



Research



Cite this article: Casey JJ, Forsyth JE, Waite TW, Cotter SL, Shearer T. 2025 Exploring natural variation in tendon constitutive parameters via Bayesian data selection and mixed effects models. *Proc. R. Soc. A* **481**: 20250034. <https://doi.org/10.1098/rspa.2025.0034>

Received: 14 January 2025

Accepted: 28 August 2025

Subject Areas:

applied mathematics, statistics, biomechanics

Keywords:

tendon, data selection, mixed effects, Bayesian inference, nonlinear elastic

Author for correspondence:

Simon L. Cotter

e-mail: simon.cotter@manchester.ac.uk

Exploring natural variation in tendon constitutive parameters via Bayesian data selection and mixed effects models

James J. Casey¹, Jessica E. Forsyth^{1,2}, Timothy W. Waite¹, Simon L. Cotter¹ and Tom Shearer¹

¹Department of Mathematics, University of Manchester, Manchester M13 9PL, UK

²School of Medicine and Population Health, University of Sheffield, Sheffield S10 2TN, UK

JJC, 0000-0001-9853-3566; JEF, 0000-0002-5839-9160; TWW, 0000-0002-8868-3146; SLC, 0000-0001-5974-7393; TS, 0000-0001-7536-5547

Combining microstructural mechanical models with experimental data enhances our understanding of the mechanics of soft tissue, such as tendons. In previous work, a Bayesian framework was used to infer constitutive parameters from uniaxial stress-strain experiments on horse tendons, specifically the superficial digital flexor tendon (SDFT) and common digital extensor tendon (CDET), on a per-experiment basis. Here, we extend this analysis to investigate the natural variation of these parameters across a population of horses. Using a Bayesian mixed effects model, we infer population distributions of these parameters. Given that the chosen hyperelastic model does not account for tendon damage, careful data selection is necessary. Avoiding ad hoc methods, we introduce a hierarchical Bayesian data selection method. This two-stage approach selects data per experiment, and integrates data weightings into the Bayesian mixed effects model. Our results indicate that the CDET is stiffer than the SDFT, probably due to a higher collagen volume fraction. The modes of the parameter distributions yield estimates of the product of the collagen volume fraction and Young's modulus as 811.5 MPa for the SDFT and 1430.2 MPa for the

© 2025 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

CDET. This suggests that positional tendons have stiffer collagen fibrils and/or higher collagen volume density than energy-storing tendons.

1. Introduction

Accurate description of the nonlinear mechanical behaviour of collagen is important for understanding and predicting the properties of a wide range of soft tissues, including arterial walls [1,2], skin [3,4] and tendons [5–7]. This knowledge is vital for designing artificial tissues for grafts and surgical interventions [8]. In this paper, we focus on a model for tendons; however, the underlying approach could be extended to other soft tissues.

Bayesian inverse methods are increasingly used in biomechanics to estimate soft tissue parameters while quantifying uncertainty. Such approaches have been applied to infer viscoelastic properties from acoustic radiation force imaging [9] and to calibrate hyper-viscoelastic models of brain tissue under varying experimental protocols [10]. A Bayesian framework has also been used for model selection and sensitivity analysis in studies of the knee meniscus [11]. These examples illustrate the value of Bayesian inference for parameter identification in complex, heterogeneous tissues.

Tendons are fibrous tissues that connect and transfer forces between, muscles and bones [5,12]. Their complex microstructure gives rise to anisotropy and nonlinear stress–strain profiles [13–15]. An example of a typical stress–strain profile for a tendon undergoing a uniaxial stretch along its longitudinal axis is given in figure 1. Tendons consist of collagen fibrils of varying lengths that assume a crimped waveform within collagen fibres. The fibres are embedded within a non-collagenous matrix (NCM), and reinforce the tendon along a preferred axis [5,16], which causes anisotropy. A schematic representation of a tendon’s microstructure is shown in figure 2. As the tendon is stretched, the fibrils begin to straighten and contribute to the stress response. Owing to the fibrils having varying lengths, their recruitment is gradual, which results in the nonlinearity of the stress profile.

The characteristic shape of the stress–strain graph has four regions: (I) toe, (II) heel, (III) linear and (IV) damage, with each region corresponding to a different physical phenomenon. The toe, region corresponds to the fibrils being slack, resulting in a stress profile that is approximately linear, as governed by the NCM. In the heel region, fibrils gradually become taut and contribute to the stress response, resulting in a nonlinear profile. In the linear region, all of the fibres have been recruited and the stress profile is primarily governed by the fibril stress response. In the final region, failure occurs, the tendon is damaged, fibrils begin to break gradually and the tendon no longer deforms elastically.

In figure 1, the stress response is idealized. In reality, fibrils may begin breaking in earlier regions (III or even II), resulting in data in region III being recorded, which is not consistent with models which neglect damage. Typically, the data are manually trimmed to a specified limit [18]. This approach tends to include a large portion of the linear region (region III), and fitting models to these data can lead to inaccurate estimates of various parameters in the presence of early fibril damage. For this reason, when fitting models to tendon stress–strain data, it is beneficial to use a more sophisticated data selection method to include as much data as possible that is valid under the assumptions of the mechanical model used, while down-weighting the contribution of data points that do not satisfy the model’s assumptions [19].

Models for the stress response of tendons often describe the deformation in terms of the physical parameters of their constituents, such as the Young’s modulus and shear modulus. These models are known as microstructural models and have been used by several authors for various soft tissues [5,20].

One of the advantages of microstructural modelling is that the model parameters have a direct physical meaning and can be measured experimentally. However, a problem that arises is that a wide range of values is often reported for the same parameter. The Young’s modulus of type 1

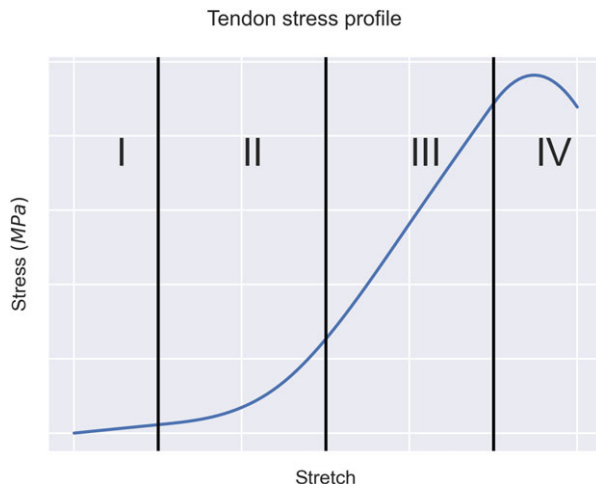


Figure 1. Idealized stress profile of a tendon. The four regions are: (I) toe, (II) heel, (III) linear and (IV) damage. In reality, damage may begin occurring in region III, or even region II, due to the shorter fibrils breaking.

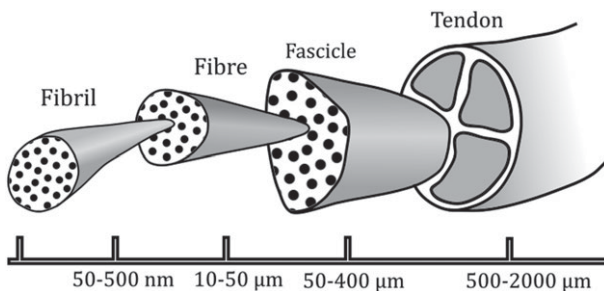


Figure 2. Microstructure of a typical tendon. Bundles of fibrils form fibres, which are embedded within fascicles, which make up the tendon. The NCM is the non-collagenous portion of the tendon. Adapted from Jull *et al.* [17].

collagen fibrils, for example, has been reported to be as low as 32 MPa [21] and as high as 2.8 GPa [22]. This spread of reported values could be due to several factors, such as differences in the ages of the samples, inter-species variation, model misspecification within the inference or even that the parameters may be poorly identifiable from the data. Owing to this, it is often not entirely clear what the credible range for a given parameter value is, nor which values are most likely. Quantification of this spread is key to understanding the natural behaviour of soft tissues and their constituents. In this paper, we tackle the problem of *intra*-species variation. That is, within a given species, individuals can yield varying values for constitutive parameters due to naturally occurring heterogeneity within the population. For context, the individuals in this study are horse tendons, and the species are the superficial digital flexor tendon (SDFT) and the common digital extensor tendon (CDET) types.

To tackle the uncertainty in the parameter values, we employ a Bayesian approach. This provides a rigorous framework with which to combine prior knowledge, mechanical models and data. Typically, inference is performed on data from a single individual from the population at a time, as in previous work [16]. In this paper, we instead use data from multiple individuals to infer the population-level variation in the constitutive parameters of a microstructural, nonlinear elastic tendon model, through a statistical mixed effects model.

Since fibril damage occurs at different points for different individuals, we are prompted to use a more sophisticated data selection technique than uniform trimming of all datasets (for example, after a certain strain has been reached [18]). We implement a hierarchical Bayesian

data selection technique, which allows us to infer which data are consistent with the chosen microstructural model [19]. We use a two-stage approach. The first stage involves data selection on a per-individual basis, leading to a re-weighting of each of the observations in the likelihood, along with an automatically chosen truncation of the data. The second stage uses the data from all individuals (the population), with re-weighted observations and truncations, in the population-level inference, which computes the population posterior distribution for the microstructural model parameters. The reason for performing the inference in two stages is that it would be prohibitively costly to simultaneously perform data selection and calculation of mixed effects posteriors. Moreover, the trimming of data with low weightings which have large discrepancies with the chosen mechanistic model improves overall model fit and reduces the bias in the overall inference [19].

We employ a mixed effects model to infer from multiple individuals simultaneously. This allows us to pool information and build knowledge of population-level distributions of the parameters in addition to the parameter distributions of each individual. Mixed effects models have previously been used to study soft tissue mechanical parameters [23,24], as has the Bayesian framework [12,16,25]; however, to the authors' knowledge, there are no examples of a Bayesian mixed effects model being used for soft tissue inference in the literature. Wang *et al.* [24] used a linear mixed effects model to infer model parameters without any Bayesian modelling, while Wang *et al.* [25] used Bayesian modelling to infer model parameters without incorporating mixed effects. Bayesian mixed effects models *have* been used across many other areas of study, however, such as in recommender systems [26], neurology [27] and pharmacokinetics [28].

