INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Lomeli LM *et al*. 2021 Optimal experimental design for mathematical models of haematopoiesis. *J. R. Soc. Interface* **18**: 20200729.
https://doi.org/10.1098/rsif.2020.0729

Received: 7 September 2020
Accepted: 4 January 2021

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
systems biology, computational biology

Keywords:
Bayesian analysis, stem cells, differential equations, feedback and feedforward regulation, haematopoiesis

Authors for correspondence:
Babak Shahbaba
e-mail: babaks@uci.edu
John S. Lowengrub
e-mail: jlowengr@uci.edu
Vladimir N. Minin
e-mail: vminin@uci.edu

†These authors contributed equally to this work.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5280484.

THE ROYAL SOCIETY PUBLISHING

Optimal experimental design for mathematical models of haematopoiesis

Luis Martinez Lomeli[1,†], Abdon Iniguez[1,†], Prasanthi Tata[2], Nilamani Jena[2], Zhong-Ying Liu[2], Richard Van Etten[1,2,3,4,5], Arthur D. Lander[1,4,5,6,7], Babak Shahbaba[1,4,8], John S. Lowengrub[1,4,5,7,9] and Vladimir N. Minin[1,4,8]

[1]Center for Complex Biological Systems, [2]Division of Hematology/Oncology, [3]Department of Biological Chemistry, [4]Center for Cancer Systems Biology, [5]Chao Family Comprehensive Cancer Center, [6]Department of Developmental and Cell Biology, [7]Department of Biomedical Engineering, [8]Department of Statistics, and [9]Department of Mathematics, University of California Irvine, Irvine, CA, USA

PT, 0000-0002-1932-9987; VNM, 0000-0002-1917-9288

The haematopoietic system has a highly regulated and complex structure in which cells are organized to successfully create and maintain new blood cells. It is known that feedback regulation is crucial to tightly control this system, but the specific mechanisms by which control is exerted are not completely understood. In this work, we aim to uncover the underlying mechanisms in haematopoiesis by conducting perturbation experiments, where animal subjects are exposed to an external agent in order to observe the system response and evolution. We have developed a novel Bayesian hierarchical framework for optimal design of perturbation experiments and proper analysis of the data collected. We use a deterministic model that accounts for feedback and feedforward regulation on cell division rates and self-renewal probabilities. A significant obstacle is that the experimental data are not longitudinal, rather each data point corresponds to a different animal. We overcome this difficulty by modelling the unobserved cellular levels as latent variables. We then use principles of Bayesian experimental design to optimally distribute time points at which the haematopoietic cells are quantified. We evaluate our approach using synthetic and real experimental data and show that an optimal design can lead to better estimates of model parameters.

1. Introduction

The haematopoietic system produces billions of mature myeloid and lymphoid blood cells from self-renewing haematopoietic stem cells (HSCs) and multipotent progenitors (MPPs) on a daily basis and facilitates massive cell increases in response to pathological stresses [1]. This system must have in place a tightly regulated feedback control mechanism at multiple levels to ensure an appropriate proportion of HSCs, MPPs, and mature cells. However, we do not yet have a good understanding of the nature of the feedback regulation and how it plays a role in cell maintenance. There has been a long-standing effort to use mathematical models in order to understand haematopoiesis under normal and diseased conditions, e.g. see [2–6]. These include ordinary differential equation (ODE) models that describe the dynamics of simplified systems (e.g. [7–12]), models that account for more realistic numbers of different cell types and branching processes [13], as well as models that account for stochasticity [14–19] and spatial dynamics in the bone marrow [20]. In many cases, models were fitted using equilibrium cell counts, or limited dynamic data, which yield point estimates for the parameters. In a few cases, uncertainties in parameter inference were considered using Bayesian methods [21–23].

With improved ability to measure cell counts, haematopoiesis researchers need guidance when designing experiments to produce such measurements. This guidance can be obtained by using established statistical tools of optimal experimental design that aim at maximizing information gain about model

parameters of interest. Within the field of systems biology, Bayesian approaches to experimental design are especially appealing since sparse data often require using prior information about some subset of model parameters [24,25]. Examples of such approaches include studies of biochemical networks, gene regulation networks and signaling pathways [26–30].

In this work, we follow the best practices of Bayesian experimental design within a new statistical framework for modelling the haematopoiesis system with feedback control and regulation. Our approach provides a rigorous hierarchical Bayesian methodology for fitting mechanistic mathematical models to empirical data. Because of their ability to account for prior information and multiple sources of uncertainty, Bayesian methods have been used increasingly for parameter inference, model selection and experimental design for complex systems in the physical and biological sciences [24,31–35]. To model the haematopoietic dynamics, we use a nonlinear ODE model that incorporates self-renewal, cell division, feedback and feedforward regulation. In particular, we track only stem and MPP cells. Nevertheless, the model is flexible enough to describe the response to external perturbations and the subsequent return to steady state. We apply our model to empirical data obtained from perturbation experiments in mice subjected to low dose radiation. We measure the numbers of HSCs and MPPs at different times in the bone marrow to investigate the recovery dynamics and infer model parameters and feedback mechanisms.

Note that the mice can not be tracked longitudinally by taking repeated measurements of cell numbers; rather, each mouse provides a single observation point because the mouse is sacrificed to extract the bone marrow. This creates a statistical challenge for quantifying the system response since each data point belongs to a different subject, for whom we do not have the baseline measurements prior to exposure to the agent; that is, the number of cells is not known initially or at any other time prior to the measurement time and thus the corresponding cell numbers are latent. To address this issue, we use a hierarchical Bayesian model where the initial cellular counts for all the subjects are treated as latent variables to be inferred and the ODE model is used to interpolate the cell numbers until the observation times. The main advantage of this approach is that it allows us to appropriately integrate and align data from multiple subjects in a coherent, statistically rigorous manner. However, the parameter inference can be sensitive to the choice of experimental design.

Determining time points to accurately and efficiently sample biological processes is a fundamental and challenging problem. Recent progress in this area has focused on machine-learning approaches, e.g. [36–38]. These general methods have been applied to analyse time series of genomic, transcriptomic and molecular data using interpolation, functional data analysis and optimization to determine sampling rates. Here, we take a complementary approach where we use a mechanistic model to fill in missing data points. This should not only enhance our understanding of the biological system, the model should also enable us to accurately extrapolate results beyond the dataset used for training.

Focusing on perturbation experiments of the haematopoietic system, we determine the timing of the measurements and the number of subjects per observation time to maximize the information gain for the parameters of our ODE model. Using the Bayesian utility theory approach, we quantify information gain about the model parameters over the space of all possible experimental designs using the Kullback–Leibler (KL) divergence utility to quantify the difference between the prior and posterior distributions of the parameters [39]. Similar approaches have been used in other areas [40–42]. This way, we are able identify the design that provides the highest expected utility, e.g. maximum information gain [27,43–51], which we call the optimal design.

To evaluate our proposed approach, we first apply it to synthetic data and show that we can identify the ODE model parameters. Next, we analyse real data from a bone marrow perturbation experiment. Investigating a finite set of experimental designs, we find that the designs with a higher number of observation times and possibly fewer subject replicates can provide better parameter estimates compared to the designs with fewer observation times even if we use a higher number of subject replicates. Also, we show that how we allocate the subjects over time matters. For example, designs with more observations at later times can provide better estimates on feedback gains, whereas designs with more observations at earlier times improve identification of cell division rates.

2. Material and methods

2.1. Experimental set-up

We consider an experimental set up where the haematopoietic system is perturbed and the results of the perturbation are observed by measuring the numbers of the cell types of interest. We primarily consider HSCs and MPPs, but other known cell types like lymphoid and myeloid progenitors (CLPs, CMPs) or mature lymphoid and myeloid cells can also be quantified experimentally. Throughout this paper, we will use simulated and real data based on the following experiment. We start with M genetically identical mice that are kept under the same laboratory conditions. Each mouse is exposed to an external perturbation, e.g. a light dose (50 cGy) of radiation, with the purpose of decreasing the number of HSCs (e.g.[52]). These mice are sacrificed at different times after irradiation and the counts of HSCs and MPPs are obtained from the bone marrow of each individual mice using flow cytometry. More information about the experimental materials and methods can be found in the electronic supplementary material.

2.2. Data generation process

We postulate that at time t_0, HSC and MPP counts come from some distribution that encapsulates normal biological variation among mice. For mathematical convenience, we assume that this distribution is lognormal, or equivalently, that log-transformed unobserved true HSC and MMP counts come from two independent normal distributions with means $\log(\mu_{\mathrm{HSC}})$ and $\log(\mu_{\mathrm{MPP}})$, respectively, and with the same variance σ_b. Denoting these latent log-transformed counts with a bivariate vector $\mathbf{u}_i$ and their means with a vector $\log(\boldsymbol{\mu})$ for each mouse $i = 1, \ldots, M$, we can write our initial condition assumption as

$$\mathbf{u}_i \sim N(\log(\boldsymbol{\mu}), \sigma_b^2 \cdot \mathbf{I}), \; i = 1, \ldots, M. \tag{2.1}$$

We are now ready to specify the distribution of observed HSC and MPP counts. First, we order mice in such a way that mice indexed by $1, \ldots, k$ correspond to animals, whose bone marrow is sampled immediately post perturbation. We assume that conditionally on the true log-counts of HSCs and MPPs, the observed log-transformed counts $\mathbf{y}_i = \log(\mathbf{y}_i^*)$, where $\mathbf{y}_i^*$ represents the raw counts, are normally distributed, with the mean being equal to the true counts and variance σ_i^2 that represents technical variation that arises due to the measurement error/noise.

