

Research



**Cite this article:** Longobardi S, Lewalle A, Coveney S, Sjaastad I, Espe EKS, Louch WE, Musante CJ, Sher A, Niederer SA. 2020 Predicting left ventricular contractile function via Gaussian process emulation in aortic-banded rats. *Phil. Trans. R. Soc. A* **378**: 20190334.  
<http://dx.doi.org/10.1098/rsta.2019.0334>

Accepted: 27 April 2020

One contribution of 16 to a theme issue 'Uncertainty quantification in cardiac and cardiovascular modelling and simulation'.

**Subject Areas:**

computational biology, computer modelling and simulation, biomedical engineering

**Keywords:**

Gaussian process, history matching, global sensitivity analysis, three-dimensional bi-ventricular model, aortic-banded rat

**Author for correspondence:**

Stefano Longobardi  
e-mail: [stefano.longobardi@kcl.ac.uk](mailto:stefano.longobardi@kcl.ac.uk)

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4977725>.

# Predicting left ventricular contractile function via Gaussian process emulation in aortic-banded rats

S. Longobardi<sup>1</sup>, A. Lewalle<sup>1</sup>, S. Coveney<sup>2</sup>, I. Sjaastad<sup>3</sup>, E. K. S. Espe<sup>3</sup>, W. E. Louch<sup>3</sup>, C. J. Musante<sup>4</sup>, A. Sher<sup>4</sup> and S. A. Niederer<sup>1</sup>

<sup>1</sup>School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK

<sup>2</sup>Insigneo Institute for in-silico Medicine and Department of Computer Science, University of Sheffield, Sheffield, UK

<sup>3</sup>Institute for Experimental Medical Research and KG Jebsen Center for Cardiac Research, University of Oslo, Oslo, Norway

<sup>4</sup>Pfizer Worldwide Research, Development and Medical, Cambridge, MA, USA

SL, 0000-0002-0111-0047; SAN, 0000-0002-4612-6982

Cardiac contraction is the result of integrated cellular, tissue and organ function. Biophysical *in silico* cardiac models offer a systematic approach for studying these multi-scale interactions. The computational cost of such models is high, due to their multi-parametric and nonlinear nature. This has so far made it difficult to perform model fitting and prevented global sensitivity analysis (GSA) studies. We propose a machine learning approach based on Gaussian process emulation of model simulations using probabilistic surrogate models, which enables model parameter inference via a Bayesian history matching (HM) technique and GSA on whole-organ mechanics. This framework is applied to model healthy and aortic-banded hypertensive rats, a commonly used animal model of heart failure disease. The obtained probabilistic surrogate models accurately predicted the left ventricular pump function ( $R^2 = 0.92$  for ejection fraction). The HM technique allowed us to fit both the control and diseased virtual bi-ventricular rat heart models to

© 2020 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.

magnetic resonance imaging and literature data, with model outputs from the constrained parameter space falling within 2 SD of the respective experimental values. The GSA identified Troponin C and cross-bridge kinetics as key parameters in determining both systolic and diastolic ventricular function.

This article is part of the theme issue 'Uncertainty quantification in cardiac and cardiovascular modelling and simulation'.

## 1. Introduction

With each beat, cardiac myocytes generate tension and relax. Cellular tension is transduced into a coordinated, global whole-heart deformation resulting in an effective, system-level pump function. The integration of cellular, tissue and organ-scale mechanisms is essential for achieving efficient transduction of work into concerted myocardial contraction and relaxation. The break down of this system of integrated mechanisms can give rise to heart failure (HF).

Cardiac biophysical models provide a useful tool for studying whole-organ contraction [1] by simulation, and can therefore be used to understand how impaired cell level function is linked with impaired organ level activity. Building a virtual representation of the real system requires tuning the cardiac model's properties and boundary conditions to experimental data. Global sensitivity analysis (GSA) studies are also needed to determine the role of cellular, tissue and haemodynamics properties on whole-organ function.

However, the high simulation cost and the multi-scale nature of the cardiac system have limited the formal parameter estimation and sensitivity analysis for this type of model. Different techniques have been employed for global parameter inference on cardiac cell models, including gradient descent [2], genetic algorithms [3], multivariate regression [4] and Markov chain Monte Carlo (MCMC) [5]. Most of these approaches are computationally intensive, and the computation burden is even higher when going from single to multiple scales.

Surrogate models (or *emulators*) can be used to replace the computationally expensive real model with a statistical representation of it, allowing rapid predictions of the model's output. Using emulators we can thus perform the high number of model evaluations required to estimate global sensitivities. In this study, we use Gaussian process emulators (GPEs), which provide confidence bounds on predictions and can therefore be used for Bayesian history matching (HM). Previously employed for fitting models of galaxy formation [6], infectious disease transmission [7] and, more recently, human atrial cell [8], HM has shown to be a valuable tool for global parameter inference.

To demonstrate the capacity of GPE and HM to fit a virtual heart model, we tune model mechanics-related parameters to data from control and left ventricular (LV) hypertrophied rat hearts. LV hypertrophy is induced surgically by constricting the ascending aorta over four to six weeks and represents a common pathway to HF development. Aortic-banded (AB) rats are commonly used as an experimental animal model for HF pathology [9].

We also seek to quantitatively characterize the AB rats' impaired LV contractile function. We therefore fit the virtual bi-ventricular rat model to changes in function that we measured or were recorded by others, using HM. Finally, we perform a GSA to understand how uncertainty in cell, tissue and haemodynamics model properties affect LV function.

The present work is organized as follows. Section 2 presents the rat experimental data and describes the three-dimensional multi-scale bi-ventricular mathematical model of rat heart contraction and the developed machine learning framework based on GPE. Section 3 shows results of the emulation and fitting processes and the GSA. Section 4 provides a discussion and §5 addresses some model limitations, while §6 closes with a brief summary. Additional details can be found in the electronic supplementary material.

## 2. Methods

### (a) Experimental data

We used control and AB rat data from experimental studies conducted by Røe *et al.* [10]. Briefly, aortic banding was performed in male Wistar rats ( $n = 29$ ), and sham-operated (SHAM) rats ( $n = 23$ ) served as controls. *In vivo* cardiac function was characterized by echocardiography six weeks after surgery. The criterion for inclusion in the AB group was posterior wall thickening (more than 1.9 mm). A cohort of rats ( $n = 8$ ,  $n = 15$  from SHAM, AB groups, respectively) underwent an examination with magnetic resonance imaging (MRI). Rats in the AB cohort showed a preserved systolic function (ejection fraction and fractional shortening). Concentric LV hypertrophy arose, as an increase in LV mass and wall thickness was observed. Moreover, evidence of reduced peak early diastolic velocity ( $e'$ ) and increased early mitral inflow velocity/mitral annular early diastolic velocity ( $E/e'$ ) confirmed an impaired diastolic function in the rats. Representative cine MRI data were obtained from a SHAM and AB rat. Each image set consisted of 47 frames. A time frame halfway through diastole was chosen for segmentation purposes to approximate the stress-free configuration.