The structure of the paper is as follows. In §2a, we introduce the microstructural model. To quantify the variability and uncertainty in the model parameters, we describe the Bayesian approach in §2b. The data we use in this study are outlined in §2c. To calculate the resulting posterior distributions, we use Markov chain Monte Carlo (MCMC) sampling. We ensure that we are inferring model parameters using data where the microstructural model is valid, for example, before a considerable amount of damage has occurred, by implementing a Bayesian data selection technique, described in §2d. To model the heterogeneity of the physical parameters across individuals in the population, a mixed effects model is used, which we discuss in §2f. We discuss our choice of priors for the parameters in §2g, and in §2h, we discuss the implementation of the Hamiltonian Monte Carlo (HMC) sampler using the Stan package [29] and also discuss sampler parameter tuning to ensure convergence of the Markov chains. The results of our modelling approach are given in §3. We conclude with a discussion in §4.

2. Methods

In this paper, we consider the inverse problem of inferring model parameters from observational data. There are many approaches to tackling inverse problems, including maximum-likelihood estimation and Bayesian maximum *a posteriori* estimation. In many circumstances, quantifying the uncertainty in parameter estimates is of interest, for example, by estimating credible ranges for the model parameters. One approach to achieving this is to use a Bayesian framework to determine model parameter posterior distributions from the data. Typically, these distributions are not analytically computable and require approximation via numerical methods such as MCMC. While MCMC is the gold standard for Bayesian inverse problems, it incurs a high computational cost due to the large number of forward model evaluations required. However, if it is computationally feasible to implement for a given model, the samples produced can be used to provide detailed statistical information, including credible regions for the model parameters [30].

We also consider the problem of selecting data for inference where a subset of the data is known not to be consistent with the model. To do this, we again employ a Bayesian framework, this time to infer the values of parameters which convey how well the model fits each data point.

These parameters can then be used to tune out data with mild discrepancies or to trim data with moderate to severe discrepancies with the model.

(a) Microstructural tendon model

To study the deformation of tendons, we use the microstructural model introduced by Haughton *et al.* [16]. We assume the tendons under consideration can be modelled as circular cylinders, with fibres oriented along their long axes. Using cylindrical polar coordinates and denoting the deformed and undeformed configurations with lower- and upper-case letters, respectively, the deformation can be described as $(r, \theta, z) = (\lambda^{-\frac{1}{2}}R, \Theta, \lambda Z)$, where λ is the longitudinal stretch. The deformation gradient is defined as the gradient of the deformed position vector $\mathbf{x}$ with respect to the undeformed coordinates, denoted

$$\mathbf{F} = \nabla \mathbf{x}, \quad (2.1)$$

where ∇ is the gradient operator with respect to the undeformed coordinates. For the prescribed uniaxial deformation, it takes the form

$$\mathbf{F} = \begin{pmatrix} \lambda^{-\frac{1}{2}} & 0 & 0 \\ 0 & \lambda^{-\frac{1}{2}} & 0 \\ 0 & 0 & \lambda \end{pmatrix}. \quad (2.2)$$

From $\mathbf{F}$, we can calculate the left Cauchy–Green strain tensor $\mathbf{B} = \mathbf{F}\mathbf{F}^T$. Denoting the undeformed fibril orientation as $\mathbf{M}$, we can calculate the deformed fibril orientation $\mathbf{m} = \mathbf{F}\mathbf{M}$ and define invariants of the deformation I_1, I_2, I_3, I_4 , and I_5 as

$$\left. \begin{aligned} I_1 &= \text{tr}(\mathbf{B}), \\ I_2 &= \frac{1}{2}(\text{tr}(\mathbf{B})^2 - \text{tr}(\mathbf{B}^2)), \\ I_3 &= \det(\mathbf{B}), \\ I_4 &= \mathbf{M} \cdot \mathbf{B}\mathbf{M} \\ I_5 &= \mathbf{M} \cdot \mathbf{B}^2\mathbf{M}. \end{aligned} \right\} \quad (2.3)$$

and

We assume that tendons are incompressible, so that $I_3 = 1$, and follow the widely used assumption that the deformation of the NCM can be described solely in terms of I_1 and that of the collagen fibrils in terms of I_4 [5,16,31]. Furthermore, we assume that the strain energy function decomposes additively into contributions from the NCM and the fibrils, each scaled by their respective volume fractions. Denoting ϕ as the collagen volume fraction, these are $1 - \phi$ and ϕ , respectively. It is assumed that the NCM is neo-Hookean with shear modulus μ , while the fibrils are assumed to be Hookean, with Young's modulus E , after the macroscale stretch reaches a critical value λ_C , which is treated as a random variable (the value of λ_C corresponds to the length of each fibril relative to the section of tendon in which it is embedded, with longer fibrils having higher values of λ_C). The critical stretch is assumed to have a triangular distribution, with lower limit a , upper limit b and mode c .

From these assumptions, the resulting strain energy function [16] is given by

$$\begin{aligned} W(I_1, I_4) = & (1 - \phi) \frac{\mu}{2} (I_1 - 3) + \phi E \left(\frac{A(I_4, a, b, c)}{2} \log I_4 + (B(I_4, a, b, c) - D(I_4, a, b, c)) \sqrt{I_4} \right. \\ & \left. + \frac{C(I_4, a, b, c)}{2} I_4 + \frac{D(I_4, a, b, c)}{2} \sqrt{I_4} \log I_4 + G(I_4, a, b, c) \right), \end{aligned} \quad (2.4)$$

where A, B, C, D and G are piecewise constant functions of I_4 , which are defined in [16]. Using this strain energy function, the only non-trivial entry of the engineering stress tensor, N , can be

calculated in terms of the applied stretch λ as

$$N(\lambda, (1 - \phi)\mu, \phi E, a, b, c) = (1 - \phi)\mu \left(\lambda - \frac{1}{\lambda^2} \right) + \frac{\phi E}{\lambda} (A(\lambda^2, a, b, c) + B(\lambda^2, a, b, c)\lambda + C(\lambda^2, a, b, c)\lambda^2 + D(\lambda^2, a, b, c)\lambda \log \lambda). \quad (2.5)$$

A full derivation of this relationship is available in [16]. For our statistical inference, we use the stress–stretch formula $N(\cdot)$ as the response function.

(i) The linear modulus

The linear modulus is the gradient of the stress in the linear region. Assuming a small contribution of the NCM to the stress response (typically $E/\mu \approx 10^6$ [18]), the linear region is dominated by the fibril stress. Under this assumption, the linear modulus, LM , can be calculated by Taylor series expanding the linear region of the fibril contribution to the engineering stress,

$$N_{\text{linear}}(\lambda) = \frac{\phi E}{\lambda} \left(-1 + 4\lambda \frac{a \log\{a/c\} + b \log\{b/c\}}{(a - b)^2} \right), \quad (2.6)$$

about the stretch at which it is being calculated, $\bar{\lambda}$, say, and taking the coefficient of the linear term, which gives

$$LM = \left. \frac{dN_{\text{linear}}(\lambda)}{d\lambda} \right|_{\lambda=\bar{\lambda}} = \frac{\phi E}{\bar{\lambda}^2}. \quad (2.7)$$

We see that in the linear region of our model, the linear modulus is proportional to ϕE , but since $\bar{\lambda} > 1$, it will always be less than ϕE . In §3b, we will compare direct estimates of the linear modulus with our inferred values of ϕE .

(b) Bayesian inference

We wish to infer the unknown model parameter values, given the data $\mathbf{y}$, using the model $N(\cdot)$. To simplify the model and to reduce the dimension of the target parameter space, we assume that the distribution of the critical stretch is symmetric, meaning that the parameter c is completely determined by a and b as $c = (a + b)/2$, and we, therefore, infer the parameters using the model $\mathcal{M}(\lambda, (1 - \phi)\mu, \phi E, a, b) = N(\lambda, (1 - \phi)\mu, \phi E, a, b, (a + b)/2)$. We denote the model parameters of interest as $\boldsymbol{\theta} = [(1 - \phi)\mu, \phi E, a, b] \in \mathbb{R}_{\geq 0}^2 \times \mathbb{R}_{\geq 1}^2$. Based on initial exploration, we concluded that the parameter ϕ was not identifiable from the data, and so we cannot infer ϕ , μ and E independently. Fortunately, measurements and estimations of ϕ can be made in a lab setting; however, this is a destructive process and therefore separate representative samples must be destroyed to measure ϕ [32].

In the Bayesian framework, our knowledge about the parameters $\boldsymbol{\theta}$ is captured via the posterior probability distribution, whose density $\pi(\boldsymbol{\theta} | \mathbf{y}; \mathcal{M})$ is obtained by combining prior knowledge with information from the observed data, $\mathbf{y}$. Prior knowledge is encoded through the choice of an appropriate prior distribution, whose probability density we denote as $\pi_0(\boldsymbol{\theta})$. The contribution from the data is captured by the likelihood, denoted $\mathcal{L}(\mathbf{y} | \boldsymbol{\theta}; \mathcal{M})$. Bayes' theorem then allows us to express the posterior density as

$$\pi(\boldsymbol{\theta} | \mathbf{y}; \mathcal{M}) \propto \mathcal{L}(\mathbf{y} | \boldsymbol{\theta}; \mathcal{M}) \pi_0(\boldsymbol{\theta}), \quad (2.8)$$

which yields an expression for the unknown density $\pi(\boldsymbol{\theta} | \mathbf{y}; \mathcal{M})$ in terms of the known densities $\mathcal{L}(\mathbf{y} | \boldsymbol{\theta}; \mathcal{M})$ and $\pi_0(\boldsymbol{\theta})$, up to a constant of proportionality. Since we only consider a single model in this study, henceforth, the dependence on the model is left implicit.

