ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Research

Article submitted to journal

Subject Areas:
osteoporosis, trabecular bone, morphometry

Keywords:
Ellipsoid Factor, SMI, plates, rods, trabecular bone, osteoporosis
1

Author for correspondence:
Alessandro Felder
e-mail: a.felder@ucl.ac.uk

The plate-to-rod transition in trabecular bone loss is elusive

A. A. Felder[1,2], S. Monzem[1,3], R. De Souza[3], B. Javaheri[1,4], D. Mills[5], A. Boyde[5] and M. Doube[1,6]

[1]Royal Veterinary College, London, UK
[2]University College London, London, UK
[3]Universidade Federal de Mato Grosso, Cuiabá, Brazil
[4]City, University of London, London, UK
[5]Queen Mary University of London, London, UK
[6]City University of Hong Kong, Kowloon, Hong Kong SAR China

Changes in trabecular micro-architecture are key to our understanding of osteoporosis. Previous work focusing on structure model index (SMI) measurements have concluded that disease progression entails a shift from plates to rods in trabecular bone, but SMI is heavily biased by bone volume fraction. As an alternative to SMI, we proposed the Ellipsoid Factor (EF) as a continuous measure of local trabecular shape between plate-like and rod-like extremes. We investigated the relationship between EF distributions, SMI and bone volume fraction of the trabecular geometry in a murine model of disuse osteoporosis as well as from human vertebrae of differing bone volume fraction. We observed a moderate shift in EF median (at later disease stages in mouse tibia) and EF mode (in the vertebral samples with low bone volume fraction) towards a more rod-like geometry, but not in EF maximum and minimum. These results support the notion that the plate to rod transition does not coincide with the onset of bone loss and is considerably more moderate, when it does occur, than SMI suggests. A variety of local shapes not straightforward to categorise as rod or plate exist in all our trabecular bone samples.

1. Introduction

The metabolic bone disease osteoporosis is a major health concern associated with high mortality rates and considerable economic costs (1; 2), likely to be

THE ROYAL SOCIETY PUBLISHING

© 2014 The Authors.

exacerbated by the increase in the proportion of elderly people in future demographics. In this disease, imbalance between osteoblastic (bone-forming) and osteoclastic (bone-resorbing) cell activity is thought to lead to lower bone turnover and relatively higher resorption than formation, and thus to a lower amount of bone (3). Lower bone mass causes reduced mechanical competence and increased fracture risk with age (4).

Bone tissue is typically classified by porosity into trabecular (porous, "spongy" bone typically found inside bone organs) and cortical (compact, "dense" bone, usually constituting the "shell" of a bone organ). The large amount of bone surface relative to bone volume in trabecular bone compared to cortical bone may make it particularly sensitive to shifts in the bone (re)modelling balance (5). Beyond the loss of bone volume fraction in the trabecular bone compartment, changes in tissue morphology may contribute to the deterioration of bone quality of osteoporotic patients. Because such osteoporosis-related changes to the trabecular bone micro-architecture form a link between the bone (re)modelling balance at a tissue level and the mechanical performance of the bone organ, they are key to our understanding of the disease.

Prominent amongst parameters considered when evaluating tissue-level morphological changes is *structure model index* (SMI) (6). SMI was designed to estimate how far a trabecular geometry may be considered rod- or plate-like (7). Evaluation of SMI across a number of data sets from human patients and animal models suggests that trabecular geometry transitions from being more plate-like to more rod-like as osteoporosis severity increases ("plate-to-rod transition")(8; 9; 10; 11; 12). However, it is well known that SMI correlates strongly with bone volume fraction, rendering dubious the comparison of SMI values between samples of considerably different bone volume fraction, such as osteoporotic samples versus healthy control samples. Furthermore, the concept of SMI is based on relative changes in surface area in response to a small dilation (a parallel offset from the surface), and relies on the fact that dilating a shape with non-negative Gaussian curvature (such as a sphere (SMI=4), a cylinder (SMI=3), or a infinite plane (SMI=1)) never decreases the surface area. This is not the case in trabecular bone, because large parts of the trabecular bone surface are hyperbolic ("saddle-like") (13), which causes local shrinking of the surface when the shape is dilated (14).

Ellipsoid Factor (EF) has been proposed as an alternative method to measure the plate-to-rod transition in trabecular bone (15). EF has since been used within and beyond bone biology, for example in bone surgical implant testing (16) and the characterisation of the trabecular bone phenotype of genetic dwarfism (17), of the primate mandible (18), of the human tibia (19), and of animal models of osteoarthritis (20), and fuel cell performance (21; 22)). Apart from the original critique of SMI (14), as far as we know, there have been no further reports of EF in osteoporotic samples in the literature.

In this study, we expand on our two previous studies on the use of EF and the putative plate-to-rod transition in osteoporosis (14; 15). Specifically, we present new EF data on trabecular bone from an animal model of disuse osteoporosis (loss of bone mass as a consequence of reduced or altogether removed loading of the bone) as well as from human second lumbar (L2) vertebral bodies from women of varying age and bone volume fraction. The aim of the study is to investigate the association between variables describing the trabecular architecture (EF and SMI) and bone health. Our EF data relies on an updated and validated implementation of EF. This implementation is written for ImageJ2, and entails an additional way to seed new ellipsoids, and the ability to average EF over several runs. We also report the effect of input parameters more in-depth (details in Supplementary Material (a)). EF is available freely as part of the latest BoneJ, a collection of ImageJ plug-ins intended for skeletal biology (23; 24).

2. Methods

(a) EF algorithm

The EF algorithm was first reported in a previous study (15) and is explained here again due to its fundamental relevance to the present study.

EF is a scalar value assigned to each foreground pixel in the three-dimensional binary image stack of interest. The EF of each pixel depends on the maximal ellipsoid that contains the pixel and that is contained in the image foreground. Denoting the semi-axis lengths of the maximal ellipsoid as a, b and c (with $a \leq b \leq c$), EF of each pixel is calculated as a difference of sorted semi-axis ratios

$$EF = \frac{a}{b} - \frac{b}{c}$$

EF is confined between -1 and 1, with -1 being very plate-like, and 1 very rod-like 1.

Ellipsoid Factor is calculated by fitting locally maximal ellipsoids into the image foreground, then iterating over the foreground pixels to find the largest ellipsoid in which each pixel is contained. Note that the locally maximal ellipsoid is generally non-unique (Supplementary Material ii).

(i) Ellipsoid fitting

First, points where a small sphere can start to grow ("seed points") are determined. Two strategies for finding seed points are provided. The first is called distance-ridge based seeding. It involves subtracting the results of a morphological opening and a closing operations on the distance transform of the input image from each other. The second is a topology-preserving skeletonisation (25). Calculating distance-ridge based seeds is computationally more efficient than skeletonisation in practice, but may overestimate the number of seeds needed to fit a particular region. However, it may miss thin features that skeletonisation preserves well (Figure 2).

After being seeded, each spherical ellipsoid grows uniformly by one user-defined increment at a time until a number of surface points equal to the user-defined "contact sensitivity" parameter hit the trabecular bone boundary (a background pixel). Surface points are chosen from a random uniform distribution on the ellipsoid surface.

When the growing ellipsoid hits the trabecular bone boundary for the first time, the vector from the ellipsoid centre to the average contact point is set as the first ellipsoid semi-axis and the ellipsoid is contracted slightly. Growth of the ellipsoid then continues in the plane orthogonal to this first semi-axis, again until the boundary is hit. This initial ellipsoid fitting is followed by a series of small random rotations, translations and dilations of the ellipsoid in an attempt to find a larger ellipsoid in the local region. Further growth directions at this point are recalculated at each iteration to be random, but in the plane defined by the mean contact vector. These attempts end if no increase in volume of the ellipsoid is found after a user set maximum number of iterations (default 50, see supplementary material (b)), or if the total number of attempts exceeds ten times the maximum iteration number. If more than half of the sampling points on the ellipsoid are outside the image boundary it is invalid, removed and ignored in further calculations.

(ii) Assign EF to each pixel and averaging over runs

Once maximal ellipsoids are found for each seed point, each foreground pixel is assigned the EF value of the largest ellipsoid that contains it, or NaN (not a number) if no ellipsoids contain that pixel. One iteration of fitting ellipsoids and assigning EF to each pixel is termed a *run*.

Ellipsoid factor is a stochastic process and therefore results can vary from run to run. The user has the option to average the outputs over several runs to smooth the results. To track how this average evolves across runs, EF reports the "filling percentage" (percentage of trabecular volume pixels that are contained by at least one ellipsoid) as well as the maximum and median EF change. The latter are estimates of how pixel EF values changed by adding the latest run. From experience on various real-life examples, we recommend averaging over 6 runs (the "repetitions" input parameter) for the final result generation. This typically reduces the median and maximum

EF change per pixel per run to less than 0.15 and 0.4, respectively (See supplementary material (c)).

(iii) EF inputs and outputs

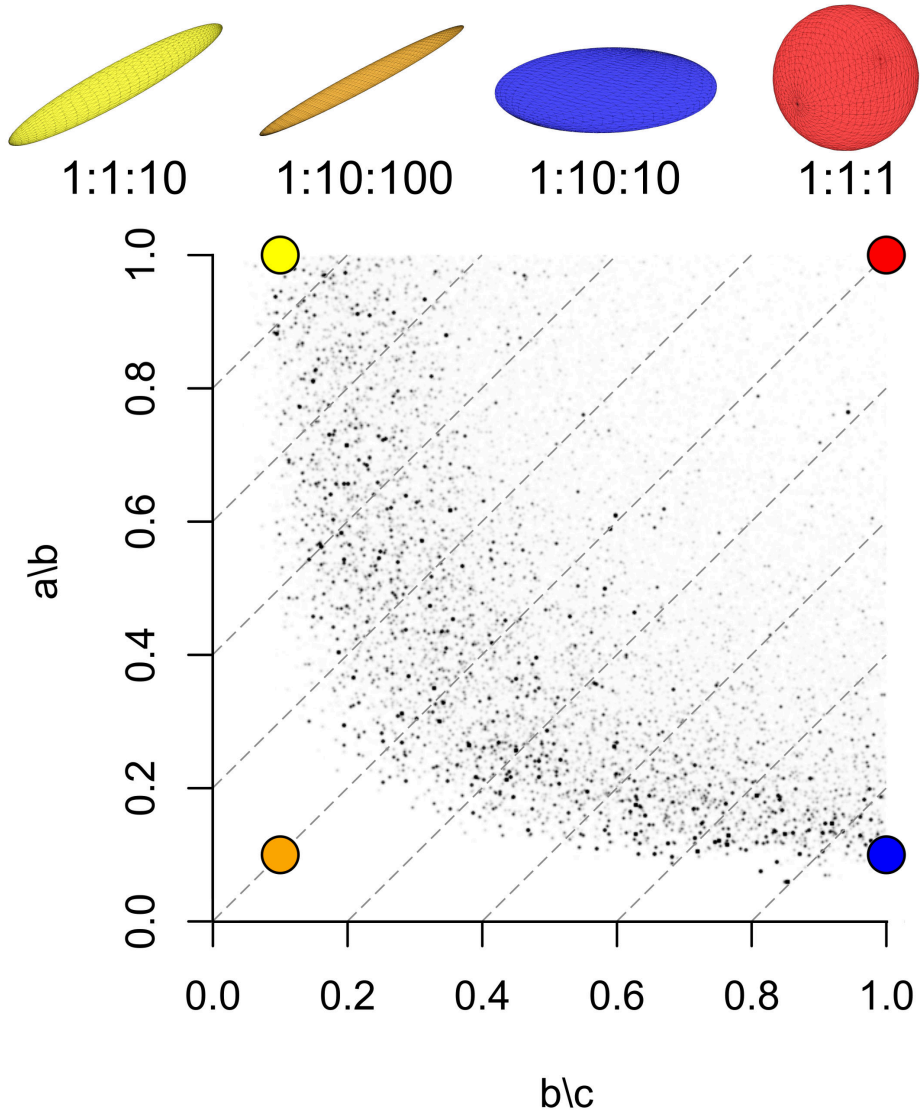

Figure 1. Ellipsoid Factor is calculated as the difference of semi-axis ratios $EF = \frac{a}{b} - \frac{b}{c}$, where $a \leq b \leq c$ are the semi axis lengths. This figure shows edge cases of possible maximal ellipsoids (top row from left: "javelin" (yellow), "surfboard", (orange), "discus" (blue), "tennis ball" (red)), their semi-axis ratios (as a:b:c), and where a pixel (large, colourful points) within such an ellipsoid would be registered on the Flinn peak plot (the Flinn peak plot shows the semi-axis ratios of each maximal ellipsoid plotted against each other). Note that the orange and the red ellipsoid have the same EF, but vastly different Flinn peak point locations. Small black points are the Flinn peak plot data from the trabecular bone of a great spotted kiwi (*Apterix hastii*)(26). Ellipsoids with the same EF value, i.e. EF isolines, are represented by the grey, dashed diagonal lines with slope 1 on the Flinn plot. Flinn peak plots of trabecular bone typically exhibit a "boomerang" shape.

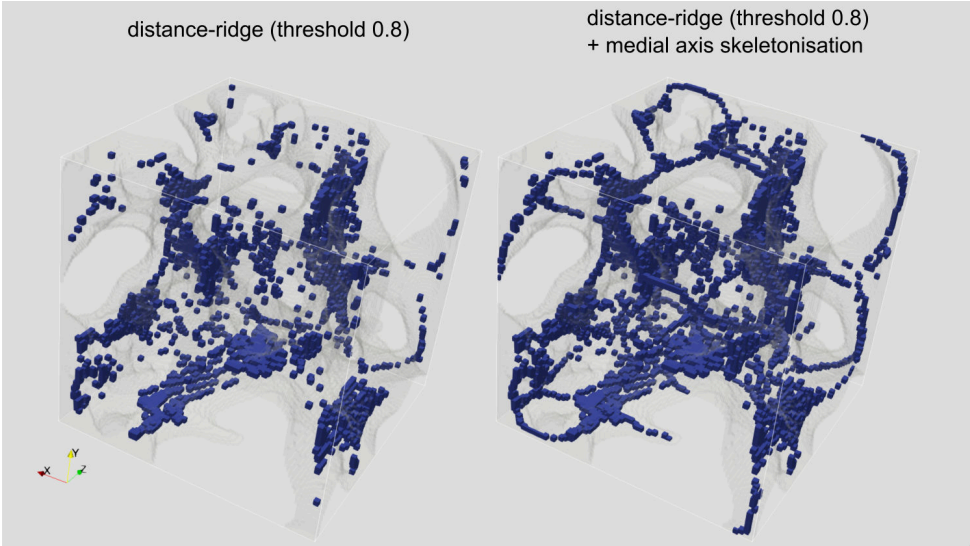

Figure 2. The trabecular surface (transparent grey) from a lesser dwarf shrew (*Suncus varilla)* femur (a test image from a previous Ellipsoid Factor study (15)) and seed points (blue) from the distance ridge (left) and from the distance ridge and the skeletonisation (right). The skeletonisation seems likely to fit thin regions better, but the distance ridge allows to seed more ellipsoids in thick regions, which can lead to higher filling percentages overall, but may also seed more ellipsoids than necessary.

Some further mathematical considerations on the shape of the distributions to be expected when calculating a difference of semi-axis ratios can be found in Supplementary Material ((d)).

In the present study, we ran Ellipsoid Factor on two data sets, with sample descriptions and statistical analysis detailed in the next two subsections. EF input parameters used for each of these studies are listed in Table 4. For both studies, we measured BV/TV and SMI, calculated descriptive statistics of the EF distribution (median, maximum, and minimum), and plotted EF histograms.

