Identifying cell-to-cell variability in internalisation using flow cytometry

Alexander P Browning[*1,2,3], Niloufar Ansari[*4], Christopher Drovandi[1,2,3], Angus Johnston[4], Matthew J Simpson[1,3], and Adrianne L Jenner[†1,3]

[1] *School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia*
[2] *ARC Centre of Excellence for Mathematical and Statistical Frontiers, QUT, Australia*
[3] *QUT Centre for Data Science, QUT, Australia*
[4] *Monash Institute of Pharmaceutical Sciences, Monash University, Australia*

November 25, 2021

Summary

Biological heterogeneity is a primary contributor to the variation observed in experiments that probe dynamical processes, such as internalisation. Given that internalisation is the primary means by which cells absorb drugs, viruses and other material, quantifying cell-to-cell variability in internalisation is of high biological interest. Yet, it is common for studies of internalisation to neglect cell-to-cell variability. We develop a simple mathematical model of internalisation that captures the dynamical behaviour, cell-to-cell variation, and extrinsic noise introduced by flow cytometry. We calibrate our model through a novel distribution-matching approximate Bayesian computation algorithm to flow cytometry data collected from an experiment that probes the internalisation of antibody by transferrin receptors in C1R cells. Our model reproduces experimental observations, identifies cell-to-cell variability in the internalisation and recycling rates, and, importantly, provides information relating to inferential uncertainty. Given that our approach is agnostic to sample size and signal-to-noise ratio, our modelling framework is broadly applicable to identify biological variability in single-cell data from experiments that probe a range of dynamical processes.

Keywords

heterogeneity; flow cytometry; endocytosis; internalisation; drug delivery; noise; mathematical model; approximate Bayesian computation

[*] These authors contributed equally.
[†] Corresponding author: adrianne.jenner@qut.edu.au

1 Introduction

Internalisation is the primary means by which cells takeup drugs, viruses, and nanoparticles [1–5]. Single-cell *in vitro* analysis of internalisation and similar dynamical processes reveals significant cell-to-cell variability in otherwise isogenic cell populations [6–12]. Such heterogeneity is ubiquitous to biology and essential to life, allowing for robust decision making, development, and adaptation of cell populations to environmental uncertainty [13–17]. From a clinical perspective, heterogeneity in drug uptake and response is considered a leading contributor to treatment inefficiencies and resistance [18, 19]. The challenges of working with potentially noisy data that often obfuscates biological variability means that it is relatively common for quantitative analysis of internalisation to neglect heterogeneity [20, 21]. Exacerbating these issues is a corresponding lack of mathematical tools that account for cell-to-cell variability and measurement noise while also providing information about the uncertainty in inferences and predictions drawn from noisy data.

Modern analysis technologies, including flow cytometry, allow the high-throughput collection of data from experiments that probe internalisation at rates exceeding a thousand cells per second (fig. 1) [22]. In a typical internalisation assay, material tagged with fluorescent probes are incubated with cells and internalised through endocytosis pathway (fig. 1a–b) [23, 24]. The fluorescence of probes can be inhibited by introducing a quencher dye, or altered due to the lower pH in early endosomes [20, 23], providing quantitative information relating to the amount of material internalised. Flow cytometry provides noisy measurements related to the total and internalised amount of material at various time points (fig. 1c–d). Since cells cannot otherwise be tracked between observation times, data comprise single-cell snapshots and individual trajectories are not available. While measurement noise introduced by the flow cytometry electronics and background autofluorescence are not insignificant sources of variation; variability in the data is primarily biological [11, 25–28]. We demonstrate this by performing an internalisation assay with a dual-labelled fluorescent probe, finding that measurements are highly correlated, suggestive of a shared source of variability.

Mathematical and statistical techniques allow quantitative analysis of transient dynamics, heterogeneity, and measurement noise. As the number of molecules internalised by each cell is relatively large, single-cell trajectories describing the relative amount of material internalised can be accurately described by deterministic models derived through kinetic rate equations. Ordinary differential equation (ODE) constrained Bayesian hierarchical and random effects models incorporate cell-to-cell variability through a parameter hierarchy where distributions parameterised by hyperparameters describe cell-level properties [29–31]. Both individual cell properties and hyperparameters are estimated during calibration of hierarchical models to data, presenting a significant computational challenge for the large sample sizes provided by flow cytometry data. In the mathematical literature, so-called heterogeneous [32] or random ODEs [33] and populations of models [34] make similar assumptions, often without assuming a parametric distribution of cell properties [9, 35, 36]. Issues presented by large sample sizes can be avoided by calibrating models using the empirical distribution of the data (through, for example, kernel density estimates) [32], an approach that provides point estimates but neglects inferential uncertainty and poses a challenge when the signal-to-noise ratio in the data is not sufficiently

