was used to find the number of cycles to failure, or fatigue life ($N_f$) and strain to failure ($\varepsilon_f$) using the point of intersection between the curve and a 0.5% line offset parallel to a line of best fit for the secondary region (Brouwers et al., 2009; Bowman et al., 1998). Third, we plotted stiffness per cycle (Fig. 4C) to identify three variables: elastic stiffness of the specimen ($K_{\text{elastic}}$), stiffness at failure ($K_f$), and rate of stiffness degradation. Elastic stiffness was calculated as the average stiffness over ±500 cycles centered about the maximum value of stiffness ($K_{\text{max}}$), which typically occurred after a few thousand cycles of loading; the stiffness at failure was defined as the stiffness measured at cycle $N_f$ and the rate of stiffness degradation ($\Delta K/\Delta N$) was calculated as the slope of a line that best fit the middle 50% of points in the range between $K_{\text{max}}$ and $K_f$. The $\Delta K/\Delta N$ parameter reflected accumulation of fatigue damage, larger values indicating a more rapid accumulation of fatigue damage per loading cycle.

In addition to these fatigue parameters, we also calculated an effective elastic tissue modulus ($E_{\text{tissue}}$, in GPa, Eq. 3). This was calculated using both the experimental ($K_{\text{elastic}}$) and computational ($K_{\text{FEA}}$) assessments of stiffness, scaling with respect to the assumed nominal value of the tissue-level modulus of 10 GPa in the finite element analysis:

$$E_{\text{tissue}} = 10 \left( \frac{K_{\text{elastic}}}{K_{\text{FEA}}} \right)$$

2.7. Alternative methods for calculating specimen-specific cyclic forces

Two published methods were also implemented to demonstrate

![Fig. 1. Machining of a parallel-planed vertebral body. A: A 3D-printed, disposable, plastic jig was designed to securely grip the specimen during machining. A vertebral body is placed into the jig using a snugly fit, metal alignment pin (gold) through the vertebral foramen (pin diameter 0.9–1 mm) and sits in a diameter-matched channel. The spinous process is oriented into the jig's shallow well, which is then filled with PMMA (blue) to secure the vertebra. B: Once the PMMA is set, the pin is removed and two plano-parallel cuts (red) through the endplates are made using a diamond-saw microtome (not shown). C: A scalpel cut (red) through the pedicles releases the planed vertebral body. In the top, right corner an enlarged view of an L5 vertebrae and the cuts applied is shown. Readers can download the stl via the supplemental material online. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)](image1)

![Fig. 2. Representative sample of a mouse vertebral body after all specimen preparation. Cranial and caudal cross-sections shown at 10-μm voxel resolution.](image2)
utility of the new method. For the first “KEXP method” (Brouwers et al., 2009; Caler and Carter, 1989; Bowman et al., 1998; Haddock et al., 2004; Lambers et al., 2013), instead of using the finite element-derived value of stiffness (KFEA), the experimental measurement KEXP was used. Following literature protocol, KEXP was found by pre-loading the specimen in compression to 1 N, cycling between 1 and 10 N 20 times, and using the slope from the force-displacement curve of the 20th cycle. All other mechanical testing details were the same as for our new method except for two differences in the KEXP method: 1) the upper platen used was a rigid design (not adjustable); and 2) the cyclic frequency was 2 Hz. However, fatigue behavior of bone has been shown to be unaltered for any loading frequency below 15 Hz (Lafferty, 1978; Yamamoto et al., 2006).

For the second “FMAX method” (McCubrey et al., 1995), Fmax was calculated using the measured strength of the adjacent L4 vertebra, obtained by monotonic compression to failure. The L5 strength was estimated as the product of the L4 strength and the ratio of minimum tissue cross-sectional areas (micro-CT) for the L4 and L5 vertebrae, and Fmax was then calculated as 50% of the L5 force (See Appendix B, Table 2 for further details). All other mechanical testing details were the same as our new method.

![Diagram of experimental test setup](Image)

**Fig. 3.** Experimental test set-up for fatigue test shown with spherically-seated upper platen and specimen after preparation.

