

# **An experimental design tool to optimise inference precision in data-driven mathematical models of bacterial infections *in vivo***

**1. Myrto Vlazaki<sup>1</sup> (\*)**

**2. David J. Price<sup>2,3</sup>**

**3. Olivier Restif<sup>1</sup>**

<sup>1</sup>Department of Veterinary Medicine, University of Cambridge  
Madingley Road, Cambridge, CB3 0ES, United Kingdom

<sup>2</sup>Centre for Epidemiology and Biostatistics, University of Melbourne  
Grattan Street, Parkville, Victoria, 3010, Australia

<sup>3</sup>The Doherty Institute for Infection and Immunity  
792 Elizabeth Street, Melbourne, Victoria, 3000, Australia

(\*) Corresponding Author e-mail address: mv382@cam.ac.uk

**Keywords:** experimental design, population dynamics, within-host dynamics, microbiology, mathematical model, parameter inference, precision, likelihood-free, Bayesian statistics, optimality criteria

## **Abstract**

The management of bacterial diseases calls for a detailed knowledge about the dynamic changes in host-bacteria interactions. Biological insights are gained by integrating experimental data with mechanistic mathematical models to infer experimentally unobservable quantities. This interdisciplinary field would benefit from experiments with maximal information content yielding high-precision inference. Here, we present a computationally efficient tool for optimising experimental design in terms of parameter inference in studies using isogenic tagged strains. We study the effect of three experimental design factors: number of biological replicates, sampling timepoint selection and number of copies per tagged strain. We conduct a simulation study to establish the relationship between our optimality criterion and the size of parameter estimate confidence intervals, and showcase its application in a range of biological scenarios reflecting different dynamics patterns observed in experimental infections. We show that in low-variance systems with low killing and replication rates, predicting high-precision experimental designs is consistently achieved; higher replicate sizes and strategic timepoint selection yield more precise estimates. Finally, we address the question of resource allocation under constraints; given a fixed number of host animals and a constraint on total inoculum size per host, infections with fewer strains at higher copies per strain lead to higher-precision inference.

## 1. Introduction

Bacterial infections remain leading causes of mortality and morbidity worldwide, accounting for > 50% of annual deaths in vulnerable demographic groups such as the elderly, young children and immunocompromised individuals, and disproportionately afflicting resource-constrained regions [1]. At the same time, the rapid emergence of antibiotic resistance [2] and tolerance [3] are swiftly incapacitating an increasing number of clinically licenced antibiotic agents. Research and development of both novel antibiotics and vaccines is slow, leaving us exposed to untreatable, chronic and recrudescing bacterial infections.

The establishment, progression and outcome of a bacterial infection in a host is shaped by the dynamics of bacterial transmission, colonisation, within-host replication, death, inter-tissue migration, and the interaction of bacteria with the immune system or administered therapeutic agents. Deciphering each of these factors individually can accelerate the transition from empirical to targeted engineered approaches in bacterial treatment and prophylaxis [4]. To aid the characterisation of these dynamics, we often combine high-resolution experimental techniques with inferential mathematical models that fill in the gaps with regards to processes that remain experimentally unobservable [5].

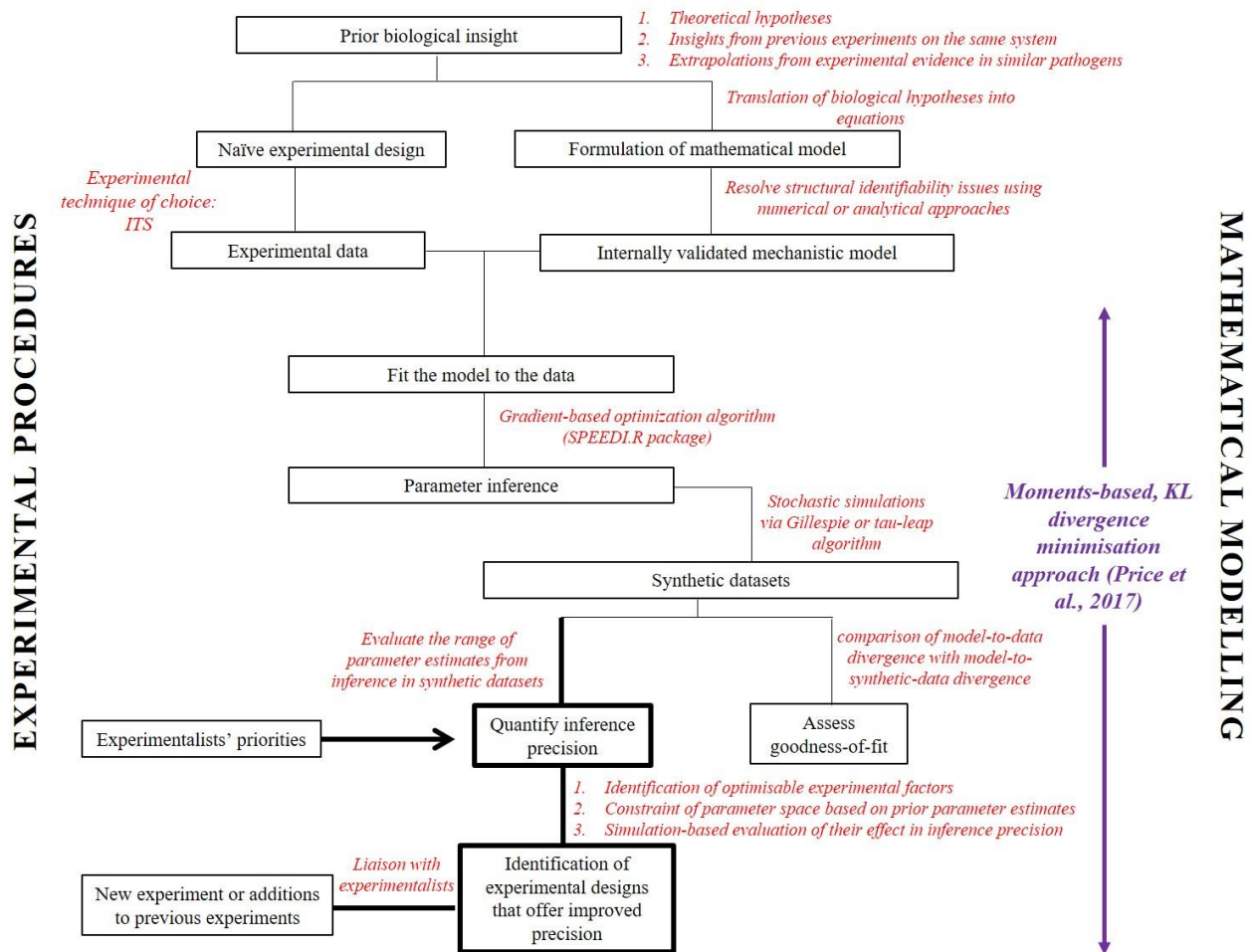
In the field of bacterial dynamics, isogenic-tagged strains (ITS) constitute one of the most widely used experimental techniques to obtain *in vivo* data about the dynamic behaviour of bacteria in the host [6]. ITS-based data have been historically paired with mathematical models to maximise the biological insight they can offer [7, 8]. By complementing experimental data with mechanistic mathematical models, we go beyond the mere characterisation of statistical differences and correlations in the data, to formulation of biological hypotheses into mathematical relations and inference-making by fitting the models to the experimental observations. Representative examples of integrating such models with ITS-based data include studies identifying bottlenecks during the colonisation and infection process [9-12], mapping out the host immune responses underpinning defence at different stages of the infection [11-12], deciphering the effects of host adaptation in subsequent infections [13-14], comparing the effects of different vaccine formulations [4] and antibiotic therapies [10, 15] on infection control, characterising the within-host dynamics of co-infections [9] and understanding the directionality of bacterial migration and *de novo* tissue colonisation *in vivo* [10, 12].

The aim of statistical inference is to estimate biological parameters as accurately and precisely as possible. Precision is typically reported in the form of confidence intervals [4, 10, 13, 15], which are instrumental to assess the evidence for or against hypotheses such as whether bacteriostatic or bactericidal effects differ at different stages in the infection process. Quantifying parameter estimate uncertainty is also important when we use a parameterised model for predictive purposes, as it allows us to predict

the range of the potential future trajectories of the infection. To enable predictions and comparisons to be as clear-cut and distinguishable as possible, it is desirable to obtain parameter estimates with the narrowest possible confidence intervals.

Although mathematical models have been increasingly used in the study of *in vivo* bacterial population dynamics, their application has until recently remained limited to inference in simple biological systems for short time intervals post-inoculation (typically in the range of 0-12 hours) largely due to computational inefficiencies [4, 10, 11, 13]. A recently developed statistical inference tool customised for making inferences based on experimental data from studies using ITS [16] has enabled inference in more complex biological systems, with larger inocula, and for the entirety of the infection timeline [15]. This moments-based, divergence minimisation approach represents a computationally efficient inference tool suitable for complex biological models, where traditional maximum likelihood approaches become too computationally expensive.