(b) Disuse osteoporosis in mouse tibiae

X-ray microtomography (XMT) scans (5 µm nominal pixel spacing) of 12 murine tibiae were obtained from an unrelated study (27) (in preparation). The animals had undergone sciatic neurectomy to the right hindlimb, inducing one-sided disuse osteoporosis. They were divided into three groups of four mice. Group 1,2, and 3 were euthanised 5, 35, or 65 days after surgery, respectively. Trabecular bone from the proximal metaphysis was segmented by drawing around the trabecular-cortical boundary using the software CTan (Bruker, Belgium).

The segmented images were denoised using a 3D median filter (radius 3 pixels) and thresholded at a pixel value of 75 (Figure 3). The thresholding value was selected visually as sensible on one sample and kept consistent across samples. As the EF distributions were unimodal and not normal in all cases, the EF median, maximum and minimum were taken as representative values for each specimen. SMI values were computed for each sample (using Hildebrand and Rüegsegger's method (7) with volume resampling 2 and mesh smoothing 0.5) using Legacy BoneJ 1.4.3 (28)(`https://github.com/bonej-org/bonej` - SMI has been discontinued by Modern BoneJ), and bone volume fraction measurements (calculated using CTan) were taken from the raw data of an unrelated study (27) (in preparation).

For each group, paired t-tests comparing EF descriptive statistics (median, maximum and minimum), SMI and bone volume fraction between control and disuse leg were performed using

the R software (29). We also performed Pearson's product-moment correlation tests for association between EF median and bone volume fraction, and between SMI and bone volume fraction across all groups (using R's `cor.test()` function). The R scripts used for this purpose can be found in an online repository (30) under `/R/paired-mouse-disuse-test.R`.

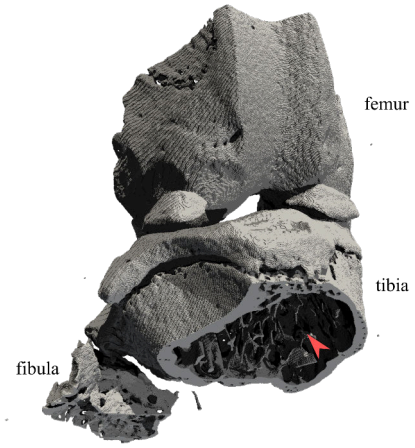

Figure 3. The right knee of one of the mouse samples rendered from a binary image. Red arrowhead points to the trabecular bone in the region of interest for this study. View is cranio-caudal with an oblique tilt towards proximal.

(c) Ellipsoid Factor in human vertebrae of varying trabecular bone volume fraction

To investigate the association of SMI and EF with human bone health, we imaged sagittal sections of 22 vertebrae from women of varying age (24-88 years old) using XMT ($30\,\mu\mathrm{m}$ pixel spacing). Pixels with a linear attenuation coefficient of more than $0.7\,\mathrm{cm}^{-1}$ were classified as bone, others as background. Cuboidal regions of interest containing trabecular bone, aligned with the image axes, were chosen manually. The vertebrae were originally collected and prepared for imaging with scanning electron microscopy in a previous study (31). This data set was interesting to the present study for two reasons. Firstly, these are the first EF numbers obtained on healthy and osteoporotic samples from humans. Secondly, they constitute a challenge for choosing reasonable EF input parameters because they are close to the resolution limit at which we can expect EF to fit the local shape well (trabecular thickness is approximately 5-8 pixels in these images). We additionally report mean and maximum trabecular thickness (Tb.Th [mm]) in these samples, in order to be able to gauge whether a change in EF might be attributed to badly-resolved trabeculae in samples with e.g. low bone volume fraction.

The age distribution of our vertebral samples was non-normal, as it was skewed to the left by the prevalence of older samples (Shapiro-Wilk test $p < 0.05$). We therefore performed a non-parametric test of association of age with bone volume fraction. All other variables of interest (bone volume fraction, EF Median, EF Maximum, EF Mode, EF Minimum, SMI, SMI+, SMI-, mean Tb. Th., maximum Tb. Th) could be assumed to follow a normal distribution (Shapiro-Wilk test $p > 0.05$). As a consequence, we used Pearson's r^2 as a measure of association between these variables and bone volume fraction in our statistical correlation tests. All statistical analysis of the vertebral samples was based on a custom script (available at (30) under `R/histo-EF-stats-vertebrae-final.R`) using the R programming language (29).

Parameter name	Description	Default values
Number of sampling vectors	Number of sampling directions used to search for contacts with the boundary.	100
Sampling increment	Increment for vector searching in pixel units.	1/2.3
Skeleton points per ellipsoid	Number of skeleton points per ellipsoid. Sets the granularity of the ellipsoid fields.	10
Contact sensitivity	Number of sampling vectors in contact with surface required to be classified as a collision.	1
Maximum iterations	Maximum fitting iterations to try improving ellipsoid fit before stopping.	50
Maximum drift	Maximum distance ellipsoid may drift from seed point. Defaults to unit pixel diagonal length	1
Repetitions	Number of separate runs over which to average EF value	1
Seed points (distance ridge)	Seed ellipsoids based on the foreground distance ridge	yes
Seed points (topology-preserving)	Seed ellipsoids based on topology-preserving skeletonisation	no
Show secondary images	Display secondary images (volume, semi-axes, semi-axis ratios, Flinn Plot).	no
Show convergence data	Display convergence data for 2 runs or more.	no

Table 1. List of EF input parameter names, brief descriptions and default values, as listed in the ellipsoid factor documentation. We suggest users to record and publish their values for these parameters as well as the BoneJ version used to enhance the reproducibility of their experiment (See Table 4 for some examples). Note that although the defaults for "skeleton points/ellipsoid" and "repetitions" are 10 and 1 respectively, these values represent "good" values to get a quick overview on an example image. Once ready to run on an entire data set, we recommend setting these to 1 and 6, respectively (See Supplementary Material (c)).

Min EF	Minimum of sample EF distribution.
Max EF	Maximum of sample EF distribution.
Median EF	Median of sample EF distribution.
Filling percentage	Percentage of foreground filled by at least one valid ellipsoid.
Number of ellipsoids found	Total number of valid ellipsoids fitted into trabecular foreground.
Median change n	Median change in EF value from run (n-1) to run (n). This indicates how well the EF algorithm converged.
Maximum change n	Maximum change in EF value from run (n-1) to run (n). This indicates how well the EF algorithm converged.

Table 2. Values written by EF into the ImageJ results table. Median change n and Maximum change n values are only shown if the "Show convergence data" input box is ticked and EF is averaged over at least 2 runs.

EF image	Image containing EF values for foreground pixels
Seed image	Binary image with ellipsoid seed points in foreground
Volume image	Image containing the volume of the locally maximal ellipsoid
ID image	Image containing the index in sorted ellipsoid list
a image	shortest semi-axis of locally maximal ellipsoids
b image	intermediate semi-axis of locally maximal ellipsoids
c image	longest semi-axis of locally maximal ellipsoids
a/b image	a/b semi-axis ratio image
b/c image	b/c semi-axis ratio image
Flinn peak plot	Plot of semi-axis ratios of locally maximal ellipsoids (y-axis: a/b, x-axis: b/c)
Flinn plot	Plot of semi-axis ratios of all (not necessarily maximal) ellipsoids fitted

Table 3. EF primary (above line) and secondary (below line) output images, with brief descriptions.

study subject	mouse tibiae	human vertebrae
description in Methods	(b)	(c)
number of vectors	100	100
sampling increment	1/2.3	0.1/2.3
seed points per ellipsoid	1	1
contact sensitivity	1	5
maximum iterations	50	50
maximum drift	1	1
number of runs	6	6
average of largest n	1	1
seed points (distance ridge)	yes	yes
seed points (topo.-preserv.)	yes	yes

Table 4. EF input parameters used for the two case studies presented in this article.

3. Results

All distributions of EF observed in images of bone were uni-modal, as seen in the histograms of Figures 4 and 13. As described earlier, we used the median, maximum and minimum (and the mode, for the vertebrae) of the distribution as a representative value to describe the distributions of local shape in these images for statistical analysis.

(a) Disuse osteoporosis in mouse tibiae

EF images and histograms for our murine samples can be seen in Figures 5 and 4, respectively. Paired one-sided t-tests (n=4) showed BV/TV and SMI values were significantly different between disuse and control limbs at the 5% level between control and disuse groups at all time points (Figures 6 and 7). Minimum and maximum EF were not statistically associated with disease state ($p > 0.05$) at any time point (Figure 8). There was no link between EF median and disuse at 5 days (paired one-sided t-test, $p > 0.05$) and 35 days (difference not normally distributed (Shapiro-Wilk $p < 0.05$), paired one-sided Wilcoxon rank sum test, $p = 0.06$), but there was a statistical difference at 65 days ($p < 0.05$). Unlike SMI, these measurements suggest therefore the presence of a small shift of about EF 0.1 occurred only after a large amount of bone had already been lost. Over all time points, there was a considerably less strong, and less statistically significant, relationship between bone volume fraction and EF median (Pearson's $r^2 = 0.25$, $p < 0.05$) than between bone volume fraction and SMI (Pearson's $r^2 = 0.81$, $p < 0.001$, Figure 9). The R-script used to perform this analysis can be found under `/R/mouse-smi-tests.R` in (30). EF filling percentage was higher than 90% for all our murine samples, although significantly differed between disuse and control at all time points (Paired t-test, $p < 0.05$).

(b) Ellipsoid Factor in human vertebrae of varying trabecular bone volume fraction

Filling percentages ranged from 74% to 97% and median change in EF between the two final runs ranged from 0.1 to 0.17 (Figure 10), which suggests EF reached a reasonable level of convergence at this point. Correlation tests showed that there was no association ($p > 0.05$) between bone volume fraction and any of the three convergence variables median change, maximum change and filling percentage, indicating that the EF algorithm did not preferentially fill the trabecular bone more completely or in a more stable way in samples with relatively low or high bone volume fraction. This was evidence for a satisfactory convergence of the EF algorithm, albeit not as complete as in the murine samples.

There was a negative association between BV/TV and age (Spearman's $\rho = -0.58$, $p = 0.004$), but not between BV/TV and mean or maximum trabecular thickness ($p > 0.05$). The latter result suggests that EF measurements are unlikely to be affected differentially by the resolution of the trabeculae in samples with lower bone volume fraction. It might also imply that lower bone volume fraction is not associated with thinner "trabeculae" (but rather "fewer"), although this was not a focus of the present study.

SMI, SMI+ and SMI- were strongly, negatively and significantly associated with bone volume fraction: Values for Pearson's r^2 were 0.47, 0.42, and 0.53, respectively, while p-values were all < 0.005 (Figure 11). SMI ranged from 1.36 to 3.11.

Median, maximum and minimum EF were not associated with bone volume fraction ($p > 0.05$, Figure 12), and there was a mild negative association between bone volume fraction and EF modal value ($r^2 = 0.2$,$p = 0.03$). Histograms of the EF distribution were occasionally skewed in either direction across all values for bone volume fraction (Figure 13). Sometimes similar EF values clustered in one region of the vertebra, while in other cases, a range of EF values could be found in all anatomical regions considered. Figure 14 shows EF images for 20 of the 22 vertebrae we analysed.

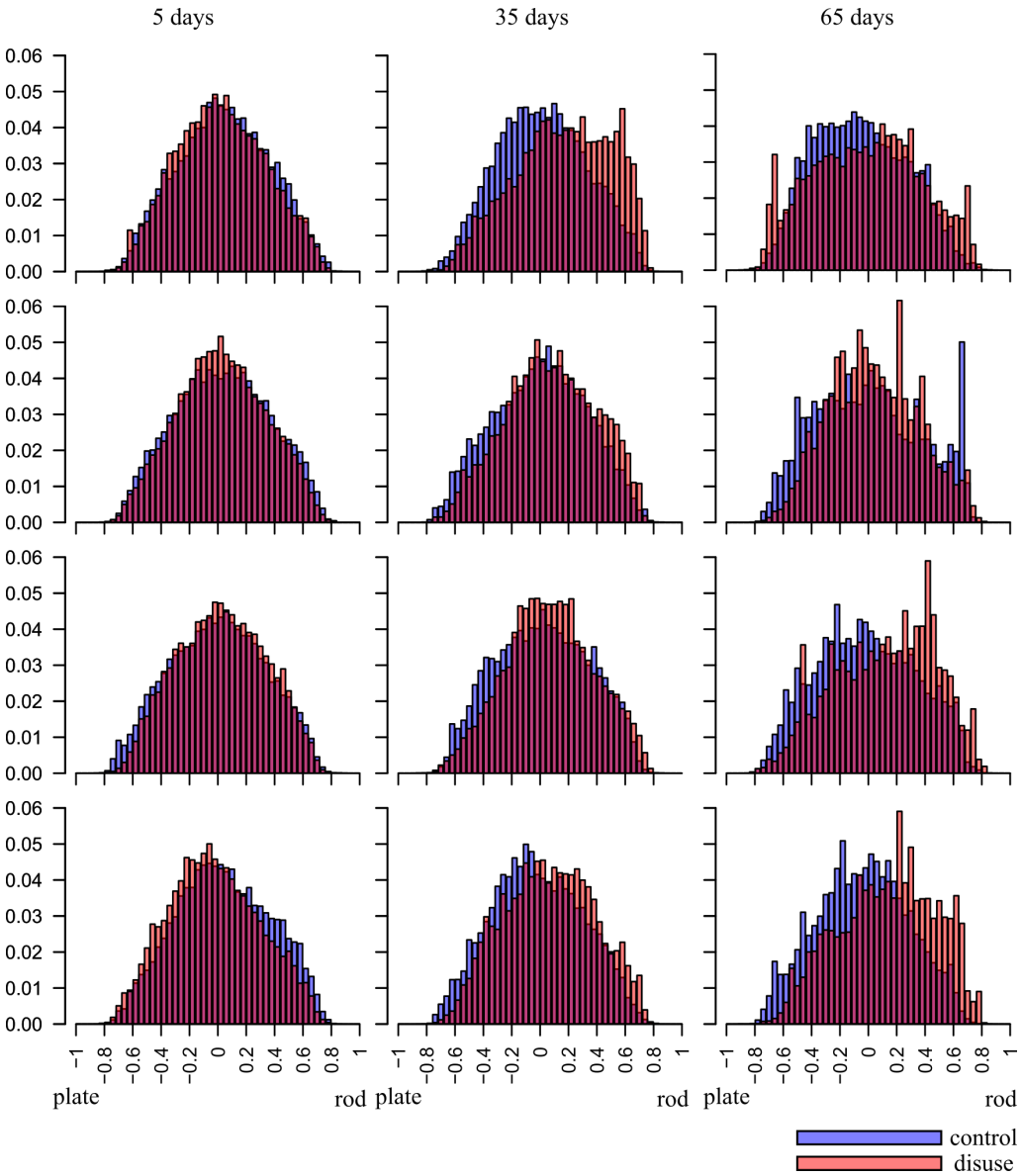

Figure 4. EF frequency histograms for each mouse, at 5 (left column), 35 (central column) and 65 (right column) days post-surgery. Large parts of the control (blue) and disuse (red) histograms overlap (purple). Paired t-tests on EF median suggest a subtle plate-to-rod-transition at 35 and 65 days, but no plate-to-rod transition despite significant bone loss at 5 days.