### (b) Rat heart model

The heart was modelled as described previously [11,12]. Briefly, the LV and RV anatomies were represented by a cubic Hermite mesh [13] fitted to manually segmented MRI images. Rule-based fibres orientation was defined with a transmural variation from  $-60^\circ$  (epicardium) to  $80^\circ$  (endocardium) [12]. Electrical activation in the heart was driven by a fixed intracellular calcium ( $\text{Ca}^{2+}$ ) stimulus applied simultaneously throughout the ventricular tissue. Previous heart models have shown that heterogeneous activation patterns have negligible impact in the small rat heart [11]. Rat ventricular myocytes'  $\text{Ca}^{2+}$  transient was described using the Gattoni *et al.* models [14] for SHAM and AB rats hearts at  $37^\circ\text{C}$  and 6 Hz pacing frequency. Rat cellular contraction was modelled using the Land *et al.* [11] model to simulate tension generation arising from the thin and thick filaments. This model simulates troponin C, tropomyosin and cross-bridge dynamics and predicts the tension generated by cardiac muscle in response to changes in intracellular  $\text{Ca}^{2+}$  concentration, sarcomere length and velocity. The passive material properties were modelled using the Guccione transversely isotropic incompressible constitutive equation [15], which is widely used across the cardiac modelling community [1,12,16,17]. The methods applied here are generalizable and could equally be applied to other myocardium constitutive equations (e.g. [18]). The spatial boundary conditions were defined by constraining the LV basal plane to lie perpendicular to the apex–base axis, with one mesh node on the interior LV wall fixed in all directions to stop rigid body motion. The pressure boundary conditions were defined by constraining the aortic pressure, which couples the heart with the systemic circulation via a three-element Windkessel model [19].

### (c) Mapping model parameters to organ scale features

We selected a finite number of key regulators (defined in table 1) for each of the employed sub-models to have a complete description of the heart mechanics, giving the following input parameter vector for the multi-scale model:

$$\mathbf{m} = (p, ap, z, c_1, ca_{50}, k_{xb}, k_{off}, T_{ref}). \quad (2.1)$$

The LV function was quantified by timings and magnitudes (defined in table 2) extracted from the volume (LVV) and pressure (LVP) transients. We selected 12 relevant quantities, giving the

**Table 1.** Input parameters of the rat bi-ventricular heart mechanics model.

label	units	definition
$p$	kPa	end-diastolic pressure
$ap$	kPa	aortic pressure
$z$	mmHg s ml <sup>-1</sup>	aortic characteristic impedance in Windkessel model
$c_1$	kPa	tissue stiffness in Guccione constitutive equation
$ca_{50}$	$\mu$ M	calcium sensitivity at resting sarcomere length
$k_{xb}$	ms <sup>-1</sup>	binding rate in cross-bridge cycling equation
$k_{off}$	ms <sup>-1</sup>	unbinding rate in Ca <sup>2+</sup> -Troponin C cooperative binding equation
$T_{ref}$	kPa	maximal cellular force in active tension development equation

**Table 2.** Output features of the rat bi-ventricular heart mechanics model.

label	units	definition	label	units	definition
EDV	$\mu$ L	end-diastolic volume	Tdiast	ms	diastolic time
ESV	$\mu$ L	end-systolic volume	PeakP	kPa	peak pressure
EF	%	ejection fraction	Tpeak	ms	time to peak pressure
IVCT	ms	isovolumetric contraction time	ESP	kPa	end-systolic pressure
ET	ms	ejection time	maxdP	kPa ms <sup>-1</sup>	max. pressure rise rate
IVRT	ms	isovolumetric relaxation time	mindP	kPa ms <sup>-1</sup>	max. pressure decay rate

following output feature vector for the multi-scale model:

$$\mathbf{d} = (\text{EDV}, \text{ESV}, \text{EF}, \text{IVCT}, \text{ET}, \text{IVRT}, \text{Tdiast}, \text{PeakP}, \text{Tpeak}, \text{ESP}, \text{maxdP}, \text{mindP}), \quad (2.2)$$

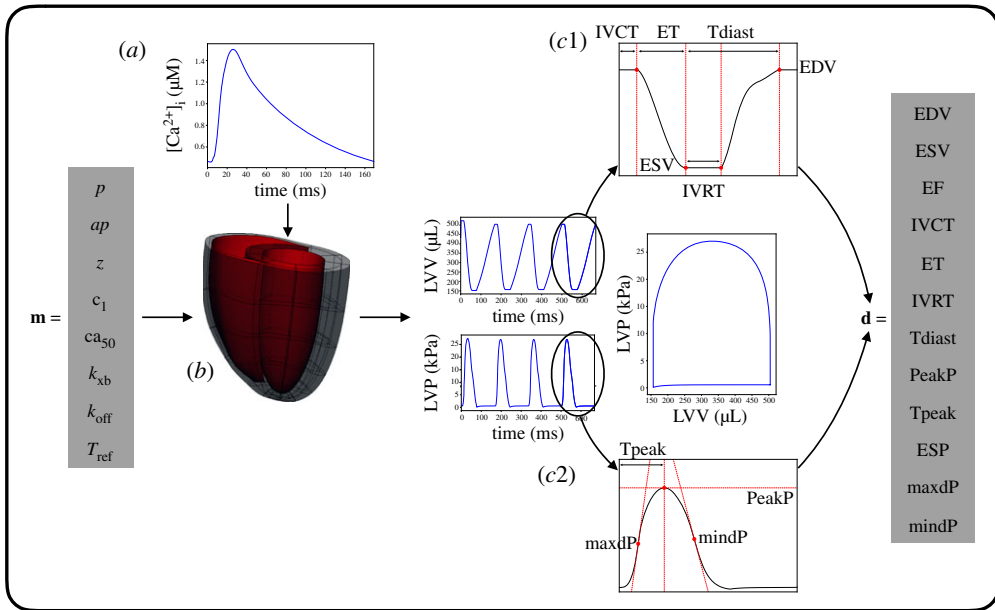
where the first seven components characterize the LVV curve, and the last five components characterize the LVP curve.