To this date, there are no published studies providing executable algorithms for extending the application of data-based mechanistic models to experimental design in microbiology studies on within-host pathogen dynamics. The availability of a computationally efficient inference method now paves the way for an early iterative integration of mathematical modelling with data collection at the experimental design stage; this can optimise the inferential capacity of models and maximise the biological insights they generate (Figure 1) [5, 16]. Model-driven experimental design can also address the long-standing challenge of resource allocation, especially with regards to the use of animals in experiments. In particular, the “Three Rs” tenet — Replacement, Reduction and Refinement [17] — calls for “Reduction Alternatives” encouraging strategies to reduce the number of animals needed to provide answers to scientific questions. Maximising the information obtained per animal has been recognised as one avenue to achieve [18]. In this paper, we show that with experimental design tools, it is possible to evaluate the information content of alternative designs and identify those with the maximum information content per animal host.



**Figure 1: Model-informed experimental design in an integrated framework of experimental procedures and mathematical modelling**

In this paper, we extend the inference methodology to improve experimental designs and provide a case study as a blueprint of how it can be applied to data from ITS-based studies. As ITS remains one of the key experimental techniques in the field of within-host bacterial dynamics [19-30], the development of dedicated tools to facilitate the integration of mathematical modelling along the experimental process can maximise biological insight by improving the statistical quality of inferences for a substantial body of research. A unique feature of ITS-based studies is the multiplicity of infections with different tagged strains within each host organism. Under the assumptions of independent action for bacteria and homogeneity in dynamics both between ITS and host organisms, both of which are assumptions commonly made in ITS-based studies [4, 11-14], the number of biological replicates is equal to the number of ITS per host multiplied by the number of hosts. Experiments produce measurements of the number of copies of each ITS in anatomical compartments of interest at the time that each host animal is killed, and inference about the within-host dynamics “fills in” the gap between the sampled timepoints.

Taking these features of ITS infections into account, we provide the first application of mathematical modelling as a tool to improve experimental design in studies of *in vivo* bacterial population dynamics using ITS. These studies commonly address questions around the dynamical composition of bacterial populations jointly shaped by environmental and biological processes such as bacterial replication, death, migration, phenotypic switch and bacterial-host interactions. We model bacterial kinetics with a continuous time Markovian process where the state variables are the number of copies of a single ITS in each anatomical compartment. Given an experimental design, the model predicts the dynamics of the lower moments (mean, variance and covariance) of the state variables. Parameter inference is based on minimising the Kullback-Liebler divergence between the predicted and observed moments (the latter being calculated from biological replicates, i.e. all ITS from all mice at each time point). First, we formulate a utility function inspired by the moments-based, divergence minimisation approach [16] that represents our criterion for optimality; the objective is to maximise this function across the design space. As we are interested in maximising global parameter inference precision, we focus on optimising the determinant of the parameter variance-covariance matrix. We use synthetic data to study its performance against the width of the corresponding 95% confidence intervals for a range of parameter sets. Finally, we illustrate its application by considering experimental designs for a variety of biological scenarios reflecting different phases of bacterial dynamics including exponential growth and decay as well as cases dominated by unidirectional inter-tissue bacterial migration.

## 2. Methods

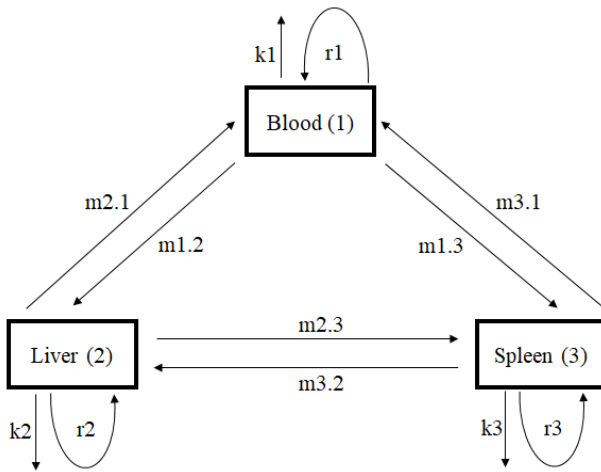
### 2.1. The ITS technique: experimental output and dataset structure

ITS represent a large class of marker-based techniques, developed over the last decade, in order to study the within-host bacterial dynamics at a population level. Libraries of isogenic tagged strains are generated by uniquely modifying the non-coding genome of bacteria with distinctive, short nucleotide sequences; the resulting ITS are genetically distinguishable, yet phenotypically identical. The ITS are mixed in equiproportional inocula and administered to the animal host. Animals are culled at predetermined time points to harvest or partly sample tissues of interest, and a total bacterial count is obtained per tissue [9, 11]. The samples are also processed to determine the ITS composition in each tissue, originally by qPCR [9-10, 12] and more recently by next-generation sequencing [8, 25]. The final structure of the dataset consists of number of copies per ITS in each tissue of interest per time point.

By obtaining snapshots of spatiotemporal data recapitulating the changing ITS composition of bacteria in different tissues, it is possible to characterise the unobserved processes that underpin the dynamics of the infection *in vivo*. Supplementing these data with compartmental mechanistic models, the rates at which these unobserved dynamical processes evolve can be quantified with optimisable precision. We proceed to show how the inference precision can be optimised for this type of experimental data, by quantifying the impact of a range of design factors that we intuitively expect to affect inference precision.

### 2.2. Conceptual Model Structure

Mechanistic models applied to ITS data typically have a compartmental structure, with each compartment representing a tissue of interest, and model parameters corresponding to the rates at which the unobserved biological processes change with time. In this paper, we use a radial model structure with three compartments, inspired by the model used in *Salmonella* Typhimurium infections [4, 11, 16], and shown as an example in Figure 2. Linear compartmental model structures or a combination are also possible, and the model can be generalised to include  $n$  compartments [16].



Parameter Name	Parameter Description
m1.2	Bacterial migration from the blood to the liver
m1.3	Bacterial migration from the blood to the spleen
m2.1	Bacterial migration from the liver to the blood
m3.1	Bacterial migration from the spleen to the blood
m2.3	Bacterial migration from the liver to the spleen
m3.2	Bacterial migration from the spleen to the liver
k1	Bacterial death in the blood
r1	Bacterial replication in the blood
k2	Bacterial death in the liver
r2	Bacterial replication in the liver
k3	Bacterial death in the spleen
r3	Bacterial replication in the spleen

**Figure 2: Conceptual structure of the compartmental model detailing the intra- and inter-tissue dynamics in a model of systemic salmonellosis. In each organ  $i$ , bacteria replicate and are killed at rates  $r_i$  and  $k_i$ , respectively, and bacteria migrate to organ  $j$  according to rates  $m_{i,j}$ .**

### 2.3. A likelihood-free inference framework for complex dynamical systems

As we are usually interested in tracking the evolution of infections over a period of days, and as bacterial infections are typically multi-organ, data collected from ITS-based experiments are often complex. Their complexity lies in (1) the length of observational time for the biological system in question, (2) the number of tissues of interest, and (3) the multiple organ-organ interactions. The ITS composition across all tissues of interest is summarised in a vector of moments (means, variances, and covariances) corresponding to each observation time. Assuming first-order dynamics for each time interval between sequential points of data collection, a simple mathematical expression that derives the moments at a future time point can be obtained and solved analytically. The moments of the experimental dataset and the calculated moments for the same time point are compared via the Kullback-Leibler divergence. Minimisation of this divergence via standard optimisation routines allows us to quantify processes underpinning the change in bacterial numbers and composition during a time interval of interest. A full set of computational tools and examples on how to use this inference framework can be freely accessed at: <https://github.com/orestif/SPEEDI.R>. For a more detailed description of the likelihood-free framework, please refer to the Supplementary Materials.

### 2.4. Parametric bootstrap to evaluate inference precision

Parameter inference would not be complete without a measure of its precision. Quantification of precision in this framework is carried out using a parametric bootstrap approach previously described [16]. Having identified the best parameter estimate by minimising the KL-divergence between the experimental and calculated moments for a range of parameter sets, we use it to simulate 200 synthetic



datasets of the same size and structure as the experimental data. The choice of 200 bootstrapped samples was motivated by our priority to maximise the computational efficiency of the suggested tool, and is in line with previous findings on the minimum adequate number of bootstrapped samples [48]. The corresponding moments are calculated, and the best parameters inferred via the same divergence-minimisation method. This process results in a range of parameter estimates, which provides us with a measure of precision regarding the inference for each parameter.

## 2.5. Deriving optimality criteria in a likelihood-free inference framework

We now turn to formulating an optimality criterion in this likelihood-free framework, inspired by the gold-standard alphabetical optimality criteria in the maximum likelihood framework [31].

In the maximum likelihood estimation (MLE) context, the Fisher Information Matrix (FIM) plays a key role in assessing parameter inference precision [32]. It summarises the amount of information contained in the experimentally observed dataset with regards to the parameter set  $\theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_n\}$ . The FIM can be analytically derived from the Hessian matrix ( $H(\theta)$ ), a matrix of 2<sup>nd</sup> order derivatives of the likelihood  $L(\theta)$  with respect to the parameters:

$$FIM = -E[H(\theta)] = -E \left[ \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \right] \quad (1)$$

For a given MLE, the FIM contains information about the curvature of the likelihood. A higher value in the corresponding entry in the FIM matrix, indicates that the estimate for that is more precise (which can be expressed as a shorter confidence interval). Furthermore, by assessing different functions of the FIM (e.g., the determinant or trace) it is possible to derive information about different aspects of the inference quality provided by each experimental design; these ways of assessing different aspects of the inference quality of a model given an experimental design are known as alphabetical optimality criteria [31]. For example, the determinant of the FIM provides an index of average size (e.g., width, area, volume) of the confidence interval for the parameters of interest (D-optimality criterion), while the trace of the FIM summarises the average variance of each parameter of interest (A-optimality criterion) [31]. Other composite functions of the FIM lead to the full range of the alphabetical optimality criteria.