(c) Data

The experimental data that we use to infer the constitutive parameters of our tendon model, which was collected by Thorpe *et al.* [33], consists of tensile tests applied to two types of tendons

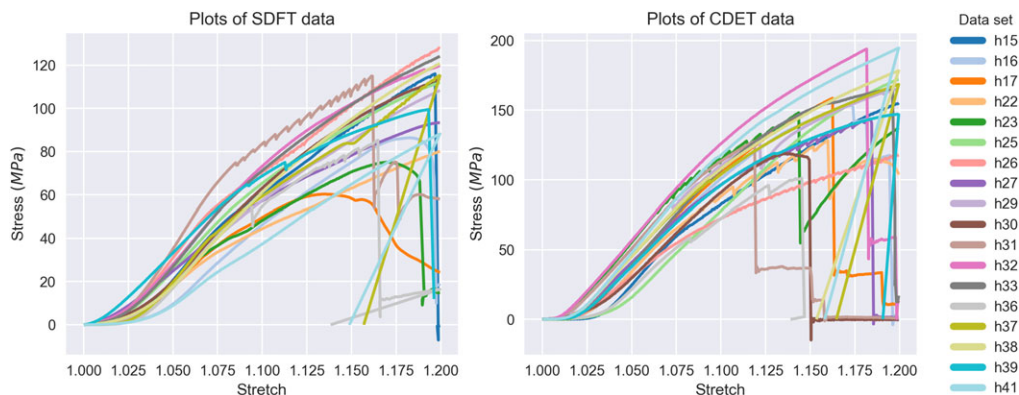


Figure 3. Plots of experimental tensile test data for the two tendon types: SDFT (left) and CDET (right). The data were manually trimmed to 20% strain (equivalent to a stretch of 1.2). Each experiment is labelled ‘hXX’, where XX are numbers that refer to the label given to each horse by Screen *et al.* [33].

harvested from 18 different horses. The two types of tendons were the SDFT and the CDET. SDFTs are energy-storing tendons, whereas CDETs are positional tendons [33]. Figure 3 shows plots of all 18 tests for the 2 tendon types, manually trimmed to 20% strain. It is clear that there is a common trend in the stress profiles of each experiment which indicates that it may be possible to use the same model of deformation to infer subject-specific model parameters for each individual tendon. Heterogeneity of the model parameters is present, as can be seen by the data having varying gradients in the linear region (region III, governed by the product ϕE), and the onset of the heel region (region II, governed by the parameters of the triangular distribution a and b). Since the data were collected by the same team using the same equipment, we assume that all experiments have the same observational noise.

Another feature of the data is heterogeneity in the suspected onset of damage between experiments, as indicated by the changing gradient, and differing failure stretches. For most experiments, the damage region appears to begin somewhere in the stretch range of 1.06–1.10, whereas a few experiments appear to have damage potentially occurring much earlier, at stretches of approximately 1.03. Clearly, trimming all of the experiments to a singular stretch value would not accurately remove the entire damage region for all tendons without also removing large portions of the elastic region for many of them. Using subjective methods to pick the point at which to trim the data for each experiment has the potential to lead to biased and/or more uncertain estimates of the model parameters, depending on whether the truncation is under- or over-zealous. We are, therefore, motivated to explore a more sophisticated methodology to quantify where the data can and cannot be well represented by the chosen mechanistic model. To do so, in the next section, we adapt a hierarchical Bayesian data selection method, as laid out in [19], to identify the regions of data where damage has occurred and the data-model discrepancy is high. These methods were first applied to unlabelled landmark matching of digital images, specifically matching biological cell clusters imaged via different modalities [34].

(d) Bayesian data selection

In previous work by Haughton *et al.* [16], a Bayesian framework was used to infer the microstructural model parameters and their associated uncertainty. In an attempt to remove data where the fibrils had started to be damaged, the data were truncated at 10% strain. In the entire population dataset shown in figure 3, however, it is evident that, in some tendons, damage initiates prior to 10% strain, and thus an elastic model would not fit all of this data well, whereas in other cases, the elastic region appears to continue beyond 10% strain and thus there would potentially be unused valid data if all curves were truncated at 10%.

In this work, we aim to use the data from all 18 of the tensile experiments on two types of horse tendon, and we are, therefore, required to select data from a total of 36 experiments. To do this in an automated and objective manner, we employ hierarchical Bayesian data selection, which enables identification of regions where there is good agreement between the model and data, thereby eliminating the need for subjective manual truncation of the data prior to model fitting. The method involves the introduction of parameters which indicate the fidelity of the model for each data point, and thus greatly increases the dimensionality of the target distribution. Unfortunately, this means that it becomes computationally infeasible to simultaneously incorporate both Bayesian data selection and the mixed effects model, which we describe in §2f. Instead, we propose a two-stage process in which we first conduct Bayesian data selection on a per-individual basis, identifying which data to trim, and re-weighting the remaining data in the likelihood. In the second stage, we use trimmed data, and the posterior means of the fidelity weights, within the mixed effects model.

(i) Definition of the likelihood

Each experiment corresponds to a single tendon and consists of a sequence of measurements of the stress, y_j , required to achieve stretch λ_j ($j=1, \dots, N$), where $N \in \mathbb{N}$ is the number of observations in the data. We assume that each observation is subject to additive, zero mean, Gaussian, independent and identically distributed noise, that is:

$$y_j = \mathcal{M}(\lambda_j, \boldsymbol{\theta}) + \eta_j, \quad \eta_j \sim \mathcal{N}(0, \sigma_{\text{obs}}^2), \quad (2.9)$$

and we let $\mathbf{y} = [y_1, \dots, y_N]^\top$. Therefore, the likelihood without data selection is given by

$$\mathcal{L}(\mathbf{y} | \boldsymbol{\theta}, \sigma_{\text{obs}}^2) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma_{\text{obs}}^2}} \exp\left(-\frac{1}{2\sigma_{\text{obs}}^2} |y_j - \mathcal{M}(\lambda_j, \boldsymbol{\theta})|^2\right). \quad (2.10)$$

We fix σ_{obs}^2 to a value that was estimated by fitting the model to very low strain data for all tendons. Then we introduce ‘fidelity’ parameters for each data point, following the naming convention of previous works [19,34], which indicate the model’s ability to represent a single observation in the data. The fidelity parameters take values $\boldsymbol{\gamma} \in (0, 1)^N$, where values close to 0 correspond to tuning out that data point’s contribution to the likelihood and values close to 1 correspond to a standard contribution to the likelihood (as in the absence of data selection). This is implemented through a modification to the likelihood:

$$\mathcal{L}(\mathbf{y} | \boldsymbol{\theta}, \sigma_{\text{obs}}^2, \boldsymbol{\gamma}) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma_{\text{obs}}^2 \gamma_j^{-1}}} \exp\left(-\frac{\gamma_j}{2\sigma_{\text{obs}}^2} |y_j - \mathcal{M}(\lambda_j, \boldsymbol{\theta})|^2\right) \quad (2.11a)$$

$$\propto \prod_{j=1}^N \frac{\sqrt{\gamma_j}}{\sigma_{\text{obs}}} \exp\left(-\frac{\gamma_j}{2\sigma_{\text{obs}}^2} |y_j - \mathcal{M}(\lambda_j, \boldsymbol{\theta})|^2\right). \quad (2.11b)$$

This approach is similar to that of power likelihoods in a generalized Bayesian framework [35], except that instead of picking a single value for the exponent of the likelihood, we have the additional flexibility of having different values of the likelihood exponent for each individual observation in the data. Our aim in hierarchical Bayesian data selection is to infer the values of the γ_j .

(ii) Data selection prior

By invoking Bayes’ rule, the posterior is given by

$$\pi(\boldsymbol{\theta}, \sigma_{\text{obs}}^2, \boldsymbol{\gamma}) \propto \pi_0(\boldsymbol{\theta}) \pi_0(\boldsymbol{\gamma}) \prod_{j=1}^N \frac{\sqrt{\gamma_j}}{\sqrt{2\pi\sigma_{\text{obs}}^2}} \exp\left(-\frac{\gamma_j}{2\sigma_{\text{obs}}^2} |y_j - \mathcal{M}(\lambda_j, \boldsymbol{\theta})|^2\right), \quad (2.12)$$

where $\pi_0(\theta)$ is the prior for the microstructural parameters θ as described in §2g, and $\pi_0(\gamma)$ is the prior on the fidelity parameters.

The prior we define on γ must be supported on $(0, 1)^N$, encode some correlation between γ_j and $\gamma_{j'}$ as a function of $|\lambda_j - \lambda_{j'}|$, dependent on a given length scale and reflect our knowledge that there is unlikely to be damage occurring within the region of data with less than 5% strain. We adopt a logit Gaussian process prior for the fidelity field, which describes the prior fidelity at all possible points in the observation space [19]. Since we are only interested in the values of the fidelity field at our observation points, this logit Gaussian process prior collapses to a logit multivariate normal prior on γ , where the logit function is applied element-wise. Therefore, the prior on the transformed fidelity parameters, $\chi_\gamma = \text{logit}(\gamma)$, is given by a multivariate normal distribution with mean μ and covariance matrix Σ_γ .

One problem that can occur if naively applying Bayesian data selection is that, in some experiments, there is some early damage and then a considerable amount of additional strain before further damage or failure occurs. In this scenario, without the application of an appropriate prior, the Bayesian data selection posterior may select the data following the initial damage, leading to low fidelity means for the early observations, which we know are unlikely to have been subject to damage. To mitigate this, we choose μ and σ_γ , which define the multivariate normal prior on the transformed fidelity parameters, such that the fidelity parameters have higher mean values for observation points with low strain.

Therefore, we allow $\mu_\gamma(\lambda)$, as well as the standard deviation, $\sigma_\gamma(\lambda)$, to vary with strain. We relate the mean and standard deviations to strain via

$$\mu_\gamma(\lambda) = A_\mu + \frac{K_\mu - A_\mu}{1 + \exp(-B(\lambda - \lambda_0))} \quad (2.13a)$$

and

$$\sigma_\gamma(\lambda) = A_\sigma + \frac{K_\sigma - A_\sigma}{1 + \exp(-B(\lambda - \lambda_0))}, \quad (2.13b)$$

where A_μ, A_σ and K_μ, K_σ are the left/right asymptotic values of μ or σ , respectively, B is the rate of decrease/increase between the two asymptotes, and λ_0 is chosen as the strain value at which we believe that damage is likely to have started to have an effect.

To construct the covariance matrix, Σ_γ , which is $N \times N$ and encodes the correlation between fidelity parameters, we use a vertical scaling of a squared exponential kernel such that

$$\Sigma_\gamma[j, j'] = \sigma_\gamma(\lambda_j) \cdot \sigma_\gamma(\lambda_{j'}) \cdot \exp\left(-\frac{|\lambda_j - \lambda_{j'}|^2}{2l^2}\right), \quad (2.14)$$

where l is the length scale of the kernel and describes the length of correlation between fidelity parameters, and $\sigma_\gamma(\cdot)$ is the standard deviation at a given strain value according to equation (2.13b). We choose $A_\mu = 4$, $K_\mu = 1$, $A_\sigma = 0.25$, $K_\sigma = 1$, $B = 50$, $\lambda_0 = 1.1$, $l = 0.05$ to ensure we have high prior mean and low variance for strains less than 5% where we expect significant damage to the tendon to be rare, and high variance (and therefore lower mean) for strains above 10% where damage is likely to have started to have an effect (see figure 4).

Without careful tuning of these parameters, we do see the data selection method fitting primarily to the yield region rather than the undamaged linear region for some of the experiments, which is unwanted. The heterogeneity of the onset of damage does make this a challenging data selection problem. However, the prior we arrive at aligns with our prior beliefs, in which we know that it is extremely unlikely for the tendon to have become damaged (and therefore not align with our model) at low strains.