### (d) The multi-scale model of rat heart contraction

Given a fixed Ca<sup>2+</sup> transient and a fixed cubic Hermite template mesh with fibres as an anatomical representation, the multi-scale model of rat heart contraction can be described by a function that maps the  $\mathbf{m}$  input parameter vector (equation (2.1)) into the  $\mathbf{d}$  output feature vector (equation (2.2)). Figure 1 shows a schematic of the multi-scale model.

### (e) Gaussian process emulation

The *in silico* rat heart biophysical model should aim to achieve the best representation of the real system. This requires tuning the model input parameters such that the model output features reflect the available experimental data. We trained GPEs on model simulations to replace the multi-scale model map (§2d) with a surrogate, fast evaluating map. Each GPE was defined as a combination of a mean function in the form of a linear regression model and a zero-mean GP (see electronic supplementary material, equations (1)–(3)). The GPE training is performed in two steps as described previously [6,20]. Firstly, the linear regression model's coefficients are fitted to the data by minimizing the residual sum of squares; secondly, the GP's hyperparameters are fitted to the residuals (data minus predicted mean) by maximization of the log-marginal likelihood function [21]. Tested mean functions were all the linear regression models (with intercept and interaction terms) with higher than first degree polynomials up to the third degree,



**Figure 1.** Multi-scale, bi-ventricular three-dimensional rat heart contraction model. Given a fixed calcium transient (a) and mesh with fibres (b), the vector  $\mathbf{m}$  of parameters representing cell, tissue and haemodynamics properties (table 1) constitutes the input for the model. LV volume and pressure transients and pressure-volume loop are obtained after a four-beats model run. The vector  $\mathbf{d}$  of LV features (table 2), extracted from the fourth-beat curves (c1–c2), constitutes the output for the model. (Online version in colour.)

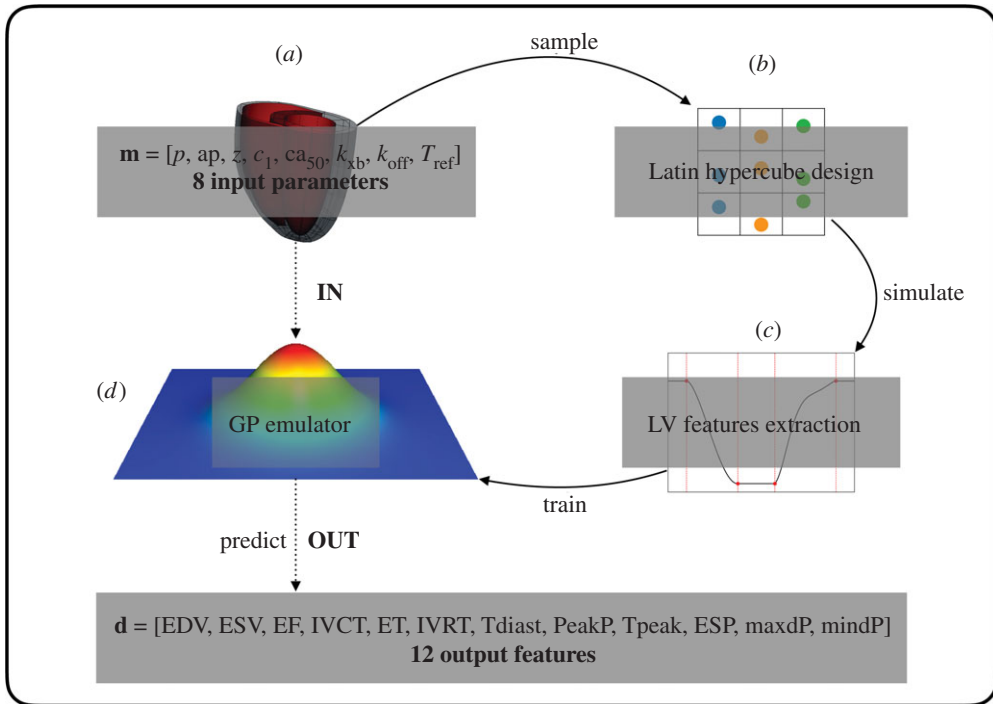
while tested GP kernels were anisotropic covariance functions among the squared exponential kernel, Matérn $_{\nu=3/2}$  and Matérn $_{\nu=5/2}$  kernels [21]. The accuracy of a GPE was evaluated using the averaged  $R^2$  (coefficient of determination) test score obtained in a five cross-validation. Briefly, one fifth of the training dataset was excluded and the GPE was trained on the remaining points. The left-out part was then used for testing the GPE accuracy, and an  $R^2$  score was calculated. This process was repeated for each of the five subsets randomly selected from the full training dataset. The final accuracy was determined by averaging the  $R^2$  scores obtained in predicting the five different left-out parts. The developed framework is outlined in figure 2.

## (f) Target values

We constrained the model parameters based on model readouts that describe key features in the LVV and LVP transients, and pressure–volume loop (table 2). In our study, the target phenotypes' mean and standard deviation came from two different sources. Specific organ-scale features unavailable in the MRI data [10] were deduced from literature experimental studies all performed on male rats at body temperature with abdominal/ascending aortic banding. In particular, EDV, ESV, ET and IVRT features were immediately available from the MRI data, while information about PeakP, maxdP and mindP features was collected from the literature. We did not have specific values for IVCT, Tdiast, Tpeak and ESP, and we chose not to match EF explicitly as this is derived from EDV and ESV. This gave us seven organ-scale phenotypes (see electronic supplementary material, Table S2) to constrain our model parameters.

## (g) Bayesian history matching

To constrain the initial mechanics-related parameter space to include only points that generate multi-scale models that are consistent with experimentally observed organ-scale LV features, we



**Figure 2.** Building the surrogate model. Input points (a) are sampled in the parameter space using a Latin Hypercube design (b). The model (*simulator*) is run using these points. The simulator output is composed of features extracted from the LV pressure and volume transients (c). The initial input matrix and the corresponding output matrix represent the training dataset for the surrogate model (*emulator*) based on Gaussian processes. The trained emulator (d) is used to predict the simulator output at a new point in the input parameter space. (Online version in colour.)

adopted the HM approach described in [8]. HM proceeds iteratively in ‘waves’. For each wave, the currently defined ‘not-ruled-out-yet’ (NROY) parameter space, defined by a large cloud of points, is tested with respect to an implausibility criterion: GPEs are evaluated for each NROY point, and an implausibility score reflecting the discrepancy between emulators’ predictions and data (see electronic supplementary material, equations (4)–(5)) is assigned to each point. Points with an implausibility score above a chosen ‘cutoff’ are deemed implausible, and the remaining ‘non-implausible’ subset becomes NROY for the next wave. The initial NROY space for the first wave is a large Latin hypercube design (LHD). For each following wave, the emulators are updated with data from simulations using a subset of points from the current NROY space, becoming more accurate and thus allowing NROY to be reduced further. Additional NROY points can be generated using the ‘cloud’ technique described in [8] if too few points remain for the next wave. The implausibility cutoff is reduced in each wave down to a predefined threshold, and HM continues until NROY converges.