An important property of the FIM is its inverse relationship with the variance-covariance matrix of  $\Theta$ , via the Cramer-Rao lower bound [33]. In 2.3. we used a parametric bootstrap approach to estimate the variance-covariance matrix of the parameter estimates; let that be  $M$ . As the inverse of  $M$  is proportional to the FIM, we can define equivalent optimality criteria by minimising functions of  $M$ , rather than maximising functions of the FIM. Analytical evaluation of the FIM is often difficult or impossible [34], but instead typically approximated numerically as we propose to do with the minimum-divergence

inference method. Here, we focus on minimising the determinant of  $M$ , which serves as the global measure of inference precision.

## 2.5. Identifying experimental designs with higher information content

The FIM summarises the information content of an experiment as a function of both the model parameters  $\Theta$ , and the design conditions under which the experiment is conducted. As such, it can be used to identify experimental designs that yield data with a higher information content. To achieve this, we first identify which aspects of the experimental process we are interested in optimising. Here, we focus on the number of biological replicates, the initial number of copies per ITS and the choice of sampling timepoints.

Next, we place sensible bounds on the parameter space to be explored. This can be done through expert elicitation, based on prior literature or understanding of the system, or by using inferences made on previous experiments on the same or a similar system.

For the scenarios we consider in Sections 3.1, 3.2.1 and 3.2.2, we randomly draw parameters from a uniform distribution  $\theta \sim U(0,1)$ . In 3.2.3, we explore 6 biological patterns of within-host dynamics commonly occurring in *in-vivo* bacterial infections, which we list below. For each of the scenarios A1, A2, B1, B2, C1, C2 and keeping the total number of replicates fixed at 500, we consider 10 parameter sets, and for each parameter set 3 numbers of copies per ITS in  $\{10,100,1000\}$  and 8 sampling time points reflecting early, mid-, late and a combination of early and late sampling in:

$t \in \{(1,4),(2,5),(10,13),(11,14),(19,23),(20,24),(1,23),(2,24)\}$ .

To provide an example of how these sampling strategies would be implemented for a time interval of 24 hours, a group of animal hosts would be euthanised at 1 hour and a separate group at 4 hours post-inoculation (1,4). In the next experimental design, a group of animal hosts would be euthanised at 2 hours and a separate group at 5 hours post-inoculation (2,5).

### A. Exponential growth (net growth = +0.2, minimal inter-tissue migration)

A1. Exponential growth with high killing and high replication rates (higher variance in data)

$$k_i \sim U(2,2.5), r_i = k_i + 0.2$$

A2. Exponential growth with low killing and low replication rates (lower variance in data)

$$k_i \sim U(0.1,0.6), r_i = k_i + 0.2$$

### B. Exponential decay (net growth = -0.2, minimal inter-tissue migration)

B1. Exponential decay with high killing and high replication rates (higher variance in data)

$$k_i \sim U(2, 2.5), r_i = k_i - 0.2$$

B2. Exponential decay with low killing and low replication rates (lower variance in data)

$$k_i \sim U(0.3, 0.8), r_i = k_i - 0.2$$

**C. Source-sink unidirectional bacterial transfer with near-zero intra-organ growth**

C1. Source-sink with high killing and high replication rates (higher variance in data)

$$k_i, r_i \sim U(2, 2.5), m_{1j} \sim U(1.5, 2)$$

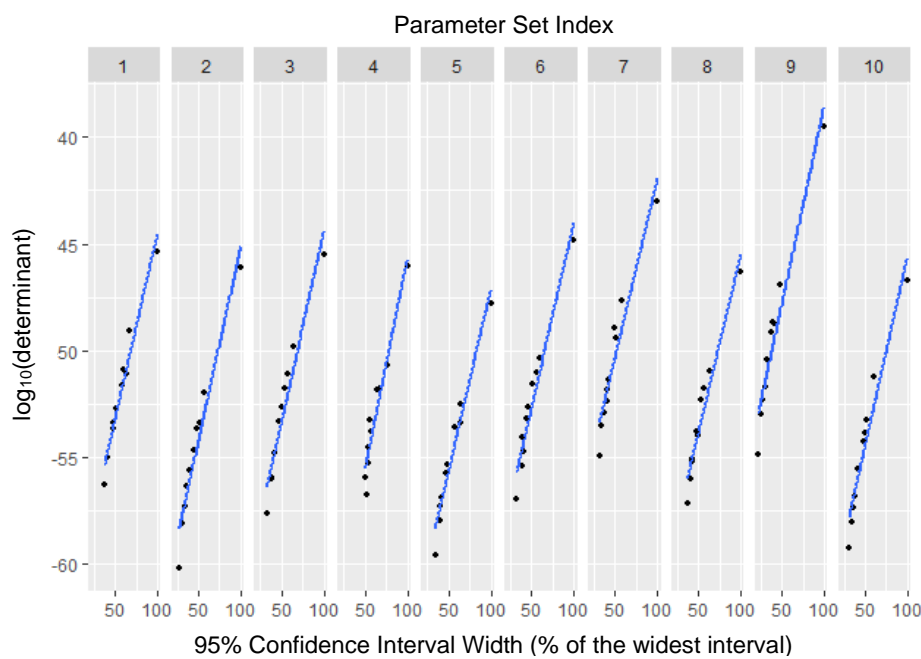
C2. Source-sink with low killing and low replication rates (lower variance in data)

$$k_i, r_i \sim U(0.1, 0.6), m_{1j} \sim U(1.5, 2)$$

### 3. Results

#### 3.1. A simulation study to correlate the determinant of the parameter variance-covariance matrix with precision in parameter inference

In this section we perform a simulation study to demonstrate that the magnitude of the determinant of the parameter variance-covariance matrix is a reliable summary statistic for the precision of inference across the entire parameter set. We start with the model structure outlined in 2.2 and generate 10 random parameter sets. We then simulate 10 synthetic datasets for each parameter set ( $10 \times 10 = 100$  in total); the synthetic datasets serve as surrogates for experimentally observed ITS-based data. For each of the 100 virtual datasets, we obtain 200 bootstrapped samples and proceed to obtain 200 sets of parameter estimates. For each set of parameter estimates, we calculate the variance-covariance matrix and obtain its determinant, which is taken to be a global measure of inference precision. Next, we obtain the 95% confidence interval for each of the 12 parameters across the 10 synthetic datasets generated for each parameter set index and report the mean of their widths across each parameter set relative to the maximum one, as a function of the corresponding determinant (Figure 3).



**Figure 3: Mapping the determinant of the variance-covariance matrix to the range of 95% confidence intervals**

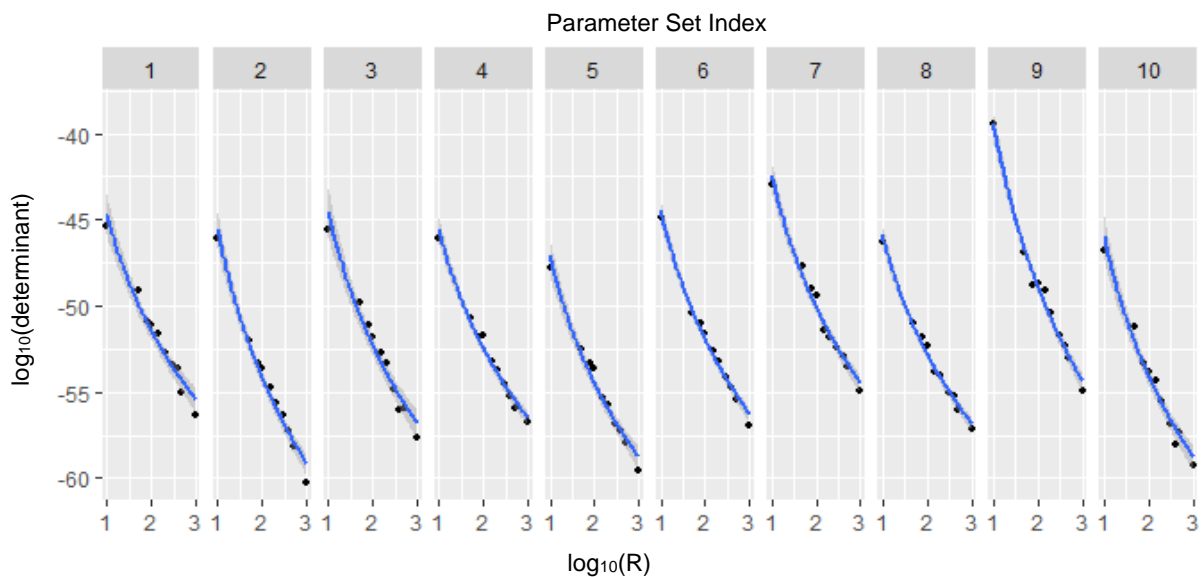
Each dot represents the mean width of the 95% confidence intervals across the 12 parameters in each synthetic dataset. A three-compartmental model is used (section 2.2) and parameters are drawn at random from a uniform distribution between 0 and 1,  $\theta \sim U(0,1)$ . For each of the randomly generated parameter sets (parameter set index 1-10), there are 10 dots representing 10 synthetic datasets simulated from that parameter set. The mean width of the 95% confidence intervals is plotted against the mean logarithm of the determinant across the 10 synthetic datasets generated by each of the parameter sets, showing a linear positive relationship between the two quantities. The mean correlation coefficient across the 10 parameter sets is 0.178 with range [0.16709, 0.19858].