(iii) Markov chain Monte Carlo to infer fidelity parameters

We implemented a Metropolis-within-Gibbs framework to sample from the data selection posterior (equation (2.12)), alternating between updating the microstructural parameters and the fidelity parameters. The random walk step-sizes were tuned for the model and fidelity parameters

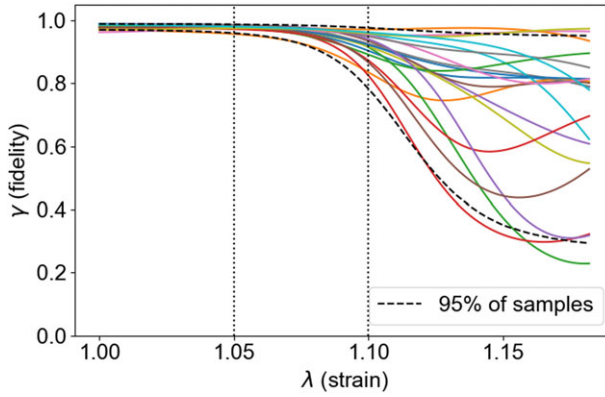


Figure 4. Twenty realizations of the logit-multivariate normal prior on the data fidelity parameters with respect to tendon strain. Dashed line represents the 95% confidence interval of the prior samples.

independently to achieve $23.4 \pm 5\%$ acceptance [36]. Three chains were initialized randomly for each dataset, with a burn-in of 2.5×10^6 iterations and 5×10^6 iterations post-burn-in.

To handle the constraints on the parameters we first transformed to an unbounded scale, giving

$$\left. \begin{aligned} \chi_\gamma &= \mathcal{T}_\gamma^{-1}(\gamma) = \text{logit}(\gamma) \in \mathbb{R}^N \\ \xi &= \mathcal{T}_\theta^{-1}(\theta) \in \mathbb{R}^4. \end{aligned} \right\} \quad (2.15)$$

and

For full details of $\mathcal{T}_\theta$ see appendix A. The posterior in the transformed space is

$$\tilde{\pi}(\xi, \chi_\gamma | \mathbf{y}) \propto \tilde{\pi}_0(\xi) \tilde{\pi}_0(\chi_\gamma) \mathcal{L}(\mathbf{y} | \mathcal{T}_\theta(\xi), \sigma_{\text{obs}}^2, \mathcal{T}_\gamma(\chi_\gamma)), \quad (2.16)$$

where

$$\tilde{\pi}_0(\xi) = \mathcal{N}(\xi; \mu, \sigma_{\text{obs}}^2) \quad (2.17a)$$

and

$$\tilde{\pi}_0(\chi_\gamma) = \mathcal{N}(\chi_\gamma; \mu_\gamma, \Sigma_\gamma), \quad (2.17b)$$

and $\mathcal{L}(\mathbf{y} | \mathcal{T}_\theta(\xi), \sigma_{\text{obs}}^2, \mathcal{T}_\gamma(\chi_\gamma))$ is as defined in equation (2.11b).

We employed the random walk Metropolis proposal within the Metropolis-within-Gibbs method to sample from the posterior distribution, with proposals made separately for the model parameters and the fidelity parameters. The covariance of the microstructural parameters was learned during sampling with the proposal covariance adjusted accordingly in an adaptive manner. The covariance matrix of the proposal distributions for the fidelity parameters was chosen to be proportional to the prior covariance matrix.

(e) Data truncation

Following successful characterization of the posterior distributions of the fidelity parameters, we need to truncate the data. The reason for this is twofold. Firstly, it is required as there are many observations with extremely small fidelity parameters whose effect on the posterior is negligible, but whose effect on the cost of computing the posterior is significant. Secondly, as has been seen in previous work, including data on the boundaries of regions with low posterior fidelity can lead to significant bias in the parameter estimates [19]. As such, we take the same approach as in [19], trimming all data from the first observation with posterior fidelity mean below 0.3. Using a threshold below 0.3 leads to significant shifts in the peaks of the parameter distributions, indicating bias caused by model-data discrepancies. Increasing the threshold above 0.3 leads to the removal of data in regions where the data-model fit is still very good, lowering the information gain and increasing uncertainties.

(f) Inferring population-level natural variation via a mixed effects model

Figure 3 clearly demonstrates that the values of key constitutive parameters vary between individuals. To account for and quantify this data heterogeneity, we assume that the i th tendon has its own individual parameters $\theta_i \in \mathbb{R}_{\geq 0}^2 \times \mathbb{R}_{\geq 1}^2$ for $i \in \{1, 2, \dots, N_e\}$ where N_e is the number of individuals within the population. In our case, $N_e = 18$ for each of the two populations, SDFT and CDET, which we model separately. We denote the truncated observed stresses of the i th individual as $\hat{\mathbf{y}}_i \in \mathbb{R}^{N_i}$ and $\hat{\lambda}_i \in \mathbb{R}^{N_i}$ as the set of stretches for which $\hat{\mathbf{y}}_i$ was measured. The associated fidelity parameters are denoted $\gamma_i \in [0.3, 1]^{N_i}$. In the following, it is assumed that the data have been pre-processed and trimmed using the fidelity threshold of 0.3 such that N_i represents the length of the data after trimming. The inferred fidelity parameters are given in §3a.

By letting $\hat{\lambda}_{ij}$ be the j th value of the stretch $\hat{\lambda}_i$ corresponding to the j th measured data point of the i -th experiment $\hat{\mathbf{y}}_{ij}$, along with the j th fidelity parameter γ_{ij} , the likelihood for the i th experiment can be expressed as

$$\mathcal{L}(\hat{\mathbf{y}}_i | \theta_i, \sigma_{\text{obs}}, \hat{\lambda}_i, \gamma_i) \propto \prod_{j=1}^{N_i} \frac{1}{\sigma_{\text{obs}}} \exp \left\{ -\frac{\gamma_{ij}}{2\sigma_{\text{obs}}^2} (\hat{\mathbf{y}}_{ij} - \mathcal{M}(\hat{\lambda}_{ij}, \theta_i))^2 \right\}. \quad (2.18)$$

Experiments on different tendons are conditionally independent given the tendon-specific parameters; therefore, the joint likelihood of all experiments is the product of the likelihoods for each individual tendon. From now on, the value of σ_{obs}^2 is chosen to be fixed with a value of 0.15 MPa² and shall be dropped from future notation.

To facilitate the specification of the random effects distribution, we transform the tendon-specific parameters to an unbounded scale using the same reparametrization as in §2d(iii), giving

$$\begin{pmatrix} \phi E_i \\ (1 - \phi)\mu_i \\ a_i \\ b_i \end{pmatrix} = \theta_i = \mathcal{T}_\theta(\xi_i) \iff \begin{pmatrix} v_i \\ \eta_i \\ \tau_i \\ \rho_i \end{pmatrix} = \xi_i = \mathcal{T}_\theta^{-1}(\theta_i). \quad (2.19)$$

We assume that the unbounded parameters are samples from a common population, and model the population with a normal distribution with mean μ_{pop} and covariance Σ_{pop} , which is a common choice for the population distribution [37–39]. The parameters ξ_i are assumed to be conditionally independent and are distributed as $\xi_i | \mu_{\text{pop}}, \Sigma_{\text{pop}} \sim \mathcal{N}(\mu_{\text{pop}}, \Sigma_{\text{pop}})$. We define $\pi_{\text{pop}}(\cdot | \mu, \Sigma)$ to be the probability density function (PDF) of a normal distribution with parameters μ and Σ such that $\pi_{\text{pop}}(\xi_i | \mu_{\text{pop}}, \Sigma_{\text{pop}})$ is the PDF of the population distribution.

The joint prior density of the unconstrained parameters $\pi(\xi_1, \dots, \xi_{N_e})$ can be expressed [40] as the conditionally independent, continuous mixture

$$\pi(\xi_1, \dots, \xi_{N_e}) = \int \pi_0^{\mu, \Sigma}(\mu_{\text{pop}}, \Sigma_{\text{pop}}) \prod_{i=1}^{N_e} \pi_{\text{pop}}(\xi_i | \mu_{\text{pop}}, \Sigma_{\text{pop}}) d\mu_{\text{pop}} d\Sigma_{\text{pop}}.$$

For conciseness, we denote the list of parameters $\xi_1, \dots, \xi_{N_e}$ as $\{\xi_i\}_{i=1}^{N_e}$, and use similar notation for other parameters.

In a Bayesian context, we can simultaneously infer the per-individual parameters and the population parameters and then marginalize out the population parameters from the distribution. We can write the joint prior distribution of the per-individual parameters and population parameters as

$$\pi(\{\xi_i\}_{i=1}^{N_e}, \mu_{\text{pop}}, \Sigma_{\text{pop}}) = \pi_0^{\mu, \Sigma}(\mu_{\text{pop}}, \Sigma_{\text{pop}}) \prod_{i=1}^{N_e} \pi_{\text{pop}}(\xi_i | \mu_{\text{pop}}, \Sigma_{\text{pop}}).$$

Finally, we must set priors for the population parameters to complete the statistical model. These are discussed in §2g and, for now, are denoted $\pi_0^{\mu, \Sigma}(\mu_{\text{pop}}, \Sigma_{\text{pop}})$. The posterior distribution can be constructed, therefore, from three parts: (1) the per-individual likelihood, (2) the

population model, (3) the priors [38]. In terms of the unconstrained parameters, the posterior distribution is given by

$$\pi(\{\xi_i\}_{i=1}^{N_e}, \mu_{\text{pop}}, \Sigma_{\text{pop}} \mid \{\hat{y}_i\}_{i=1}^{N_e}, \{\mathbf{y}_i\}_{i=1}^{N_e}, \sigma_{\text{obs}}) \propto \prod_{i=1}^{N_e} \mathcal{L}(\hat{y}_i \mid \mathcal{T}_{\theta}(\xi_i), \hat{\lambda}_i, \mathbf{y}_i, \sigma_{\text{obs}}) \cdot \prod_{i=1}^{N_e} \pi_{\text{pop}}(\xi_i \mid \mu_{\text{pop}}, \Sigma_{\text{pop}}) \cdot \pi_0^{\mu, \Sigma}(\mu_{\text{pop}}, \Sigma_{\text{pop}}). \quad (2.20)$$

To obtain the posterior distribution in terms of the constrained, untransformed parameters θ_i , the parameters ξ_i are transformed back to the constrained scale, which requires a Jacobian adjustment of the posterior distribution in equation (2.20) due to the transform $\mathcal{T}_{\theta}(\cdot)$, resulting in the posterior distribution of the untransformed parameters:

$$\pi(\{\theta_i\}_{i=1}^{N_e}, \mu_{\text{pop}}, \Sigma_{\text{pop}} \mid \{\hat{y}_i\}_{i=1}^{N_e}, \{\mathbf{y}_i\}_{i=1}^{N_e}, \sigma_{\text{obs}}) \propto \prod_{i=1}^{N_e} \mathcal{L}(\hat{y}_i \mid \theta_i, \hat{\lambda}_i, \mathbf{y}_i, \sigma_{\text{obs}}) \cdot \prod_{i=1}^{N_e} \pi_{\text{pop}}(\mathcal{T}_{\theta}^{-1}(\theta_i) \mid \mu_{\text{pop}}, \Sigma_{\text{pop}}) \left| \frac{d\mathcal{T}_{\theta}^{-1}(\theta_i)}{d\theta_i} \right| \cdot \pi_0^{\mu, \Sigma}(\mu_{\text{pop}}, \Sigma_{\text{pop}}). \quad (2.21)$$

To understand the population variability in the mechanical parameters, we compute the posterior predictive distribution of the parameters θ^* for a future randomly selected tendon from the same population. This takes into account both population variability of the parameters and our uncertainty about that variability due to limited data.