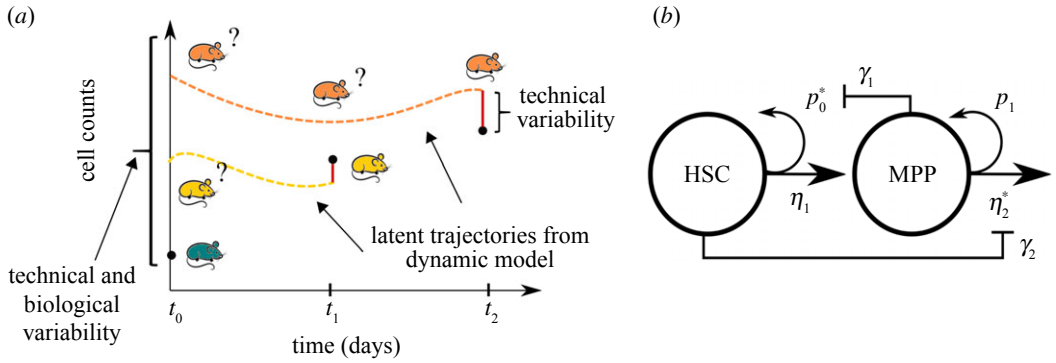

Figure 1. Illustration of the proposed latent variables approach and mechanistic model. (*a*) Description of the proposed latent variables approach. Each mouse's cell counts are observed only once at the time of the mouse sacrifice and bone marrow extraction. Cell counts before perturbation (e.g. low dose radiation) are allowed to be different among mice due to normal biological variation. We model this by assuming each data point at time t_0 to be subject to technical and biological variability. At times greater than t_0, we assume each data point has a latent trajectory subject to technical variability (shown in dashed). These latent trajectories are modelled using our mechanistic ODE model subject to these initial conditions. (*b*) The ODE lineage model consisting of HSC and MPP compartments. HSCs and MPPs have the ability to self-renew with probabilities p_0^* and p_1 and divide at rates η_1 and η_2^*, where the stars indicate that the corresponding parameters are subject to feedback regulation. The HSCs self-renewal probabilities are negatively regulated by the MPPs and the MPPs division rates are negatively regulated by the HSCs.

For the mice that are sacrificed right after the perturbation experiment, this assumption translates into the following conditional distribution:

$$\mathbf{y}_i(t_0) \mid \mathbf{u}_i \sim N(\mathbf{u}_i, \sigma_t^2 \cdot \mathbf{I}), \quad i = 1, \ldots, k. \tag{2.2}$$

The unconditional distribution for the HSC and MPP counts at the initial time t_0 is derived from equations (2.1) and (2.2) as

$$\mathbf{y}_i(t_0) \sim N(\log(\boldsymbol{\mu}), (\sigma_t^2 + \sigma_b^2) \cdot \mathbf{I}), \quad i = 1, \ldots, k. \tag{2.3}$$

To model cell counts measured at time points after the initial time t_0, we assume that the cellular population levels in each mouse start from latent initial conditions and follow deterministic latent trajectories according to a mechanistic process model. The latent population trajectories evolve for all times $0 < t \leq t_j$ until the moment t_j when the mouse is harvested and the cell counts are measured (with noise) (figure 1*a*). The conditional distribution of the observed HSC and MPP cell counts after the initial time given latent initial conditions is

$$\mathbf{y}_j(t_j) \mid \mathbf{u}_j \sim N(\log(\mathbf{x}(\mathbf{u}_j, \boldsymbol{\theta}, t_j)), \sigma_t^2 \cdot \mathbf{I}), \quad j = k+1, \ldots, M, \tag{2.4}$$

where $\mathbf{x}(\mathbf{u}_j, \boldsymbol{\theta}, t_j)$ is a bivariate vector of HSC and MPP counts for mouse j that started with latent counts $\mathbf{u}_j$ and evolved according to some process with parameters $\boldsymbol{\theta}$ up to time t_j. Note that equation (2.4) does not include the biological noise term explicitly since it is already included in the model for initial cell counts. This approach implies that the HSC and MPP trajectories are latent and can be observed cross-sectionally only once, in contrast to typical longitudinal studies, where repeated measures are taken from a cohort of animals/subjects followed over time.

2.3. Mechanistic model of the mean process

The dynamical model for latent trajectories is based on classic cell lineage models for describing the growth of hierarchically organized tissues [8,53,54] where cells are arranged in a lineage starting with HSCs that are followed by more differentiated cells downstream. Although there are many cell types in the branched lineage that describes the haematopoietic system (e.g. [55,56]), we focus here only on the least differentiated types: the HSCs and MPPs that have the simple hierarchical relationship shown in figure 1*b*. This is because the experiments suggest that the cell compartments downstream are largely unaffected by irradiation (data not shown) and thus we assume that the downstream cells do not significantly influence the HSC and MPP dynamics. Furthermore, for simplicity, we do not distinguish between the different types of HSC and MPP cells, which eliminates the need for considering branching. Of course, the model can easily be extended to include more cell types and branching (e.g. [13,20]) although this would require more data for proper parameter estimation.

In the mathematical model (figure 1*b*), we assume that the HSCs and MPPs have the ability to divide at the rates η_1 and η_2^*, respectively, and to undergo self-renewal with probabilities p_0^* and p_1, where the stars denote that the corresponding parameters are subject to feedback regulation. Let x_{HSC} and x_{MPP} be the numbers of HSCs and MPPs, respectively, then their dynamics can be modelled using a system of ODE

$$\left.\begin{aligned} x'_{\mathrm{HSC}} &= (2p_0^* - 1)\eta_1 x_{\mathrm{HSC}} \\ \text{and} \quad x'_{\mathrm{MPP}} &= 2(1 - p_0^*)\eta_1 x_{\mathrm{HSC}} + (2p_1 - 1)\eta_2^* x_{\mathrm{MPP}}, \end{aligned}\right\} \tag{2.5}$$

where $' = \mathrm{d}/\mathrm{d}t$. The factor 2 in the first equation and in the first term in the second equation arises from the fact that p_0^* of HSCs remain HSCs after division, while $(1 - p_0^*)$ of HSCs differentiate and produce two MPP cells. Similarly, in the second equation, the fraction $(1 - p_1)$ of MPPs also produce two more specialized cells but they are not included in the model here since the more differentiated cells are largely unaffected by radiation. This model is a simplified version of a general model of lineage progression under feedback regulation introduced in [53,57].

The self-renewal probabilities and division rates should be subject to feedback regulation. Single cell RNA sequencing data (scRNA-seq) can be used to identify putative feedback loops and sender and receiver cells. We re-analysed data from a scRNA-seq study of normal haematopoiesis that identified many interesting cell clusters whose transcriptomes suggested pairwise combinations of cells expressing feedback ligands and their receptors [58]. Although the study did not cleanly separate out different kinds of early stem/progenitor cells, the early stem/progenitors did cluster into two groups, perhaps representing HSCs and MPPs. In these two groups, we were able to recognize several ligands and receptors (such as ANGPT1 and CCL3 and their receptors), although we could not be certain about the sender and receiver cell types.

Following [59], we hypothesized that ANGPT1 is secreted by HSCs and negatively regulates MPP division rates and that CCL3 is produced by MPPs and negatively regulates HSC self-

renewal [60]. These hypotheses will be tested in future work. The feedback on HSC self-renewal can be modelled using a simple Hill function

$$p_0^* = \frac{p_0}{1 + \gamma_1 x_{\mathrm{MPP}}}, \tag{2.6}$$

where p_0 is the unregulated self-renewal probability and γ_1 is the feedback gain. Note that we have implicitly assumed that the concentration of the negatively regulating biomolecule (e.g. CCL3) is proportional to the cell population. This approximation assumes that spatial variation of the biomolecule can be neglected and is similar to the those used in [8,13,53]. Although the HSC division rate η_1 could be subject to similar feedback as p_0^* [53], we assume for simplicity that this rate is constant.

The negative feedforward loop on MPP division rates can be modelled as

$$\eta_2^* = \frac{\overline{\eta}_2}{1 + \gamma_2 x_{\mathrm{HSC}}}, \tag{2.7}$$

where $\overline{\eta}_2$ is the unregulated MPP division rate and γ_2 is the feedforward gain. In principle, the MPP self-renewal probability p_1 should also be subject to feedback regulation. Assuming this regulation arises from more differentiated cell types (e.g. [8,13,53]), we can assume that p_1 is constant because the number of differentiated cells is roughly constant, as mentioned above. Further, $p_1 < 0.5$ since the MPP should not be able to fully self-renew. Therefore, we may rewrite the system in equation (2.5) as

$$\left.\begin{aligned} x'_{\mathrm{HSC}} &= (2p_0^* - 1)\eta_1 x_{\mathrm{HSC}} \\ \text{and} \quad x'_{\mathrm{MPP}} &= 2(1 - p_0^*)\eta_1 x_{\mathrm{HSC}} - \eta_2^* x_{\mathrm{MPP}}, \end{aligned}\right\} \tag{2.8}$$

where

$$\eta_2^* = \frac{\eta_2}{(1 + \gamma_2 x_{\mathrm{HSC}})}, \tag{2.9}$$

and $\eta_2 = (1 - 2p_1)\overline{\eta}_2$. Note that p_1 and $\overline{\eta}_2$ can not be identified simultaneously, therefore we focus instead on the combined rate η_2.