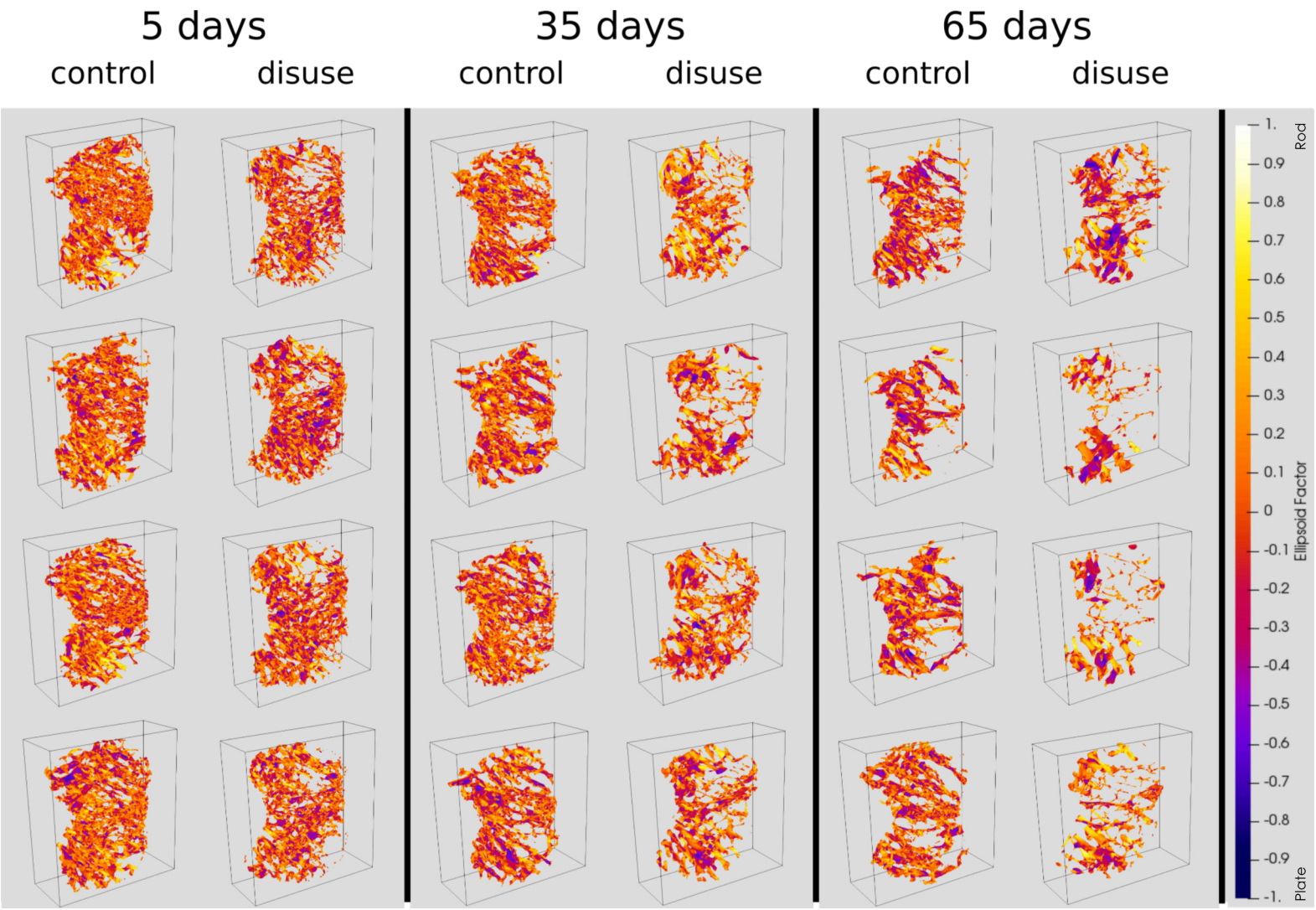

Figure 5. EF renders for each mouse limb over the three time points. The positive Y-direction is towards posterior and the positive z-direction is towards proximal. The X-direction is towards lateral for the control limbs, and towards medial for the disuse limbs. Bone loss appears to start medially.

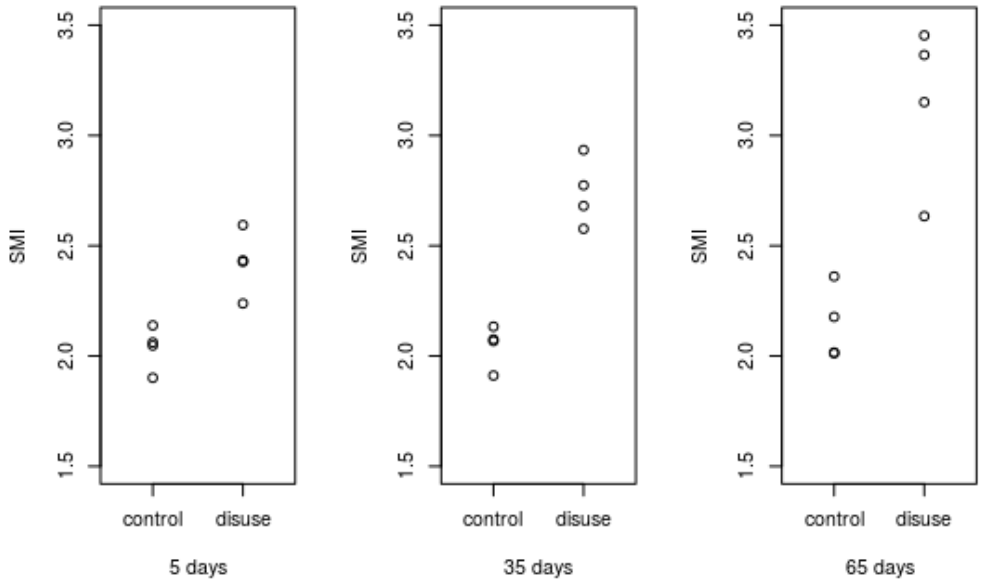

Figure 6. SMI at each time point for control and disuse mouse limbs. SMI is significantly different between control and disuse limb at all time points ($p < 0.05$).

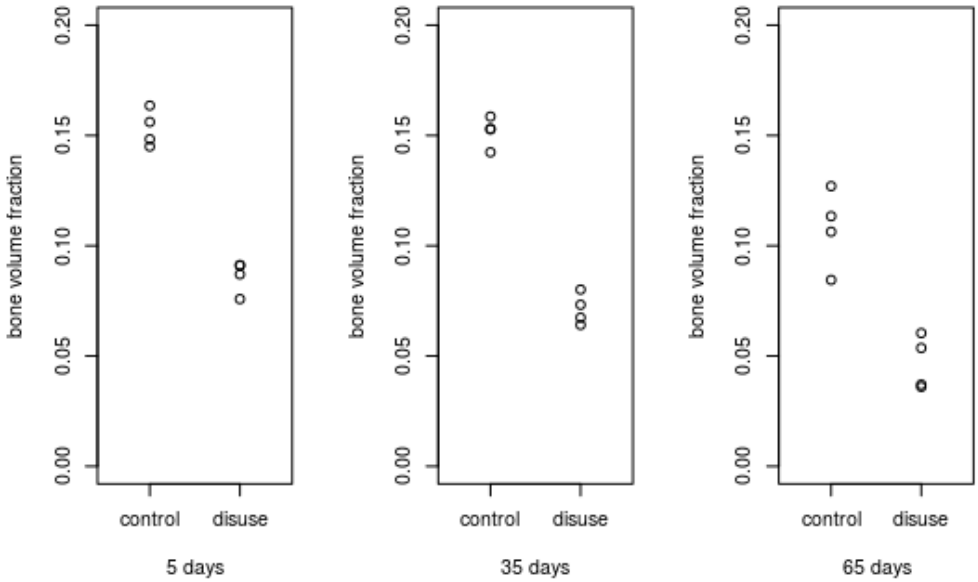

Figure 7. Bone volume fraction at each time point for control and disuse mouse limbs. Bone volume fraction is significantly different between control and disuse limb at all time points ($p < 0.05$).

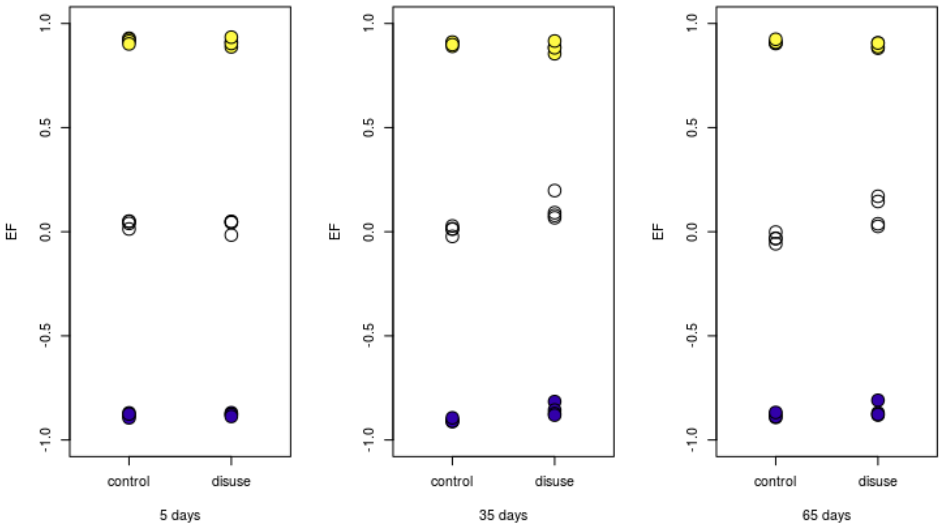

Figure 8. EF median (white), maximum (yellow) and minimum (blue) at each time point for control and disuse mouse limbs. EF median is significantly different ($p < 0.05$) between disuse and control limbs only at 65 days.

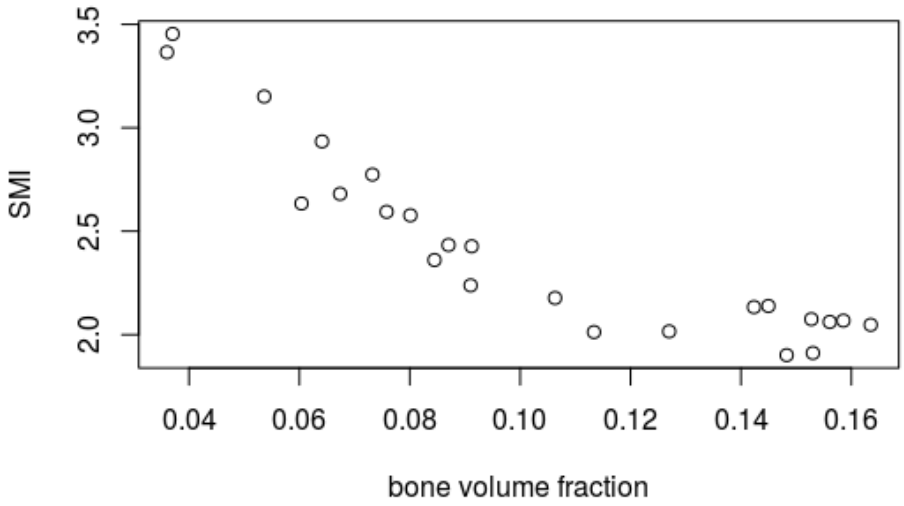

Figure 9. SMI value plotted against bone volume fraction in murine trabecular bone samples. Bone volume fraction and SMI values are strongly, negatively correlated (Pearson's squared moment-product correlation $r^2 = 0.81$, $p < 0.001$).

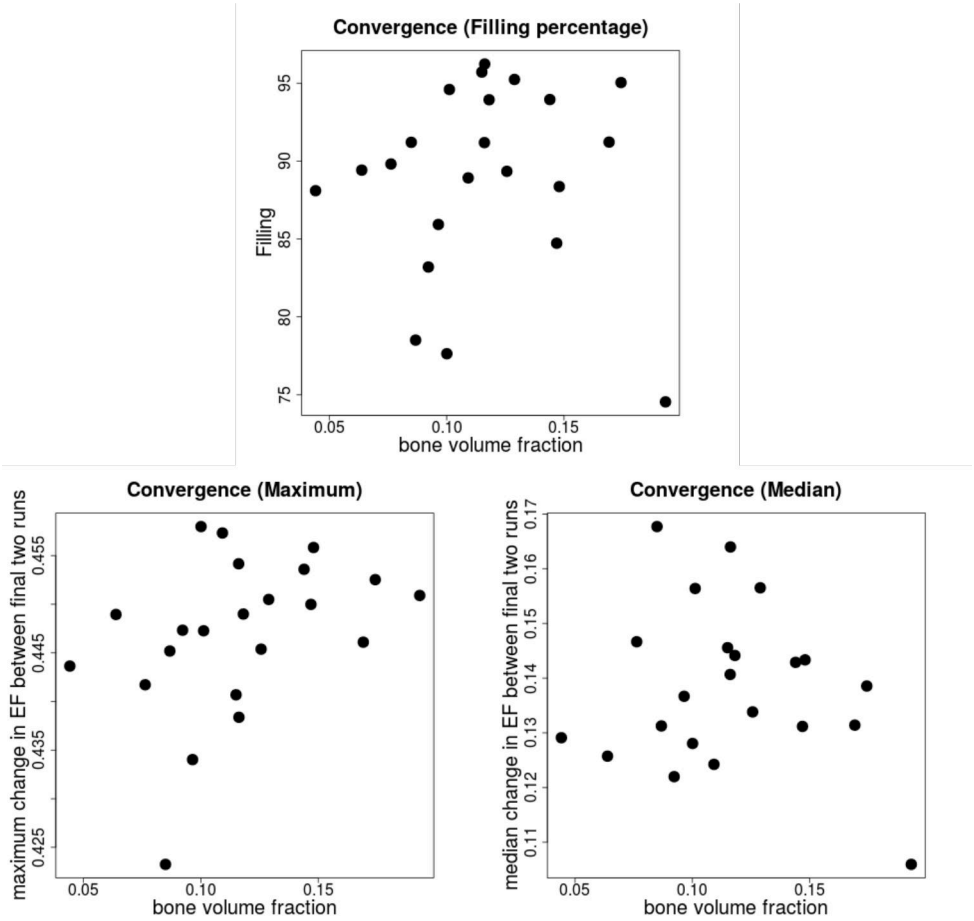

Figure 10. EF convergence parameters (Filling percentage, maximum and median EF change between two final runs) plotted against bone volume fraction for our human vertebral samples. There were no statistical associations between the variables, showing that all samples were equally likely to have a high filling percentage, independent of volume fraction.

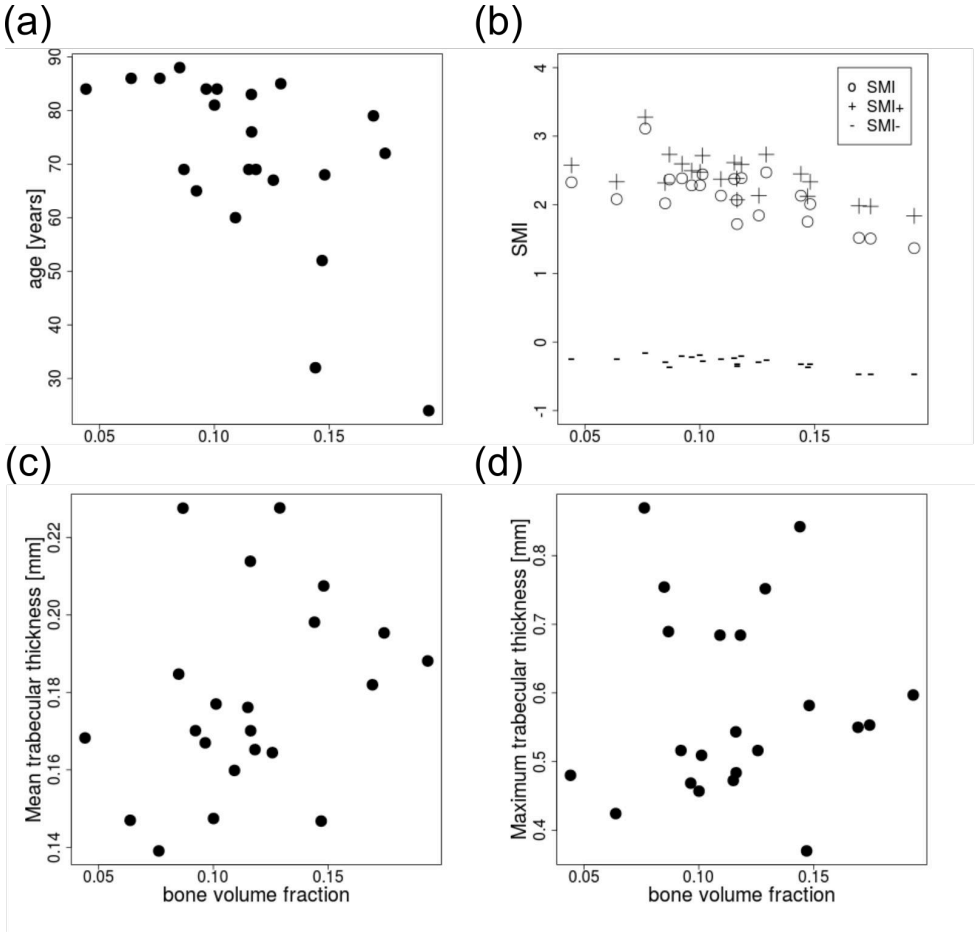

Figure 11. Age, SMI, mean and maximum trabecular thickness versus bone volume fraction for human vertebral samples. Bone volume fraction was correlated with age (a; Spearman's $\rho = -0.58,\ p = 0.004$) and SMI, SMI+ and SMI- (b; Pearson's r^2 is 0.47, 0.42, and 0.53, respectively; $p < 0.005$). Mean (c) and maximum (d) Tb.Th. did not correlate with bone volume fraction ($p > 0.05$).