(g) Mixed effects model prior selection

Eliciting priors for use in a Bayesian context is often a non-trivial task requiring expert knowledge of the domain to inform the choice of distributions and the parameters of those distributions [41–43]. To evaluate the posterior density equation (2.20), we must first set priors for the parameters μ_{pop} , and Σ_{pop} . Furthermore, following common practices [37,39], we do not infer the covariance matrix directly but rather a scale matrix, $\mathbf{S}$, and correlation matrix, $\mathbf{C}$, such that Σ_{pop} can be written as

$$\Sigma_{\text{pop}} = \mathbf{S}\mathbf{C}\mathbf{S}.$$

The scale matrix, $\mathbf{S}$, is a diagonal matrix containing the population standard deviations for each component of ξ_i . The Cholesky factorization of Σ_{pop} is, therefore, $\mathbf{S}\mathbf{L}_C$, where $\mathbf{L}_C$ is the Cholesky factor of the correlation matrix, $\mathbf{C}$.

Given the posterior distribution in equation (2.21) and the above discussion, we must choose the prior distribution $\pi_0^{\mu, S, C}(\mu_{\text{pop}}, \mathbf{S}, \mathbf{C})$. We assume that the population parameters μ_{pop} , $\mathbf{S}$, and $\mathbf{C}$ are independent and, therefore, the prior distribution decomposes into the product of the priors of each parameter $\pi_0^{\mu}(\mu_{\text{pop}})\pi_0^S(\mathbf{S})\pi_0^C(\mathbf{C})$.

In previous work [16], priors were set on the model parameters, θ , using distributional parameters informed by values found in the literature. Due to the hierarchical model used in this paper, we are unable to set priors directly on the θ_i as these parameters are governed by the population distribution. Instead, priors that were previously set on the model parameters are now placed on $\mu_{\text{pop}} = [\nu_{\text{pop}}, \eta_{\text{pop}}, \tau_{\text{pop}}, \rho_{\text{pop}}]^T$. Following the previous work, a transformation is used based on the natural bounds of the parameters, as described in appendix A.

Conveniently, the transformation used in the previous work, which was used to boost computational efficiency by matching the support of the parameters and the MCMC transition kernel is also an invertible transformation, as given in equation (2.19), which is required for the mixed effects model.

We set the following independent priors on the entries of μ_{pop} :

$$\nu_{\text{pop}} \sim \mathcal{N}(1.05309738, 1.30927056^2),$$

$$\eta_{\text{pop}} \sim \mathcal{N}(6.83672018, 0.47191773^2),$$

$$\tau_{\text{pop}} \sim \mathcal{N}(-3.80045123, 0.64387023^2)$$

and

$$\rho_{\text{pop}} \sim \mathcal{N}(-3.59771868, 0.7310165^2).$$

The prior distributional parameters were derived using values from Silver [44], Purslow [45] and Goh [46]. We reduced the variance of the prior distributions in comparison to the previous study [16] due to regions of high probability density of the priors coinciding with relatively unphysical parameter ranges of the model parameters in that study.

Setting priors for a correlation matrix is non-trivial. Many distributions have been proposed as priors for correlation matrices [43], such as the Jeffreys prior, per-entry uniform priors and a per-entry Gaussian process prior [47]. An increasingly common choice for correlation priors is the Lewandowski–Kurowicka–Joe (LKJ) distribution [48,49] with parameter $\kappa > 0$. We set an LKJ(1) prior on the correlation matrix, which is the default recommendation of the brms R package [50]. Finally, for the entries of the scale matrix, S , we set half Student's t prior with three degrees of freedom, mean zero and unit scale, in line with the brms defaults [50].

Since our data appear to be very informative about the parameters in our model, the posterior is not overly sensitive to the choice of priors used here. The prior was chosen to cover values from the literature, including those collagen from skin rather than tendons, and to be overdispersive, so as to cover all possibilities and prevent an overinformative prior.

(h) Posterior sampling implementation

To characterize the mixed effects posterior distribution [equation \(2.21\)](#), we use the Stan library [29]. Below, we discuss the specific configuration of Stan used for the inference. The sampling algorithm used in Stan is based on the HMC algorithm [51] and its extension, the No-U-Turn sampler (NUTS) [52].

In classical HMC sampling, the proposal vector θ' obtained by integrating Hamilton's equations is conditionally accepted using the Metropolis acceptance ratio. In comparison, the HMC+NUTS algorithm builds a binary tree along the integrated Hamiltonian by integrating forwards and backwards with respect to the 'time' variable. Rather than using a Metropolis acceptance step, the proposed state θ' is taken to be the point along the integrated trajectory that is furthest away from the initial state where the integration began, where integrating the Hamiltonian any further would cause the path to 'turn back' on itself [52]. The number of points generated along the trajectory is controlled by a parameter $j_{\text{max}} \geq 1$ that limits the maximum depth of the tree generated to be no more than $2^{j_{\text{max}}} - 1$. In Stan, this parameter is called `max_tree_depth` and defaults to 10.

Stan automatically tunes the leapfrog parameters and the Euclidean metric during an adaptive warm-up period before sampling. More information about these parameters, and what they mean in the context of the HMC+NUTS algorithm can be found in the Stan user manual [29]. The user is able to control the adaptation of the integrator parameters by changing the values of user-facing parameters called `step_size` and `adapt_delta`. The `step_size` parameter defines the starting guess for the integration step size before the adaptation phase. The parameter `adapt_delta` acts as a surrogate for a target Metropolis acceptance ratio and controls the adaptation of the number of integration steps, the estimation of the Euclidean metric and the integration step size.

If the HMC+NUTS parameters are not set properly or are unable to adapt to optimal values, then the integrator will accrue enough errors that the integrated Hamiltonian trajectory will not follow the true trajectory. This divergence from the true trajectory is reported to the user, and divergences indicate poor exploration of the state space and the simulation cannot be trusted.

With the posterior distribution we are sampling from in this study, the default parameter values resulted in divergences, poor effective sample sizes and non-convergence to the posterior as reflected by the split $\hat{R}$ -hat statistic. Instead, we used the following values for the parameters: `step_size` = 0.01, `adapt_delta` = 0.99, and `max_tree_depth` = 14. To interface with Stan, we used `cmdstanr` [53] with a modified Stan script generated from `brms` [50].

For our analysis, we generated 4000 samples per chain, with 10 chains running in parallel, for a total of 40,000 posterior samples. The average sampling time per chain was approximately 60 hr between both runs using a server with an Intel(R) Xeon(R) CPU E5-4627 v2 @ 3.30 GHz (32 core) with 64 GB memory, with each chain utilizing one core.

3. Results

We provide numerical results for the data selection method, introduced in §2d, in §3a. Using these results, we truncate the data as explained in §2e, and feed the truncated data into the mixed effects inference. Results for the mixed effects model, using the data fed from the data selection method, are given in §3b.

(a) Data selection results

Each individual tendon sample dataset was fitted using the microstructural model with data selection. Stress–strain data were extracted up to the point of maximum stress as the model is elastic and, therefore, not capable of describing subsequent decreases in stress that are associated with damage. The model and fidelity parameters were inferred using the MCMC methodology outlined in §2d(iii). Here, we focus on the fidelity parameter marginal means, which are subsequently used within the mixed effects model. The majority of samples exhibited fidelity parameter profiles which started at values close to unity and decreased to zero at some point between 5% and 10% strain (see figures 5 and 6a). Some fidelity profiles, however, exhibited a secondary, smaller increase beyond 10% strain. This was simply due to the model and data crossing at larger strains. It was evident that the model had already diverged from the data by this point and the increase was simply an artefact of model extrapolation (see figure 6b).

Interestingly, the fidelity parameters of all the tendons reduced to values below 0.3 before 10% strain, indicating a significant reduction in the contribution of the data to the likelihood beyond this point. This is shown in figure 7 where the stress–strain profiles of the SDFTs and CDETs have been trimmed to the point where the fidelity parameters first go below 0.3. There is considerable variability in the strain at which the fidelity parameters first reduce below 0.3, in both the SDFTs and CDETs. This further supports our approach of using data selection and the inference of fidelity parameters as opposed to homogeneous manual trimming of the data.

(b) Mixed effects model results

Next, we use the truncated data and fidelity parameter posterior means to infer the population-level distributions of the constitutive parameters. We used the MCMC methods described in §2h to sample from the posterior distribution of the mixed effects model as outlined in §2f. The mechanical model consists of two elastic parameters, $(1 - \phi)\mu$ and ϕE , and two structural parameters, a and b . Owing to the nature of the transformation $T_\theta(\cdot)$ and the priors, the posterior predictives for the parameters are best understood through their modes and not their means. This is because the posterior distributions are heavy tailed and may not necessarily have finite means or variances.