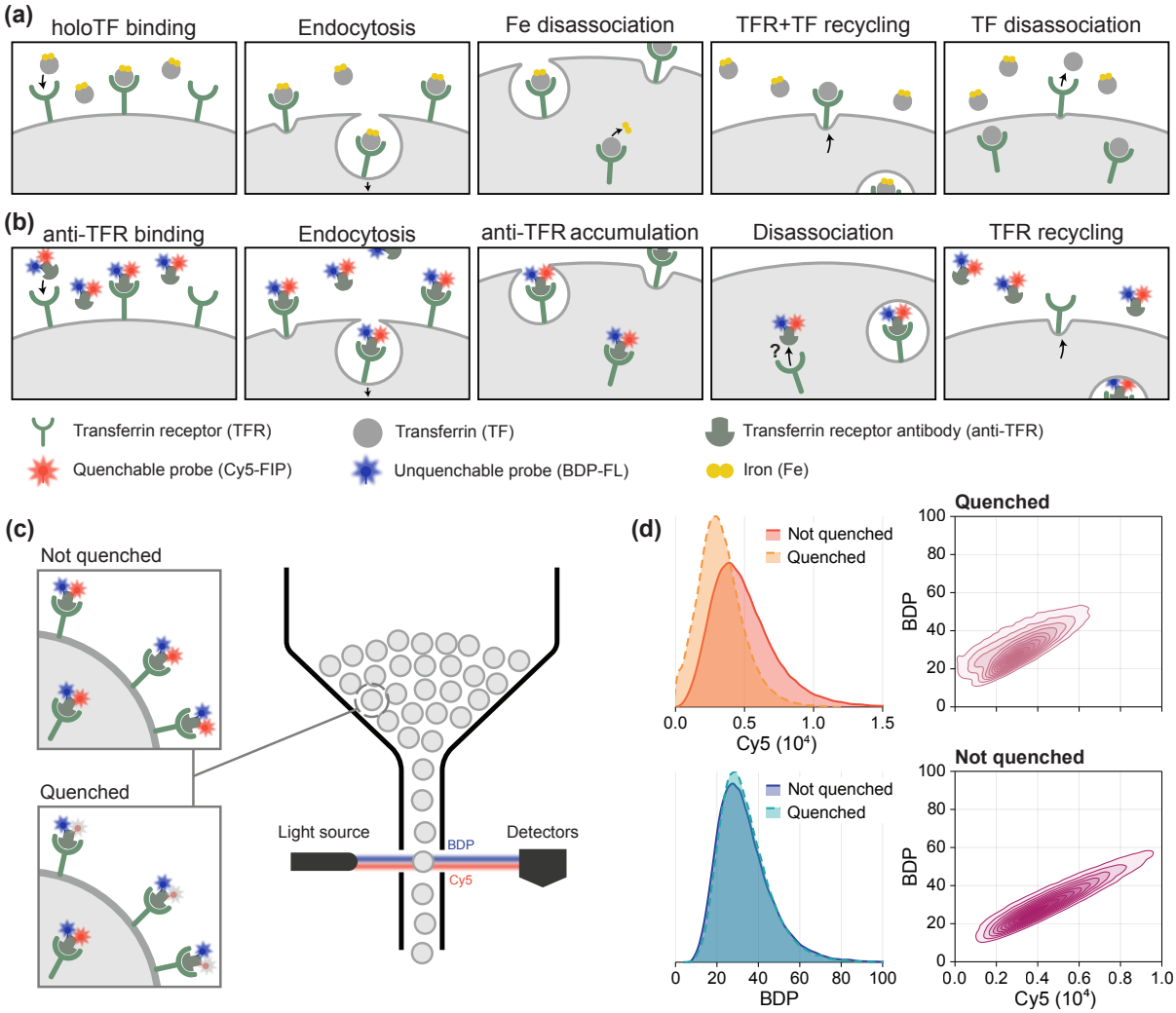

Figure 1. Internalisation dynamics and corresponding experimental assay. (a) Internalisation of transferrin, a protein responsible for the uptake of iron by cells. Iron-saturated transferrin binds to receptors on the cell surface that are internalised through *endocytosis*. Inside the cell, iron disassociates and the transferrin-receptor complex recycles to the cell surface. (b) A corresponding internalisation assay. Anti-transferrin receptor antibody (anti-TFR) labelled with a dual fluorescent probe replaces iron-loaded transferrin and is internalised through the usual pathway. Experimental observation suggest that a small proportion of labelled antibody disassociates inside the cell, allowing receptor recycling and the accumulation of antibody inside the cell. (c) A quencher dye inhibits fluorescence from the probe labelled with a Cy5 (red) fluorophore, providing information relating to the proportion of antibody that has internalised. Single-cell measurements of fluorescence from both probes is measured using flow cytometry. (d) Data obtained $t = 10\,\text{min}$ after antibody are introduced. Since variability in the data is predominantly biological, data from each probe are highly correlated.

high.