![Graphs of force-displacement, strain, and stiffness](Image)

**Fig. 4.** Three plots generated from fatigue testing of a specimen. A: Force-displacement shown for three cycles as the test progresses; an early cycle, the cycle of maximum stiffness (dmax is displacement of this cycle), and the cycle defined as failure (df is displacement of this cycle). B: Strain per cycle was plotted for total strain, creep strain, and damage strain (Haddock et al., 2004). Damage strain was only plotted from the cycle at which stiffness was the greatest (Nmax) and began to decline. Fatigue life, Nf, and strain to failure, εf, are found from the total strain curve (Bowman et al., 1998). C: Stiffness per cycle displaying the maximum stiffness, Kmax, from which Kelastic was calculated, and stiffness at failure, Kf. The rate of stiffness degradation is the slope of the solid red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
2.8. Statistical analysis

The means and standard deviations of all outcome measurements were compared between the new method versus the two literature methods. To compare means, a one-way ANOVA was conducted, with a Tukey-Kramer post-hoc test applied to evaluate significant differences (at \( p \leq 0.05 \)). To compare standard deviations, a Bartlett test was conducted. Standard deviations were considered significantly different with a F-ratio > 1.0 and \( p \leq 0.05 \). Also, linear regression was used to compare the relationships between maximum cyclic loading force and log of fatigue life (equivalent to a traditional S-N fatigue curve) between the three methods. All statistics were performed using JMP (v13.0, SAS Institute).

3. Results

Of all three methods tested, the new \((K_{\text{FEA}})\) method produced the lowest variation in fatigue life. The mean value of the applied force was the same for all groups (22 N), but the standard deviation was lowest for the new \((K_{\text{FEA}})\) method \(( \pm 0.77 \text{ N})\), 8-fold and 7-fold lower than the \(K_{\text{EXP}}\) and \(F_{\text{MAX}}\) methods \((\text{F-ratio} = 5.5; p < 0.01; \text{Fig. 5A})\), respectively. Because of these variations in applied force for the literature methods, the fatigue life displayed a negative S-N type relation with the maximum cyclic loading force and log of fatigue life \((\text{equivalent to a traditional S-N fatigue curve})\) between the three methods. All statistics were performed using JMP (v13.0, SAS Institute).

The means and standard deviations of all outcome measurements were compared between the new method versus the two literature methods. To compare means, a one-way ANOVA was conducted, with a Tukey-Kramer post-hoc test applied to evaluate significant differences (at \( p \leq 0.05 \)). To compare standard deviations, a Bartlett test was conducted. Standard deviations were considered significantly different with a F-ratio > 1.0 and \( p \leq 0.05 \). Also, linear regression was used to compare the relationships between maximum cyclic loading force and log of fatigue life (equivalent to a traditional S-N fatigue curve) between the three methods. All statistics were performed using JMP (v13.0, SAS Institute).

The consistency of our strain-to-failure results with those from other studies supports the validity of our new approach. While the fatigue life will depend on the level of loading, strain-to-failure is emerging as a more uniform outcome across many different cyclic loading experiments. Our strain-to-failure \((\text{mean} \pm \text{SD})\) value of \(3.20 \pm 0.80\%\), from cyclic compression testing of mouse vertebrae, is consistent with...
cyclic compression results for other species: 3.36 ± 2.13% for human trabecular bone (Haddock et al., 2004), 2.85% ± 0.66% for bovine trabecular bone (Bowman et al., 1998), and 4.19 ± 1.52% for rat whole-bone vertebra (Brouwers et al., 2009). A similar value in cyclic compressive strain-to-failure between these studies and ours might be expected for two reasons. First, the bone material, primarily the collagen component which dominates the plasticity and therefore strain-to-failure of the bone (Burstein et al., 1975), is similar amongst the species listed. And second, others have reported on relatively uniform values for the strain-to-failure independent of apparent density, applied loads, or temperature (Bowman et al., 1998; Keaveny et al., 2001; Keaveny et al., 1994c). While the absolute value of strain to failure will be influenced by any machine compliance, the general level of agreement and small degree of scatter suggests that bone tissue does indeed fail at a narrow range of strain under cyclic loading, and that our protocol was successful in reducing the variation in strain across different specimens.