(i) Posterior predictives for model parameters

The posterior predictives of the model parameters are compared for the SDFTs and CDETs in figure 8. For the CDET distributions, ϕE has a peak at 1430.2 MPa, whereas the corresponding peak for the SDFT lies at 811.5 MPa. In addition, the CDET population distribution for $(1 - \phi)\mu$

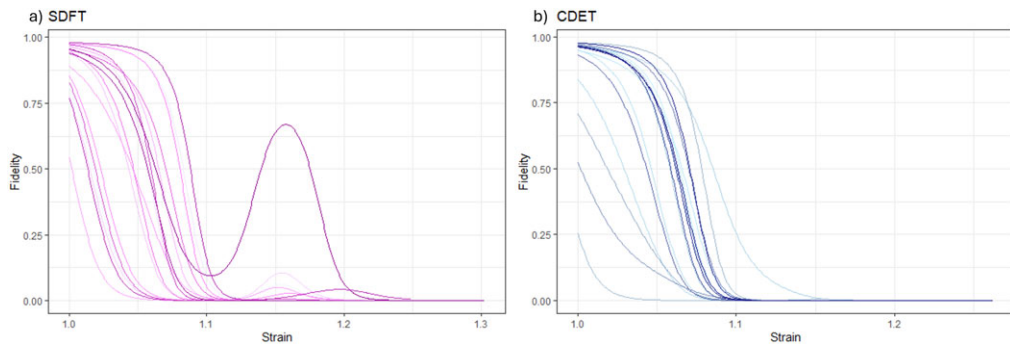


Figure 5. Posterior means of the fidelity parameters against strain for (a) SDFT and (b) CDET tendon samples. Each line represents a fidelity profile for a single tendon.

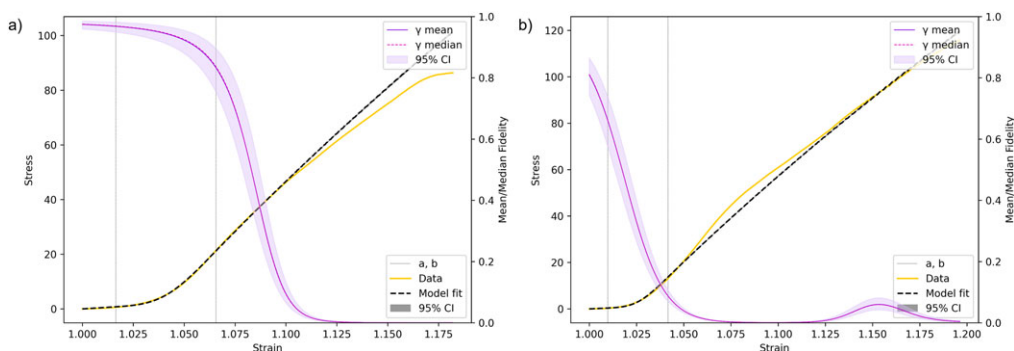


Figure 6. Comparison of the model fit for two SDFTs and the fidelity parameter profiles. (a) A tendon (H16 SDFT) where the fidelity parameters decrease with strain and (b) a tendon (H15 SDFT) where the fidelity parameters decrease but then increase again at a larger strain due to model-data crossing.

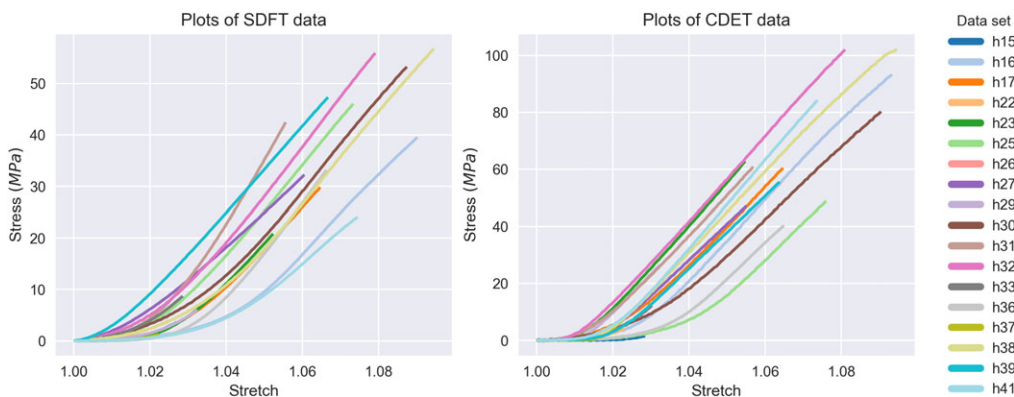


Figure 7. Plots of experimental tensile test data for SDFT (left) and CDET (right), trimmed automatically using Bayesian data selection and a fidelity threshold of 0.3.

has density closer to zero in comparison to the SDFT population distribution. Both of these quantities include the non-identifiable structural parameter ϕ . There are two possible causes, which in some combination could give rise to the differences that we see in the mechanical

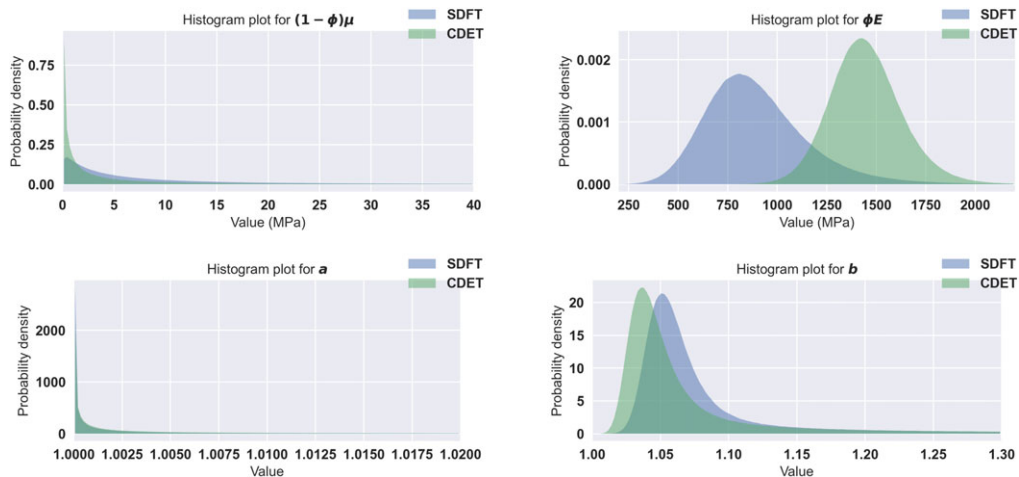


Figure 8. Comparison of the marginal posterior predictive distribution for each model parameter.

parameter distributions. The CDET could contain fibres that are stiffer than those found in the SDFT and/or the CDET could contain a larger volume fraction of fibrils than the SDFT.

For the structural parameters, a and b , the distributions for the SDFT tend to favour higher values than those of the CDET. The parameter, b , has peak probability density at a value of 1.05 for the SDFT, a larger value than the CDET peak density value of 1.04, and the spread of values is larger for the SDFT than the CDET.

These differences in peak probability density relate to the biological functions of the tendons. As the SDFT is an energy-storing tendon that extends and contracts to facilitate locomotion [33], the higher values for the structural parameters reflect the larger range of strains the tendon operates over and the correspondingly longer collagen fibril lengths. The lower values of the mechanical parameters indicate lower stiffness, complementing the functionality of the tendon. The higher stiffness and lower values for structural parameters of CDET reflect their relative inextensibility, which facilitates the transfer of force between muscle and bone.

Note that these distributions do not solely represent the natural variation of the parameters in the population but also the uncertainty in these distributions due to the relatively small number of individuals that have been used in the inference. As such, it is likely that these distributions are overdispersed compared to the true underlying distributions. It is infeasible to disentangle this uncertainty from the natural variation of the parameters, but with the addition of further data, the posterior would probably contract further.

The tangent modulus in the linear region of a tendon's stress–strain curve is often reported as an important mechanical quantity [33]. As a naive estimate, the linear modulus might be used as an approximation of ϕE ; however, as we showed in §2a(i), an exact calculation of the linear modulus of our model shows that it will be strictly less than ϕE . To explore whether the linear modulus can be used as a reasonable estimate of ϕE , we compare our posterior distributions to distributions of the linear modulus of the respective tendon type. To define the linear region, for the upper bound, we used the cut-off stretch as calculated in §3a, and for the lower bound, the posterior mean of samples of b from *independent* inferences of each dataset was used. For the SDFT, 12 of the 18 datasets had valid detected linear regions, where the lower bound is less than the upper bound, and 16 of the 18 data for the CDET had valid detected linear regions. From the data with valid linear regions, we used linear regression to find the gradient of a straight line fit to the data as the linear modulus. Then, motivated by ϕE being strictly non-negative, and the lognormal-like priors as described in §2g along with the parameter transform described in appendix A, in figure 9, we plot lognormal distributions for the linear moduli with mean and variance equal to the mean and variance of the linear moduli calculated from the regression data.

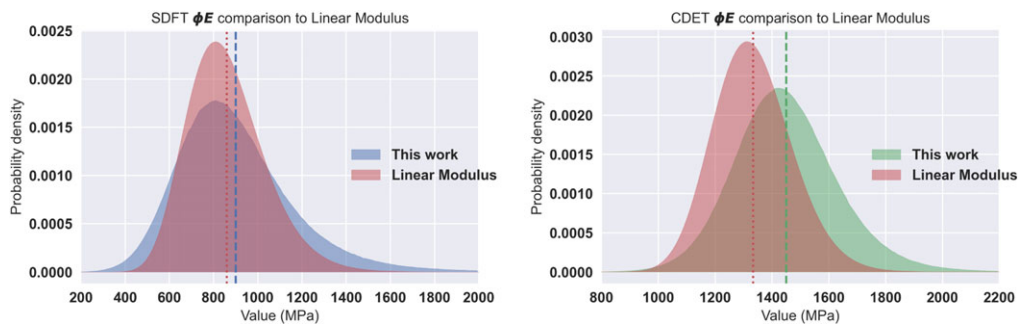


Figure 9. Estimates of the posterior distributions of ϕE for the SDET (left, blue) and CDET (right, green) compared with lognormal distributions with prescribed means and variances equal to the mean and variance of the linear modulus as calculated from the data (both, red). The means of each distribution are indicated with vertical lines (linear modulus: dotted; SDET/CDET: dashed) in the corresponding colours for their respective distributions.

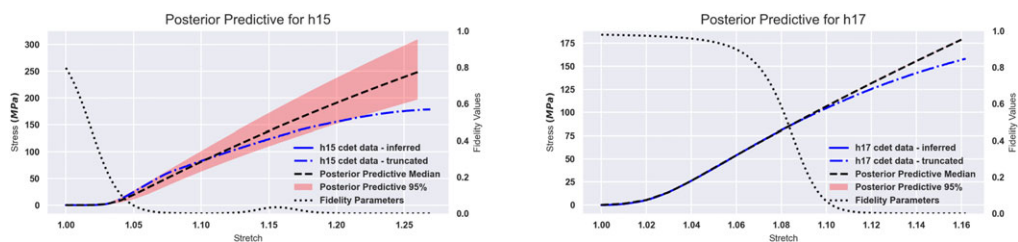


Figure 10. Posterior predicts for the CDET datasets h15 and h17. For h15, there are not many data points as determined by the fidelity threshold chosen, whereas h17 has many data points. The difference in inferential power is seen by the 95% spread of the predictive. The large spread is due to the lack of information about the mechanical parameters $(1 - \phi)\mu$ and ϕE .