### (h) Sobol sensitivity analysis

To assess how uncertainty in the model’s output can be attributed to different sources of uncertainty in the model input factors, we performed a variance-based sensitivity analysis. Variance-based methods of probabilistic sensitivity analysis quantify the sensitivity of the output to the model inputs in terms of a reduction in the variance of the output. We measured this reduction by estimating the Sobol sensitivity indices [22] through the Saltelli method [23] using the SALib Python library [24]. We followed the simple approach of estimating first-, second-order and total effect indices by neglecting the emulator’s uncertainty and using the emulator’s

mean as a point-wise approximation of the model's output. The employed emulators were trained on model simulations from the whole input parameter space before performing HM (first wave). Given the initial parameters' ranges, quasi-random low-discrepancy Sobol sequences were produced and the emulators were evaluated at points belonging to these sequences in order to obtain sensitivity indices estimates.

### 3. Results

#### (a) Model emulation

Eight LHDs of 1024 input parameter points each were simulated and the successfully completed simulations were collected to form the training dataset, with final size of 825 and 850 points for the SHAM and AB models, respectively. The low number of successful runs relates to a combination of failure of the mechanics simulations converging or completing a full cardiac cycle, which may happen if the contraction is insufficient to reach the aortic pressure. One GPE was trained for each output feature, for a total of 12 GPEs. The obtained cross-validation  $R^2$  test scores (GPEs' accuracy) are reported in the electronic supplementary material, Table S1. Emulator evaluation at a new parameter set took approximately 1.2 s against a full simulator single-core run of approximately 4 h, for a total gained speedup of 12 000-fold.

The trained GPEs were firstly employed to perform an entirely *in silico* GSA. A subset of seven GPEs (specifically the ones predicting the features for which we had experimental values to match) was then used in combination with the related experimental target values to fit the rat models.

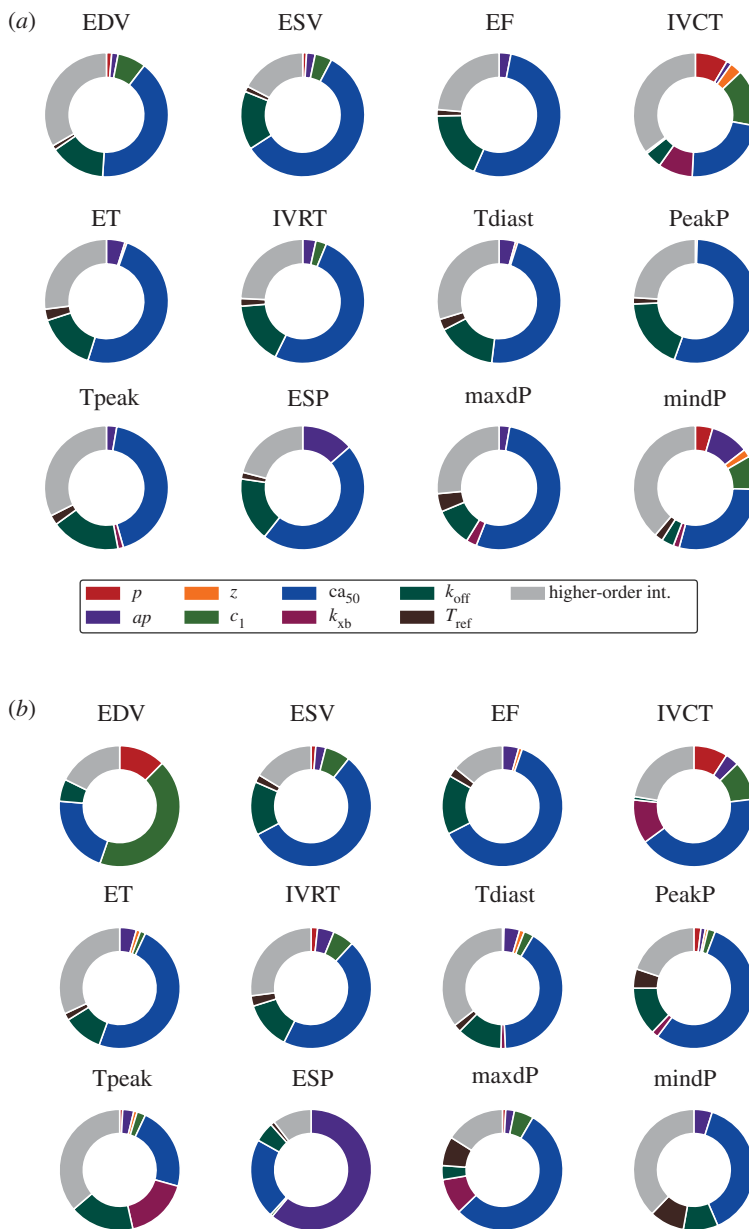
#### (b) Model global sensitivity analysis

Figure 3 shows how input parameters contribute to explaining each LV feature's total variance in SHAM and AB rat models. The  $ca_{50}$  parameter is the most influential across all the LV features. The  $k_{\text{off}}$ ,  $T_{\text{ref}}$  and  $ap$  parameters also play a key role. All these statements are true for both rat phenotypes. By contrast,  $c_1$  is more important for AB than for SHAM, explaining a fraction of the total variance of 9 out of 12 LV features (in SHAM,  $c_1$  is important for only five features). In particular, this parameter has a significant contribution in explaining end-diastolic volume in AB (and not in SHAM). The  $k_{\text{xb}}$  parameter affects the AB more than the SHAM features, albeit only moderately (5 out of 12 against 4 out of 12, respectively). Finally, we observe an increased contribution of  $ap$  in explaining end-systolic pressure when going from SHAM to AB. A complete description of all the first- and second-order parameters' interactions can be found in the electronic supplementary material.