Our findings also indicate that implementing a pivotable loading platen to ensure uniform contact is important for accuracy of results, at least for these small-animal bones. We observed a reduced $K_{\text{elastic}}$ and fatigue life for the $K_{\text{EXP}}$ method compared to the new method, even though the specimens used in both protocols had similar size and microarchitecture (Appendix B; Table 1). We attribute these inconsistent results to the rigid platen design used in the $K_{\text{EXP}}$ method testing. Even with rigorous specimen preparation in place, we found that perfectly plano-parallel loading surfaces on small bones cannot always be achieved. And without an adjustable platen, the load can be localized to an area that comes into contact with the platen first, producing an ineffective stress concentration. This localized load can result in reduced stiffness and the stress concentration in reduced fatigue life.

There are some limitations to this study. First, the bone tissue used in this study was shared tissue, which did not allow for random assignment of animals across all groups, and thus it is possible specimens differed across groups. However, the similarity in size and structure of specimen across all groups as per micro-CT analysis (Appendix B; Table 1) suggests the only source of appreciable heterogeneity could be in tissue material properties, which is unlikely because the animals were of the same strain, sex, and approximate age. Second, the initial strain values of ($\varepsilon_{\text{max}} = 0.5\%$) used for our new method may be specific to this particular protocol, which assumed an effective tissue modulus of 10 GPa in the finite element analyses. The assumed value was chosen empirically by trial and error on some preliminary test specimens, such that the resulting $F_{\text{max}}$ values were approximately half of the directly measured ultimate force. To use this method in other animal models, large or small (e.g. rat, bovine, monkey), similar preliminary testing to determine the value of $\varepsilon_{\text{max}}$ is suggested, particularly if it is suspected that the effective tissue modulus might differ from mouse vertebrae. We recommend using an elastic strain level that is associated with about half of the ultimate strength. Third, the values of the effective tissue elastic modulus reported here are specific to this experiment. Parameters such as temperature, loading frequency, and use of different platens, can all contribute to machine compliance and variation between tests, thus the effective tissue modulus should not be considered a true material property. Placing an extensometer across the platens directly beside the specimen ends can minimize machine compliance effects. However, it is best to use the effective tissue modulus as a relative measurement to compare between untreated and treated specimens within an experiment, but not across studies (e.g. between mouse and rat bones). Lastly, as with any laboratory-based biomechanical testing, the loading configuration does not mimic exact in vivo conditions. Our removal of the endplate and loading via a rigid platen versus loading via intact disc and endplate is a clear limitation. Thus, while our results do not describe expected in vivo behavior, our protocol does produce highly repeatable results, which should help elucidate changes in bone quality characteristics by aging, disease, or treatment.

The main application of our new method is to provide greater statistical power in using small-animal experiments to mechanically assess relevant aspects of bone quality associated with aging, disease, or treatment. A small sample size and scatter associated with the fatigue testing present challenges for detecting statistically significant effects unless those effects are large. For example, Brouwers et al. (Brouwers et al., 2009) found a 37.5% difference in steady-state creep rate and 15.7% difference in loss of stiffness with zoledronate treatment in an osteoporotic rat model, but could not report a detectable significance due to high variation in the data and loss of some samples. Our $K_{\text{EXP}}$ method was similar to their method, which produced standard deviations in fatigue life (reported as “time-to-failure”) up to the same magnitude as the mean value. Our new method should reduce that range up to 8-fold. Furthermore, unlike in the Brouwers literature protocol, we had no loss of specimens, enabled by repeatable specimen preparation (e.g. 3D-printed support jigs for machining plano-parallel surfaces) and uniform loading of specimen surfaces (e.g. spherically-seated platens), which further improves statistical power, and can be implemented for monotonic testing or fatigue testing of small-animal vertebrae. That said, our new method is a time intensive process for specimen preparation, micro-CT, image processing, and finite element analysis. However, the extra effort improves the ability to detect small differences across treatment groups without using a large number of animals, which is advantageous both ethically and also logistically when sample sizes are inherently limited (e.g. space flight experiments). Thus, use of our new method may improve the ability of future studies to elucidate mechanically relevant aspects of bone quality associated with aging, disease, or treatment.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bonr.2018.10.002.