### 3.2. Determinants of precision inference in experimental design

Having established the determinant of the parameter variance-covariance matrix as a reliable correlate to inference precision in 3.1, here we study the effect of 3 design factors on the quality of inference precision in ITS-based experiments, namely: the number of biological replicates, the number of copies per ITS and the choice of sampling time points. First, in 3.2.1, we focus on number of biological replicates ( $R$ ); under the assumption of homogeneous inter-host immunological responses [4, 11-14],  $R$  is equal to the number of animal hosts multiplied by the number of ITS per host. In 3.2.2, we introduce the biological limitation of a constraint on the total inoculum size ( $I$ ) per host;  $I$  is equal to the number of ITS per host multiplied by the number of copies per ITS. Finally, in 3.2.3, we study the mixed effect of experimental designs with different numbers of copies per ITS and different sampling timepoints.

#### 3.2.1. Number of biological replicates ( $R$ )

In 3.2.1, we use the same simulation-based approach and parameter sets as in 3.1. to evaluate the effect of the number of biological replicates on the magnitude of the determinant of the parameter variance-covariance matrix. At this stage, the number of biological replicates ( $R$ ) is equal to the animal hosts multiplied by any number of ITS per host, without constraint on the total inoculum size per animal host. As shown in Figure 4, there is a directional effect between the determinant of the parameter variance-covariance matrix generated by parametric bootstrap and the total number of replicates. As the number of total replicates increases ( $x$ -axis), the quality of parametric inference increases (reflected by lower values in the determinant on the  $y$ -axis).



**Figure 4: Inference precision as a function of total number of replicates**

For each parameter set, 8 experimental designs are considered with number of biological replicates  $R \in \{5, 10, 20, 50, 100, 200, 500, 10000\}$ . The  $\log_{10}$  of the determinant of the parameter variance-covariance matrix is plotted against  $\log_{10}(R)$  to yield a consistent pattern of exponential decay across the 10 parameter sets. A three-compartmental model is used (section 2.2) and parameters are drawn at random from a uniform distribution between 0 and 1,  $\theta \sim U(0, 1)$ .

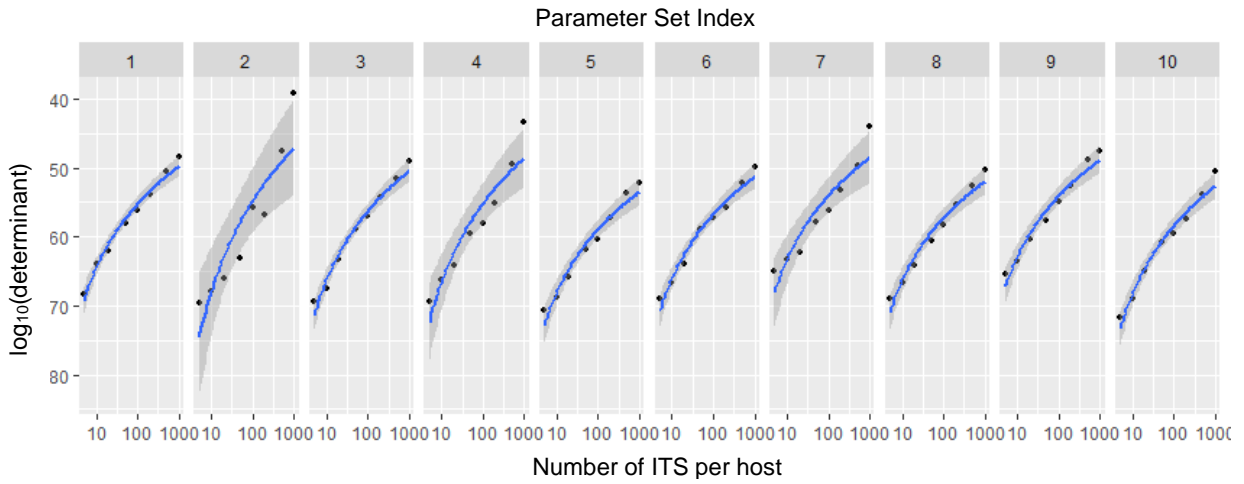
### 3.2.2. Distribution of replicates in animal hosts with a constraint on total inoculum size per host ( $I$ )

Here, we move a step further from 3.2.1 and introduce the biological constraint of total inoculum size per host ( $I$ ). In ITS-based experimental infections, there is a target total inoculum  $I$  achieved by multiplying the number of different ITS used by the number of copies per ITS. As detailed in Table 1, we consider experimental designs with the total inoculum size fixed at 10,000 CFU; this is consistent with the orders of magnitude of inocula used in ITS studies [4, 8, 11-12]. We also keep the number of biological replicates fixed at 1000 biological replicates to minimise the stochastic noise in experiments with small replicate sizes, and consider 8 possible combinations between the total number of host animals per sampling time point and the total number of ITS per host animal.

Experimental Design Index	Total number of host animals per sampling time point	Total number of ITS per host animal	Mean number of CFU per ITS			Total number of replicates (identical simulations)	Total inoculum size
			Compartment 1	Compartment 2	Compartment 3		
1	200	5	2000	1	1	1000	10000
2	100	10	1000	1	1	1000	10000
3	50	20	500	1	1	1000	10000
4	20	50	200	1	1	1000	10000
5	10	100	100	1	1	1000	10000
6	5	200	50	1	1	1000	10000
7	2	500	20	1	1	1000	10000
8	1	1000	10	1	1	1000	10000

**Table 1: A matrix tabulating the 8 experimental designs arising from different combinations of total number of host animals per sampling time point and total number of ITS per host animal, with fixed total number of replicates and fixed total inoculum size. Fixed variables are highlighted in grey, and variable design factors in yellow.**

In Figure 5, we plot the determinant of the parameter variance-covariance matrix as a function of the total number of ITS per host animal. There is a consistent positive relationship between the two variables, indicating that lower numbers of ITS per host yield lower determinant values, which reflect higher quality inference in terms of precision. Our analysis shows that given a fixed number of biological replicates and total inoculum size per host, experimental designs with higher numbers of hosts and lower numbers of ITS per host are predicted to yield higher precision estimates.



**Figure 5: Inference precision as a function of number of ITS per host**

For each parameter set, 8 experimental designs are considered with number of ITS per host in  $\{5,10,20,50,100,200,500,10000\}$ . The  $\log_{10}$  of the determinant of the parameter variance-covariance matrix is plotted against the number of ITS per host to yield a consistent pattern of exponential growth across the 10 parameter sets. A three-compartmental model is used (section 2.2) and parameters are drawn at random from a uniform distribution between 0 and 1,  $\theta \sim U(0,1)$ .

### 3.2.3. The choice of sampling time points in designs with different numbers of copies per ITS

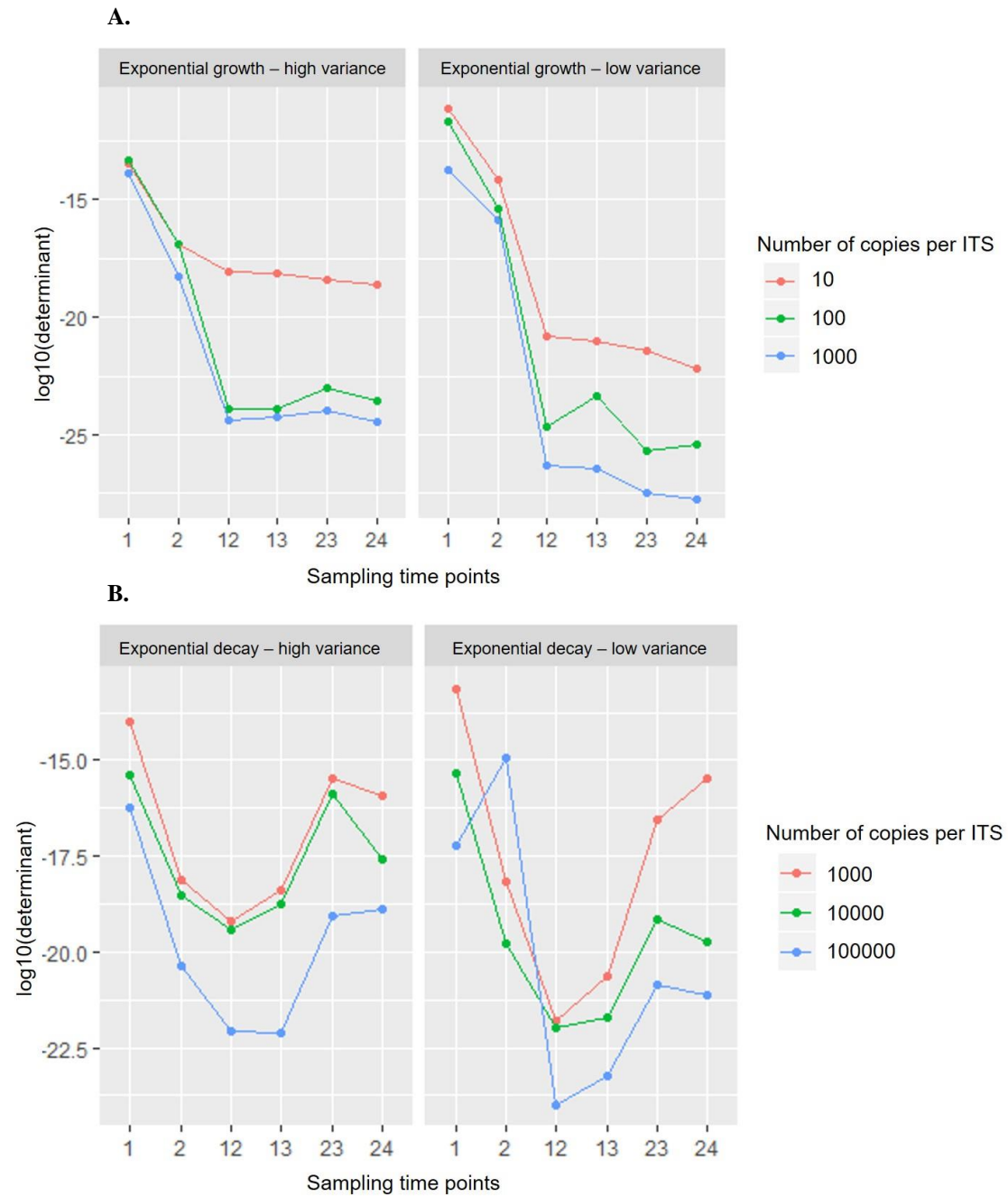
In 3.2.2. we found that given the presence of biologically realistic constraints on total inoculum size per host and total number of replicates, infections with fewer ITS at higher copy number per ITS consistently perform better in terms of precision inference. As a result, achieving high precision comes with the cost of using more animal hosts. In this section, we introduce a third design factor, the choice of sampling time points during the infection process. We seek to address the question: can we choose sampling time points that maximise the precision inference, allowing us to discount on the number of copies per ITS and, by extension, the number of animal hosts we need to inoculate?