#### (c) Model fitting

HM is illustrated in figure 4a for the SHAM and AB rat models. The eight-dimensional input parameter space is plotted as a two-dimensional projection for each pair of parameters. An initial set (lightest colour variant) consisting of 400 000 points sampled with an LHD in the input parameter space was tested against an *implausibility criterion* (see electronic supplementary material, equation (5)) and became progressively constrained to smaller regions (darker colour variants) as the HM process went forward. The initial parameters' one-dimensional ranges were determined after a literature search which included both experimental and modelling papers (see electronic supplementary material, Tables S3–S4). At each successive wave, 256 points uniformly sampled from the non-implausible region of the previous wave were simulated to build the new emulators' training dataset along with the previously used training dataset, while the remaining points from the non-implausible region of the previous wave plus 50 000 new points sampled in the same region constituted the testing set. The threshold constant used

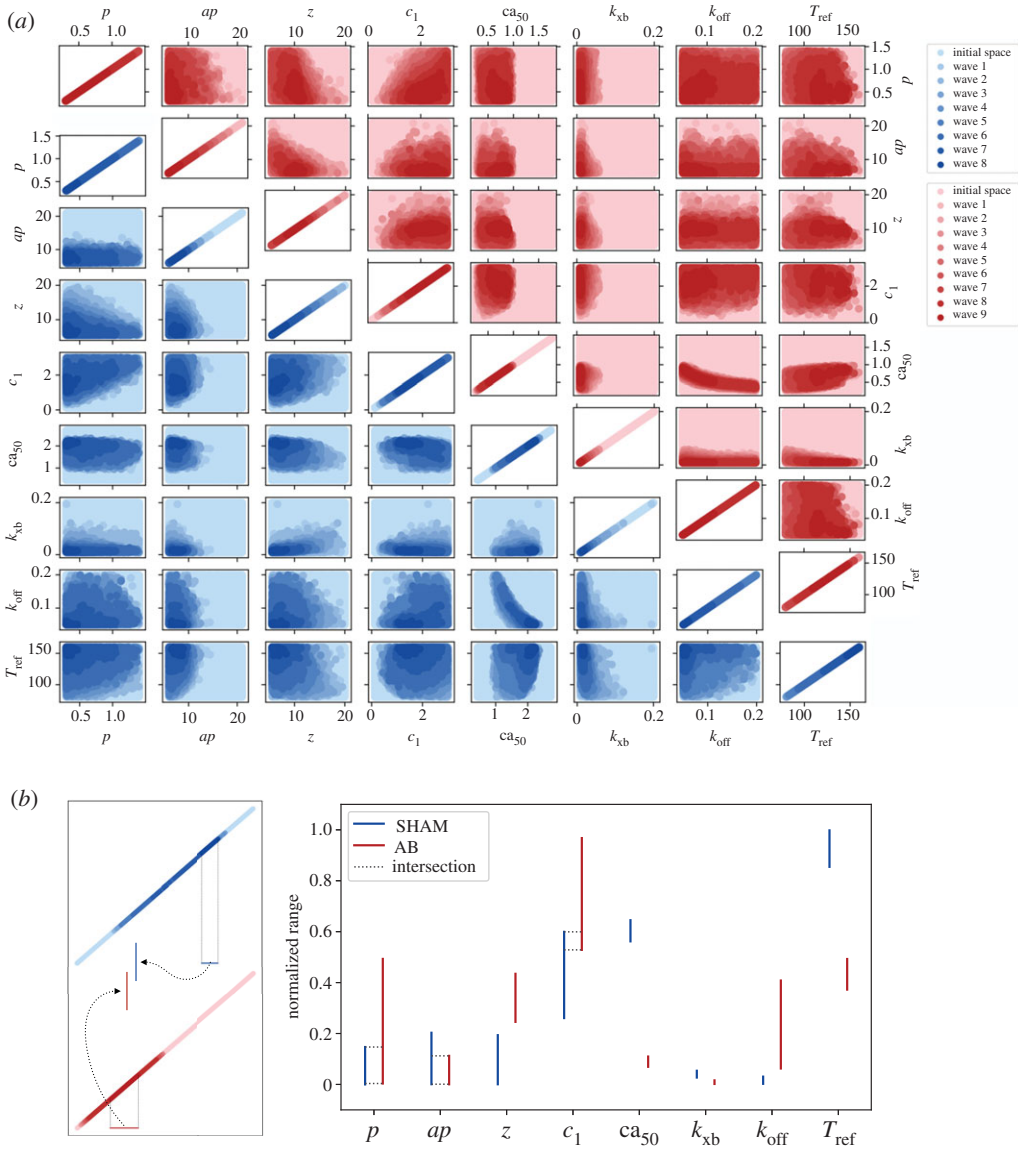




**Figure 3.** The impact of cell, tissue and boundary conditions properties on organ-scale phenotypes in SHAM (*a*) and AB (*b*) models. The contribution of each parameter is represented by the sum of its first- and (when present) second-order effects. For each LV feature, higher-order interactions (coloured in grey) are represented by the sum of total effects minus the sum of all first- and second-order effects. Parameters' and features' labels are explained in tables 1–2. (Online version in colour.)

for discerning whether or not a point was to be discarded (respectively, implausible and non-implausible point), started from 5.5 and 5.0 in the SHAM and AB models' tuning, respectively, and was progressively reduced by 0.5 units in each wave, until reaching the commonly used value of 3.0 (e.g. [6]). The choice of these specific starting values ensured that a non-empty non-implausible set was obtained after the first wave (see the electronic supplementary material for additional information on the cutoff initial tuning). The process was terminated when the





**Figure 4.** (a) High-dimensional input parameter space reduction during HM. SHAM HM (blue, lower diagonal axes) completed within eight waves while AB HM (red, upper diagonal axes) completed within nine waves. (b) Input parameters' one-dimensional ranges after completing HM procedure. Comparison between SHAM and AB rat models' fittings. For each input parameter, the displayed intervals are scaled according to the initial range for that specific parameter. When SHAM and AB intervals' intersection is non-empty, this is also provided with dotted lines. Parameters' labels are explained in table 1. (Online version in colour.)

fraction of non-implausible points became small (namely 0.1% and 1.1% of the total testing set for SHAM and AB models, respectively), completing SHAM and AB models' fitting in eight and nine waves, respectively. A total of 4352 model simulations (without considering the two initial waves, this is equivalent to a total of  $15 \text{ waves} \times 256 \text{ points}$ ) and 1 550 000 emulator evaluations ( $2 \text{ LHDs} \times 400\,000 \text{ points} + 15 \text{ waves} \times 50\,000 \text{ points}$ ) were performed.

Although HM does not ensure the reduction in individual parameters' ranges but rather a whole high-dimensional space reduction [8], the diagonal plots in figure 4a suggest that all

the parameters underwent reduction (to different degrees) along their one-dimensional domain, both in the SHAM and AB models. We therefore analysed whether these parameters had been constrained to lie in spaces which differ according to the rat phenotype. This phenomenon was observed in five of the eight total parameters, as highlighted in figure 4*b*. We can see that  $z$ ,  $ca_{50}$ ,  $k_{xb}$ ,  $k_{off}$  and  $T_{ref}$  take values in completely separated spaces for the SHAM and AB models. The biggest space separation is observed for  $ca_{50}$  and  $T_{ref}$ .