Acknowledgements

This research was supported by a NASA Science and Technology Research Fellowship (NSTRF) NNX14AM56H to MMP, and guidance from NSTRF research collaborator, Dr. Jean D. Sibonga. Additional experimental resources were provided through NASA Space Biology PECASE to JSA. Computational resources were made available through the National Science Foundation via XSEDE, Grant TG-MCA00N019 to TMK. We also thank Shannon Emerzian and Ryan Scott for help with specimen preparation.

Conflicts of interest

All authors certify that there are no conflicts of interest related to the work presented in this manuscript.

Appendix A. Image processing

Image processing of the original micro-CT images was conducted prior to finite element analysis for each specimen. First, the image stack was reoriented to align the specimen with the z-axis (DataViewer v1.5.2.4, Bruker). Second, the reoriented images were thresholded using the ‘Default’ Autothreshold algorithm in ImageJ (ImageJ 1.51h, Java 1.6.0_24). The final stack was used to build the finite element model.

A.1. Finite element analysis

To determine the computational stiffness, high-resolution linear finite element analysis was performed. Each 10-μm cubic voxel in the images was
converted into an eight-noded brick element to create a finite element model of the entire vertebral body. All bone elements were assigned the same elastic, homogeneous and isotropic material properties: elastic modulus 10 GPa and Poisson’s ratio of 0.3. These values are based on experimental testing of core samples of trabecular bone from human vertebrae (Bevill and Keaveny, 2009). A varied modulus based on tissue mineral density and thus micro-CT greyscale was not used in this methodology due to underestimation of apparent modulus using this approach (Kaynia et al., 2015). Each model was uniaxially compressed to 2% apparent strain, the input was displacement which varied depending on specimen height. Typical bone reaction force, which was then used to calculate computational sti

\[ K_{\text{FEA}} = \frac{\text{Reaction force}}{\text{Input displacement}} \]  

(4)

**Appendix B**

**Table 1**

The \( K_{\text{EXP}} \) and \( K_{\text{FEA}} \) method specimen structural parameters indicate there were no significant differences in specimen size and shape. There were also no significant differences in the mean or standard deviations (F-ratio = 1.65; \( p = 0.20 \)) of computed cyclic force between these two groups. The computed cyclic loading force for \( K_{\text{EXP}} \) group was retrospectively found to demonstrate the difference in standard deviation of cyclic loading force computed with the \( K_{\text{EXP}} \) method compared to \( K_{\text{FEA}} \) method (F-ratio = 10.9; \( p < 0.01 \)). The structural parameters from specimens of the \( \text{FMAX} \) method are not compared to \( K_{\text{EXP}} \) and \( K_{\text{FEA}} \) method groups because of a difference in resolution of these micro-CT scans (20.7 \( \mu \)m). However, the stiffness and fatigue life measurements are the same for the \( \text{FMAX} \) and \( K_{\text{FEA}} \) methods, which suggest the structural parameters for the \( \text{FMAX} \) method specimens were similar to those used in the \( K_{\text{EXP}} \) and \( K_{\text{FEA}} \) methods.

| Measure                  | \( K_{\text{EXP}} \) (n = 6) | \( K_{\text{FEA}} \) (n = 5) | \( p \)  |
|--------------------------|-------------------------------|-------------------------------|---------|
| Min tissue area (mm\(^2\)) | 0.97 ± 0.19                  | 0.96 ± 0.09                  | 0.92    |
| Avg tissue area (mm\(^2\))  | 1.71 ± 0.33                  | 1.59 ± 0.11                  | 0.47    |
| BV/TV (%)                 | 34.0 ± 1.00                  | 33.0 ± 1.00                  | 0.08    |
| Tb.BV/TV (%)              | 22.0 ± 3.00                  | 23.0 ± 1.00                  | 0.83    |
| Tb.N (1/mm)               | 4.94 ± 0.42                  | 5.30 ± 0.32                  | 0.15    |
| Tb.Th (\( \mu \)m)        | 47.1 ± 3.26                  | 44.5 ± 2.58                  | 0.19    |
| Tb.Sp (\( \mu \)m)        | 198 ± 17.8                   | 185 ± 14.1                   | 0.24    |
| Computed cyclic loading force (N) | 20.9 ± 1.54                  | 21.2 ± 0.78                  | 0.73    |
| Cyclic loading force (N)   | 22.5 ± 6.78                  | 21.2 ± 0.78                  | 0.68    |