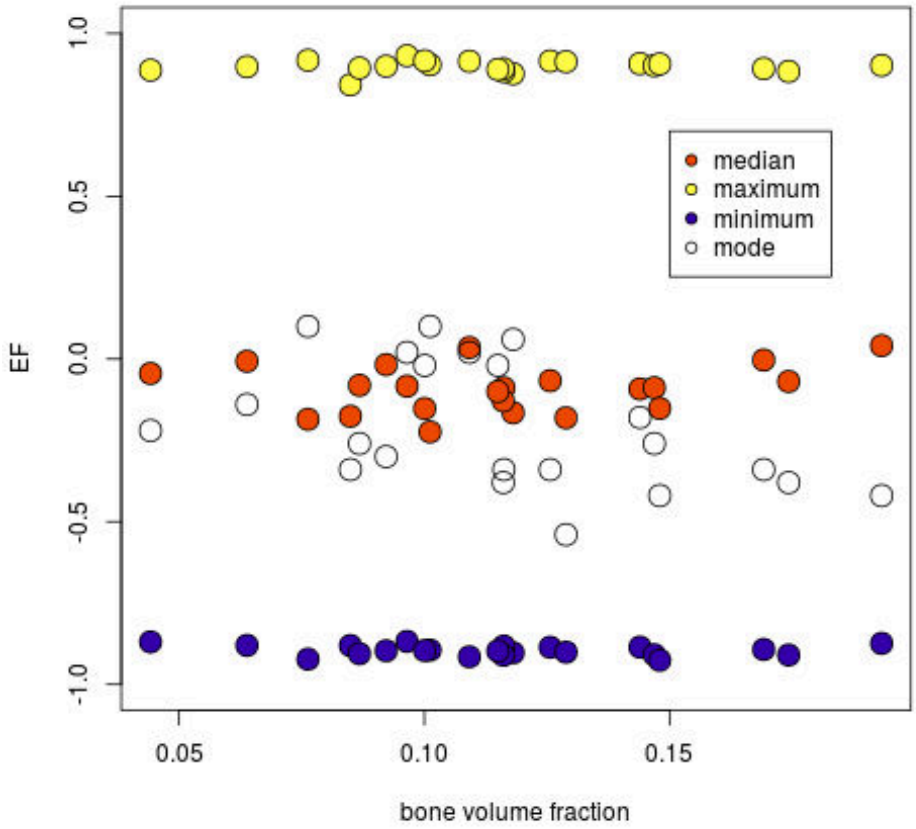

Figure 12. EF distribution parameters plotted against bone volume fraction in human vertebral samples. Only the mode of the distribution was mildly associated with bone volume fraction; Median, maximum and minimum were not.

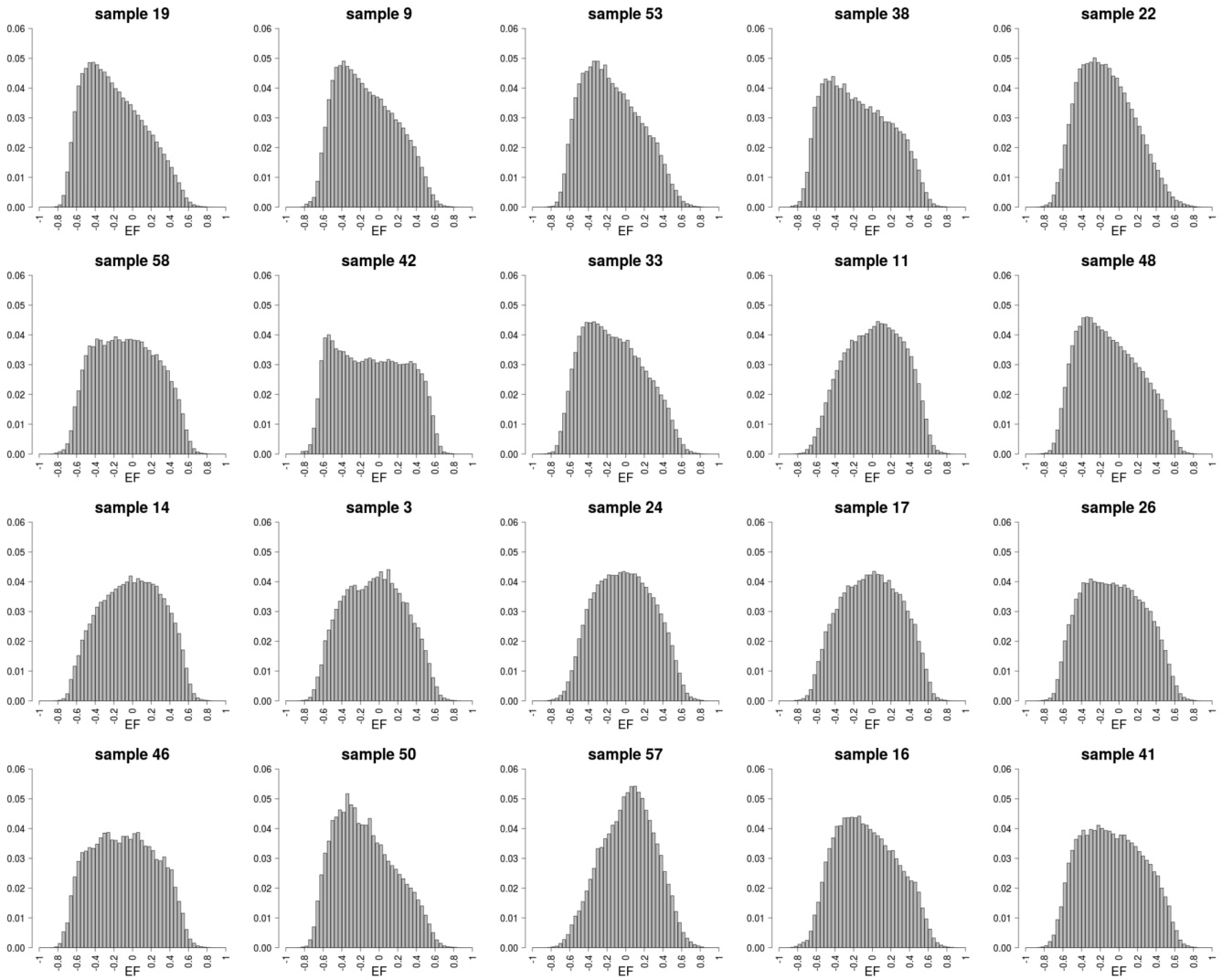

Figure 13. EF histograms of human vertebral trabecular bone samples, sorted from top-left to bottom-right by bone volume fraction. Images of the samples are shown in the same order in 14. EF histograms are all unimodal, and skew in either direction. The mode, but not the median, of the EF distribution was associated with bone volume fraction. Samples 49 and 12 not shown due to the lack of a reasonable way to arrange 22 figures on a page (there were several other samples with the same volume fraction), but were included in the analysis.

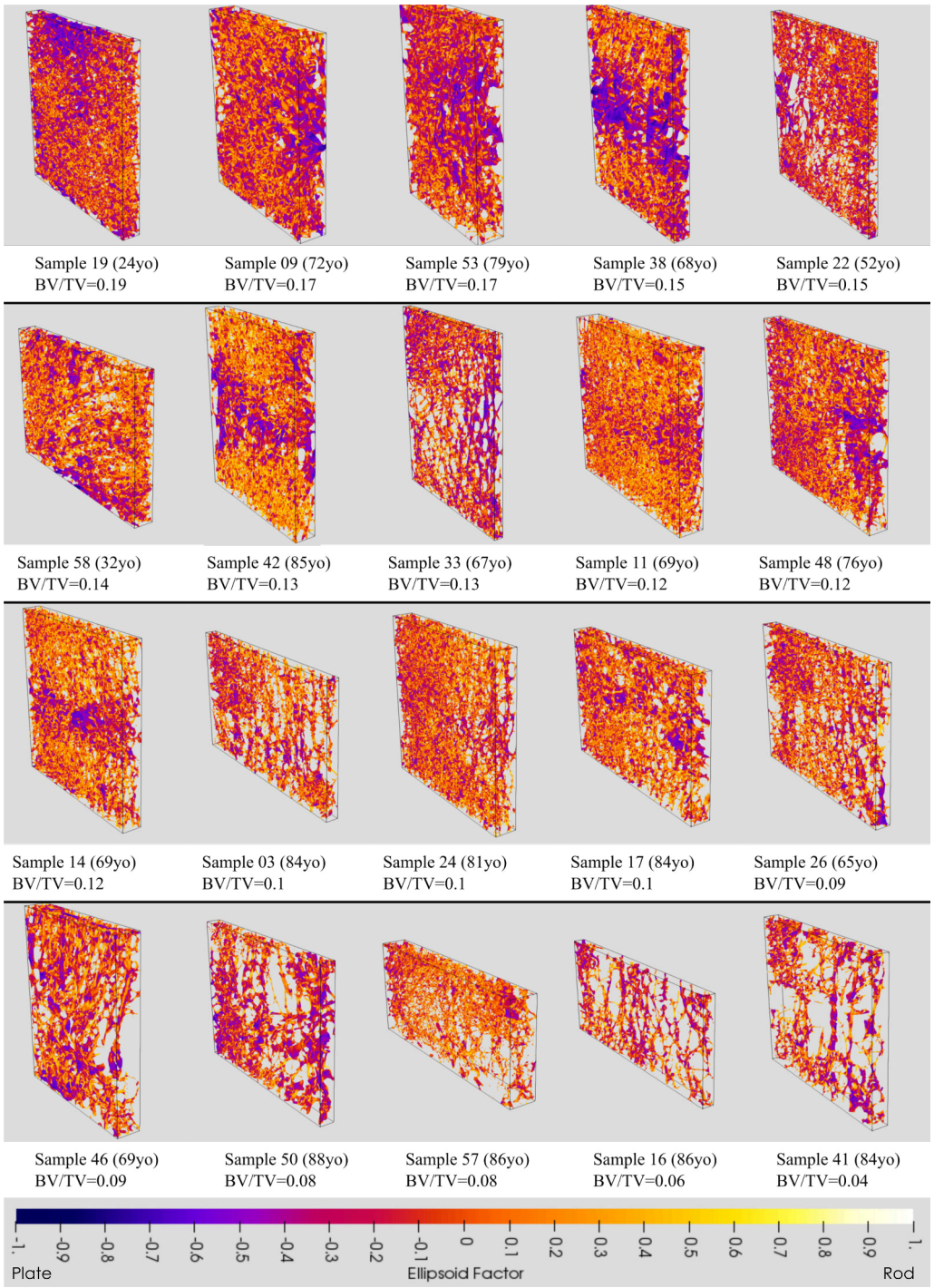

Figure 14. EF image renders of the vertebrae studied, sorted from top-left to bottom-right by bone volume fraction. Yellow pixels indicate a more rod-like, and blue pixels a more plate-like local shape, with orange indicating a shape on a continuum between plates and rods. Histograms of the EF values are shown in the same order in 13. All samples display a range of EF values. In some samples, pixels of similar EF value seem to cluster in the same region (e.g sample 42), while in others there seems to be a mix of EF values in all regions (e.g. sample 33). Samples 49 and 12 not shown due to the lack of a reasonable way to arrange 22 figures on a page (there were several other samples with the same volume fraction), but were included in the analysis. Slices are parasagittal with superior towards the top of the images.

4. Discussion

We measured Ellipsoid Factor distributions in trabecular bone from healthy and unloaded mouse tibiae and from human vertebrae. Only on some occasions, EF supported the presence of a small shift towards a more rod-like geometry linked with decreases in bone volume fraction. SMI, on the other hand, suggested the presence of a drastic plate to rod transition whenever a difference in bone volume fraction was found. EF distributions in the samples from both species we investigated in the present study were consistently uni-modal.

In the murine samples, bone loss happened shortly after surgery in one condyle, but EF median changed only later during disease progression. This suggests that local shape changes in the trabecular bone may be delayed with respect to the initial loss of bone. The strong interdependence between SMI and bone volume fraction is misleading in this case, as it supports an immediate change in local trabecular shape that implies a geometry that is more rod-like than a perfect rod (SMI>3, tending towards spherical where SMI = 4) at the latest time point. EF median was mildly and only just significantly associated with bone volume fraction across all samples. We interpret this as a possible subtle tendency towards a more rod-shaped local geometry in some samples, which would be impossible to glean from observing SMI alone. Minimum and maximum EF values are not different in healthy and osteoporotic murine samples, underlining that very plate- and very rod-like structures co-exist in all samples.

Similarly, in the human vertebra samples, only the mode of the distribution correlated with bone volume fraction, highlighting that any changes in local shape linked to a decrease in bone volume fraction are subtle. Considerable variability in local shape can be seen in the EF images of the vertebral samples. Some of the samples agree with the results of a descriptive anatomical study of human 4th lumbar vertebral bodies, which characterised the trabecular geometry as central plates and braces surrounded superiorly and inferiorly by a honeycomb of rods (32).

In this study, we further presented some recommendations for suitable default parameters for EF (Table 1), based on the convergence behaviour of EF reported in Supplementary Material (c).

(a) What is the mechanical relevance of plates and rods in cancellous bone?

Considering trabecular bone a "cellular solid" (33) gave rise to the idea that plates and rods contribute to mechanical performance. Theoretical, idealised models of open-cell and closed-cell porous solids predicted a dependence of the stiffness and strength on the square and the cube of the characteristic length r, respectively. In a seminal study for the concept of rods and plates in trabecular bone, Gibson analysed previous data from this perspective and showed that these models were consistent with a transition from open-cell to closed-cell mechanical behaviour at a bone volume fraction of 0.2 (34). This is further evidence that attempting to measure rods and plates in trabecular bone is not independent of the amount of bone present (contrary to what was stated in the original SMI study (7)). The bone volume fraction in our samples was below 20%, where the influence of hyperbolic parts of the surface and negative SMI are less than in samples with greater BV/TV (14), so it would be interesting to compare EF in samples with bone volume fraction above and below this value in the future.

The mechanical environment has a strong effect on bone size and shape at an organ and tissue level (e.g. (35; 36; 37), for a review, see (38)), but Frost's mechanostat may not be the main driver of trabecular adaptation within the life of an individual (5). Across species, trabecular bone micro-structure scales as a function of animal size and is likely to behave differently in small animals compared to large animals (39).

Changes in local shape may indicate preferential osteoclastic resorption and/or osteoblastic formation in certain areas of bone. Qualitative descriptions based on scanning electron micrographs of human lumbar vertebrae suggest defective and or slowed bone formation and mineralisation, as well as decoupling of resorption and formation as characteristic of the osteoporotic trabecular geometry at a length scale below the one investigated in the present

study (40). Resorption cavities in human fourth lumbar vertebrae may occur most often near trabecular nodes, with the next most common location plate-like trabeculae (41). The study gives no details on how plates, rods, nodes and "fenestrations" are characterised. It would be interesting to correlate SMI and EF results with such observational studies in the future.