We finally investigated whether the constrained input parameter space was mapped by the simulator into model readouts that matched experimental observations. For this purpose, we sampled and simulated 1024 non-implausible points in the HM final wave's parameter space of both SHAM and AB models. This was performed using rejection sampling, i.e. randomly selecting new points with components in the final wave's reduced parameter ranges and checking their implausibility measure, eventually discarding the points with implausibility measure above the current cutoff value. Of these, only 264 and 178 points led to a successfully completed simulation for SHAM and AB models, respectively. After extracting the LV features from the simulator outputs, we examined their distributions around the related experimentally observed mean values (see electronic supplementary material, Table S2).

Figure 5*a*.1–2 shows that points belonging to the reduced input parameter space reproduce the majority of the emerging organ-scale LV features except ESV, maxdP and mindP which distribute far from the experimentally observed mean values. The simulated LVV and LVP transients which the features were extracted from are also reported (figure 5*b*.1–2). The resulting pressure-volume loops are provided in figure 5*c*.1–2. By comparing the highlighted average curves, we can notice smaller end-diastolic and end-systolic volumes for AB, which in turn still preserve the ejection fraction observed in the control. Diastolic-time is increased in AB due to an increase in the isovolumetric relaxation time. LV peak pressure is visibly higher in AB.

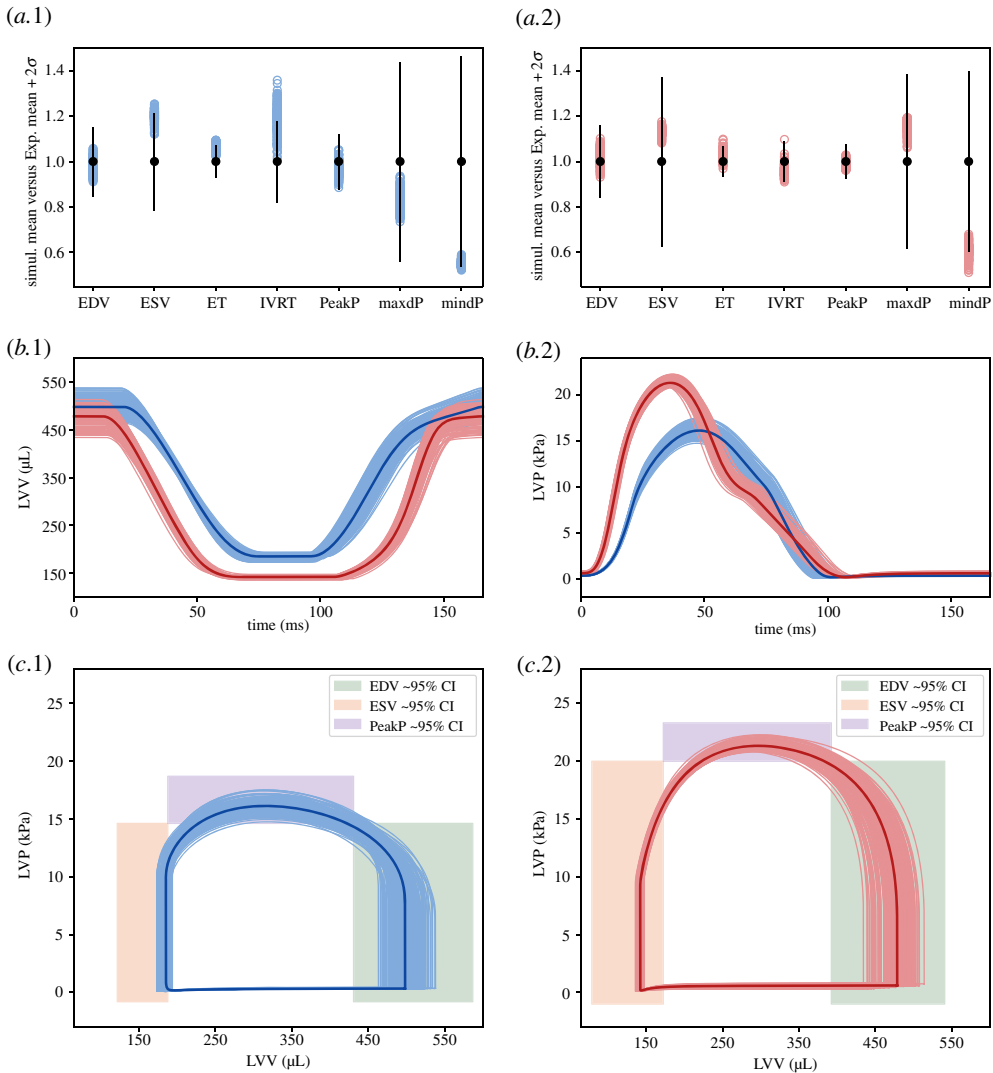
## 4. Discussion

In this study, we have shown how GPEs can be used to rapidly and robustly tune a multi-scale rat heart model to organ-scale functional measurements and characterize the global sensitivity of the model. This is the first example, to the best of our knowledge, of the use of HM to constrain multi-scale cardiac mechanics models. Previous attempts to tune cardiac mechanics model parameters have sequentially fitted parameters [25] or fitted a maximum of two to four parameters using deterministic approaches [12]. Here we fitted eight parameters concurrently. This means that error is evenly distributed across parameters and does not accumulate in the latter parameters as occurs in sequential fitting. We showed that active and passive parameters can be fitted simultaneously, avoiding potential confounding effects of slow relaxation, as occurs in HF models [16]. In addition, by providing bounds on model parameters (therefore an estimate of parameter uncertainty), the HM technique is compatible with verification, validation and uncertainty quantification (VVUQ) methods [26] supported by the ASME V&V40 standards and the FDA.<sup>1</sup>

HM produced different constraints for the SHAM and AB models, even if starting from the same hypercube in the input parameter space. Interestingly, the separation across the two rat phenotypes in the one-dimensional parameters' ranges occurs mostly for parameters describing active tension development at the cellular level. We interpret this result as a confirmation that impaired whole-organ function in the AB rat is linked to altered properties at the cell level.

We have performed the first GSA of cardiac mechanics. Previous models have performed local sensitivity analysis [27], or attempted GSA on simplified cardiac cell models [28]. Our GSA approach highlighted that LV function is mostly influenced by cellular-level properties, consistent with our HM results.

<sup>1</sup>Assessing credibility of computational modeling through verification and validation.