**Table 2**

Structural parameters and resulting forces and fatigue life for L4 and L5 vertebrae used in testing the “\( \text{FMAX} \) method”. Max Stress is calculated from the resulting maximum force found from monotonic testing of the L4 (i.e. \( \sigma = F_{\text{L4}}/A_{\text{L4}} \)). This maximum stress is a material property and assumed to be the same for the adjacent L5 bone. The cyclic loading force applied to the L5 was derived from this stress (i.e. \( F_{\text{max}} = 0.5(\sigma A_{\text{L5}}) \)), labeled “50% of L5 Max Force” in the table. The range in applied force calculated led to a larger standard deviation in fatigue life when compared to the new method \( K_{\text{FEA}} \).

| Specimens (L4 or L5) | L4 Min Tissue Area (mm\(^2\)) | L4 Max Force (MPa) | Max Stress (MPa) | L5 Min Tissue Area (mm\(^2\)) | 50% of L5 Max Force (N) | L5 \( N_f \) |
|----------------------|-------------------------------|-------------------|-----------------|-------------------------------|------------------------|----------|
| DX.2                 | 1.69                          | 44.1              | 26.1            | 1.36                          | 17.8                   | 374,418  |
| DX.3                 | 1.42                          | 50.2              | 35.3            | 1.75                          | 30.9                   | 46,516   |
| DX.5                 | 1.23                          | 37.5              | 30.5            | 1.18                          | 18.0                   | 476,045  |
| DX.6                 | 2.33                          | 47.3              | 20.3            | 1.92                          | 19.5                   | 172,800  |
| DX.7                 | 1.58                          | 65.5              | 41.4            | 1.29                          | 26.7                   | 98,217   |
| Mean ± SD            | 1.65 ± 0.4                    | 48.9 ± 9.3        | 30.7 ± 7.3      | 1.50 ± 0.3                    | 22.6 ± 5.3             | 233,600 ± 164,700 |

**References**

Adams, M.F., Bayraktar, H.H., Keaveny, T.M., Papadopoulos, P., 2004. Ultrascalable implicit finite element analyses in solid mechanics with over a half a billion degrees of freedom, supercomputing, 2004. In: Proceedings of the ACM/IEEE SC2004 Conference. IEEE, pp. 34.

Bevill, G., Keaveny, T.M., 2009. Trabecular bone strength predictions using finite element analysis of micro-scale images at limited spatial resolution. Bone 44 (4), 579–584.

Bouxsein, M.L., 2003a. Mechanisms of osteoporosis therapy: a bone strength perspective. Clin. Cornerstone (Suppl. 2), S13–S21.

Bouxsein, M.L., 2003b. Bone quality: where do we go from here? Osteoporos. Int. 14 (Suppl. 5), S118–S127.

Burstein, A.H., Zika, J.M., Heiple, K.G., Klein, L., 1975. Contribution of collagen and mineral to the elastic-plastic properties of bone. J. Bone Joint Surg. Am. 57 (7), 956–961.

Burstein, A.H., Zika, J.M., Heiple, K.G., Klein, L., 1975. Contribution of collagen and mineral to the elastic-plastic properties of bone. J. Bone Joint Surg. Am. 57 (7), 956–961.

Calcif. Tissue Int. 97 (3), 213–228.

Farr, J.N., Kholos, S., 2016. Determinants of bone strength and quality in diabetes mellitus in humans. Bone 82, 28–34.

Fyhrie, D.P., Christiansen, B.A., 2015. Bone material properties and skeletal fragility. Calcif. Tissue Int. 97 (3), 213–228.

Farr, J.N., Kholos, S., 2016. Determinants of bone strength and quality in diabetes mellitus in humans. Bone 82, 28–34.

Caler, W.E., Carter, D.R., 1989. Bone creep-fatigue damage accumulation. J. Biomech. 22 (6–7), 625–635.

Haddack, S.M., Yeh, O.C., Mummaneni, P.V., Rosenberg, W.S., Keaveny, T.M., 2004. Similarity in the fatigue behavior of trabecular bone across site and species. J. Biomech. 37 (2), 181–187.