(b) Measures of local shape beyond SMI and EF

Individual trabecula segmentation (ITS) has been proposed as a method to classify the local shape of trabecular bone as rods and plates (42). ITS is based on a decomposition of the trabecular geometry into surfaces and curves (43), with subsequent assignment of all foreground pixels to one of these surfaces and curves based on a measure of vicinity and orientation (44). ITS has been measured in biopsies of hip replacement patients with inter-trochanteric fractures (45). Compared to cadaveric controls, these fracture patients had lower ITS plate bone volume fraction, but equal ITS rod bone volume fraction, as well as lower stiffness moduli and lower overall bone volume fraction (BV/TV). We find it interesting that ITS-measured plate volume fraction correlates with stiffness in these studies. However, we note that ITS-measured axial volume fraction is also (often more strongly) correlated to stiffness than plate volume fraction. It is clear that, at equal bone volume fraction, bone that is less aligned to the direction in which stiffness is measured will behave in a more compliant manner than bone that is more strongly aligned to this direction (46). We therefore suggest that, in the ITS studies, the driving factor for these observations may not be a change in local plate/rod shape, but rather a change in local alignment to the axes in which stiffness is measured. It would be interesting to compare ITS and EF results in the future.

Another method that decomposes trabecular bone into rods and plates was developed but validated only on objects with non-negative Gaussian curvature(47). Applying it to human vertebral samples suggested that three parameters of micro-architecture (two relating to the supposed rod elements) explained 90% of bone stiffness, the same amount of variation in sample stiffness explained by apparent bone volume fraction alone (48). However, all three of these parameters had a significant and strong correlation with bone volume fraction, and this study therefore does not constitute evidence for geometrical changes in the trabecular compartment driving mechanical properties beyond the loss of material. Fatigue failure of trabecular bone may further be related to elements oriented transversely to the main loading direction, which have little effect on stiffness and strength (49).

(c) Limitations and Future work

Ellipsoid Factor is a useful addition to the many geometrical and topological quantities that are routinely measured in trabecular bone, some of which depend on each other, as we have shown here. Ellipsoid Factor is at least designed to be a priori independent of bone volume fraction, the most important descriptor of trabecular bone mechanical properties (46; 50). The lengths of the ellipsoid semi-axes a,b,c as half-thickness, half-width, and half-length trabecular variables could be seen as an extension to measuring trabecular thickness alone.

The fact that EF is generally non-unique (Supplementary material ii) is at least in theory a limitation of EF. We doubt that ellipsoids found by EF in real-life trabecular structures will often be in this situation in practice.

The samples we consider in this paper are cross-sectional, which unfortunately precludes us from following the trabecular architecture of a single individual over time. Ellipsoid Factor, like all other measures of trabecular micro-architecture, requires a sufficient resolution of the individual geometrical features to minimise artefacts such as noise and partial volume effect. Where resolution is insufficient for EF to run on a binarised image, it might be possible to locate the trabecular boundary using fuzzy edge detection (and therefore circumventing the need for precise thresholding), as is done in the tensor scale algorithm (51; 52; 53). The current EF software is designed in such a way as to make an approach based on fuzzy boundary detection

straightforward. Very small trabeculae may be routinely missed by XMT altogether, but dealing with this limitation was outside the scope of this study.

Ellipsoid Factor is a complex algorithm, with several input parameters that need to be tailored to the application. We believe that this is also an advantage in some ways, as it will force users to better understand the methods they are using. We encourage users to ask questions can be asked on the ImageJ forum (https://forum.image.sc/). Despite its complexity, an advantage of EF is that it reduces local shape down to a single number per pixel. Important information on the subtlety of trabecular local shape is lost due to this simplification and users are encouraged to view and interpret the Flinn peak plot because it is a more complete, but more complex, representation of the local shapes present in their sample (Figure 1). The Flinn plot may require more advanced statistics, for 2-D, non-independent response variables, to rigorously compare sampled groups. It might be possible to improve the performance of EF in the future by transferring some parallel computations onto the graphics card (54).

The analyses in the present study focus on median, maximum and median EF, and therefore treat the distribution of local shapes as a whole. This approach does not take into account the orientation of the ellipsoids, or whether pixels of a certain EF cluster around particular anatomical regions. Such considerations might be interesting to explore in the future, as would combinations of EF measurements with other measures of trabecular local shape.

Note that because EF was designed to be independent of scale (as it is a difference of ratios), it may assign the same EF to a "small", mechanically unimportant and a "large", load-bearing trabecula of the same local shape. Features are weighted by their size in the histogram and Flinn peak plot, because each pixel contained by an ellipsoid is assigned the EF of that ellipsoid. Consequently large and mechanically important features containing numerous pixels make a greater contribution to the final analysis, and small features may have a negligible effect. On the other hand, this intentional feature of EF may be useful to detect situations were a pure change in shape (without differences bone volume fraction or local thickness) is responsible for altered mechanical properties. Torres et al. (49) found that the size, shape and orientation of small off-axis members can be critical for the mechanical performance of porous materials. The maximal ellipsoids' EF, orientation and volume could provide this additional information.

Further avenues of future research could investigate how well EF characterises curved trabecular bone, and understanding whether characteristic combinations of semi-axis ratios $\frac{a}{b}$ and $\frac{b}{c}$ for an individual or a group exist that are not immediately recognised by looking at the semi-axis ratio difference, but which may emerge in the Flinn peak plot. Extracting the ellipsoid semi-axis directions might also provide interesting information about the orientation of local shapes. For example, the local orientation may provide clues about the dominant loading environment in different regions of the trabecular compartment.

5. Conclusion

Our investigations suggest that local shape in trabecular bone is not straightforward to decompose into rods and plates, and that a wealth of shapes across the plate-rod continuum exist in any sample. Our data support the presence of a slight tendency of the trabecular geometry to have higher EF in osteoporotic samples, possibly as a consequence of a cell-driven re-organisation that is delayed in respect to the initiation of bone loss. This transition, where it occurs, is considerably more subtle than SMI values suggest.

Acknowledgments

The authors acknowledge the reviewers for their comments that helped improve the manuscript. The authors further thank Phil Salmon (Bruker Micro-CT), Andy Pitsillides (RVC) and the wider RVC Skeletal Biology Group for helpful discussions on trabecular bone, as well as Richard Domander (RVC) and Curtis Rueden (University of Wisconsin-Madison) for valuable help and support in working with ImageJ2. We also thank Yu-Mei Chang for advice with statistical analysis,

and Eva Herbst for critically reading the manuscript. This research was supported by a BBSRC Project Grant to MD (BB/P006167/1).

Author contributions

MD had the idea for Ellipsoid Factor and designed the overall project. AAF and MD implemented the ImageJ code. AAF wrote the custom R, IJ and Python helper scripts for the analysis(30), did the statistical analysis, made the figures, and wrote the manuscript draft. RS performed the surgery and dissected the mouse tibiae. BJ imaged the murine samples. SM segmented the trabecular compartment of the mouse tibiae, and measured bone volume fraction in the mouse samples under the guidance of BJ. DM and AB prepared and imaged the human vertebrae. All authors read the manuscript, provided feedback and approved the final version of the manuscript.

Competing interests

MD was a member of the Editorial Board of Royal Society Open Science at the time of submission and was not involved in the assessment of this submission.

Data availability

Vertebral XMT scans (https://figshare.com/projects/Assessment_of_Bone_Quality_in_Osteoporosis_-_XMT_and_SEM/76962) and segmented XMT of mouse trabecular bone (https://figshare.com/projects/Segmented_trabecularbone_from_microCT_scans_of_mice_with_one-sided_neurectomy_to_hindlimb/79583) are available on the data sharing repository figshare. Some vertebral scans can additionally be explored as 3d renderings on SketchFab: https://sketchfab.com/alexjcb/collections/vertebrae-sections.

The BoneJ source code can be found on Github https://github.com/bonej-org/BoneJ2, with installation instructions at https://imagej.net/BoneJ2#Installation, and is archived within the Zenodo repository (https://doi.org/10.5281/zenodo.1427262). Custom Python, R and IJ Macro scripts (30) for the analyses presented here have also been archived within Zenodo (https://doi.org/10.5281/zenodo.3556577).

Ethics

The image data were obtained and re-used from unrelated experiments in which animal procedures and human samples were used with appropriate ethical approval. The use of animals in the unrelated study was carried out in accordance with the Animals (Scientific Procedures) Act 1986, an Act of Parliament of the United Kingdom, approved by the Royal Veterinary College Ethical Review Committee and the United Kingdom Government Home Office, and followed ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines. Human second lumbar vertebral body samples were obtained via the European Union BIOMED I study "Assessment of Bone Quality in Osteoporosis".

References

1 Williamson S, Landeiro F, McConnell T, Fulford-Smith L, Javaid MK, Judge A, Leal J. 2017 Costs of fragility hip fractures globally: a systematic review and meta-regression analysis. *Osteoporosis International* **28**, 2791–2800.

2 Weisenthal B, Chotai S, Sivaganesan A, Hills J, Devin CJ. 2018 Healthcare burden of osteoporosis. *Seminars in Spine Surgery* **30**, 2–7.

3 Kanis JA. 1996 Estrogens, the menopause, and osteoporosis. *Bone* **19**, 185S–190S.

4 Nielson CM, Marshall LM, Adams AL, LeBlanc ES, Cawthon PM, Ensrud K, Stefanick ML, Barrett-Connor E, Orwoll ES. 2011 BMI and fracture risk in older men: The osteoporotic fractures in men study (MrOS). *Journal of Bone and Mineral Research* **26**, 496–502.

5 Yang H, Xu X, Bullock W, Main RP. 2019 Adaptive changes in micromechanical environments of cancellous and cortical bone in response to in vivo loading and disuse. *Journal of Biomechanics* **89**, 85–94.

6 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. *Journal of Bone and Mineral Research* **25**, 1468–1486.

7 Hildebrand T, Rüegsegger P. 1997 Quantification of Bone Microarchitecture with the Structure Model Index. *Computer Methods in Biomechanics and Biomedical Engineering* **1**, 15–23.

8 Akhter MP, Lappe JM, Davies KM, Recker RR. 2007 Transmenopausal changes in the trabecular bone structure. *Bone* **41**, 111–116.

9 Borah B, Dufresne TE, Chmielewski PA, Johnson TD, Chines A, Manhart MD. 2004 Risedronate preserves bone architecture in postmenopausal women with osteoporosis as measured by three-dimensional microcomputed tomography. *Bone* **34**, 736–746.

10 Glatt M, Pataki A, Evans GP, Hornby SB, Green JR. 2004 Loss of vertebral bone and mechanical strength in estrogen-deficient rats is prevented by long-term administration of zoledronic acid. *Osteoporosis International* **15**, 707–715.

11 Martín-Fernández M, Martínez E, Díaz-Curiel M, Guede D, Caeiro JR, De la Piedra C. 2014 Effects of PTH (1–84) on bone quality in a validated model of osteoporosis due to androgenic deprivation. *The Aging Male* **17**, 42–50.

12 Patsch JM, Kiefer FW, Varga P, Pail P, Rauner M, Stupphann D, Resch H, Moser D, Zysset PK, Stulnig TM, Pietschmann P. 2011 Increased bone resorption and impaired bone microarchitecture in short-term and extended high-fat diet–induced obesity. *Metabolism* **60**, 243–249.

13 Jinnai H, Watashiba H, Kajihara T, Nishikawa Y, Takahashi M, Ito M. 2002 Surface curvatures of trabecular bone microarchitecture. *Bone* **30**, 191–194.

14 Salmon PL, Ohlsson C, Shefelbine SJ, Doube M. 2015 Structure Model Index Does Not Measure Rods and Plates in Trabecular Bone. *Frontiers in Endocrinology* **6**.

15 Doube M. 2015 The Ellipsoid Factor for Quantification of Rods, Plates, and Intermediate Forms in 3D Geometries. *Frontiers in Endocrinology* **6**.

16 Fürst D, Senck S, Hollensteiner M, Esterer B, Augat P, Eckstein F, Schrempf A. 2017 Characterization of synthetic foam structures used to manufacture artificial vertebral trabecular bone. *Materials Science and Engineering: C* **76**, 1103–1111.

17 Colombo A, Stephens NB, Tsegai ZJ, Bettuzzi M, Morigi MP, Belcastro MG, Hublin JJ. 2018 Trabecular Analysis of the Distal Radial Metaphysis during the Acquisition of Crawling and Bipedal Walking in Childhood: A Preliminary Study. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*.

18 Coiner-Collier S, Vogel ER, Scott RS. 2018 Trabecular Anisotropy in the Primate Mandibular Condyle Is Associated with Dietary Toughness. *The Anatomical Record* **301**, 1342–1359.

19 Du J, Brooke-Wavell K, Paggiosi MA, Hartley C, Walsh JS, Silberschmidt VV, Li S. 2019 Characterising variability and regional correlations of microstructure and mechanical competence of human tibial trabecular bone: An in-vivo HR-pQCT study. *Bone* **121**, 139–148.

20 Ketola JH, Karhula SS, Finnilä MAJ, Korhonen RK, Herzog W, Siltanen S, Nieminen MT, Saarakkala S. 2018 Iterative and discrete reconstruction in the evaluation of the rabbit model of osteoarthritis. *Scientific Reports* **8**, 1–10.

21 Zenyuk IV, Parkinson DY, Connolly LG, Weber AZ. 2016 Gas-diffusion-layer structural properties under compression via X-ray tomography. *Journal of Power Sources* **328**, 364–376.

22 Shum AD, Parkinson DY, Xiao X, Weber AZ, Burheim OS, Zenyuk IV. 2017 Investigating Phase-Change-Induced Flow in Gas Diffusion Layers in Fuel Cells with X-ray Computed Tomography. *Electrochimica Acta* **256**, 279–290.

23 Domander R, Felder A, Doube M. 2021 BoneJ2 - refactoring established research software [version 1; peer review: awaiting peer review]. *Wellcome Open Research* **6**.

24 Richard Domander, Michael Doube, Curtis Rueden, Alessandro Felder, Mark Hiner, Jan Eglinger. 2020 bonej-org/BoneJ2: styloid. .

25 Lee TC, Kashyap RL, Chu CN. 1994 Building Skeleton Models via 3-D Medial Surface/Axis Thinning Algorithms. *CVGIP: Graph. Models Image Process.* **56**, 462–478.

26 Doube M, Kłosowski MM, Hutchinson J, Shefelbine SJ. 2018 X-ray microtomography images of trabecular bone from the femoral head and condyle of 18 avian, 72 mammalian and one crocodilian species.. p. 38242272154 Bytes. Artwork Size: 38242272154 Bytes Publisher: Figshare.

27 Javaheri B, Monzem S, De Souza RL, Pitsillides AA. 2020 A mouse model of disuse by sciatic neuroctomy saturates longer in trabecular than cortical bone and deteriorates the cortex along the entire tibial length. *(in preparation)*.

28 Doube M, Kłosowski MM, Arganda-Carreras I, Cordelières FP, Dougherty RP, Jackson JS, Schmid B, Hutchinson JR, Shefelbine SJ. 2010 BoneJ: Free and extensible bone image analysis in ImageJ. *Bone* **47**, 1076–1079.

29 R Core Team R: A language and environment for statistical computing. https://www.R-project.org/.

30 Felder A. 2019 alessandrofelder/EF-helper-scripts: add null case (zenodo). .

31 Boyde A, Jones SJ. 1995 Assessment of Quality of Bone in Osteoporosis: BIOMED I Project. *Clinical Rheumatology* **14**, 596–602.

32 Jayasinghe J, Jones S, Boyde A. 1994 Three-dimensional photographic study of cancellous bone in human fourth lumbar vertebral bodies. *Anatomy and Embryology* **189**.