The nonlinear feedback and feedforward loops in the above model enable the tight control of growth, the establishment of equilibria that are robust to large changes in parameter values and rapid regeneration of equilibia after perturbations by external stimuli (e.g. [8,53]), although the regeneration dynamics can exhibit oscillatory behaviour. Finally, all the ODE model parameters can be grouped in the vector $\boldsymbol{\theta} = (p_0, \eta_1, \eta_2, \gamma_1, \gamma_2)$.

2.4. Hierarchical Bayesian framework

The goal of our statistical framework is to link the dynamic mathematical model with empirical data. We use a latent variables approach where the ODE model allows us to interpolate all the latent trajectories for the measurements that are missing before a mouse is harvested. According to the assumed data generation process described above, equations (2.3) and (2.4) define the likelihood function as the product of normal densities at time zero and at later times

$$\begin{aligned} p(\mathbf{y} \mid \Theta) = &\prod_{i=1}^{k} N(\mathbf{y}_i \mid \log \boldsymbol{\mu}, \sigma_t^2 + \sigma_b^2) \\ &\times \prod_{j=k+1}^{M} N(\mathbf{y}_j \mid \log \mathbf{x}(\mathbf{u}_j, \boldsymbol{\theta}, t_j), \sigma_t^2), \end{aligned} \tag{2.10}$$

where the vector $\boldsymbol{\Theta} = (\boldsymbol{\theta}, \mathbf{u}_{k+1:M}, \sigma_b^2, \sigma_t^2, \boldsymbol{\mu})^t$ includes all the parameters and latent variables of the hierarchical model, and $\mathbf{u}_{k+1:M}$ is a vector of latent initial cell counts of mice that are sacrificed after t_0. Using a Bayesian approach, we provide measures of uncertainty to all model parameters by calculating the posterior distribution of all the parameters

$$p(\Theta \mid \mathbf{y}) \propto p(\mathbf{y} \mid \Theta) p(\Theta). \tag{2.11}$$

We assume *a priori* independence among the model parameters, resulting in the following prior distribution decomposition:

$$p(\Theta) = p(\boldsymbol{\theta}) \cdot p(\mathbf{u}) \cdot p(\boldsymbol{\mu}) \cdot p(\sigma_b^2) \cdot p(\sigma_t^2). \tag{2.12}$$

Each of the parameters $\theta_i \in \boldsymbol{\theta}$ has a lognormal prior: $\log \theta_i \sim N(m_{\theta_i}, \sigma_{\theta_i}^2)$, with the exception of p_0 that requires a logit transformation: $\mathrm{logit}((p_0 - (1/2))2) \sim N(m_{p_0}, \sigma_{p_0}^2)$, since it is constrained between 0.5 and 1. Here, $p(\mathbf{u})$ represents the mice initial conditions according to equation (2.1), and $p(\boldsymbol{\mu})$ represents the prior on the mean of the initial conditions. Finally, $p(\sigma_b^2)$ and $p(\sigma_t^2)$ represent the prior distributions on the biological and technical noises correspondingly. We describe how we approximate the posterior distribution in equation (2.11) via Markov chain Monte Carlo (MCMC) in the Computational Implementation section below.

2.5. Experimental design

In the laboratory, there are multiple variables that can be modified when an experiment is performed. Each combination of these variables defines an experimental design, which belongs to the set of all possible designs, D. The optimal experimental design goal is to discriminate between all the possible designs by using a suitable utility metric $U(\mathbf{y}, d)$ that quantifies the amount of information gain about the model parameters for a dataset $\mathbf{y}$ collected under the design $d \in D$. The optimal design is determined by comparing the expected utility, where the expectation is taken over all datasets, $\mathbf{y}$, and model parameters, $\boldsymbol{\Theta}$, for the design d:

$$u(d) = \mathbb{E}_{\Theta,\mathbf{y}}[U(\mathbf{y}, d)] = \int_{\mathbf{y}} \int_{\Theta} U(\mathbf{y}, d) p(\mathbf{y}|\Theta, d) p(\Theta) d\Theta d\mathbf{y}. \tag{2.13}$$

For the utility metric $U(\mathbf{y}, d)$, we use the KL divergence function that quantifies the information gain as the difference between the prior and posterior distributions of the model parameters:

$$U(\mathbf{y}, d) = \int_{\Theta} \log\left(\frac{p(\Theta|\mathbf{y}, d)}{p(\Theta)}\right) p(\Theta|\mathbf{y}, d) d\Theta. \tag{2.14}$$

This utility has its origin in information theory and it is equivalent to the mutual information gain [26,29]. We take a similar approach to calculate the marginal expected utility for each parameter $\theta_j \in \Theta$ individually

$$U_j(\mathbf{y}, d) = \int_{\theta_j} \log\left(\frac{p_j(\theta_j|\mathbf{y}, d)}{p(\theta_j)}\right) p(\theta_j|\mathbf{y}, d) d\theta_j. \tag{2.15}$$

Alternative utility functions have also been proposed in recent years including the inverse of the determinant of the posterior covariance [47], the quadratic loss for obtaining point estimates of the parameters [25], and the total separation utility for model discrimination [61]. Some of these utilities avoid the calculation of the evidence or marginal likelihood, $p(y \mid d)$, which is computationally challenging and it is required in equations (2.14), (2.15). However, we are able to use a thermodynamics integration method for the estimation of the marginal likelihood from posterior samples, known as *bridge sampling* [62], allowing us to overcome this computational obstacle. Therefore, in this work, we focus in the KL divergence due to its rigorous justification, computational feasibility and intuitive interpretation.

2.6. Computational implementation

2.6.1. Estimating $p(\Theta|\mathbf{y})$

We approximate the posterior distribution in equation (2.11) using the No U-Turn sampler (NUTS) algorithm implemented

in a statistical computing software platform called Stan [63–65]. NUTS is a generalization of Hamiltonian Monte Carlo (HMC), which explores the parameter space by solving a Hamiltonian system involving the parameters of interest (defining the potential energy of the system) and a set of auxiliary (fictitious) momentum variables (defining the kinetic energy of the system). At each iteration of the MCMC simulation, the solution of this system generates a new proposal sample from the posterior distribution, which is then subject to a Metropolis accept/reject decision to ensure the sampler converges to the right target distribution. The speed of convergence depends on a set of tuning parameters including the step size, ϵ, and the number of steps, L, to obtain the numerical solution of Hamiltonian dynamics. NUTS automatically finds a combination of ϵ and L for faster convergence of HMC. It is especially useful in the case when parameters dependencies creates a complex geometry that makes the posterior sampling difficult. For our hierarchical model, we show in §3 that we can successfully obtain samples from the posterior distribution (2.11).

2.6.2. Estimating $U(\mathbf{y}, d)$

Each design $d \in D$ specifies the timing of the measurements and the number of mice replicates in a haematopoiesis experiment. When the design is fixed, we can simulate data by first sampling one set of true parameters from their corresponding prior distributions, and then sampling a synthetic dataset, $\mathbf{y}$, that contains the cellular population records for HSCs and MPPs according to the assumed data generating model. At the initial time t_0, the initial conditions for the latent trajectories are determined by sampling from two independent normal distributions with means log (μ_{HSC}) and log (μ_{MPP}) and equal variances $\sigma_b^2 + \sigma_t^2$. From the initial conditions, the ODE model (2.8) is forward-solved until the time when a trajectory is observed (mouse is sacrificed and the cell counts are observed). Then a random bivariate vector is generated from the normal distribution with the mean equal to the ODE solution and variance–covariance matrix $\sigma_t^2\mathbf{I}$. This process is repeated for all the latent trajectories to obtain the dataset $\mathbf{y}$.

Given a generated dataset, we obtain MCMC samples from the posterior distribution. The posterior samples are used to compute the value of the KL utility function in equation (2.13) by Monte Carlo integration using all the MCMC iterations $i = 1, \ldots, n$:

$$\begin{aligned} U(\mathbf{y}, d) \approx \widehat{U}(\mathbf{y}, d) &= \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{p(\Theta^{(i)}|\mathbf{y}, d)}{p(\Theta^{(i)})}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{p(\mathbf{y}|\Theta^{(i)}, d)p(\Theta^{(i)})}{p(\mathbf{y}|d)}\right) - \log p(\Theta^{(i)}) \\ &= \frac{1}{n}\sum_{i=1}^{n} \log p(\mathbf{y}|\Theta^{(i)}, d) - \log p(\mathbf{y}|d), \end{aligned} \tag{2.16}$$

where $\Theta^{(i)}$ is the ith sample from the posterior distribution which is evaluated at the prior and posterior densities. Note that the KL ratio in equation (2.16) requires an estimate of the marginal likelihood $p(\mathbf{y} \mid d)$, which we obtain by Bridge Sampling. The marginal KL utility value for each parameter θ_j is also approximated via Monte Carlo

$$U_j(\mathbf{y}, d) \approx \widehat{U}_j(\mathbf{y}, d) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{\widehat{p}(\theta_j^{(i)}|\mathbf{y}, d)}{p(\theta_j^{(i)})}\right), \tag{2.17}$$

where $\widehat{p}(\theta_j|\mathbf{y}, d)$ is an approximation for the true marginal posterior density and it is estimated by a Gaussian kernel density estimator using the posterior samples [66,67]. This step is required since the analytical marginal posterior density is generally not available. Note that the kernel density estimation can only be applied to the marginal utility calculation since this approach is known to provide poor approximations for multiple-dimensional problems as in the case of the joint utility calculation.