To answer this question, we consider 6 biological scenarios that encompass common patterns in within-host dynamics of bacterial disease, described in 2.6. In particular, we look at scenarios of exponential decay and growth with either high or low variance, as well as cases with unidirectional bacterial migration from one organ to the other.

We start with sampling at a single observation point. In the likelihood-free inference framework, inference using a single observation time point is only possible for models whose number of parameters is smaller than or equal to the number of moments that summarise the ITS distributions. We first consider a simple, two-compartment model structure with unidirectional bacterial flow from the first to the second compartment, and bacterial death and replication in both (5 parameters in total). Figure 6 shows that unlike in the scenarios of exponential growth, in the scenarios of exponential decay intermediate sampling times (12- and 13-hours post-inoculation) yield higher-precision inference, whereas early or late sampling times provide less information on average. This is likely due to either minute changes in the bacterial population in the very early stages (1-2 hours), or bacterial populations having died out during the later stages (23-24 hours), at which point we can only get bounds on the parameter estimates that are consistent with extinction "up to" that time (Figure 6A). This pattern differs in the growth example (Figure 6B), where early time points will be subject to greater variation in the population sizes. Later time points will have allowed sufficient time for the populations to establish/replicate whereby the variation between subpopulations will be less impactful on the inference.

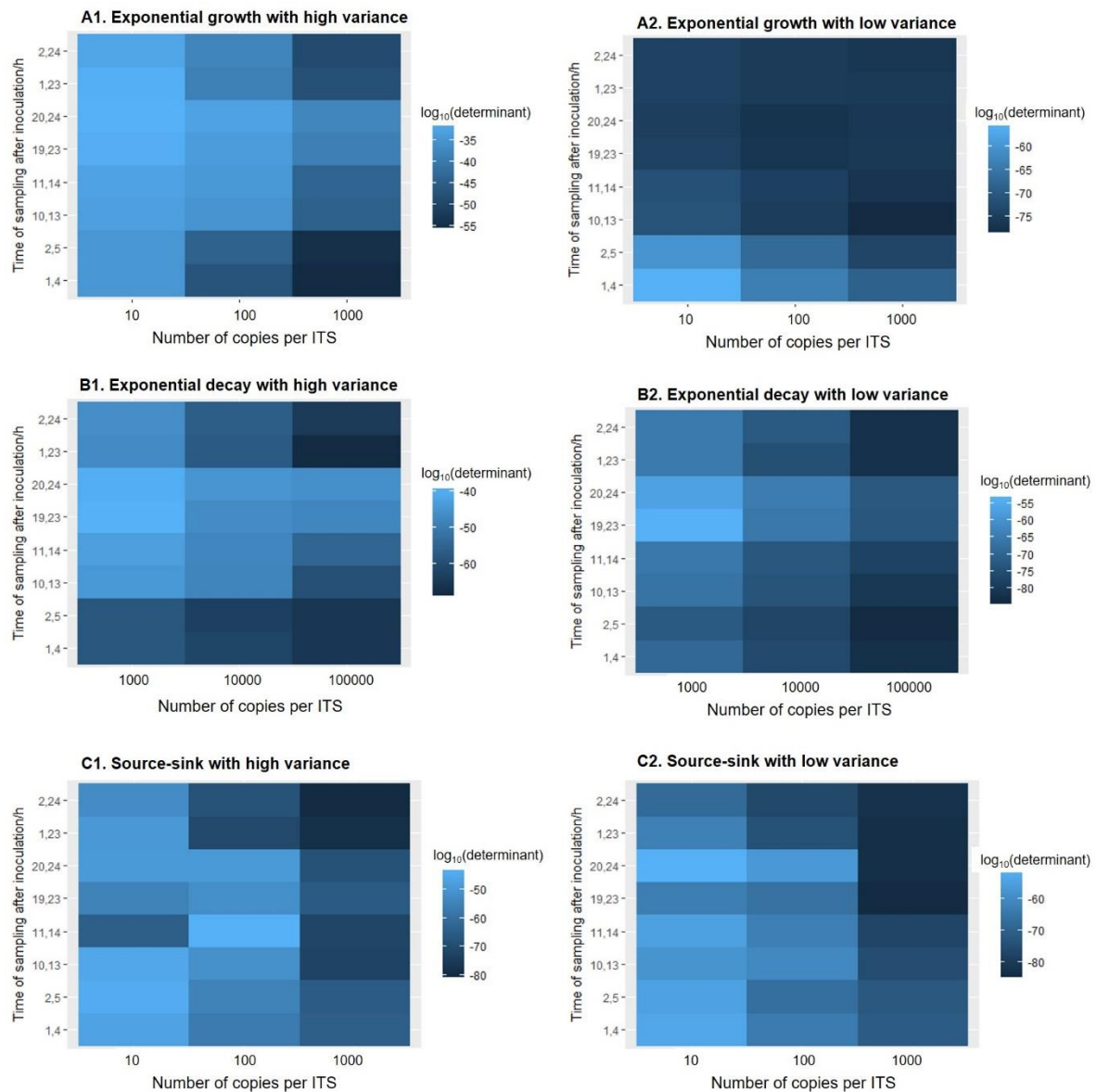
In Figure 7, we present the results of the same analysis performed using a combination of sampling time points for a four-compartment model. The precision across all experimental designs considered in Figure 7 is consistently higher than in Figure 6, as combining data two observation times increases the information content overall. Finally, experimental infections with higher number of copies per ITS perform consistently better in terms of inference precision (Figures 6 and 7). Assuming a constraint on the total inoculum size per host, this translates into designs involving a larger number of hosts per sampling time point.





**Figure 6: Scatter plots showing the mean  $\log_{10}(\text{determinant})$  across 10 parameter sets under 4 biological scenarios with single sampling time points for a three-compartment model**

For the exponential growth scenarios, later sampling yields the highest-precision inference (Panel A). For the exponential decay scenarios, there is a U-shaped relationship between the time of sampling and inference precision, with intermediate sampling time points being the most informative. A two-compartmental model with intra-organ bacterial growth and death, and bidirectional bacterial migration is used. Parameters are drawn at random from a uniform distribution between 0 and 1,  $\theta \sim U(0,1)$ .



**Figure 7: Heatmap plots showing the mean  $\log_{10}(\text{determinant})$  across 10 parameter sets under 6 biological scenarios for a range of experimental designs with variable sampling time points and starting number of copies per ITS for a four-compartment model**

The first column of panels (A1, B1, C1) illustrates biological scenarios with high killing and replication rates, while the second column (A2, B2, C2) corresponds to biological scenarios with low killing and replication rates. Darker blue shades represent experimental designs yielding higher inference precision, while lighter blue shades represent designs with lower inference precision. For all experimental designs across all scenarios (A1-C2), higher numbers of copies per ITS improve the precision inference. For low-variance scenarios (low replication, low killing), the inference precision is globally better and the number of high precision experimental designs is larger than for high-variance scenarios. A three-compartmental model is used (section 2.2) and parameters are drawn at random from distributions specified in section 2.5.

### 3.3. Strategies to reduce the number of animal hosts

In 3.2.3, we showed that infections with higher numbers ITS per host tend to correlate with higher inference precision consistently across the sampling time points we considered. Under a fixed inoculum size per host, this translates into the cost of requiring more animal hosts, which incurs primarily ethical, but also financial costs. The analysis in 3.2.3 assumed a fixed number of biological replicates of 500.