33 Gibson LJ, Ashby MF. 1997 *Cellular solids: structure and properties*. Cambridge solid state science series. Cambridge: Cambridge University Press 2 edition.

34 Gibson L. 1985 The mechanical behaviour of cancellous bone. *Journal of Biomechanics* **18**, 317–328.

35 Lanyon LE. 1972 In vivo bone strain recorded from thoracic vertebrae of sheep. *Journal of biomechanics* **5**, 277IN7279–278IN8281.

36 De Souza RL, Matsuura M, Eckstein F, Rawlinson SCF, Lanyon LE, Pitsillides AA. 2005 Non-invasive axial loading of mouse tibiae increases cortical bone formation and modifies trabecular organization: A new model to study cortical and cancellous compartments in a single loaded element. *Bone* **37**, 810–818.

37 Carriero A, Pereira A, Wilson A, Castagno S, Javaheri B, Pitsillides A, Marenzana M, Shefelbine S. 2018 Spatial relationship between bone formation and mechanical stimulus within cortical bone: Combining 3D fluorochrome mapping and poroelastic finite element modelling. *Bone Reports* **8**, 72–80.

38 Meakin LB, Price JS, Lanyon LE. 2014 The Contribution of Experimental *in vivo* Models to Understanding the Mechanisms of Adaptation to Mechanical Loading in Bone. *Frontiers in Endocrinology* **5**.

39 Doube M, Klosowski MM, Wiktorowicz-Conroy AM, Hutchinson JR, Shefelbine SJ. 2011 Trabecular bone scales allometrically in mammals and birds. *Proceedings of the Royal Society B: Biological Sciences* **278**, 3067–3073.

40 Jayasinghe JAP, Jones SJ, Boyde A Scanning electron microscopy of human lumbar vertebral trabecular bone surfaces. p. 10.

41 Goff MG, Slyfield CR, Kummari SR, Tkachenko EV, Fischer SE, Yi YH, Jekir MG, Keaveny TM, Hernandez CJ. 2012 Three-dimensional characterization of resorption cavity size and location in human vertebral trabecular bone. *Bone* **51**, 28–37.

42 Liu XS, Sajda P, Saha PK, Wehrli FW, Bevill G, Keaveny TM, Guo XE. 2008 Complete Volumetric Decomposition of Individual Trabecular Plates and Rods and Its Morphological Correlations With Anisotropic Elastic Moduli in Human Trabecular Bone. *Journal of Bone and Mineral Research* **23**, 223–235.

43 Saha PK, Chaudhuri BB, Dutta Majumder D. 1997 A new shape preserving parallel thinning algorithm for 3D digital images. *Pattern Recognition* **30**, 1939–1955.

44 Liu XS, Sajda P, Saha PK, Wehrli FW, Guo XE. 2006 Quantification of the Roles of Trabecular Microarchitecture and Trabecular Type in Determining the Elastic Modulus

of Human Trabecular Bone. *Journal of Bone and Mineral Research* **21**, 1608–1617. _eprint: https://asbmr.onlinelibrary.wiley.com/doi/pdf/10.1359/jbmr.060716.

45 Wang J, Zhou B, Parkinson I, Thomas CDL, Clement JG, Fazzalari N, Guo XE. 2013 Trabecular Plate Loss and Deteriorating Elastic Modulus of Femoral Trabecular Bone in Intertrochanteric Hip Fractures. *Bone Research* **1**, 346–354.

46 Musy SN, Maquer G, Panyasantisuk J, Wandel J, Zysset PK. 2017 Not only stiffness, but also yield strength of the trabecular structure determined by non-linear µFE is best predicted by bone volume fraction and fabric tensor. *Journal of the Mechanical Behavior of Biomedical Materials* **65**, 808–813.

47 Stauber M, Müller R. 2006 Volumetric spatial decomposition of trabecular bone into rods and plates—A new method for local bone morphometry. *Bone* **38**, 475–484.

48 Stauber M, Rapillard L, Lenthe GHv, Zysset P, Müller R. 2006 Importance of Individual Rods and Plates in the Assessment of Bone Quality and Their Contribution to Bone Stiffness. *Journal of Bone and Mineral Research* **21**, 586–595.

49 Torres AM, Trikanad AA, Aubin CA, Lambers FM, Luna M, Rimnac CM, Zavattieri P, Hernandez CJ. 2019 Bone-inspired microarchitectures achieve enhanced fatigue life. *Proceedings of the National Academy of Sciences* **116**, 24457–24462.

50 Maquer G, Musy SN, Wandel J, Gross T, Zysset PK. 2015 Bone Volume Fraction and Fabric Anisotropy Are Better Determinants of Trabecular Bone Stiffness Than Other Morphological Variables. *Journal of Bone and Mineral Research* **30**, 1000–1008.

51 Saha PK, Wehrli FW. 2004 A robust method for measuring trabecular bone orientation anisotropy at in vivo resolution using tensor scale. *Pattern Recognition* **37**, 1935–1944.

52 Xu Z, Saha PK, Dasgupta S. 2012 Tensor scale: An analytic approach with efficient computation and applications. *Computer Vision and Image Understanding* **116**, 1060–1075.

53 Saha PK, Liu Y, Chen C, Jin D, Letuchy EM, Xu Z, Amelon RE, Burns TL, Torner JC, Levy SM, Calarge CA. 2015 Characterization of trabecular bone plate-rod microarchitecture using multirow detector CT and the tensor scale: Algorithms, validation, and applications to pilot human studies: Trabecular bone plate-rod microarchitecture using MD-CT. *Medical Physics* **42**, 5410–5425.

54 Haase R, Royer LA, Steinbach P, Schmidt D, Dibrov A, Schmidt U, Weigert M, Maghelli N, Tomancak P, Jug F, Myers EW. 2020 CLIJ: GPU-accelerated image processing for everyone. *Nature Methods* **17**, 5–6.

55 Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW. 2017 ImageJ2: ImageJ for the next generation of scientific image data. *arXiv:1701.05940 [cs, q-bio]*. arXiv: 1701.05940.

56 Domander R, Doube M, Rueden C, Felder A, Hiner M, Eglinger J. 2018 Bonej-Org/Bonej2: Cuneiform Experimental Patch2. .

57 Boutroy S, Bouxsein ML, Munoz F, Delmas PD. 2005 *In Vivo* Assessment of Trabecular Bone Microarchitecture by High-Resolution Peripheral Quantitative Computed Tomography. *The Journal of Clinical Endocrinology & Metabolism* **90**, 6508–6515.

58 Saha PK, Xu Y, Duan H, Heiner A, Liang G. 2010 Volumetric Topological Analysis: A Novel Approach for Trabecular Bone Classification on the Continuum Between Plates and Rods. *IEEE Transactions on Medical Imaging* **29**, 1821–1838.

59 Bischoff S, Kobbelt L. 2002 Ellipsoid decomposition of 3D-models. In *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on* pp. 480–488. IEEE.

Supplementary material

(a) Implementation and software design

Results presented in this study were obtained with a development version of BoneJ (commit 0ce1c5eba). We have verified that the differences obtained with the initial "styloid" release of BoneJ (24) are approximately 0.15 on average, which is what is to be expected from the stochasticity of the EF algorithm.

The latest Ellipsoid Factor implementation adheres to the principles of modern ImageJ - ImageJ2 (55), dividing the the execution of the algorithm into small, modular and re-usable parts ("ops", in our case: seed-point finding, ellipsoid fitting) combined into a high-level ImageJ plugin ("command", in our case, the Ellipsoid Factor command, part of BoneJ2 (56)). Installation instructions for BoneJ2 can be found at https://imagej.net/BoneJ2#Installation.

The modularity of the seed point finding allows easy switching between the two existing seed point finding strategies (distance-ridge-based and topology-preserving). Similarly, the modularity of ellipsoid fitting allows future extension to e.g. fuzzy edge detection of trabecular surface points on low-resolution grey scale images (which may be relevant for *in-vivo* HR-pQCT images of human trabecular bone with (comparatively) low resolution (57); for a related algorithm, see (51; 53; 58)) or a surface-based ellipsoid fitting strategy (59).

The two modules are decoupled from each other, so a change in one will not affect the usability of the other.

(b) Sensitivity to max iteration parameter

Figure S1 shows that the difference in EF for different values of "max iterations" was small (other input parameters being equal). Numerical experiments showed that for the emu and shrew test images, an ellipsoid that had not improved its volume for approximately 40 iterations was equally likely to not be fitted well than it was to be fitted well. We therefore chose 50 as a reasonable, conservative default value for "max iterations".

(c) Sensitivity to some input parameters

We tested the sensitivity of the EF image on variations in number of seeds (skip ratio), volume-weighted averaging, number of sampling vectors and number of runs. This was done on three test images, already used in the previous EF study (15). The tests showed consistently that adding more seeds was important to achieve a high filling percentage, and that an increased number of runs (about 6) was necessary for good convergence (median change per run < 0.15). More seeds and more runs come at the cost of a longer run-time, however. We therefore recommend setting a low skip-ratio and averaging over 6 runs to users for EF experiments once they are satisfied with the other settings.

(d) Some mathematical considerations on Ellipsoid Factor

(i) How is a difference of ratios of sorted random variable triplets distributed?

As seen in the main document text, EF is calculated as the difference of sorted semi-axis ratios of the fitted ellipsoids. The ellipsoid fitting is stochastic, which means that the ellipsoid semi-axis lengths will be randomly distributed, but their distribution is *a priori* unclear. It may therefore be useful to know how the difference of ratios of sorted random triplets is distributed for some known defined random distributions: the normal distribution (Figure S5) and the uniform distribution (Figure S6).

These simple examples caution against the undiscerning interpretation of EF values. However, it seems unlikely that a,b,c follow either of the distributions above. We can simulate this to an

extent with a gamma distribution (Figure S7). This also shows that EF values distribute in an approximately triangular fashion.

Finally, we expect that a structure clearly divisible into rod and plate like parts would have the properties displayed in Figure S8, i.e. the middle radius clustering into two clusters near where the smallest radius and the largest radius are clustered.

This subsection shows that one difference of sorted ratio distribution does not imply a unique underlying ellipsoid distribution: we have seen two ways of getting triangular distributions. It also tells us that one way of getting a bimodal distribution is have bimodal semi-axis ratios.

The R script used to perform these numerical experiments and plot their results can be found at (30) under `/R/null-case-EF.R`.

(ii) Non-uniqueness of EF

Note that this locally maximal ellipsoid is not unique. For example, theoretically, one could have two axis-aligned ellipsoids with (x,y,z)-axis lengths of (1,1,9) and (3,3,1) respectively. Both would have volume 9, but very different EF. Assuming these ellipsoids are locally maximal, pixels within the first ellipsoid would have $\mathrm{EF} = 1 - \frac{1}{9} = \frac{8}{9}$ (i.e. rod-like) whereas pixels within the other would have $\mathrm{EF} = \frac{1}{3} - 1 = -2/3$ (i.e. rather plate-like (Figure S9)). The pixels in the intersection of the two ellipsoid would have non-unique EF. If these ellipsoids were found in the same EF run, which one would be chosen to be maximal would depend on which one would be added to the list of ellipsoids first (which will vary due to the stochastic nature of EF, but also because ellipsoid fitting is implemented in parallel!). Java's sorting algorithm will not influence this order, as it guarantees that the order of equal elements in the list is not changed. If these ellipsoids were found in separate runs, the EF values would be averaged for the pixels in the region of intersection. This seems sensible to us, as that area is both part of a very rod- and a very plate-like structure.

(iii) Sensitivity to some more input parameters

We used vertebral sample 57 to investigate how the sampling increment ("step"), the contact sensitivity and a semi-axis length filter (see next paragraph) affected the EF distribution. The results are for contact sensitivity values of 1 and 5 are shown in Figure S10 and S11, respectively.

Small ellipsoids may get caught within a thin feature when growing, especially when the step-size is close to some stair-case-like feature and the contact-sensitivity is low (Supplementary Material iii). For this reason, we built an option to remove ellipsoids whose longest semi-axis c is smaller than a user-defined threshold (in pixel units, but does not have to be an integer number) into our code. We refer to this as the minimum valid longest semi-axis length filter.

We observed that at contact sensitivity 5 S11, reducing the step size was enough to remove spurious small ellipsoids, and using larger filters did not really affect the EF distribution. We therefore removed the minimum valid longest semi-axis filter in the interest of reducing an already high number of input parameters, and ran the vertebral measurements with contact sensitivity 5 and the lowest step size.

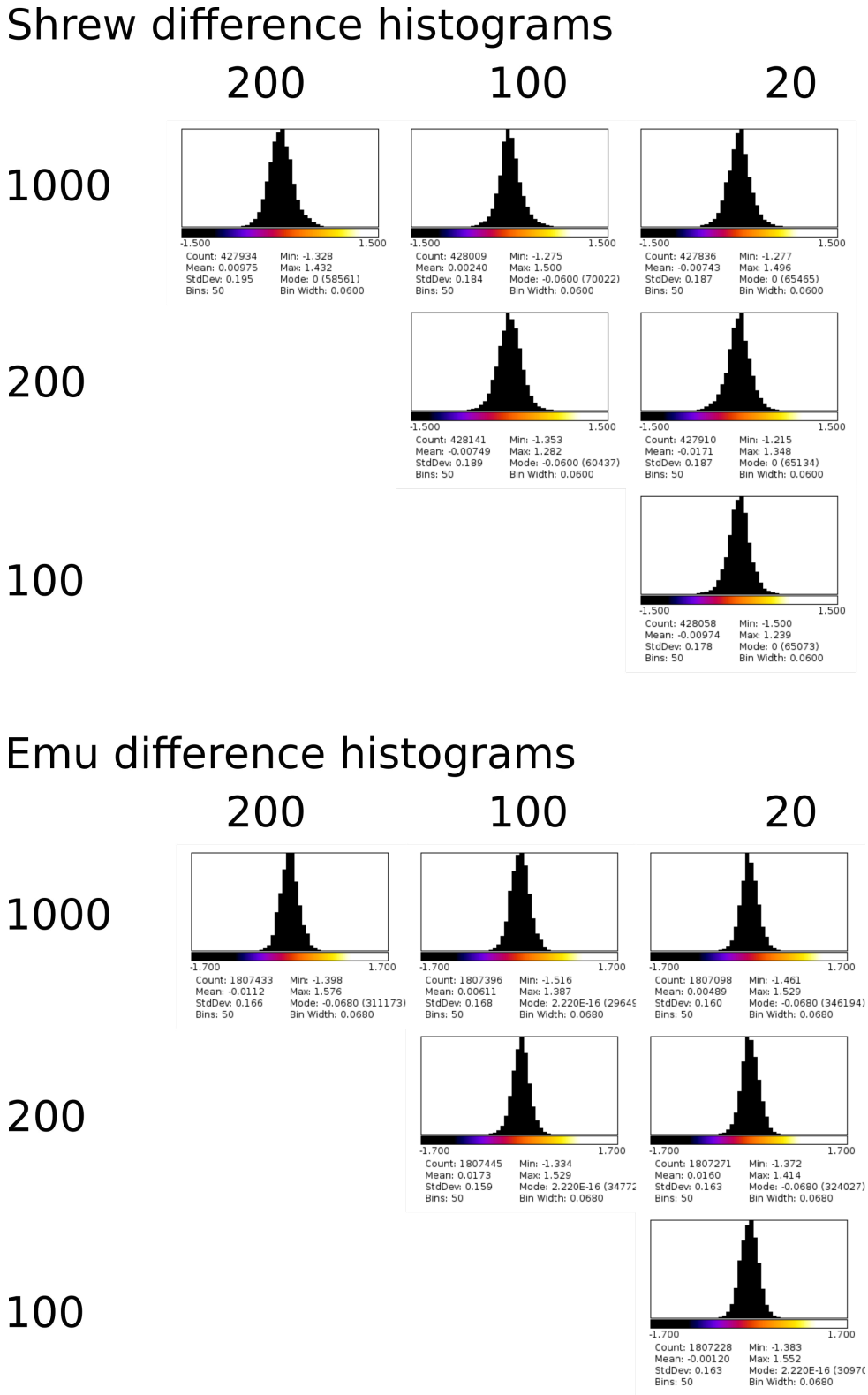

Figure S1. Histograms of the difference between EF outputs for the "max iteration" value on the x- and on the y-axis. A higher "max iteration value" did not alter the EF distribution much.