The means of the plotted distributions are shown as vertical lines. As anticipated, using the linear modulus as a direct estimate of ϕE underestimates its value compared to the inferences from our model; however, there is considerable overlap in the assumed distribution of the linear modulus and our posterior distributions for ϕE .

(ii) Posterior predictives for the stress

To visualize the fit and the amount of data included in the inference in context, we considered the posterior predictive distribution for the response of each tendon. This can be understood as the marginal posterior distribution of new stresses y_i^* observed in a hypothetical replication of the experiment for the i th tendon at the same stretch levels as in the original data, and can be simulated by evaluating the microstructural model $\mathcal{M}(\cdot)$ at the observed stretches λ_i , using posterior samples for the model parameters θ_i .

Figure 10 shows posterior predictives for the response of two CDET individuals that have differing numbers of data points after being trimmed using our data selection method. We see that, for data with fidelity parameters that do not decay until well within region III of the stress response, there is a very tight predictive with a small standard deviation about the median. By contrast, where much of the data have been removed in the selection process, the inference results in a diffuse predictive with a high standard deviation. The reduction in the amount of inferred data included in the inference greatly increases the uncertainty in the posterior predictive distributions of the parameters in experiments with small N_i . However, the data selection methodology enabled us to fit our model to regions where the probability of the data

being consistent with the chosen model was highest, ultimately giving a truer reflection of the constitutive parameter values.

4. Discussion

In this paper, we introduced a two-stage process for producing higher-quality inferences from stress–strain tests to failure, where the mechanical model that we fitted to the data does not account for damage or failure. Several studies have proposed methods to identify the start of the yield region, often relying on curve-fitting or localized gradient estimates. Techniques such as moving-average smoothing of tangent moduli [33], spline fitting for yield point detection [54] and polynomial fitting to identify inflection points [55] offer practical means of identifying which data should be used for inference, though they rely on heuristic choices. Others have noted that there can be challenges in parameter identifiability, particularly when relying on uniaxial data alone [56].

The first stage of our approach involved inferring fidelity parameters which measure how consistent the model is with the data. By using a Gaussian process prior for the underlying fidelity field, we were able to learn the different regions of observation space for which the model is not valid. After inferring fidelity parameters independently for each dataset, a fidelity threshold was chosen which was used to trim each dataset automatically, and the observational noise value for each observation was scaled by its fidelity value, effectively tuning out data points which are less consistent with the model. As used in previous work [19], we chose a threshold of 0.3.

Using the Bayesian data selection method, we obtained higher-quality inferences which accurately trimmed the data on a per-individual basis, to minimize bias and maximize information quality. This was necessary because of the heterogeneity of the data between experiments, leading to the model being a valid representation of the data for different strain ranges in each case. This research provides a basis for improving the predictive power of results derived from Bayesian inferences and point estimates as our data selection method is agnostic to the statistical method chosen in the second stage. It is clear from figures 3 and 7 that capturing the heterogeneity in damage is important in the modelling process and ignoring it will skew parameter estimates away from their true values. It is important to note the additional computational cost of our approach; an additional inference was required for each of the tendon experiments to appropriately trim the data. However, these can be conducted in parallel, and on modern computer infrastructures this can be achieved straightforwardly. The additional cost of the data selection is still relatively small in comparison with the cost of characterizing the Bayesian mixed effects posterior, which has a large number of correlated parameters that make mixing of the Markov chain challenging.

Following the data selection stage, the trimmed and re-weighted data were fed into a Bayesian mixed effects statistical model which was used to infer population-level parameters for the mechanical model. We found that the product of the collagen volume fraction and collagen fibril Young's modulus, ϕE , has posterior modal values of 811.5 MPa for the SDFT and 1430.2 MPa for the CDET. We found that CDETs are stiffer than SDFTs, due either to having stiffer fibrils, a higher collagen volume fraction or both. We also found that the SDFTs have longer fibrils on average.

The better a model is at representing data, the more of the data can be used to infer its parameters. For example, using models which attempt to model damage or failure of the tendon [57,58] will increase the number of data points that are valid under the assumptions of the model, and therefore less data will be lost to trimming. The use of a more computationally costly mechanical model, however, will add to the computational complexity of characterizing the posterior distribution. In addition, it is very likely that however sophisticated and flexible the chosen model is, the experimental data will still require some level of selection prior to analysis to arrive at accurate parameter estimates. Combining sophisticated data selection methodology with mixed effects models enabled us to analyse the natural variability of parameters which define the physical properties of horse tendons more accurately than previous approaches.

Data accessibility. The code used to implement the methods explained in §2d,f is available [59]. The raw data [46] used in this study are available at: <https://qmro.qmul.ac.uk/xmlui/handle/123456789/13395>.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. J.J.C.: investigation, methodology, software, validation, visualization, writing—original draft and writing—review and editing; J.E.F.: funding acquisition, investigation, methodology, software, visualization, writing—original draft and writing—review and editing; T.W.W.: methodology, supervision and writing—review and editing; S.L.C.: conceptualization, methodology, project administration, supervision, writing—original draft and writing—review and editing; and T.S.: conceptualization, methodology, project administration, supervision and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. J.F. was supported by the Engineering and Physical Sciences Research Council (EPSRC) (grant no. EP/W522466/1). J.C. was supported by the Engineering and Physical Sciences Research Council (EPSRC) (grant no. EP/T517823/1).

Appendix A. Parameter transform

In [16], a transformation of the model parameters θ was used to improve the computational efficiency of the random walk Metropolis–Hastings algorithm. The transformation arises from considering the natural bounds on the model parameters. The parameters, E and μ , are greater than 0, and ϕ is a parameter constrained to the interval $[0, 1]$; therefore, $(1 - \phi)\mu$ and ϕE are non-negative. Thus, a logarithmic transform is appropriate. The structural parameter, a , is greater than 1, and b is greater than a , so we consider the parameters $a - 1$ and $b - a$ which are both greater than 0 and such that their logarithms exist on the entire real line. The transform $T_\theta(\cdot)$ and its inverse $T_\theta^{-1}(\cdot)$ is therefore given by

$$T_\theta^{-1}(\theta) = \xi = \begin{pmatrix} \nu \\ \eta \\ \tau \\ \rho \end{pmatrix} = \begin{pmatrix} \ln\{(1 - \phi)\mu\} \\ \ln\{\phi E\} \\ \ln\{a - 1\} \\ \ln\{b - a\} \end{pmatrix} \iff \theta = T_\theta(\xi) = \begin{pmatrix} \exp\{\nu\} \\ \exp\{\eta\} \\ \exp\{\tau\} + 1 \\ \exp\{\rho\} + \exp\{\tau\} + 1 \end{pmatrix}.$$

References

- Holzappel GA, Ogden RW. 2010 Constitutive modelling of arteries. *Proc. R. Soc. A* **466**, 1551–1597. (doi:10.1098/rspa.2010.0058)
- Mishani S, Belhouf-Fakir H, Lagat C, Jansen S, Evans B, Lawrence-Brown M. 2021 Stress distribution in the walls of major arteries: implications for atherogenesis. *Quant. Imag. Med. Surg.* **11**, 3494–3505. (doi:10.21037/qims-20-614)
- Limbert G. 2019 Constitutive modelling of skin mechanics. In *Skin Biophysics: From Experimental Characterisation to Advanced Modelling* (ed. G Limbert), pp. 19–76. Cham, Switzerland: Springer International Publishing. (doi:10.1007/978-3-030-13279-8_2)
- Pissarenko A, Ruestes CJ, Meyers MA. 2020 Constitutive description of skin dermis: through analytical continuum and coarse-grained approaches for multi-scale understanding. *Acta Biomater.* **106**, 208–224. (doi:10.1016/j.actbio.2020.01.026)
- Shearer T. 2015 A new strain energy function for the hyperelastic modelling of ligaments and tendons based on fascicle microstructure. *J. Biomech.* **48**, 290–297. (doi:10.1016/j.jbiomech.2014.11.031)
- Shearer T. 2015 A new strain energy function for modelling ligaments and tendons whose fascicles have a helical arrangement of fibrils. *J. Biomech.* **48**, 3017–3025. (doi:10.1016/j.jbiomech.2015.07.032)
- Shearer T, Parnell WJ, Lynch B, Screen HR, David Abrahams I. 2020 A recruitment model of tendon viscoelasticity that incorporates fibril creep and explains strain-dependent relaxation. *J. Biomech. Eng.* **142**, 071003. (doi:10.1115/1.4045662)
- Theodossiou SK, Schiele NR. 2019 Models of tendon development and injury. *BMC Biomed. Eng.* **1**, 32. (doi:10.1186/s42490-019-0029-5)