2.6.3. Estimating $u(d)$

The overall mean utility in equation (2.13) is approximated by averaging dataset-specific utilities over N simulated data

$$u(d) \approx \widehat{u}(d) = \frac{1}{N}\sum_{k=1}^{N} \widehat{U}(\mathbf{y}_k, d), \tag{2.18}$$

where the utility $\widehat{U}(\mathbf{y}_k, d)$ is approximated using equation (2.16). For individual parameters, the mean utility value is computed similarly

$$u_j(d) \approx \widehat{u}_j(d) = \frac{1}{N}\sum_{k=1}^{N} \widehat{U}_j(\mathbf{y}_k, d), \tag{2.19}$$

where $\widehat{U}_j(\mathbf{y}_k, d)$ is estimated using equation (2.17).

2.6.4. Finding the optimal design

We consider a finite grid of designs $d_1, d_2, \ldots, d_M$ that are relevant for the haematopoiesis experiment. The optimal experimental design is determined by computing the expected utilities (2.18) and (2.19) for all the designs and finding a design with the maximum utility. We illustrate the optimal design process in figure 2.

2.6.5. Computational implementation

The software developed for this project and an illustrative demonstration notebook can be found on the repository https://github.com/luisdm1/BayesOptimalDesign.

3. Results

3.1. Successful parameter identification

Because our proposed mechanistic model is complex, we first need to ensure that the parameters in the model are identifiable. This is illustrated in figure 3 using synthetic data. Here, we have generated a dataset, where the HSCs and MPPs are observed at 7 consecutive days with seven replicates for each day (49 mice in total). We have estimated the five ODE parameters, (p_0, η_1,η_2,γ_1, γ_2), the two initial conditions, (μ_{HSC} and μ_{MPP}), the two error terms, (σ_t and σ_b), and the 84 latent trajectories (93 parameters in total). For this, we have used Stan [64,65] to obtain the posterior distribution of the model parameters. As we can see in figure 3, there are significant changes from prior to posterior distributions. Furthermore, we are able to recover the true parameter values within the 95% posterior probability intervals.

To show that parameter identification is achieved under this design (7 days × 7 replicates), we have generated 60 synthetic datasets and obtained the posterior distribution for the model parameters. Using these results, we examine the coverage of the model parameters by calculating the percentage of the times the true parameter values are included in the 95% credible intervals (electronic supplementary material, equation (A-4)). Additionally, we use the width of these intervals as a metric for the model precision (electronic supplementary material, equation (A-6)). We determine the relative bias for the MCMC simulations by using the posterior medians as point estimates and calculate the normalized residuals using the true parameter values (electronic supplementary material, equation (A-3)). The estimated mean relative bias, mean relative width, and coverage of the 95%

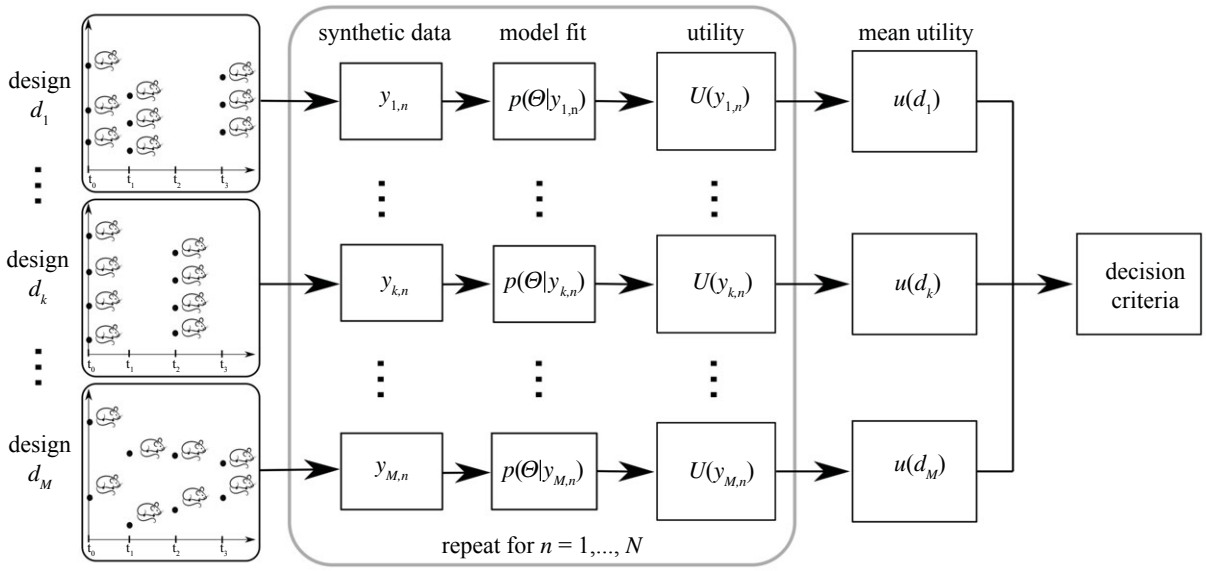

Figure 2. Flow chart for computational implementation. Given a design d_k, we sample a set of true parameter values from the priors to obtain a synthetic dataset generated from our mechanistic model using the latent variables approach. Using the synthetic data, we obtain the posterior distribution and we calculate the utility value associated with the nth synthetic dataset. Once we have repeated this process N times, we calculate the mean utility for design d_k. If we do this for M designs, we are able to use this process to choose the design with the optimal mean utility based on a decision criterion that typically depends on several constraints. For example, these constraints can vary from minimizing the number of days or having a limit on the number of mice used.

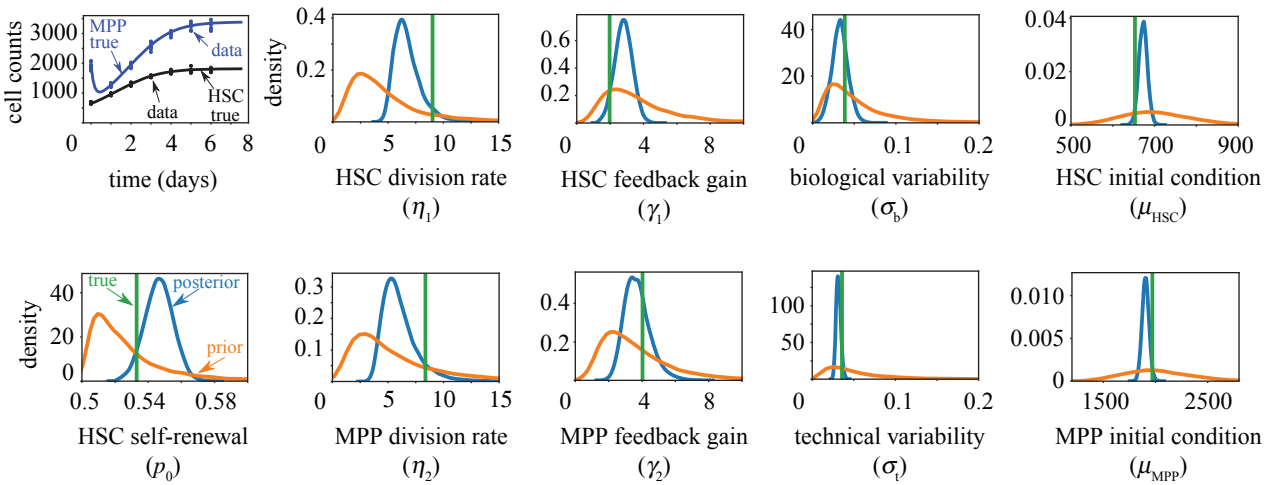

Figure 3. Parameter identification for a synthetic dataset. Sampling from our prior distributions (see text), we generate a synthetic dataset from our ODE solution (*top left*). Solid curves correspond to ODE solutions for MPPs (blue) and HSCs (black). Symbols (data) correspond to HSC and MPP cell numbers from the ODEs but with the addition of technical and biological noise following our latent variables framework (e.g. equations (2.3) and (2.4)). The data correspond to observations on 7 consecutive days with seven mice replicates per day. Using this synthetic data, in the remaining graphs we show the prior (orange) and posterior (blue) distributions for the ODE parameters and the latent HSC and MPP initial cell numbers together with the exact parameters and initial conditions used in the ODE model (green). The ODE parameters, prior distributions and 95% credible intervals are shown in electronic supplementary material table A-2.

credible intervals for all the parameters in the 7 days × 7 replicates design are shown in table 1.

As we can see, most model parameters have good coverage probabilities. Also, except for γ_2 and σ_b, the mean relative bias tends to be small.

Tables A-3 and A-4 in the electronic supplementary material provide more simulation results based on other designs illustrated in figure 2. In general, our results show that the true model parameters are included within the 95% credible intervals. We also observe that designs with more data points allow for narrower posteriors, as quantified by the mean relative width, but higher bias is observed in some parameters.

Next, we focus on quantifying information gain using the Bayesian utility theory described in the previous section. In particular, we are interested in finding the optimal experimental set-up that provides the highest information gain for inferring the model parameters.

3.1.1. Low dose radiation targets HSCs for cell death

As a proof of concept, we use our hierarchical model to fit and obtain posterior distributions from preliminary data from a bone marrow perturbation experiment targeting stem cells.