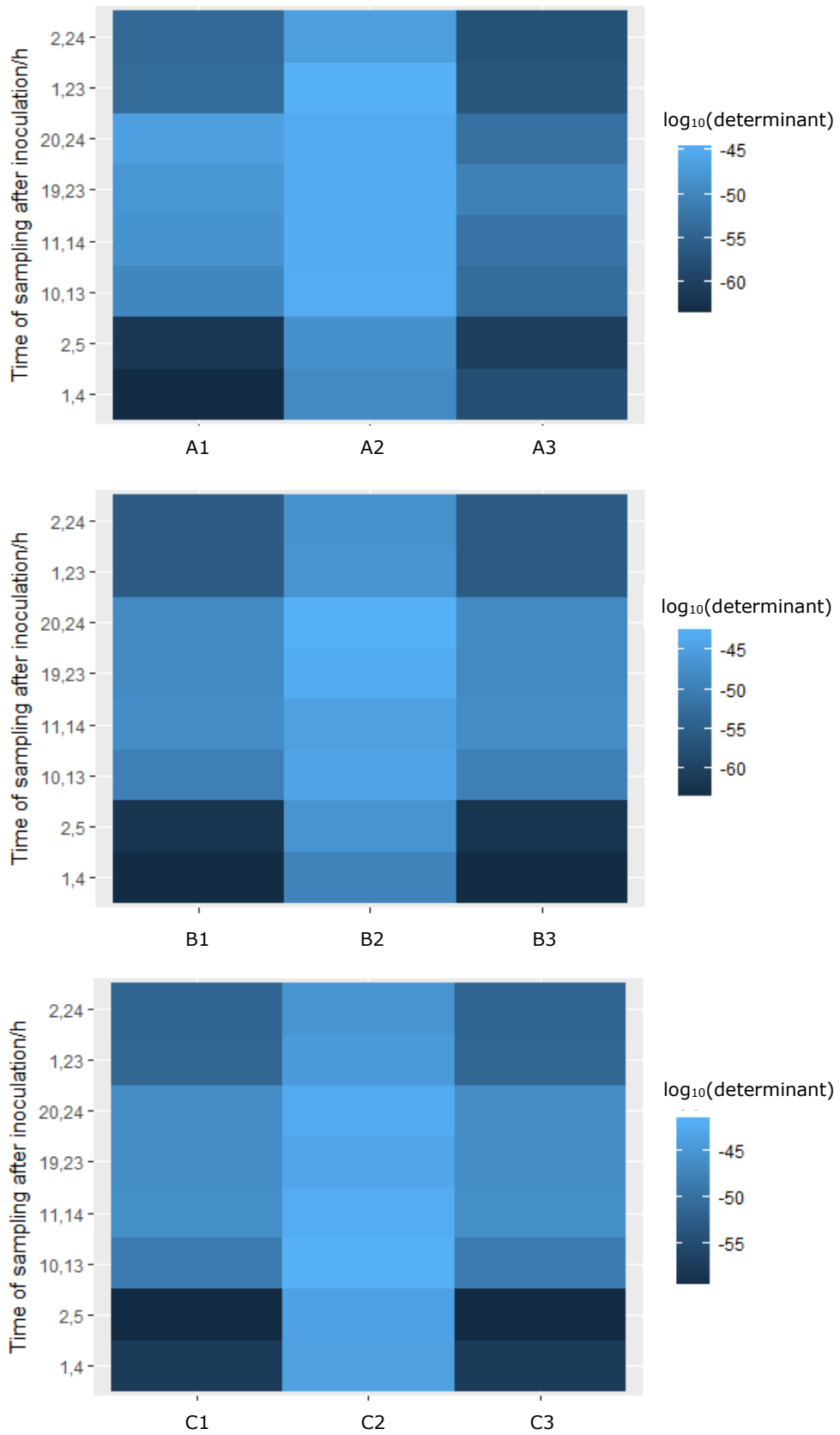
Here, we consider the interaction between the total number of biological replicates (number of animal hosts multiplied by number of ITS per host) and the number of ITS per host, the latter being inversely correlated with the initial number of copies per ITS. Taking the case of a system with exponential growth dynamics and high variance as an example, we consider the experimental designs tabulated in Table 2. Each of A to C include experimental designs with a decreasing number of animal hosts per timepoint. From our previous analysis, the first design in each category A-C is expected to yield the highest precision, as it requires a high number of animal hosts and a high number of initial copies per ITS. In alternative designs 2 and 3 there is a ten-fold decrease in the number of animal hosts, with design #2 having a higher number of biological replicates combined with a low number of copies per ITS and design #3 having a low number of biological replicates combined with a high number of copies per ITS.

Experimental Design Index	Total number of host animals per sampling time point	Total number of ITS per host animal	Mean number of CFU per ITS			Total number of biological replicates (identical simulations)	Total inoculum size
			Compartment 1	Compartment 2	Compartment 3		
A1	200	5	2000	1	1	1000	10000
A2	20	50	200	1	1	1000	10000
A3	20	5	2000	1	1	100	10000
B1	100	5	2000	1	1	500	10000
B2	10	50	200	1	1	500	10000
B3	10	5	2000	1	1	50	10000
C1	50	10	1000	1	1	500	10000
C2	5	100	100	1	1	500	10000
C3	5	10	1000	1	1	50	10000

**Table 2:** Experimental designs to evaluate the relative effects of two design factors (total number of replicates and initial number of copies per ITS) on inference precision

In Figure 8, we comparatively plot the mean  $\log_{10}(\text{determinant})$  for the 3 classes of experimental designs outlined in Table 3, across a range of sampling time points. A three-compartmental model is used (section 2.2) and parameters are drawn at random from distributions specified in section 2.5. While

experimental design index 2 and 3 use the same number of animal hosts, the latter performs significantly better and very close to experimental design index 1, which involves 10 times more animal hosts. Overall, our analysis shows, that given a constraint on the number of animal hosts, higher inference precision is achieved by using experimental infections with large initial numbers of ITS even at lower number of biological replicates; the effect of the initial number of ITS is stronger than that of number of biological replicates in determining inference precision.



**Figure 8:** Heatmap plots showing the mean  $\log_{10}(\text{determinant})$  across 10 parameter sets for experimental designs 1-3 in classes A-C with variable number of biological replicates and initial number of copies per ITS.

#### 4. Discussion

We have presented a novel, computationally efficient approach to information-based experimental design for ITS-based data, with the aim of maximising inference precision for the within-host dynamics of bacterial infections. Inspired by the canonical D-optimality criterion, we chose the determinant of the parameter variance-covariance matrix, obtained via a parametric bootstrap, as a criterion for global parameter inference precision in a likelihood-free inference framework previously developed by Price *et al.* [16]. We used a simulation study to illustrate that the magnitude of the determinant effectively captures the mean width of confidence intervals across parameter sets. As the determinant is a global measure of variance, we correlated it with the mean width of the confidence interval across the entire parameter set (12 parameters in our model) and demonstrated the strong correlation between the value of the determinant and the average width of the confidence interval.

We proceeded to show that the replicate size, initial number of copies per ITS and sampling timepoint selection all have an impact on inference precision in range of general biological cases representative of bacterial dynamics typically observed during experimental infections. Biological scenarios with low replication and killing rates are characterised by low variance in the distribution of copies per [16]; for the entirety of experimental designs considered for these low-variance scenarios, the inference precision is globally higher than for higher-variance scenarios. Furthermore, there is greater redundancy in experimental designs yielding similar levels of high-precision inference for lower-variance scenarios, entailing that prior intuitive knowledge about the dynamics of a system can inform experimental microbiologists about how much focus they should place on identifying unique experimental designs with the highest information content. Finally, we found that for a fixed total number of replicates and assuming inter-host immunological homogeneity, infections with lower numbers of ITS per host at higher initial copies per ITS tend to give higher-level inference precision for both high- and low-variance scenarios.

The benefit of our approach is that it can be used not only to identify a single optimal experimental design, but also to rank experimental designs, thus providing flexibility to experimentalists to choose those that reflect their own priorities. For example, assuming that the reduction in the number of animal hosts is a priority, one can dismiss the globally optimal experimental design that requires using a 10-fold higher number of animal hosts and can choose the design with the second highest precision, which can be achieved by strategically choosing the sampling time points while keeping the number of animal hosts low. Finally, we looked at the interaction between two design factors (total number of replicates and initial copies per ITS) under a constraint on number of animal hosts, which is one of the key priorities in terms of animal welfare, logistical and financial considerations alike. We show that in this scenario, favouring an experimental design with lower number of biological replicates and fewer ITS

per host yields consistently higher inference precision than the alternative allocation of resources with high number of biological replicates and more ITS per host. This consideration becomes particularly relevant as individual ITS detection in a multi-ITS sample has become more advanced, moving away from qPCR with a detection threshold of a <10 ITS per host tissue [11-12] to next generation sequencing able to detect tens or hundreds of ITS per host tissue sample [9, 29].