**Figure 5.** Simulator runs using the HM last wave's parameter points as an input. (a) Obtained LV features' (empty, coloured dots) distributions around experimental mean values (filled, black dots) for SHAM (.1) and AB (.2). 2 SD confidence intervals are shown as vertical straight lines centred in their respective mean value. All the displayed values (including confidence intervals) are normalized by the respective experimental mean values. Features' labels are explained in table 2. (b) Simulated LV volume (.1) and pressure (.2) curves which the features in (a) were extracted from, for SHAM (blue) and AB (red). The average curves are displayed in darker colour variant. (c) Resulting SHAM (.1) and AB (.2) pressure-volume loops with approximately 95% EDV, ESV, PeakP features' experimental confidence intervals. (Online version in colour.)

## 5. Limitations

We have developed a detailed biophysical rat heart model. However, this model is inherently a simplification of the underlying system, with four specific limitations.

Firstly, the boundary conditions do not account for the pericardium that may be important in constraining cardiac mechanics [29]. We have not accounted for potential spatial variations in cellular properties and  $\text{Ca}^{2+}$  transients, and the latter homogeneously activates contraction throughout ventricular walls. Extending the sensitivity analysis to include  $\text{Ca}^{2+}$  transient features would allow us to estimate if this has a major effect on our results.

Secondly, HM provides a bounded region of non-implausible parameter sets. Parameter bounds do not define parameter distributions. Including an MCMC parameter fit using the HM bounds as priors would extend this method to estimate likely parameter distributions as opposed to parameter bounds.

The heterogeneity of the data used to constrain the models is another limitation. Specifically, pressure measurements were not available for the rats used as an animal model and were therefore collected and averaged over similar experimental studies. Related to this, we only used a single representative anatomy for the healthy rats' cohort and single representative anatomy for the AB rats' cohort: ejection time and isovolumetric relaxation time features were constrained according to single measurements coming from these two anatomies' segmentations.

The fourth limitation concerns the GPE, HM and GSA implementations. In §2e, we have fitted the mean function separately and then trained the GP on the residuals, following the GPE–HM approach applied previously [20,30]. However, different emulation strategies exist (see e.g. [31], where the linear regression model parameters and the GP hyperparameters are jointly optimized; [8] uses this GP formulation for HM), and may impact the implausibility measure. To test if this choice of GP formulation impacted our final results we have performed a preliminary comparison with the Oakley and O'Hagan [31] emulation strategy (see electronic supplementary material). This suggests that HM can potentially converge to the same reduced parameter space, however, the optimal choice of GP for fitting cardiac models requires further research.

For both HM and GSA, we used independent GPEs for each output. This univariate approach provided us the flexibility to tune the hyperparameters and choose basis functions for each output independently. However, it did not account for potential correlations in outputs, which could be accounted for with a multi-output strategy, e.g. [32], although this assumes common hyperparameters across all outputs. A multi-variate implausibility measure could be constructed from multi-output GPEs, which could potentially assist in ruling out more implausible parameter space.

In the implausibility measure (electronic supplementary material, equation (4)) calculation, we omitted the model discrepancy term, for which we did not have any estimate available. To completely replace animal models with virtual representations of them, further experts knowledge will be needed to quantify the difference between the *in silico* model and the real world system that it represents.

Finally, global sensitivity indices were estimated using the mean of the available emulator, ignoring the uncertainty of the emulator itself. However, the emulator's uncertainty was small, therefore, the estimated sensitivity indices were still able to capture the parameters' influence on the model outputs (see electronic supplementary material for an example on ejection fraction feature). This can be further improved in future studies by looking for alternative estimations of the sensitivity indices which directly make use of the emulator mean and variance.

## 6. Conclusion

We have shown that GPE can effectively be used to robustly and rapidly constrain multi-scale cardiac models from organ-scale measurements. The obtained virtual bi-ventricular rat hearts are a valuable system to be used for further investigation of HF. Moreover, we demonstrated that our GPE-based GSA approach enables the computationally efficient identification of key components in three-dimensional ventricular mechanics, giving a deeper insight into the link between cell, tissue, shape and boundary conditions properties and organ-scale function.

**Data accessibility.** This article has no additional data.

**Authors' contributions.** S.A.N., C.J., A.L. and A.S. conceived the study, supervised the project, and contributed to and edited the manuscript. S.C. provided code, advised on H.M. and edited the manuscript. I.S., E.K.S.E. and W.E.L. provided and consulted on the experimental data. S.L. implemented the software, ran the simulations, analysed the results and lead the manuscript writing.

**Competing interests.** A.S. and C.J.M. are employees of Pfizer Inc.

**Funding.** This research is supported by the EPSRC (EP/P01268X/1, EP/M012492/1, NS/A000049/1, EP/R045046/1), the BHF (SP/18/6/33805, PG/19/44/34368), the National Institute for Health Research (NIHR) KCL Biomedical Research Centre, and Pfizer, Inc.