In this experiment, low dose radiation (50 cGy) was applied to a number of mice following [68] where it was claimed that

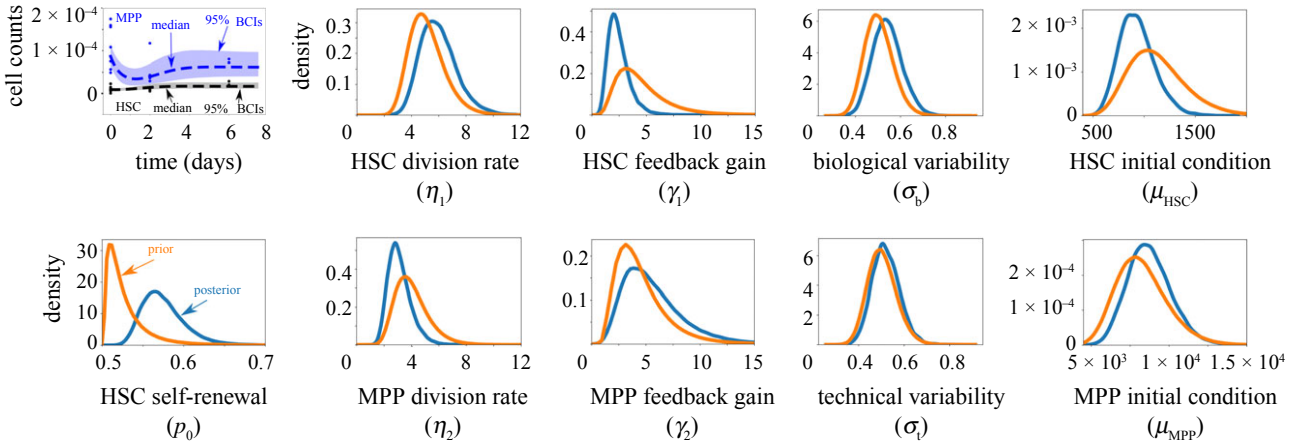

Figure 4. Parameter estimation for a real dataset. *Top left.* Preliminary data (symbols) obtained from mice that were irradiated with 50 cGy and fits of our hierarchical model to this dataset where we show the median (dashed) of the posterior distribution for the ODE solutions and the 95% Bayesian credible intervals (bands). *Remaining graphs.* Plots of the prior and posterior distributions for the fitted ODE parameters and the mean of the latent HSC and MPP initial conditions. Additional details on the priors and posterior distributions can be found on electronic supplementary material, table A-1.

Table 1. Metrics for 60 simulations using seven mice replicates during 7 days. The relative bias, relative width and coverage are calculated according to equations (A-3), (A-6) and (A-4) shown in the electronic supplementary material.

7 × 7 design	p_0	η_1	η_2	γ_1	γ_2	σ_t	σ_b	μ_1	μ_2
mean relative bias	0.00	0.10	0.05	−0.08	−0.47	0.07	−0.31	−0.03	0.05
mean relative width	0.04	0.60	0.72	1.34	2.52	0.36	1.98	0.10	0.11
coverage	0.97	0.94	0.94	1.00	0.89	1.00	1.00	0.94	0.86

low dose radiation decreases HSC numbers. This is consistent with what we observed as well, see figure A-2 in the electronic supplementary material. To observe the system dynamics in response to this perturbation, the HSC and MPP cell numbers were obtained shortly after irradiation and at two other time points (days 0, 2 and 6). At each time point, some of the mice were sacrificed, their bone marrows were extracted and sorted, and the cell numbers were determined by flow cytometry (see Experimental Methods in the electronic supplementary material). While there were 13 mice in total, seven mice were sacrificed at day 0, four were sacrificed on day 2 and two were sacrificed on day 6. Note the large variability in the data, particularly in the MPP numbers on figure A-1, electronic supplementary material. The data suggest that the HSCs and MPPs return to equilibrium within one week.

Figure 4 (top left) shows the fits of the mechanistic model to the data for HSCs (black) and MPPs (blue). In particular, the medians (dashed) of the posterior distribution of the ODE solutions are shown together with the 95% Bayesian credible regions [69] (shaded). The ODE fit seems reasonable based on the small amount of experimental data. However, only the HSCs initial conditions comparing experiment versus control data (electronic supplementary material, figure A-1) showed a significant shift between the prior and posterior distributions (electronic supplementary material, figure A-2). All the other parameters did not show a substantial shift in the posterior distributions compared to the priors (figure 4; electronic supplementary material, table A-1, figure A-3), suggesting there is little information gained from this data. Moreover, the Bayes Factor (e.g. [69]) between a model with feedback (equation (2.8)) compared to the same model without feedback regulation (equation (2.8) with γ_1 and γ_2 identically zero) is equal to $BF = 1.18$, which indicates that there is no strong evidence in favour of the model with feedback according to this limited dataset.

3.1.2. Utility grid search

Since we know that the model parameters are identifiable using a sufficient amount of synthetic data, the results in the previous section highlight the importance of finding a proper experimental design to maximize information gain in the radiation experiment. To this end, we explore different experimental set-ups by varying the number of mice collected per day and the timing of the measurements. That is, each design represents some number of mice observed at some sampling frequency. We consider a finite number of experimental designs (70) to include a varying amount of mice over different observation days. Our design space is defined as the following. The first observation day starts at day 0 right after radiation, and more times are added until day 6 since it was shown in our preliminary perturbation experiment that the system returns to equilibrium in less than a week. We also assume that the number of mice observed per day could be 3, 4, 5, 6 or 7.

Exploring this finite experimental design space requires calculating the expected utilities for which we use 60 different synthetic datasets per design (2.18), (2.19). As shown in figure 5, the expected utilities show an increasing trend when the number of replicates and the frequency of sampling increase. The values in this figure correspond to a fold change with respect to the baseline design, three mice observed

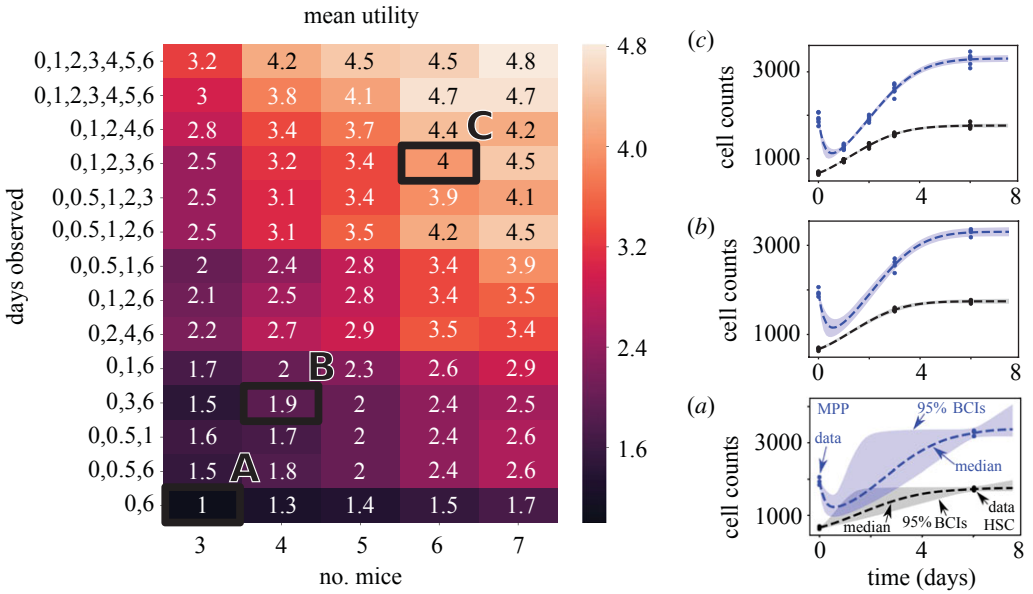

Figure 5. Mean utility values for sampled designs. *Left column.* The mean utility heat map across all model parameters following the simulation process presented in figure 2. Utility values correspond to a fold change to the minimum utility value, which corresponds to three mice observed at days 0 and 6. The higher utility fold changes correspond to a greater information gain. The boxes labelled A, B and C correspond to three designs with increasing mean utility value. *Right column.* The medians and 95% Bayesian credible intervals (bands) of the posterior distributions for the ODE solutions using the same synthetic dataset for the labelled designs. The parameter values additional details on the prior and posterior distributions are available in electronic supplementary material, table A-2.

at day 0 and 6, with minimum utility. To better understand the scale, we choose several designs (boxed), compare their mean utilities, and plot the corresponding ODE credible regions. The plots provide a visual explanation on how higher information gain, as quantified by the mean utility values, correspond to designs with higher number of data points, which in turn provide narrower credible regions (figure 5). Note that higher mean utillity can be interpreted as lower uncertainty regarding the system dynamics.

3.1.3. Parameter utility

The expected utility provides an overall metric for information gain by averaging over all model parameters. Alternatively, to quantify the amount of information provided by the observed data with respect to a specific model parameter, we can use the individual parameter utilities. This way, we can better understand how data affects identification of certain model parameters.

Additionally, obtaining individual parameter utilities allows us to find a specific design that is more informative for a parameter of interest. For example, if we focus on SC division (η_1), the estimated parameter utilities suggest that more observations should be allocated to earlier days. Looking at the top four design rows for η_1 in figure 6, the marginal mean utility value does not change even though we are adding more mice and more observation days. This shows that for η_1, the main contribution to information gain is coming from days 0 to 3. Similarly, if we focus on the HSC self-renewal probability feedback parameter, γ_1, we observe that adding data points at later times will lead to higher utilities. For the same parameter, we observe a lower marginal mean utility for designs where day 6 is not included. We also note that only a small amount of data is needed to identify the initial conditions, μ_1 and μ_2. For the full set of parameter utilities, please refer to figure A-5 in the electronic supplementary material.