Hernandez, C.J., Keaveny, T.M., 2006. A biomechanical perspective on bone quality.
Keaveny, T.M., Borchers, R.E., Gibson, L.J., Hayes, W.C., 1993. Theoretical analysis of the experimental artifact in trabecular bone compressive modulus. J. Biomech. 26 (4–5), 599–607.

Keaveny, T.M., Wachtel, E.F., Ford, C.M., Hayes, W.C., 1994a. Differences between the tensile and compressive strengths of bovine tibial trabecular bone depend on modulus. J. Biomech. 27 (9), 1137–1146.

Keaveny, T.M., Guo, X.E., Wachtel, E.F., McMahon, T.A., Hayes, W.C., 1994b. Trabecular bone exhibits fully linear elastic behavior and yields at low strains. J. Biomech. 27 (9), 1127–1136.

Keaveny, T.M., Wachtel, E.F., Guo, X.E., Hayes, W.C., 1994c. Mechanical behavior of damaged trabecular bone. J. Biomech. 27 (11), 1309–1318.

Keaveny, T.M., Pinilla, T.P., Crawford, R.P., Kopperdahl, D.L., Lou, A., 1997. Systematic and random errors in compression testing of trabecular bone. J. Orthop. Res. 15 (1), 101–110.

Keaveny, T.M., Morgan, E.F., Niebur, G.L., Yeh, O.C., 2001. Biomechanics of trabecular bone. Annu. Rev. Biomed. Eng. 3, 307–333.

Kruzic, J.J., Ritchie, R.O., 2008. Fatigue of mineralized tissues: cortical bone and dentin. J. Mech. Behav. Biomed. Mater. 1 (1), 3–17.

Lafferty, J.F., 1978. Analytical model of the fatigue characteristics of bone. Aviat. Space Environ. Med. 49 (1 Pt. 2), 170–174.

Lambers, F.M., Bouman, A.R., Rimnac, C.M., Hernandez, C.J., 2013. Microdamage caused by fatigue loading in human cancellous bone: relationship to reductions in bone biomechanical performance. PLoS One 8 (12), e83662.

McCubrey, D.A., Cody, D.D., Peterson, E.L., Kuhn, J.L., Flynn, M.J., Goldstein, S.A., 1995. Static and fatigue failure properties of thoracic and lumbar vertebral bodies and their relation to regional density. J. Biomech. 28 (8), 891–899.

Moekolde, L., 1995. Assessing bone quality—animal models in preclinical osteoporosis research. Bone 17 (4 Suppl), 343S–352S.

Nalla, R.K., Kinney, J.H., Ritchie, R.O., 2003. Mechanistic fracture criteria for the failure of human cortical bone. Nat. Mater. 2 (3), 164–168.

Odgaard, A., Linde, F., 1991. The underestimation of Young's modulus in compressive testing of cancellous bone specimens. J. Biomech. 24 (8), 691–698.

Odgaard, A., Hvid, I., Linde, F., 1989. Compressive axial strain distributions in cancellous bone specimens. J. Biomech. 22 (8–9), 829–835.

Ominksky, M.S., Stouch, B., Schroeder, J., Pyrah, I., Stolina, M., Smith, S.Y., Kostenuik, P.J., 2011. Denosumab, a fully human RANKL antibody, reduced bone turnover markers and increased trabecular and cortical bone mass, density, and strength in ovariectomized cynomolgus monkeys. Bone 49 (2), 162–173.

Rubin, M.R., Patsch, J.M., 2016. Assessment of bone turnover and bone quality in type 2 diabetic bone disease: current concepts and future directions. Bone Res. 4, 16001.

Seeman, E., Delmas, P.D., 2006. Bone quality—the material and structural basis of bone strength and fragility. N. Engl. J. Med. 354 (21), 2250–2261.

Turner, C.H., Burr, D.B., 1993. Basic biomechanical measurements of bone: a tutorial. Bone 14 (4), 595–608.

Vestergaard, P., 2007. Discrepancies in bone mineral density and fracture risk in patients with type 1 and type 2 diabetes—a meta-analysis. Osteoporos. Int. 18 (4), 427–444.

Yamamoto, E., Paul Crawford, R., Chan, D.D., Keaveny, T.M., 2006. Development of residual strains in human vertebral trabecular bone after prolonged static and cyclic loading at low load levels. J. Biomech. 39 (10), 1812–1818.