9. Zhao X, Pelegri AA. 2016 A Bayesian approach for characterization of soft tissue viscoelasticity in acoustic radiation force imaging. *Int. J. Numer. Methods Biomed. Eng.* **32**, e02741.
10. Teferra K, Brewick PT. 2019 A Bayesian model calibration framework to evaluate brain tissue characterization experiments. *Comput. Methods Appl. Mech. Eng.* **357**, 112604. (doi:10.1016/j.cma.2019.112604)
11. Elmukashfi E, Marchiori G, Berni M, Cassiolas G, Lopomo NF, Rappel H, Girolami M, Barrera O. 2022 Model selection and sensitivity analysis in the biomechanics of soft tissues: a case study on the human knee meniscus. *Adv. Appl. Mech.* **55**, 425–511. (doi:10.1016/bs.aams.2022.05.001)
12. Wang JHC. 2006 Mechanobiology of tendon. *J. Biomech.* **39**, 1563–1582. (doi:10.1016/j.jbiomech.2005.05.011)
13. Haken R, Blümich B. 2000 Anisotropy in tendon investigated in vivo by a portable NMR scanner, the NMR-MOUSE. *J. Magn. Reson.* **144**, 195–199. (doi:10.1006/jmre.2000.2040)
14. Böl M, Ehret AE, Leichsenring K, Ernst M. 2015 Tissue-scale anisotropy and compressibility of tendon in semi-confined compression tests. *J. Biomech.* **48**, 1092–1098. (doi:10.1016/j.jbiomech.2015.01.024)
15. Hoffmeister BK, Handley SM, Wickline SA, Miller JG. 1996 Ultrasonic determination of the anisotropy of Young's modulus of fixed tendon and fixed myocardium. *J. Acoust. Soc. Am.* **100**, 3933–3940. (doi:10.1121/1.417246)
16. Haughton J, Cotter SL, Parnell WJ, Shearer T. 2022 Bayesian inference on a microstructural, hyperelastic model of tendon deformation. *J. R. Soc. Interface* **19**, 20220031. (doi:10.1098/rsif.2022.0031)
17. Jull G, Moore A, Falla D, Lewis J, McCarthy C, Sterling M. 2015 *Grieve's modern musculoskeletal physiotherapy*. Amsterdam, The Netherlands: Elsevier Health Sciences.
18. Shearer T, Thorpe CT, Screen HRC. 2017 The relative compliance of energy-storing tendons may be due to the helical fibril arrangement of their fascicles. *J. R. Soc. Interface* **14**, 20170261. (doi:10.1098/rsif.2017.0261)
19. Cotter SL. 2024 Hierarchical Bayesian data selection. *ACM Trans. Prob. Mach. Learn.* **1**, article number 7 (1–32). (doi:10.1145/3699721)
20. Ciarletta P, Micera S, Accoto D, Dario P. 2006 A novel microstructural approach in tendon viscoelastic modelling at the fibrillar level. *J. Biomech.* **39**, 2034–2042. (doi:10.1016/j.jbiomech.2005.06.025)
21. Graham JS, Vomund AN, Phillips CL, Grandbois M. 2004 Structural changes in human type I collagen fibrils investigated by force spectroscopy. *Exp. Cell Res.* **299**, 335–342. (doi:10.1016/j.yexcr.2004.05.022)
22. Svensson RB, Hansen P, Hassenkam T, Haraldsson BT, Aagaard P, Kovanen V, Krogsgaard M, Kjaer M, Magnusson SP. 2012 Mechanical properties of human patellar tendon at the hierarchical levels of tendon and fibril. *J. Appl. Physiol.* **112**, 419–426. (doi:10.1152/japplphysiol.01172.2011)
23. Wang S *et al.* 2019 Bayesian inference-based estimation of normal aortic, aneurysmal and atherosclerotic tissue mechanical properties: from material testing, modeling and histology. *IEEE Trans. Biomed. Eng.* **66**, 2269–2278. (doi:10.1109/TBME.2018.2886681)
24. Wang YN, Lee K, Shofer JB, Ledoux WR. 2017 Histomorphological and biochemical properties of plantar soft tissue in diabetes. *The Foot* **33**, 1–6. (doi:10.1016/j.foot.2017.06.001)
25. Wang P, Yan Z, Du Z, Fu Y, Liu Z, Qu S, Zhuang Z. 2023 A Bayesian method with nonlinear noise model to calibrate constitutive parameters of soft tissue. *J. Mech. Behav. Biomed. Mater.* **146**, 106070. (doi:10.1016/j.jmbbm.2023.106070)
26. Condliff MK, Lewis DD, Madigan D, Posse C. 1999 Bayesian mixed-effects models for recommender systems. In *ACM SIGIR 99 Workshop on Recommender Systems: Algorithms and Evaluation*. vol. 99, pp. 23–30. Berkeley, CA: ACM Press.
27. Ziegler G, Penny WD, Ridgway GR, Ourselin S, Friston KJ, Alzheimer's Disease Neuroimaging Initiative. 2015 Estimating anatomical trajectories with Bayesian mixed-effects modeling. *Neuroimage* **121**, 51–68. (doi:10.1016/j.neuroimage.2015.06.094)
28. Patoux A, Bleyzac N, Boddy A, Doz F, Rubie H, Bastian G, Maire P, Canal P, Chatelut E. 2001 Comparison of nonlinear mixed-effect and non-parametric expectation maximisation modelling for Bayesian estimation of carboplatin clearance in children. *Eur. J. Clin. Pharmacol.* **57**, 297–303. (doi:10.1007/s002280100306)

29. Stan Development Team. 2024 Stan Modeling Language Users Guide and Reference Manual, version 2.34. See <https://mc-stan.org>
30. Brooks S, Gelman A, Jones G, Meng XL. 2011 *Handbook of Markov Chain Monte Carlo*. London: Chapman and Hall/CRC.
31. Holzapfel GA, Gasser TC, Ogden RW. 2000 A new constitutive framework for arterial wall mechanics and a comparative study of material models. *J. Elast.* **61**, 1–48. (doi:10.1023/A:1010835316564)
32. Goh KL, Holmes DF, Lu YH, Kadler KE, Purslow PP. 2018 Age-related dataset on the mechanical properties and collagen fibril structure of tendons from a murine model. *Sci. Data* **5**, article number 180140. (doi:10.1038/sdata.2018.140)
33. Thorpe CT, Udeze CP, Birch HL, Clegg PD, Screen HR. 2012 Specialization of tendon mechanical properties results from interfascicular differences. *J. R. Soc. Interface* **9**, 3108–3117. (doi:10.1098/rsif.2012.0362)
34. Forsyth JE, Al-Anbaki AH, Plusa B, Cotter SL. 2023 Unlabelled landmark matching via Bayesian data selection, and application to cell matching across imaging modalities. *Stat. Comput.* **33**, 100. (doi:10.1007/s11222-023-10259-7)
35. Holmes CC, Walker SG. 2017 Assigning a value to a power likelihood in a general Bayesian model. *Biometrika* **104**, 497–503.
36. Gelman A, Gilks WR, Roberts GO. 1997 Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7**, 110–120. (doi:10.1214/aoap/1034625254)
37. Betancourt M. 2020 Hierarchical Modeling—github pages. See https://betanalpha.github.io/assets/case_studies/hierarchical_modeling.html (Accessed 24 December 2023).
38. Lee SY. 2022 Bayesian nonlinear models for repeated measurement data: an overview, implementation, and applications. *Mathematics* **10**, article number 898. (doi:10.3390/math10060898)
39. Gelman A, Hill J. 2006 *Data analysis using regression and multilevel/hierarchical models*. Analytical Methods for Social Research. Cambridge, UK: Cambridge University Press.
40. De Finetti B. 1972 *Probability, induction and statistics: the art of guessing*. Wiley Series in Probability and Statistics: Probability and Statistics Section Series. Chichester, UK: J. Wiley.
41. Stefan AM, Katsimpokis D, Gronau QF, Wagenmakers EJ. 2022 Expert agreement in prior elicitation and its effects on Bayesian inference. *Psychon. Bull. Rev.* **29**, 1776–1794. (doi:10.3758/s13423-022-02074-4)
42. Mikkola P *et al.* 2024 Prior Knowledge Elicitation: The Past, Present, and Future. *Bayesian Analysis* **19**(4), 1129–1161. (December 2024) (doi:10.1214/23-BA1381)
43. Gelman A, Simpson D, Betancourt M. 2017 The prior can often only be understood in the context of the likelihood. *Entropy* **19**, article number 555. (doi:10.3390/e19100555)
44. Silver FH, Seehra GP, Freeman JW, DeVore D. 2002 Viscoelastic properties of young and old human dermis: A proposed molecular mechanism for elastic energy storage in collagen and elastin. *J. Appl. Polym. Sci.* **86**, 1978–1985. (doi:10.1002/app.11119)
45. Purslow PP. 2009 The shear modulus of connections between tendon fascicles. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, pp. 134–136. (doi:10.1109/TIC-STH.2009.5444520).
46. Goh KL, Holmes DF, Lu HY, Richardson S, Kadler KE, Purslow PP, Wess TJ. 2008 Ageing changes in the tensile properties of tendons: influence of collagen fibril volume fraction. *J. Biomech. Eng.* **130**, 021011. (doi:10.1115/1.2898732)
47. Roininen L, Lehtinen M, Lasanen S, Orispää M, Markkanen M. 2011 Correlation priors. *Inverse Probl. Imag.* **1**, 167–184. (doi:10.3934/ipi.2011.5.167)
48. Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D. 2013 *Bayesian data analysis*, 3rd edn. Chapman & Hall/CRC Texts in Statistical Science. London, UK: Taylor & Francis.
49. Lewandowski D, Kurowicka D, Joe H. 2009 Generating random correlation matrices based on vines and extended onion method. *J. Multivariate Anal.* **100**, 1989–2001. (doi:10.1016/j.jmva.2009.04.008)
50. Bürkner PC. 2017 BRMS: an R package for bayesian multilevel models using stan. *J. Stat. Softw.* **80**, 1–28. (doi:10.18637/jss.v080.i01)
51. Duane S, Kennedy A, Pendleton BJ, Roweth D. 1987 Hybrid Monte Carlo. *Phys. Lett. B* **195**, 216–222. (doi:10.1016/0370-2693(87)91197-X)
52. Hoffman MD, Gelman A. 2014 The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**(1), 1593–1623.

53. Gabry J, Cesnovar R, Johnson A. 2023 *CmdStanr: R Interface to 'CmdStan'*. See <https://mc-stan.org/cmdstanr/>, <https://discourse.mc-stan.org>.
54. Peloquin JM, Santare MH, Elliott DM. 2016 Advances in quantification of meniscus tensile mechanics including nonlinearity, yield, and failure. *J. Biomech. Eng.* **138**, 021002. (doi:10.1115/1.4032354)
55. Goh KL, Holmes D, Lu Y, Purslow PP, Kadler K, Bechet D, Wess TJ. 2012 Bimodal collagen fibril diameter distributions direct age-related variations in tendon resilience and resistance to rupture. *J. Appl. Physiol.* **113**, 878–888. (doi:10.1152/japplphysiol.00258.2012)
56. Safa BN, Santare MH, Ethier CR, Elliott DM. 2021 Identifiability of tissue material parameters from uniaxial tests using multi-start optimization. *Acta Biomater.* **123**, 197–207. (doi:10.1016/j.actbio.2021.01.006)
57. Gregory J, Hazel AL, Shearer T. 2021 A microstructural model of tendon failure. *J. Mech. Behav. Biomed. Mater.* **122**, 104665. (doi:10.1016/j.jmbbm.2021.104665)
58. Hamedzadeh A, Gasser TC, Federico S. 2018 On the constitutive modelling of recruitment and damage of collagen fibres in soft biological tissues. *Eur. J. Mech. A. Solids* **72**, 483–496. (doi:10.1016/j.euromechsol.2018.04.007)
59. Casey JJ, Forsyth J. 2025 JJ-Casey/TendonMEDS-Code: Initial Release (v1.0.0). Zenodo. (doi:10.5281/zenodo.16929076)