As an example on how to use the parameter utilities for decision making, we can assume we have been given a finite budget of 20 mice for our perturbation experiment. Looking at the boxed *b*, *c* and *d* marginal mean utilities in figure 6, we can compare how the allocation of 20 mice could affect information gain for η_1 and γ_1. As mentioned previously, designs with more data points near the initial condition provide higher utilities for η_1 but lower utilities on γ_1. Design C provides a reasonable compromise between these two alternatives. For the full set of plots comparing prior and posterior distributions, please refer to figure A-9 in the electronic supplementary material. In general, we have found that given a fixed number of mice, designs with fewer mice at more time points tend to provide better results. Furthermore, as we obtain more data, we see to reach a point of saturation, see figure A-6 in the electronic supplementary material for more design comparisons.

4. Discussion

We have presented a new method to find the optimal experimental design for inferring parameters in complex, mechanistic mathematical models where the experimental data are temporal but not longitudinal. That is, each data point in time corresponds to a different experimental subject. Our method incorporates Bayesian utility theory in a hierarchical latent variables framework where the mechanistic model is used to predict the unobserved temporal dynamics. We quantified the amount of information gained from a specific experimental design as the difference between the prior and posterior distributions of model parameters using a utility metric based on the KL divergence function [24,39]. Furthermore, by calculating marginal utilities, our proposed framework also allows us to identify the information gained for each parameter from a specific design. By searching over the space of possible experimental designs, the best design is the one that maximizes the expected information gain

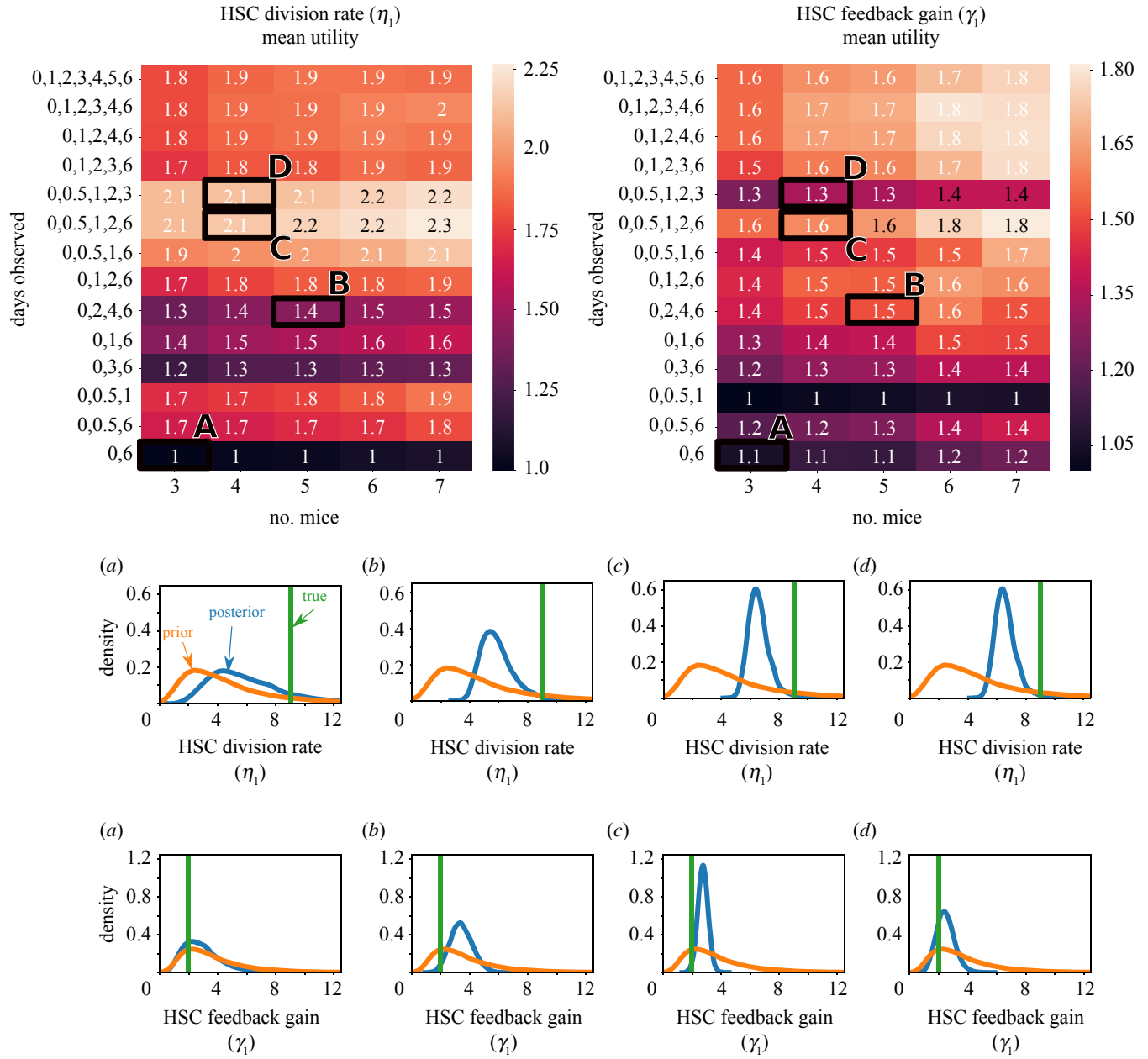

Figure 6. Individual utilities for the HSC division rate and feedback gain. *Top row.* Two mean utility heat maps for the HSC division rate η_1 and self-renewal feedback gain γ_1. The mean utility values are relative to the design with the lowest mean utility value for in each heat map. Four designs (labelled A–D) are selected with increasing mean utility values and different the sampling frequencies. *Middle and bottom rows.* Plots of the prior and posterior distributions for η_1 and γ_1 for the selected experimental designs, as labelled. The true values are shown in green.

subject to specific constraints, e.g. fixing the total number of experimental subjects or the total number of observations. Our method can be applied to a wide range of problems in systems biology and life sciences where multiple individuals are tracked over time and it is important to distinguish between subject to subject variation and measurement error. Also, our latent variable formulation can be used in any experiment where the data source is lost after it is measured, but a mathematical model is available to interpolate the missing trajectories.

We applied this framework to infer parameters in mathematical models of haematopoiesis in mice that incorporate feedback and feedforward regulation of self-renewal and division rates of HSC and MPP cells using experimental data on cell counts collected in mice. To obtain a cell count, each mouse had to be sacrificed and the bone marrow extracted and analysed. In the experiments, mice were subjected to low doses of radiation, which decreased the number of HSCs in the bone marrow. After about 7 days, the HSC and MPP cell counts returned to normal.

We considered a finite grid of possible designs relevant for the radiation experiments where the control parameters were the number of mice observed at each time point and the time points at which the cell counts were obtained. Generally, we found that for designs with a fixed number of subjects, having more observation times with fewer mice replicates leads to higher information gain measured by the expected utility values. Additionally, we found that the amount of information gain from a specific design can vary significantly across parameters. For example, division rates are better informed by designs that include more measurements at early times, right after the radiation was applied, but the feedback parameters are better identified by designs that have more measurements at later times closer to the equilibrium.

As illustrated in this paper, by using MCMC simulation, Bridge Sampling, and Monte Carlo approximation, we have been able to compute the required expected utilities for our optimal experimental design task. For more complex models, however, these methods might require a significant amount of time and computational resources for their implementation. Faster sampling methods [70–73], efficient numerical solvers for differential equations [74,75], and a more efficient exploration of the design space [48,76] can provide a significant reduction in computation time.

From the modelling perspective, our framework can be extended to stochastic formulations of the haematopoietic system with explicit feedback terms that generalize well-known branching process models [23]. Incorporating such an approach would enable us to model the cellular stochasticities and the heterogeneity of the system components directly. However, fitting stochastic models tends to be difficult since the likelihood function is generally not available and approximate Bayesian methods have to be used. Future research directions could involve developing better implementations of stochastic process models for this problem. Our method could also be extended to incorporate data on more differentiated cell types in order to provide insight into the response dynamics for more realistic models of the haematopoietic system with additional feedback mechanisms. To this end, we can include new perturbations, such as the depletion of specific differentiated cells, into our approach to experimental design.

Our approach can be extended to incorporate richer sources of data such as serially sampled barcoded single cells, which allows for lineage tracking and fate determination. Also, if the destruction of the data source can be avoided, repeated measurements of covariates and cell counts can provide a better understanding of marginal and population-level responses of haematopoiesis regulation by allowing the model parameters to change across individuals [77].

The hierarchical framework presented in this paper can significantly improve mathematical modelling of haematopoiesis by determining the required experiments to validate theoretical predictions and to obtain measures of uncertainty of model components. In particular, our approach can be used for testing more complex feedback regulation and control mechanisms. Finally, our work can motivate new collaborations between biologists, data scientists and mathematicians to develop a unified framework for hypothesis generation, modelling and experimental validation [78].

Data accessibility. The dataset used in this work can be found in the repository https://github.com/luisdm1/BayesOptimalDesign.

Authors' contributions. B.S., J.S.L., R.V.E. and V.N.M. designed the study. A.I., A.L. and J.S.L. developed the mechanistic, ordinary differential equation model. L.M.L., B.S. and V.N.M. developed the Bayesian framework. L.M.L and A.I. developed the numerical implementation of the framework, performed the simulations and analysed the numerical data. P.T., N.J., Z.-Y.L. and R.V.E. designed the mice radiation experiments, which were performed by P.T., and analysed the experimental data. L.M.L., A.I., A.L., P.T., R.V.E., B.S., J.S.L. and V.N.M. wrote the manuscript. L.M.L. and A.I. contributed equally to this work.