Discussing the assumptions and limitations of the likelihood-free inference framework used as the basis for the experimental design tool presented here, the moments-based, divergence-minimisation inference framework makes the simplifying assumption that these distributions are multivariate normal across all tissues of interest and sampling time points; this assumption facilitates the derivation of the expressions for the moments. Given that the shape of the true distribution is different for every tissue and sampling time point, and that one would need to approximate each of those individually, the simplifying choice of a multivariate normal distribution seems sensible. Furthermore, the first- and second- order moments of these distributions were used as sufficient summary statistics. Although it could be argued that higher order moments would capture additional features of the distributions, the added complexity is not justified by the resolution of the available data and the previous simplifying assumption about the structure of the distributions. As shown in [16], there is excellent agreement between both the inferences and the width of the confidence intervals obtained when using the maximum-likelihood framework and the likelihood-free framework used here; as a result, the use of summary statistics and the assumption of multivariate normal distributions do not affect the quality of inference. Finally, the estimated precision as identified by the optimal design is valid under the assumption that the model used to find the optimal design is a sufficiently accurate characterisation of the true underlying mechanism.

In terms of the bootstrapped samples, we limited their number to 200 per experimental design considered in this study, striking a balance between reducing inter-simulation variation (Monte Carlo error) and limiting the applicability of the tool due to long running times. Earlier work in [16] shows that even at 100 experiments (or bootstrapped samples), the Gillespie and moments-based approaches both lead to more and more precise estimates as the number of biological replicates increases, while at 500-1000 biological samples, the variance in the parameter estimates converges to 0. Hence, by large, the source of the Monte Carlo error is not the number of bootstrapped samples, but the number of biological replicates, which is one of the design factors considered in this study. Furthermore, our choice of 200 bootstrapped samples was further motivated by a study in [48], who developed a computationally efficient bootstrapping algorithm in the field of phylogenetics and used it to formulate testing criteria for determining the minimum number of bootstrapped samples needed to infer confidence values; in most cases, a size of 100-500 bootstrapped samples was adequate. Finally, we considered at least 10 parameter sets for each experimental design, obtaining consistent results in terms of how the experimental designs ranked relative to each other across the 10 individual datasets; this

consistent pattern further confirms that our choice of bootstrapped sample size is not a major source of Monte Carlo error.

As exploratory research is gaining momentum and the competition for funding is increasing, optimal resource allocation is emerging as a first step in the process of experimental setup conceptualisation [35], while also affecting the welfare of animal hosts which is further reflected in the principle of the 3Rs – replacement, reduction and refinement [17, 36]. *In-vivo* studies in population dynamics microbiology have typically not used experimental design tools until now. For example, despite the sustained wide application of the ITS technique, questions around optimising experimental design and resource allocation have remained formally unaddressed, with the exception of past studies having used small-scale simulation-based approaches to determine a suitable inoculum size [4] or distribution of ITS per host [4, 12]. Limited examples of efforts to develop experimental design tools for population-level data in dynamical systems can be found in other microbiological fields like systems biology [37-42], ecology [43-44, 49], and food science [45-46].

The apparent under-representation of optimal design algorithms in population dynamics microbiology is largely attributable to the lack of computationally efficient inference methods that cater for mathematical models with intractable likelihoods. Such models are often necessary to capture even simplified representations of biological systems which become increasingly more complex, as infections progress [16]. Here, we have provided a blueprint for the use of experimental design tools in complex systems in population dynamics microbiology. We addressed replicate size, sampling time point selection, and the initial number of copies per ITS in the inoculum as design factors commonly considered by experimentalists when trying to optimise multifactorial resource allocation to achieve a desired level of inference precision. In terms of future work, our framework for optimising experimental design can be further extended to serve a wider range of experimentalists' goals beyond maximising inference precision. Other common goals include improving experimental designs to enable model selection, which can aid in assessing competing biological hypotheses [47], as well as minimising the correlations between parameters, which enables a more clear-cut discrimination between often correlated processes, such as bacterial replication and death [4, 7]. In this paper, we focused on improving the overall inference precision and focused on the determinant of the parameter variance-covariance matrix as the utility function of choice. However, since our likelihood-free experimental design tool is based on an approximation of the FIM, it could accommodate other utility functions such as the trace or the eigenvalue of the parameter variance-covariance matrix, which correspond to the A- [50] and E- optimality criteria [51], respectively.



**Data Availability**

Sample code required to produce parameter estimates and compare experimental designs is available at:

<https://github.com/myrtovlazaki/OED>

**Conflict of interest**

The authors have no conflicts of interest to declare.

**Funding**

MV is jointly funded by a Newnham College Major Studentship and a Vergottis Award from the Cambridge Trust. OR is funded by the ALBORADA Trust. OR and DJP acknowledge research Grant BB/M020193 from the BBSRC.

## References

1. World Health Organization. World Health Statistics 2019. [https://www.who.int/gho/publications/world\\_health\\_statistics/2019/EN\\_WHS\\_2019\\_Main.pdf?ua=1](https://www.who.int/gho/publications/world_health_statistics/2019/EN_WHS_2019_Main.pdf?ua=1) (01 October 2019, date last accessed).
2. Hofer, U. The cost of antimicrobial resistance. 2018. *Nature Reviews Microbiology*, 17(1), 3–3. DOI: 10.1038/s41579-018-0125-x
3. Brauner, A., Fridman, O., Gefen, O., & Balaban, N. Q. 2016. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology*, 14(5), 320–330. DOI: 10.1038/nrmicro.2016.34
4. Coward, C., Restif, O., Dybowski, R., Grant, A. J., Maskell, D. J., & Mastroeni, P. 2014. The Effects of Vaccination and Immunity on Bacterial Infection Dynamics In Vivo. *PLoS Pathogens*, 10(9), e1004359. DOI: 10.1371/journal.ppat.1004359
5. Restif, O., Thakar, J., & Harvill, E. T. 2009. Analysis with mathematical models provides insights into infectious diseases. *Microbe*, 4(4), 176-182. DOI: 10.1128/microbe.4.176.1
6. Abel, S., Abel zur Wiesch, P., Davis, B. M., & Waldor, M. K. (2015). Analysis of Bottlenecks in Experimental Models of Infection. *PLoS Pathogens*, 11(6), e1004823. DOI: <https://doi.org/10.1371/journal.ppat.1004823>
7. Vlazaki, M., Huber, J., & Restif, O. 2019. Integrating mathematical models with experimental data to investigate the within-host dynamics of bacterial infections. *Pathogens and Disease*, 77(8). DOI: <https://doi.org/10.1093/femspd/ftaa001>
8. Rossi, O., Vlazaki, M., Kanvatirth, P., Restif, O., & Mastroeni, P. 2020. Within-host spatiotemporal dynamic of systemic salmonellosis: Ways to track infection, reaction to vaccination and antimicrobial treatment. *Journal of microbiological methods*, 176 106008. DOI: 10.1016/j.mimet.2020.106008
9. Abel S, Abel zur Wiesch P, Chang HH, Davis BM, Lipsitch M, Waldor MK. 2015. Sequence tag-based analysis of microbial population dynamics. *Nat Methods*. 12(3):223-226. DOI: 10.1038/nmeth.3253
10. Kaiser, P., Regoes, R. R., Dolowschiak, T., Wotzka, S. Y., Lengefeld, J., Slack, E., Grant, A. J., Ackermann, M., & Hardt, W.-D. 2014. Cecum Lymph Node Dendritic Cells Harbor Slow-Growing Bacteria Phenotypically Tolerant to Antibiotic Treatment. *PLoS Biology*, 12(2), e1001793. DOI: 10.1371/journal.pbio.1001793
11. Grant, A. J., Restif, O., McKinley, T. J., Sheppard, M., Maskell, D. J., & Mastroeni, P. 2008. Modelling within-Host Spatiotemporal Dynamics of Invasive Bacterial Disease. *PLoS Biology*, 6(4), e74. DOI: 10.1371/journal.pbio.0060074
12. Kaiser, P., Slack, E., Grant, A. J., Hardt, W.-D., & Regoes, R. R. 2013. Lymph Node Colonization Dynamics after Oral Salmonella Typhimurium Infection in Mice. *PLoS Pathogens*, 9(9), e1003532. DOI: 10.1371/journal.ppat.1003532
13. Dybowski, R., Restif, O., Goupy, A., Maskell, D. J., Mastroeni, P., & Grant, A. J. 2015. Single passage in mouse organs enhances the survival and spread of Salmonella enterica. *Journal of The Royal Society Interface*, 12(113), 20150702. DOI: 10.1098/rsif.2015.0702