## References

- Niederer SA, Campbell KS, Campbell SG. 2019 A short history of the development of mathematical models of cardiac mechanics. *J. Mol. Cell. Cardiol.* **127**, 11–19. (doi:10.1016/j.yjmcc.2018.11.015)
- Dokos S, Lovell NH. 2004 Parameter estimation in cardiac ionic models. *Prog. Biophys. Mol. Biol.* **85**, 407–431. (doi:10.1016/j.pbiomolbio.2004.02.002)
- Groenendaal W, Ortega FA, Kherlopian AR, Zygmunt AC, Krogh-Madsen T, Christini DJ. 2015 Cell-specific cardiac electrophysiology models. *PLoS Comput. Biol.* **11**, e1004242. (doi:10.1371/journal.pcbi.1004242)
- Sarkar AX, Sobie EA. 2010 Regression analysis for constraining free parameters in electrophysiological models of cardiac cells. *PLoS Comput. Biol.* **6**, e1000914. (doi:10.1371/journal.pcbi.1000914)
- Johnstone RH, Chang ETY, Bardenet R, de Boer TP, Gavaghan DJ, Pathmanathan P, Clayton RH, Mirams GR. 2016 Uncertainty and variability in models of the cardiac action potential: can we build trustworthy models?. *J. Mol. Cell. Cardiol.* **96**, 49–62. (doi:10.1016/j.yjmcc.2015.11.018)
- Vernon I, Goldsteiny M, Bowerz RG. 2010 Galaxy formation: a Bayesian uncertainty analysis. *Bayesian Anal.* **5**, 619–670. (doi:10.1214/10-BA524)
- Andrianakis I, Vernon IR, McCreesh N, McKinley TJ, Oakley JE, Nsubuga RN, Goldstein M, White RG. 2015 Bayesian history matching of complex infectious disease models using emulation: a tutorial and a case study on HIV in Uganda. *PLoS Comput. Biol.* **11**, e1003968. (doi:10.1371/journal.pcbi.1003968)
- Coveney S, Clayton RH. 2018 Fitting two human atrial cell models to experimental data using Bayesian history matching. *Prog. Biophys. Mol. Biol.* **139**, 43–58. (doi:10.1016/j.pbiomolbio.2018.08.001)
- Camacho P, Fan H, Liu Z, He J-Q. 2016 Small mammalian animal models of heart disease. *Am. J. Cardiovasc. Dis.* **6**, 70–80.
- Røe ÅT, Aronsen JM, Skårdal K, Hamdani N, Linke WA, Danielsen HE, Sejersted OM, Sjaastad I, Louch WE. 2017 Increased passive stiffness promotes diastolic dysfunction despite improved  $\text{Ca}^{2+}$  handling during left ventricular concentric hypertrophy. *Cardiovasc. Res.* **113**, 1161–1172. (doi:10.1093/cvr/cvx087)
- Land S, Niederer SA, Aronsen JM, Espe EK, Zhang L, Louch WE, Sjaastad I, Sejersted OM, Smith NP. 2012 An analysis of deformation-dependent electromechanical coupling in the mouse heart. *J. Physiol.* **590**, 4553–4569. (doi:10.1113/jphysiol.2012.231928)
- Lewalle A, Land S, Carruth E, Frank LR, Lamata P, Omens JH, McCulloch AD, Niederer SA, Smith NP. 2018 Decreasing compensatory ability of concentric ventricular hypertrophy in aortic-banded rat hearts. *Front. Physiol.* **9**, 37. (doi:10.3389/fphys.2018.00037)
- Lamata P, Niederer S, Nordsletten D, Barber DC, Roy I, Hose DR, Smith N. 2011 An accurate, fast and robust method to generate patient-specific cubic Hermite meshes. *Med. Image Anal.* **15**, 801–813. (doi:10.1016/j.media.2011.06.010)
- Gattoni S, Røe ÅT, Aronsen JM, Sjaastad I, Louch WE, Smith NP, Niederer SA. 2017 Compensatory and decompensatory alterations in cardiomyocyte  $\text{Ca}^{2+}$  dynamics in hearts with diastolic dysfunction following aortic banding. *J. Physiol.* **595**, 3867–3889. (doi:10.1113/JP273879)
- Guccione JM, McCulloch AD, Waldman LK. 1991 Passive material properties of intact ventricular myocardium determined from a cylindrical model. *J. Biomech. Eng.* **113**, 42–55. (doi:10.1115/1.2894084)
- Xi J *et al.* 2013 The estimation of patient-specific cardiac diastolic functions from clinical measurements. *Med. Image Anal.* **17**, 133–146. (doi:10.1016/j.media.2012.08.001)
- Land S, Niederer SA, Smith NP. 2012 Efficient computational methods for strongly coupled cardiac electromechanics. *IEEE Trans. Biomed. Eng.* **59**, 1219–1228. (doi:10.1109/TBME.2011.2112359)

18. Holzapfel GA, Ogden RW. 2009 Constitutive modelling of passive myocardium: a structurally based framework for material characterization. *Phil. Trans. R. Soc. A* **367**, 3445–3475. (doi:10.1098/rsta.2009.0091)
19. Westerhof N, Elzinga G, Sipkema P. 1971 An artificial arterial system for pumping hearts. *J. Appl. Physiol.* **31**, 776–781. (doi:10.1152/jappl.1971.31.5.776)
20. Vernon I, Liu J, Goldstein M, Rowe J, Topping J, Lindsey K. 2018 Bayesian uncertainty analysis for complex systems biology models: emulation, global parameter searches and evaluation of gene functions. *BMC Syst. Biol.* **12**, 1. (doi:10.1186/s12918-017-0484-3)
21. Rasmussen CE, Williams CKI. 2005 *Gaussian processes for machine learning* (Adaptive Computation and Machine Learning). New York, NY: The MIT Press.
22. Sobol' IM. 2001 Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Math. Comput. Simulat.* **55**, 271–280. (doi:10.1016/S0378-4754(00)00270-6)
23. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. 2010 Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Commun.* **181**, 259–270. (doi:10.1016/j.cpc.2009.09.018)
24. Herman J, Usher W. 2017 SALib: an open-source python library for sensitivity analysis. *J. Open Source Softw.* **2**, 97. (doi:10.21105/joss.00097)
25. Wang VY, Lam H, Ennis DB, Cowan BR, Young AA, Nash MP. 2009 Modelling passive diastolic mechanics with quantitative MRI of cardiac structure and function. *Med. Image. Anal.* **13**, 773–784. Includes Special Section on the 12th International Conference on Medical Imaging and Computer Assisted Intervention (doi:10.1016/j.media.2009.07.006)
26. Patten RD, Hall-Porter MR. 2009 Small animal models of heart failure. *Circ. Hear. Fail.* **2**, 138–144. (doi:10.1161/CIRCHEARTFAILURE.108.839761)
27. Sher AA, Wang K, Wathen A, Maybank PJ, Mirams GR, Abramson D, Noble D, Gavaghan DJ. 2013 A local sensitivity analysis method for developing biological models with identifiable parameters: application to cardiac ionic channel modelling. *Future Gen. Comput. Syst.* **29**, 591–598. (doi:10.1016/j.future.2011.09.006)
28. Pathmanathan P, Cordeiro JM, Gray RA. 2019 Comprehensive uncertainty quantification and sensitivity analysis for cardiac action potential models. *Front. Physiol.* **10**, 721. (doi:10.3389/fphys.2019.00721)
29. Strocchi M *et al.* 2020 Simulating ventricular systolic motion in a four-chamber heart model with spatially varying robin boundary conditions to model the effect of the pericardium. *J. Biomech.* **101**, 109645. (doi:10.1016/j.jbiomech.2020.109645)
30. Salter JM, Williamson D. 2016 A comparison of statistical emulation methodologies for multi-wave calibration of environmental models. *Environmetrics* **27**, 507–523. (doi:10.1002/env.2405)
31. Oakley JE, O'Hagan A. 2004 Probabilistic sensitivity analysis of complex models: a Bayesian approach. *J. R. Stat. Soc. B (Stat. Methodol.)* **66**, 751–769. (doi:10.1111/j.1467-9868.2004.05304.x)
32. Conti S, Gosling JP, Oakley JE, O'Hagan A. 2009 Gaussian process emulation of dynamic computer codes. *Biometrika* **96**, 663–676. (doi:10.1093/biomet/asp028)