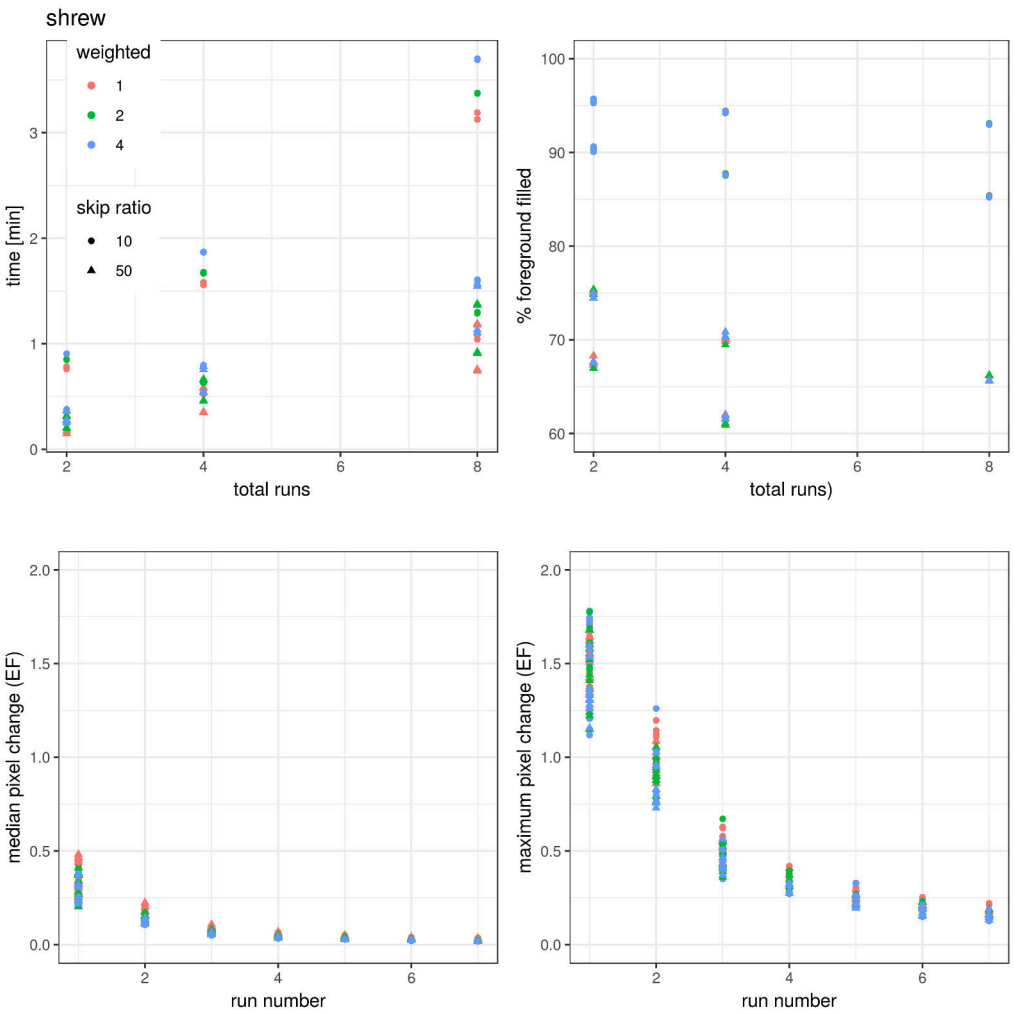

Figure S2. Convergence parameter sensitivity of the shrew test image to variation in skip ratio (shape of symbols), volume-weighted average over several local ellipsoids (colour), the number of vectors (10 and 100 were run, distinguishable only by location on plot) and the number of runs averaged over (x-axis). Within run volume-weighted averaging had little effect on any convergence parameter. Skipping over 5 times more seeds reduced the run-time of the algorithm by roughly a factor of 2, but reduced the filling percentage by 15-20%. Increasing the number of sampling vectors roughly doubles the run-time, and increases the filling percentage by about 5%. Median and maximum changes per additional run in EF value per pixel required 6 runs to be less than 0.1 and 0.4 respectively, which we defined as satisfactorily converged. These statements are also true for two other test images, emu (Figure S3) and rods_plates (Figure S4).

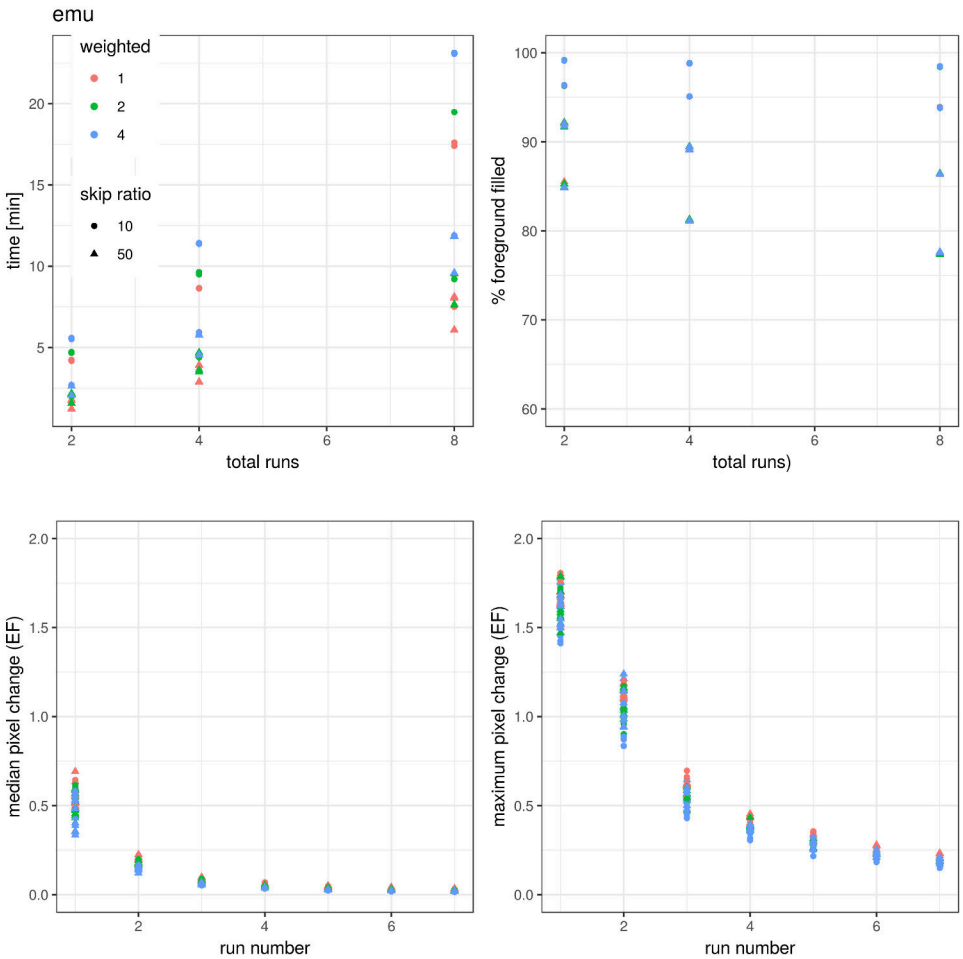

Figure S3. Please refer to the caption of Figure S2.

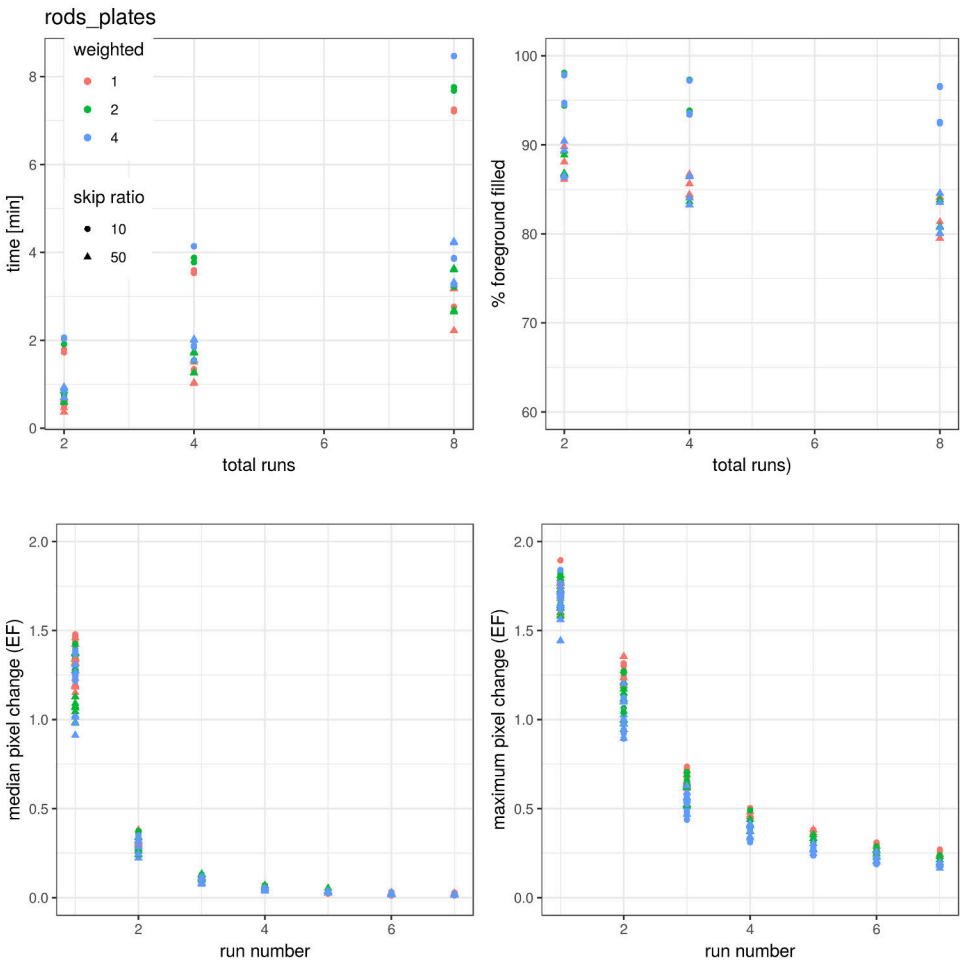

Figure S4. Please refer to the caption of Figure S2.

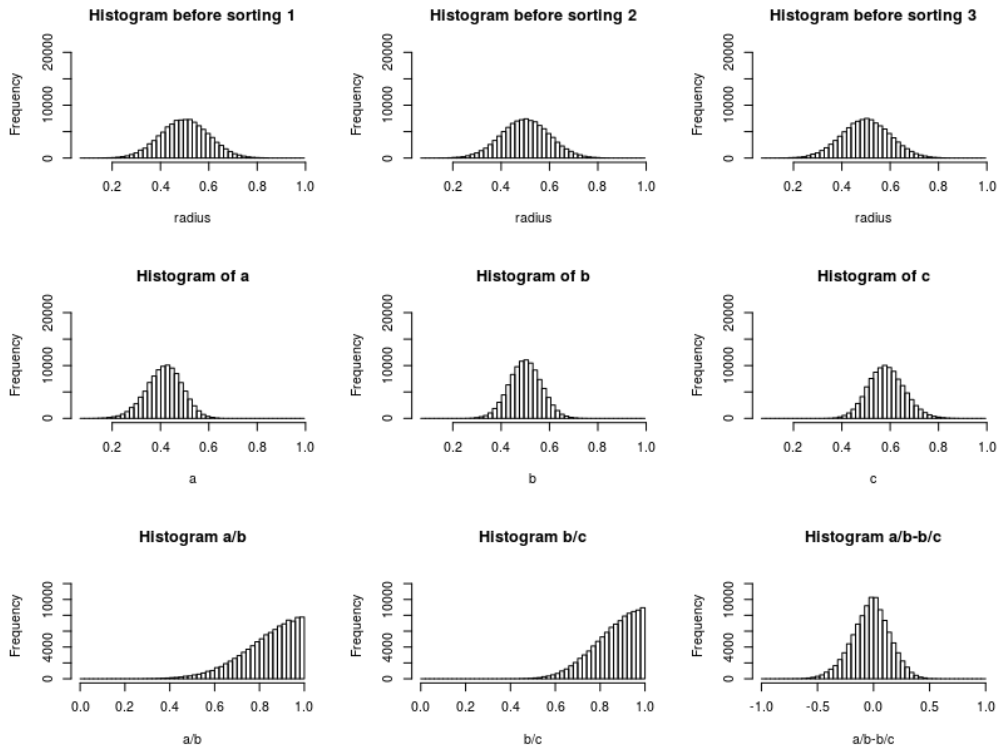

Figure S5. Numerical experiment with normally distributed radii (mean=0.5, standard deviation=0.1) (first row). Sorting causes a to have a lower mean, and c to have a higher mean (second row). The ratio of sorted, normally distributed triplets is heavily skewed to the left. The difference of sorted ratios appears triangular (third row).

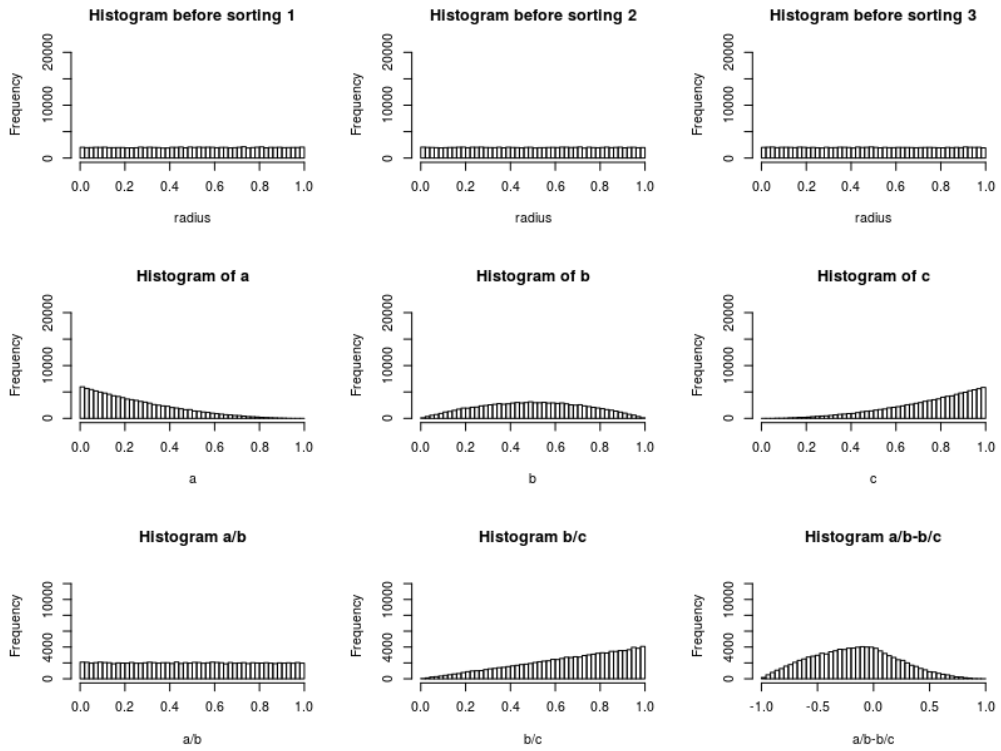

Figure S6. Numerical experiment with uniformly distributed radii on [0,1] (first row). Sorting causes a to be heavily skewed to the right, b to be distributed according to what resembles a parabola and c to be heavily skewed to the left (second row). The ratio of sorted, uniformly distributed triplets is either uniformly distributed (a/b) or heavily skewed to the left (b/c) depending on whether the larger radius is in the numerator or the denominator. The difference of sorted ratios resembles a shark fin, and is skewed to the right.