14. Dybowski, R., Restif, O., Price, D. J., & Mastroeni, P. 2017. Inferring within-host bottleneck size: A Bayesian approach. *Journal of Theoretical Biology*, 435, 218–228. DOI: 10.1016/j.jtbi.2017.09.011
15. Vlazaki, M., Rossi, O., Price, D. J., McLean, C., Grant, A. J., Mastroeni, P., & Restif, O. 2020. A data-based mathematical modelling study to quantify the effects of ciprofloxacin and ampicillin on the within-host dynamics of *Salmonella enterica* during treatment and relapse. *Journal of The Royal Society Interface* DOI: 10.1098/rsif.2020.0299
16. Price, D. J., Breuzé, A., Dybowski, R., Mastroeni, P., & Restif, O. 2017. An efficient moments-based inference method for within-host bacterial infection dynamics. *PLOS Computational Biology*, 13(11), e1005841. DOI: 10.1371/journal.pcbi.1005841
17. Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*. London, UK: Universities Federation for Animal Welfare; 1959.
18. Fenwick, N., Griffin, G., & Gauthier, C. 2009. The welfare of animals used in science: how the "Three Rs" ethic guides improvements. *The Canadian veterinary journal = La revue vétérinaire canadienne*, 50(5), 523–530.
19. Melton-Witt, J. A., Rafelski, S. M., Portnoy, D. A., & Bakardjiev, A. I. 2011. Oral Infection with Signature-Tagged *Listeria monocytogenes* Reveals Organ-Specific Growth and Dissemination Routes in Guinea Pigs. *Infection and Immunity*, 80(2), 720–732. DOI: 10.1128/iai.05958-11
20. Schwartz, D. J., Chen, S. L., Hultgren, S. J., & Seed, P. C. 2011. Population Dynamics and Niche Distribution of Uropathogenic *Escherichia coli* during Acute and Chronic Urinary Tract Infection. *Infection and Immunity*, 79(10), 4250–4259. DOI: 10.1128/iai.05339-11
21. Walters, M. S., Lane, M. C., Vigil, P. D., Smith, S. N., Walk, S. T., & Mobley, H. L. T. 2012. Kinetics of Uropathogenic *Escherichia coli* Metapopulation Movement during Urinary Tract Infection. *MBio*, 3(1). DOI: 10.1128/mbio.00303-11
22. Lowe, D. E., Ernst, S. M. C., Zito, C., Ya, J., & Glomski, I. J. 2013. *Bacillus anthracis* Has Two Independent Bottlenecks That Are Dependent on the Portal of Entry in an Intranasal Model of Inhalational Infection. *Infection and Immunity*, 81(12), 4408–4420. DOI: 10.1128/iai.00484-13
23. Gerlini, A., Colomba, L., Furi, L., Braccini, T., Manso, A. S., Pammolli, A., Wang, B., Vivi, A., Tassini, M., van Rooijen, N., Pozzi, G., Ricci, S., Andrew, P. W., Koedel, U., Moxon, E. R., & Oggioni, M. R. (2014). The Role of Host and Microbial Factors in the Pathogenesis of Pneumococcal Bacteraemia Arising from a Single Bacterial Cell Bottleneck. *PLoS Pathogens*, 10(3), e1004026. DOI: 10.1371/journal.ppat.1004026
24. Lim, C. H., Voedisch, S., Wahl, B., Rouf, S. F., Geffers, R., Rhen, M., & Pabst, O. 2014. Independent Bottlenecks Characterize Colonization of Systemic Compartments and Gut Lymphoid Tissue by *Salmonella*. *PLoS Pathogens*, 10(7), e1004270. DOI: 10.1371/journal.ppat.1004270
25. Lam, L. H., & Monack, D. M. (2014). Intraspecies Competition for Niches in the Distal Gut Dictate Transmission during Persistent *Salmonella* Infection. *PLoS Pathogens*, 10(12), e1004527. DOI: 10.1371/journal.ppat.1004527
26. Maier, L., Diard, M., Sellin, M. E., Chouffane, E.-S., Trautwein-Weidner, K., Periaswamy, B., Slack, E., Dolowschiak, T., Stecher, B., Loverdo, C., Regoes, R. R., & Hardt, W.-D. 2014.

- Granulocytes Impose a Tight Bottleneck upon the Gut Luminal Pathogen Population during *Salmonella Typhimurium* Colitis. *PLoS Pathogens*, 10(12), e1004557. DOI: 10.1371/journal.ppat.1004557
27. Rego, R. O. M., Bestor, A., Štefka, J., & Rosa, P. A. 2014. Population Bottlenecks during the Infectious Cycle of the Lyme Disease Spirochete *Borrelia burgdorferi*. *PLoS ONE*, 9(6), e101009. DOI: 10.1371/journal.pone.0101009
  28. Gonzalez, R. J., Weening, E. H., Lane, M. C., & Miller, V. L. 2015. Comparison of Models for Bubonic Plague Reveals Unique Pathogen Adaptations to the Dermis. *Infection and Immunity*, 83(7), 2855–2861. DOI: <https://doi.org/10.1128/iai.00140-15>
  29. Zhang, T., Abel, S., Abel zur Wiesch, P., Sasabe, J., Davis, B. M., Higgins, D. E., & Waldor, M. K. 2017. Deciphering the landscape of host barriers to *Listeria monocytogenes* infection. *Proceedings of the National Academy of Sciences*, 114(24), 6334–6339. DOI: 10.1073/pnas.1702077114
  30. Martin, C. J., Cadena, A. M., Leung, V. W., Lin, P. L., Maiello, P., Hicks, N., Chase, M. R., Flynn, J. L., & Fortune, S. M. 2017. Digitally Barcoding *Mycobacterium tuberculosis* Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *MBio*, 8(3). DOI: 10.1128/mbio.00312-17
  31. Box, G. E. P. 1982. Choice of Response Surface Design and Alphabetic Optimality. *Utilitas Mathematica*, 21B: 11-55.
  32. Chaloner, K., & Verdinelli, I. 1995. Bayesian Experimental Design: A Review. *Statistical Science*, 10(3), 273-304.
  33. Casella, G., & Berger, R. L. 2002. *Statistical inference*. Belmont, CA: Duxbury.
  34. Spall JC. 2005. *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*, volume 65. John Wiley & Sons.
  35. O'Malley MA. 2007. Exploratory experimentation and scientific practice: metagenomics and the proteorhodopsin case. *Hist Philos Life Sci* 29(3):337–360.
  36. Sneddon, L. U., Halsey, L. G., & Bury, N. R. 2017. Considering aspects of the 3Rs principles within experimental animal biology. *The Journal of Experimental Biology*, 220(17), 3007–3016. DOI: 10.1242/jeb.147058
  37. Mdluli, T., Buzzard, G. T., & Rundell, A. E. 2015. Efficient Optimization of Stimuli for Model-Based Design of Experiments to Resolve Dynamical Uncertainty. *PLOS Computational Biology*, 11(9), e1004488. DOI: 10.1371/journal.pcbi.1004488
  38. Sinkoe, A., & Hahn, J. 2017. Optimal Experimental Design for Parameter Estimation of an IL-6 Signaling Model. *Processes*, 5(4), 49. DOI: 10.3390/pr5030049
  39. Braniff, N., Richards, A., & Ingalls, B. (2019). Optimal Experimental Design for a Bistable Gene Regulatory Network. *IFAC-PapersOnLine*, 52(26), 255–261. DOI: 10.1016/j.ifacol.2019.12.267
  40. White, A., Tolman, M., Thames, H. D., Withers, H. R., Mason, K. A., & Transtrum, M. K. 2016. The Limitations of Model-Based Experimental Design and Parameter Estimation in Sloppy Systems. *PLOS Computational Biology*, 12(12), e1005227. DOI: 10.1371/journal.pcbi.1005227

41. Liepe, J., Filippi, S., Komorowski, M., & Stumpf, M. P. H. 2013. Maximizing the Information Content of Experiments in Systems Biology. *PLoS Computational Biology*, 9(1), e1002888. DOI: 10.1371/journal.pcbi.1002888
42. Steiert, B., Raue, A., Timmer, J., & Kreutz, C. 2012. Experimental Design for Parameter Estimation of Gene Regulatory Networks. *PLoS ONE*, 7(7), e40052. DOI: 10.1371/journal.pone.0040052
43. Zhang, J. F., Papanikolaou, N. E., Kypraios, T., & Drovandi, C. C. 2018. Optimal experimental design for predator–prey functional response experiments. *Journal of The Royal Society Interface*, 15(144), 20180186. DOI: 10.1098/rsif.2018.0186Moffat et al., 2020
44. Moffat, H., Hainy, M., Papanikolaou, N. E., & Drovandi, C. 2020. Sequential experimental design for predator–prey functional response experiments. *Journal of The Royal Society Interface*, 17(166), 20200156. DOI: 10.1098/rsif.2020.0156
45. Garre, A., Peñalver-Soto, J. L., Esnoz, A., Iguaz, A., Fernandez, P. S., & Egea, J. A. 2019. On the use of in-silico simulations to support experimental design: A case study in microbial inactivation of foods. *PLOS ONE*, 14(8), e0220683. DOI: 10.1371/journal.pone.0220683
46. Price, D. J., Bean, N. G., Ross, J. V., & Tuke, J. 2018. Designing group dose-response studies in the presence of transmission. *Mathematical Biosciences*, 304, 62–78. DOI: 10.1016/j.mbs.2018.07.007
47. Vanlier, J., Tiemann, C. A., Hilbers, P. A., & van Riel, N. A. 2014. Optimal experiment design for model selection in biochemical networks. *BMC Systems Biology*, 8(1), 20. DOI: 10.1186/1752-0509-8-20.
48. Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E., & Stamatakis, A. 2010. How Many Bootstrap Replicates Are Necessary? *Journal of Computational Biology*, 17(3), 337–354. DOI: 10.1089/cmb.2009.0179
49. Overstall, A., & McGree, J. 2020. Bayesian Design of Experiments for Intractable Likelihood Models Using Coupled Auxiliary Models and Multivariate Emulation. *Bayesian Analysis*, 15(1), 103–131. DOI: 10.1214/19-ba1144
50. Chernoff, Herman. Locally Optimal Designs for Estimating Parameters. *Ann. Math. Statist.* 24. 1953. no. 4, 586--602. DOI: 10.1214/aoms/1177728915.
51. Ehrenfeld, E. 1955. On the efficiency of experimental design, *Annals of Math. Stat.*, 26247–255.