Competing interests. We declare we have no competing interests.

Funding. The authors acknowledge support from NSF grant nos. DMS-1936833 (L.M.L., A.I., R.V.E., A.L., B.S., J.S.L. and V.M.) and DMS-1714973 (J.S.L.). In addition, A.L., B.S., J.S.L., V.M. acknowledge support from DMS-1763272 and the Simons Foundation (594598QN) for a NSF-Simons Center for Multiscale Cell Fate Research. R.V.E., A.L., B.S., J.S.L., V.M. also thank the National Institutes of Health for partial support through grant nos. 1U54CA217378-01A1 for a National Center in Cancer Systems Biology at UC Irvine and P30CA062203 for the Chao Family Comprehensive Cancer Center at UC Irvine. Finally, L.M.L. acknowledges the support from the UC-MEXUS CONACYT and Fulbright-Garcia Robles doctoral fellowships.

References

1. Rieger MA, Schroeder T. 2012 Hematopoiesis. *Cold Spring Harb. Perspect. Biol.* **4**, a008250. (doi:10.1101/cshperspect.a008250)
2. Corey SJ, Kimmel M, Leonard JN 2014 *A systems biology approach to blood*, vol. 844. Advances in Experimental Medicine and Biology. New York, NY: Springer.
3. Höfer T, Barile M, Flossdorf M. 2016 Stem-cell dynamics and lineage topology from *in vivo* fate mapping in the hematopoietic system. *Curr. Opin. Biotechn.* **39**, 150–156. (doi:10.1016/j.copbio.2016.04.001)
4. Pujo-Menjouet L. 2016 Blood cell dynamics: half of a century of modelling. *Math. Model. Nat. Phenom.* **11**, 92–115. (doi:10.1051/mmnp/201611106)
5. MacLean AL, Lo Celso C, Stumpf MP. 2017 Concise review: stem cell population biology: insights from hematopoiesis. *Stem Cells* **35**, 80–88. (doi:10.1002/stem.2508)
6. Fornari C, O'Connor LO, Yates JWT, Cheung SYA, Jodrell DI, Mettetal JT, Collins TA. 2018 Understanding hematological toxicities using mathematical modeling. *Clin. Pharmacol. Ther.* **104**, 644–654. (doi:10.1002/cpt.1080)
7. Michor F, Hughes TP, Iwasa Y, Branford S, Shah NP, Sawyers CL, Nowak MA. 2005 Dynamics of chronic myeloid leukaemia. *Nature* **435**, 1267. (doi:10.1038/nature03669)
8. Marciniak-Czochra A, Stiehl T, W Jäger ADHo, Wagner W. 2009 Modeling of asymmetric cell division in hematopoietic stem cells-regulation of self-renewal is essential for efficient repopulation. *Stem Cells Dev.* **18**, 377–386. (doi:10.1089/scd.2008.0143)
9. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, Reth M, Höfer T, Rodewald H-R. 2015 Fundamental properties of unperturbed haematopoiesis from stem cells *in vivo*. *Nature* **518**, 542. (doi:10.1038/nature14242)
10. Craig M, Humphries AR, Mackey MC. 2016 A mathematical model of granulopoiesis incorporating the negative feedback dynamics and kinetics of G-CSF/neutrophil binding and internalization. *Bull. Math. Biol.* **78**, 2304–2357. (doi:10.1007/s11538-016-0179-8)
11. Glauche I, Kuhn M, Baldow C, Schulze P, Rothe T, Liebscher H, Roy A, Wang X, Roeder I. 2018 Quantitative prediction of longterm molecular response in TKI treated CML? Lessons from an imatinib versus dasatinib comparison. *Sci. Rep.* **8**, 12330. (doi:10.1038/s41598-018-29923-4)
12. Mahadik B, Hannon B, Harley BAC. 2019 A computational model of feedback-mediated hematopoietic stem cell differentiation *in vitro*. *PLoS ONE* **14**, 1–21. (doi:10.1371/journal.pone.0212502)
13. Manesso E, Teles J, Bryder D, Peterson C. 2013 Dynamical modelling of haematopoiesis: an integrated view over the system in homeostasis and under perturbation. *J. R. Soc. Interface* **10**, 20120817. (doi:10.1098/rsif.2012.0817)
14. Roeder I, Loeffler M. 2002 A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. *Exp. Hematol.* **30**, 853–861. (doi:10.1016/S0301-472X(02)00832-9)
15. Dingli D, Traulsen A, Pacheco JM. 2007 Stochastic dynamics of hematopoietic tumor stem cells. *Cell Cycle* **6**, 461–466. (doi:10.4161/cc.6.4.3853)
16. Kimmel M. 2014 Stochasticity and determinism in models of hematopoiesis. In *A Systems Biology Approach to Blood* (eds SJ Corey, M Kimmel, JN Leonard), Advances in Experimental Medicine and Biology, pp. 79–97. New York, NY: Springer.
17. Rozhok AI, Salstrom JL, DeGregori J. 2016 Stochastic modeling reveals an evolutionary mechanism

underlying elevated rates of childhood leukemia. *Proc. Natl Acad. Sci. USA* **113**, 1050–1055. (doi:10.1073/pnas.1509333113)