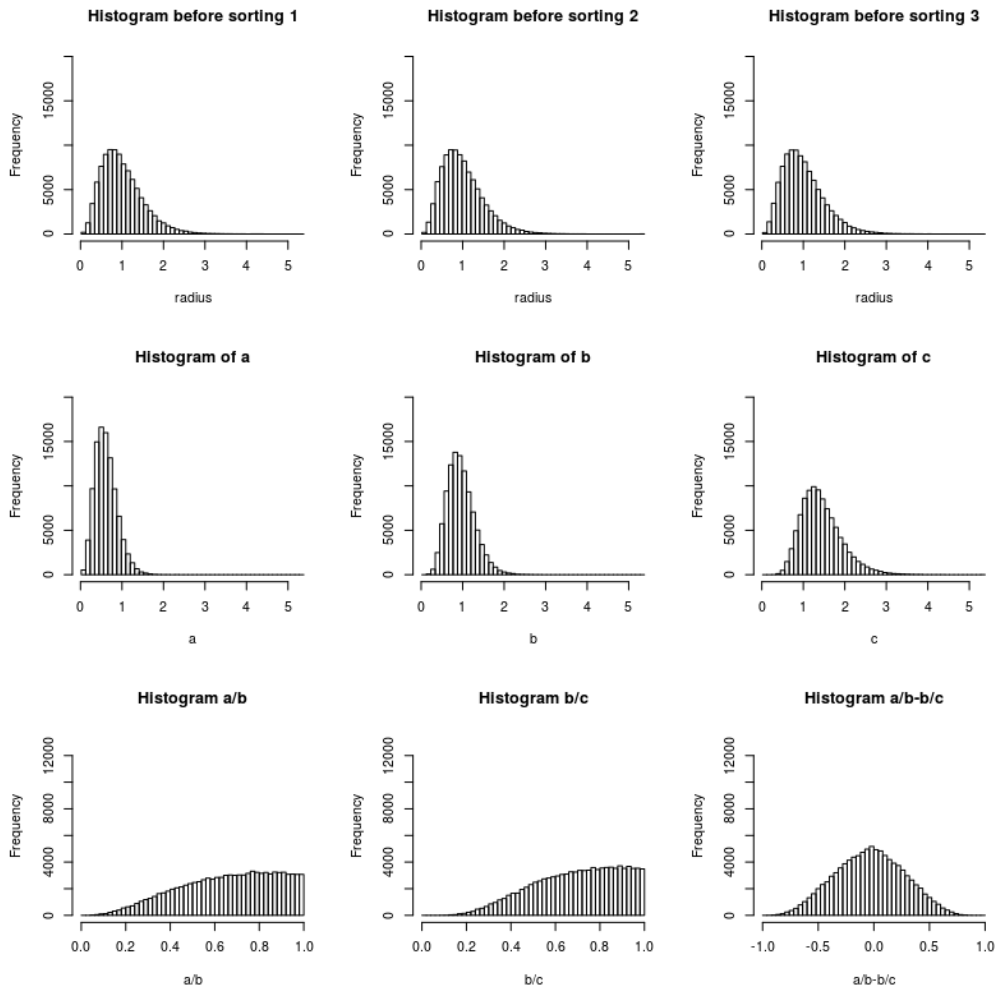

Figure S7. Numerical experiment with radii according to Gamma(k=4,theta=1)/4 (first row). Sorting causes the median of a to decrease, and the median of c to increase, but the distribution shape stays the same (second row). The ratio of sorted, gamma-distributed triplets is heavily skewed to the left and their difference appears triangular.

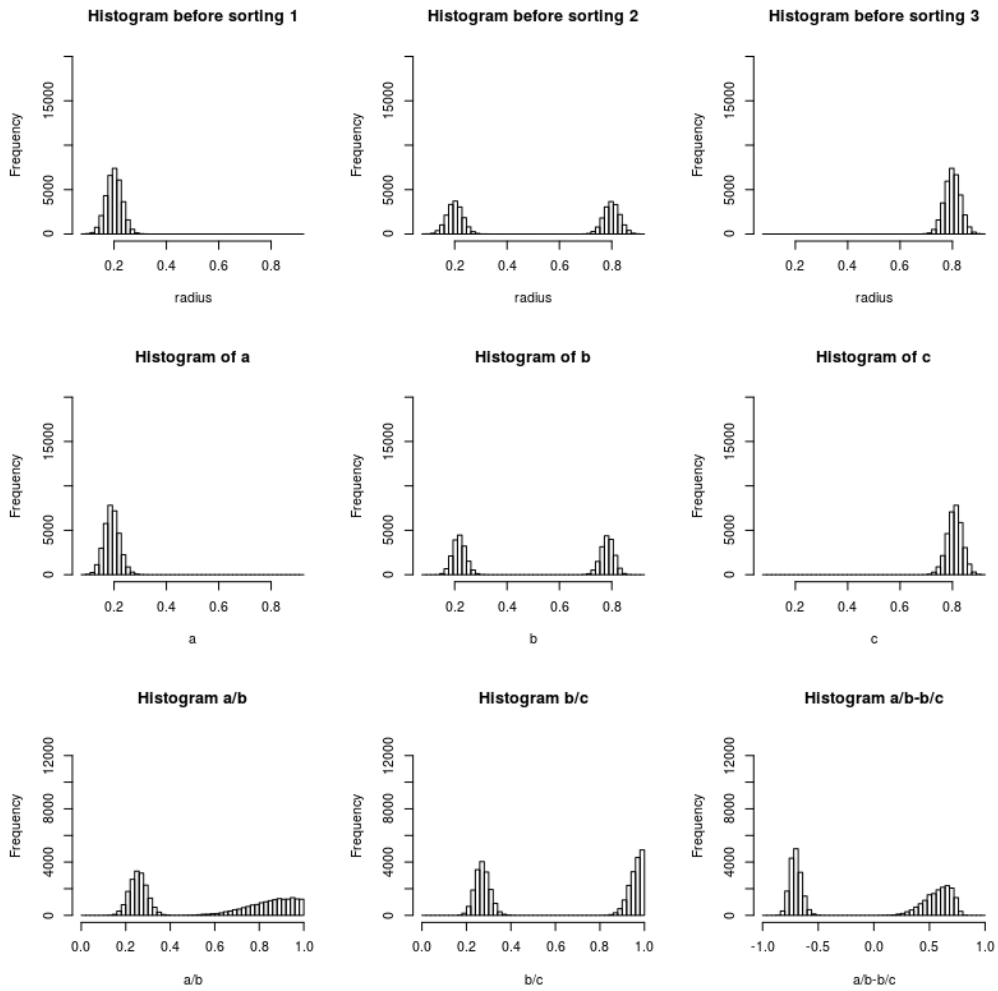

Figure S8. Numerical experiment with bimodal b radius.

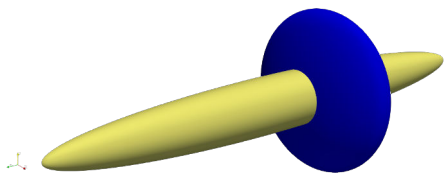

Figure S9. Two ellipsoids of same volume, but different axes as an example to show that EF can be non-unique. The yellow ellipsoid has (x,y,z)-semi-axis lengths of (1,1,9), while the blue ellipsoid has (x,y,z)-semi-axis lengths of (3,3,1). In this case, the pixels in the intersection of the two ellipsoids will have non-unique EF (assuming the two ellipsoids are locally maximal).

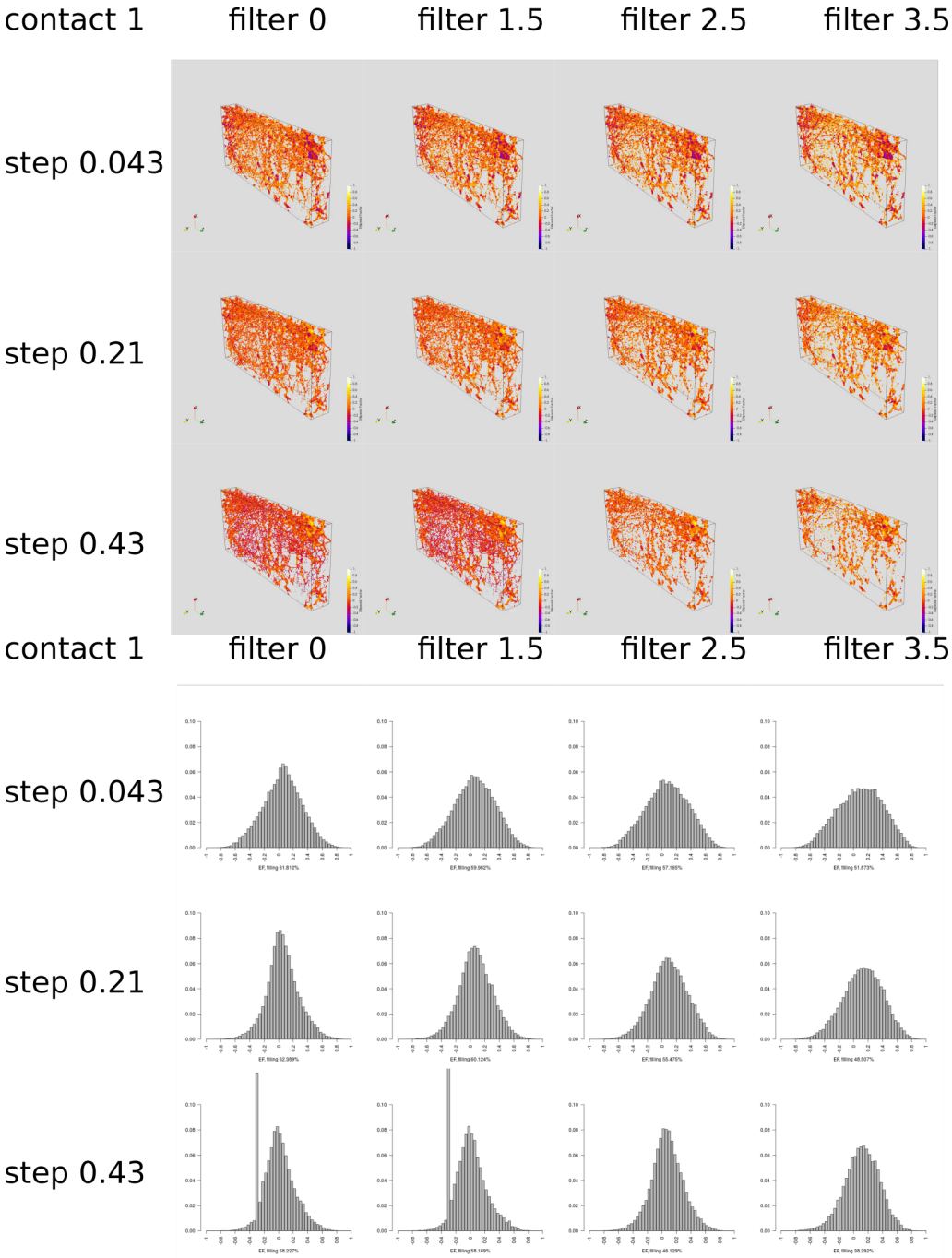

Figure S10. Renders and histograms of the EF image of vertebral sample 57 with a contact sensitivity of 1 and various step sizes and semi-axis filters. A large step size causes comparatively small features to fill up with ellipsoids that have taken only a few steps in every direction (e.g. 3,3 and 2. These ellipsoid cause the spurious mode in the corresponding histogram. This effect is reduced by using a smaller step size or a larger filter. The larger filter changes the distribution, however.

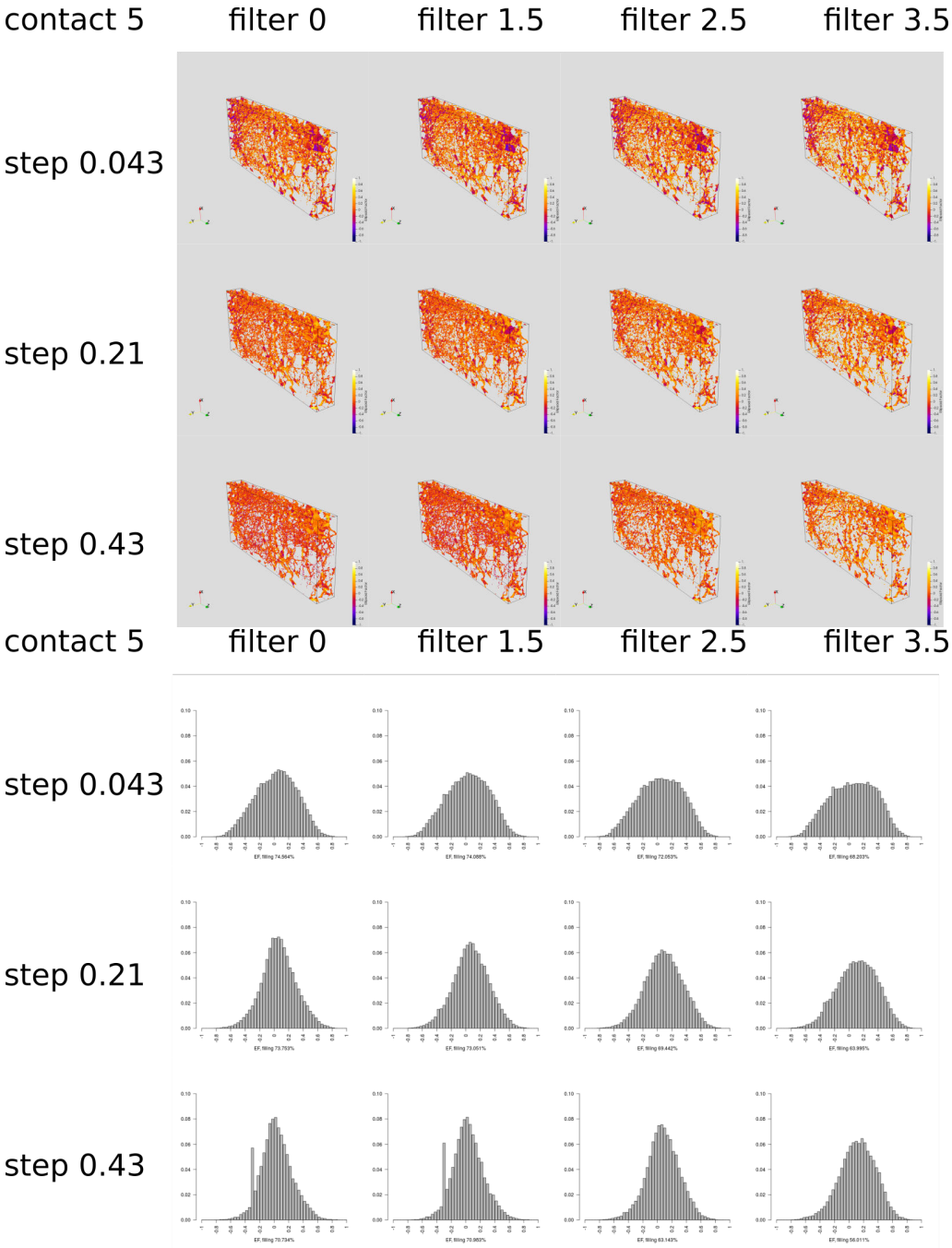

Figure S11. Renders and histograms of the EF image of vertebral sample 57 with a contact sensitivity of 5 and various step sizes and semi-axis filters. A large step size causes comparatively small features to fill up with ellipsoids that have taken only a few steps in every direction (e.g. 3,3 and 2. These ellipsoid cause the spurious mode in the corresponding histogram. This effect is reduced by using a smaller step size or a larger filter.