18. Jäkel F, Worm O, Lange S, Mertrelsmann R. 2018 A stochastic model of myeloid cell lineages in hematopoiesis and pathway mutations in acute myeloid leukemia. *PLoS ONE* **13**, e204393. (doi:10.1371/journal.pone.0204393)
19. Xu J, Wang Y, Guttorp P, Abkowitz JL. 2018 Visualizing hematopoiesis as a stochastic process. *Blood Adv.* **2**, 2637–2645. (doi:10.1182/bloodadvances.2018023705)
20. Krinner A, Roeder I, Loeffler M, Scholz M. 2013 Merging concepts-coupling an agent-based model of hematopoietic stem cells with an ode model of granulopoiesis. *BMC Syst. Biol.* **7**, 117. (doi:10.1186/1752-0509-7-117)
21. Golinelli D, Guttorp P, Abkowitz JA. 2006 Bayesian inference in a hidden stochastic two-compartment model for feline hematopoiesis. *Math. Med. Biol.* **23**, 153–172. (doi:10.1093/imammb/dql008)
22. Fong Y, Guttorp P, Abkowitz J. 2009 Bayesian inference and model choice in a hidden stochastic two-compartment model of hematopoietic stem cell fate decisions. *Ann. Appl. Stat.* **3**, 1696. (doi:10.1214/09-AOAS269)
23. Xu J, Koelle S, Guttorp P, Wu C, Dunbar C, Abkowitz JL, Minin VN. 2019 Statistical inference in partially observed stochastic compartmental models with application to cell lineage tracking of *in vivo* hematopoiesis. *Ann. Appl. Stat.* **13**, 2091–2119. (doi:10.1214/19-AOAS1272)
24. Ryan EG, Drovandi CC, McGree JM, Pettitt AN. 2016 A review of modern computational algorithms for Bayesian optimal design. *Int. Stat. Rev.* **84**, 128–154. (doi:10.1111/insr.12107)
25. Chaloner K, Verdinelli I. 1995 Bayesian experimental design: a review. *Stat. Sci.* **10**, 273–304. (doi:10.1214/ss/1177009939)
26. Liepe J, Filippi S, Komorowski M, Stumpf MPH. 2013 Maximizing the information content of experiments in systems biology. *PLoS Comput. Biol.* **9**, e1002888. (doi:10.1371/journal.pcbi.1002888)
27. Silk D, Kirk PDW, Barnes CP, Toni T, Stumpf MPH. 2014 Model selection in systems biology depends on experimental design. *PLoS Comput. Biol.* **10**, e1003650. (doi:10.1371/journal.pcbi.1003650)
28. Vanlier J, Tiemann CA, Hilbers PAJ, van Riel NAW. 2012 A Bayesian approach to targeted experiment design. *Bioinformatics* **28**, 1136–1142. (doi:10.1093/bioinformatics/bts092)
29. Busetto AG, Hauser A, Krummenacher G, Sunnäker M, Dimopoulos S, Ong CS, Stelling J, Buhmann JM. 2013 Near-optimal experimental design for model selection in systems biology. *Bioinformatics* **29**, 2625–2632. (doi:10.1093/bioinformatics/btt436)
30. Overstall AM, Woods DC, Parker BM. 2020 Bayesian optimal design for ordinary differential equation models with application in biological science. *J. Am. Stat. Assoc.* **115**, 583–598. (doi:10.1080/01621459.2019.1617154)
31. Wilkinson DJ. 2007 Bayesian methods in bioinformatics and computational systems biology. *Brief. Bioinformatics* **8**, 109–116. (doi:10.1093/bib/bbm007)
32. Gomez-Ramirez J, Sanz R. 2013 On the limitations of standard statistical modeling in biological systems: a full Bayesian approach for biology. *Prog. Biophys. Mol. Biol.* **113**, 80–91. (doi:10.1016/j.pbiomolbio.2013.03.008)
33. MacLean AL, Harrington HA, Stumpf MPH, Byrne HM. 2016 Mathematical and statistical techniques for systems medicine: the wnt signaling pathway as a case study. *Methods Mol. Biol.* **1386**, 405–439.
34. Sverchkov Y, Craven M. 2017 A review of active learning approaches to experimental design for uncovering biological networks. *PLoS Comput. Biol.* **13**, e1005466. (doi:10.1371/journal.pcbi.1005466)
35. Oden JT. 2018 Adaptive multiscale predictive modeling. *Acta Numer.* **27**, 353–450. (doi:10.1017/S096249291800003X)
36. Rosa BA, Zhang J, Major IT, Chen J. 2012 Optimal time point sampling in high-throughput gene expression experiments. *Bioinformatics* **28**, 2773–2781. (doi:10.1093/bioinformatics/bts511)
37. Kleyman M, Sefer E, Nicola T, Espinoza C, Chhabra D, Hagood JS, Kaminski N, Ambalavanan N, Bar-Joseph Z. 2017 Selecting the most appropriate time points to profile in high-throughput studies. *Elife* **6**, e18541. (doi:10.7554/eLife.18541)
38. Ezer D, Keir J. 2019 Nitpicker: selecting time points for follow-up experiments. *BMC Bioinf.* **20**, 166. (doi:10.1186/s12859-019-2717-5)
39. Kullback S, Leibler RA. 1951 On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86. (doi:10.1214/aoms/1177729694)
40. Dehideniya MB, Drovandi CC, McGree JM. 2018 Optimal Bayesian design for discriminating between models with intractable likelihoods in epidemiology. *Comput. Stat. Data Anal.* **124**, 277–297. (doi:10.1016/j.csda.2018.03.004)
41. Zhang JF, Papanikolaou NE, Kypraios T, Drovandi CC. 2018 Optimal experimental design for predator–prey functional response experiments. *J. R. Soc. Interface* **15**, 20180186. (doi:10.1098/rsif.2018.0186)
42. Han C, Chaloner K. 2004 Bayesian experimental design for nonlinear mixed-effects models with application to HIV dynamics. *Biometrics* **60**, 25–33. (doi:10.1111/j.0006-341X.2004.00148.x)
43. Huan X, Marzouk YM. 2013 Simulation-based optimal Bayesian experimental design for nonlinear systems. *J. Comput. Phys.* **232**, 288–317. (doi:10.1016/j.jcp.2012.08.013)
44. Müller P, Parmigiani G. 1995 Optimal design via curve fitting of Monte Carlo experiments. *J. Am. Stat. Assoc.* **90**, 1322–1330.
45. Wakefield J. 1994 An expected loss approach to the design of dosage regimens via sampling-based methods. *J. R. Stat. Soc.: Ser. D (The Statistician)* **43**, 13–29. (doi:10.2307/2348929)
46. Palmer JL, Müller P. 1998 Bayesian optimal design in population models for haematologic data. *Stat. Med.* **17**, 1613–1622. (doi:10.1002/(SICI)1097-0258(19980730)17:14<1613::AID-SIM867>3.0.CO;2-C)
47. Drovandi CC, McGree JM, Pettitt AN. 2013 Sequential Monte Carlo for Bayesian sequentially designed experiments for discrete data. *Comput. Stat. Data Anal.* **57**, 320–335. (doi:10.1016/j.csda.2012.05.014)
48. Müller P. 1999 Simulation-based optimal design. *Bayesian Stat.* **25**, 509–518. (doi:10.1080/01621459.1995.10476636)
49. Biegler L *et al.* 2011 *Large-scale inverse problems and quantification of uncertainty*, vol. 712. Chichester, UK: John Wiley & Sons.
50. Müller P, Berry DA, Grieve AP, Krams M. 2006 A Bayesian decision-theoretic dose-finding trial. *Decis. Anal.* **3**, 197–207. (doi:10.1287/deca.1060.0079)
51. Lomeli LM, Iniguez A, Shahbaba B, Lowengrub HS, Minin V. 2020 Optimal experimental design for mathematical models of hematopoiesis. arXiv 2004.09065.
52. Stewart FM, Zhong S, Wuu J, Hsieh C-C, Nilsson SK, Quesenberry PJ. 1998 Lymphohematopoietic engraftment in minimally myeloablated hosts. *Blood* **91**, 3681–3687. (doi:10.1182/blood.V91.10.3681.3681_3681_3687)
53. Lander AD, Gokoffski KK, Wan FYM, Nie Q, Calof AL. 2009 Cell lineages and the logic of proliferative control. *PLoS Biol.* **7**, 1–1. (doi:10.1371/journal.pbio.1000015)
54. Buzi G, Lander AD, Khammash M. 2015 Cell lineage branching as a strategy for proliferative control. *BMC Biol.* **13**, 13. (doi:10.1186/s12915-015-0122-8)
55. Dharampuriya PR, Scapin G, Wong C, Wagner KJ, Cillis JL, Dhavanit S. 2017 Tracking the origin, development, and differentiation of hematopoietic stem cells. *Curr. Opin. Cell Biol.* **49**, 108–115. (doi:10.1016/j.ceb.2018.01.002)
56. Brown G, Tsapogas P, Ceredig R. 2018 The changing face of hematopoiesis: a spectrum of options is available to stem cells. *Immunol. Cell Biol.* **96**, 898–911. (doi:10.1111/imcb.12055)
57. Inigiuez A. 2019 Mathematical modeling of malignant myelopoiesis: optimal experimental design and targeted therapy. PhD thesis, University of California, Irvine.
58. Olsson A, Venkatasubramanian M, Chaudhri VK, Aronow BJ, Salomonis N, Singh H, Grimes HL. 2016 Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698. (doi:10.1038/nature19348)
59. Arai F, Hirao A, Ohmura M, Sato H, Matsuoka S, Takubo K, Ito K, and Koh J-Y. 2004 Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell* **118**, 149–161. (doi:10.1016/j.cell.2004.07.004)
60. Staversky RJ, Byun DK, Georger MA, Zaffuto BJ, Goodman A, Becker MW, Calvi LM, Frisch BJ. 2018 The chemokine CCL3 regulates myeloid differentiation and hematopoietic stem cell numbers. *Sci. Rep.* **8**, 14691. (doi:10.1038/s41598-018-32978-y)
61. McGree J, Drovandi CC, Pettitt AN. 2012 A sequential Monte Carlo approach to the sequential

design for discriminating between rival continuous data models. See https://eprints.qut.edu.au/53813/1/smc_continuous.pdf.

62. Gronau QF *et al.* 2017 A tutorial on bridge sampling. *J. Math. Psychol.* **81**, 80–97. (doi:10.1016/j.jmp.2017.09.005)
63. Hoffman M, Gelman A. 2011 The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. arXiv, 1111.4246. (http://arxiv.org/abs/1111.4246).
64. Carpenter B *et al.* 2017 Stan: a probabilistic programming language. *J. Stat. Softw.* **76**. (doi:10.18637/jss.v076.i01)
65. Stan Development Team. PyStan: the Python interface to Stan. Version 2.17.1.0. 2018.
66. Jones E, Oliphant T, Peterson P. 2001 SciPy: Open source scientific tools for Python.
67. Scott DW. 2015 *Multivariate density estimation: theory, practice, and visualization*. Hoboken, NJ: John Wiley & Sons.
68. Stewart FM, Zhong S, Wuu J, Hsieh C-C, Nilsson SK, Quesenberry PJ. 1998 Lymphohematopoietic engraftment in minimally myeloablated hosts. *Blood* **91**, 3681–3687. (doi:10.1182/blood.V91.10.3681.3681_3681_3687)
69. Silva DS, Skilling J. 2006 *Data analysis: a Bayesian tutorial*. Oxford, UK: Oxford University Press.
70. Shahbaba B, Lomeli LM, Chen T, Lan S. 2019 Deep Markov chain Monte Carlo. arXiv, 1910.05692. See https://arxiv.org/abs/1910.05692.
71. Efendiev Y, Datta-Gupta A, Ginting V, B Mallick XMa. 2005 An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resour. Res.* **41**, W12423. (doi:10.1029/2004WR003764)
72. Drovandi CC, Tran M-N. 2018 Improving the efficiency of fully Bayesian optimal design of experiments using randomised quasi-Monte Carlo. *Bayesian Anal.* **13**, 139–162. (doi:10.1214/16-BA1045)
73. Beskos A, Girolami M, Lan S, Farrell PE, Stuart AM. 2017 Geometric MCMC for infinite-dimensional inverse problems. *J. Comput. Phys.* **335**, 327–351. (doi:10.1016/j.jcp.2016.12.041)
74. Calderhead B, Girolami M, Lawrence ND. 2009 Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes. In *Advances in neural information processing systems*, pp. 217–224. See https://proceedings.neurips.cc/paper/2008/hash/07563a3fe3bbe7e3ba84431ad9d055af-Abstract.html.
75. Chkrebtii OA, Campbell DA, Calderhead B, Girolami MA. 2016 Bayesian solution uncertainty quantification for differential equations. *Bayesian Anal.* **11**, 1239–1267. (doi:10.1214/16-BA1017)
76. Ryan EG, Drovandi CC, Thompson MH, Pettitt AN. 2014 Towards Bayesian experimental design for nonlinear models that require a large number of sampling times. *Comput. Stat. Data Anal.* **70**, 45–60. (doi:10.1016/j.csda.2013.08.017)
77. Davidian M. 2017 *Nonlinear models for repeated measurement data*. London, UK: Routledge.
78. Navlakha S, Bar-Joseph Z. 2011 Algorithms in nature: the convergence of systems biology and computational thinking. *Mol. Syst. Biol.* **7**, 546. (doi:10.1038/msb.2011.78)