In this study, we develop a mathematical model of internalisation that captures cell-to-cell variability by describing cell properties—specifically, the number of receptors, the internalisation rate, and the recycling rate of each cell—as jointly distributed random variables. To describe non-biological sources of variability from flow cytometry measurements of an internalisation assay, we couple the dynamical model to a probabilistic observation process that captures autofluorescence and measurement noise. We take a Bayesian approach to parameter estimation and develop a novel approximate Bayesian computation (ABC) [37–39] algorithm that matches distributional information from flow cytometry measurements. This approach is agnostic to sample size and the signal-to-noise ratio and provides point parameter estimates and information relating to inferential uncertainty.

We demonstrate our approach by studying heterogeneity in the internalisation of anti-transferrin receptor (anti-TFR) antibody in C1R cells, a human B lymphoblastoid line. Data comprise potentially noisy flow cytometry measurements from an internalisation assay developed in our previous work, specific hybridisation internalisation probe (SHIP) (fig. 1b–c) [20,40]. Measurements are collected from anti-TFR antibody dual labelled with BDP-FL and FIP-Cy5 fluorophores. We take measurements both with and without a quencher due, which inhibits fluorescence of FIP-Cy5 without affecting the BDP-FL signal. Therefore, we obtain jointly distributed data that comprises noisy measurements of the total and internalised amount of antibody in each cell (fig. 1d). Snapshots are collected from samples that are incubated with antibody saturated medium for various periods of time to provide noisy measurements of both the total and internalised amounts of antibody present on each cell. We are able to identify key sources of biological variability and provide predictions that give insight into how the uptake of material varies between cells. Importantly, our approach to parameter inference enables us to quantify the uncertainty in inferences made, allowing us to provide experimental design guidance.

2 Results

Dynamical model of antibody internalisation

We describe the internalisation of antibody and the recycling of receptors using a compartment model. Given the concentration of antibody in the surrounding medium is sufficiently high, we assume that the association rate of antibody to free receptors on the cell surface is much higher than the kinetic rates of internalisation and recycling (supplementary material). Therefore, we describe the number of antibody-receptor complexes on the cell surface, S, and that endocytosed, E. Before incubation in antibody-saturated medium, T endocytosed receptors are bound to transferrin. In the experiments, we observe that the total concentration of antibody continues to increase, consistent with receptor recycling and antibody accumulation inside the cell which we capture by describing the number of free endocytosed antibody, F. These assumptions give

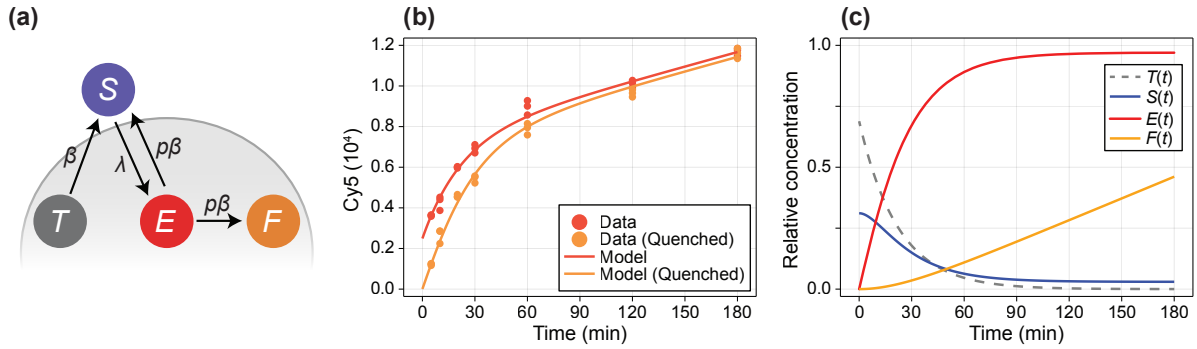

Figure 2. Dynamical model of internalisation and recycling matches experimental data. (a) The dynamical model describes the relative concentration of internal, transferrin bound receptors, T (grey); surface antibody-bound receptors, S (blue); internal antibody bound receptors, E (red); and internal free antibody, F (orange). (b) MFI of Cy5 fluorescence measurements for samples that are not quenched (red) and those that are (orange) at various time points. The dynamical model is calibrated using maximum likelihood estimation, with the solution shown (solid curve). (c) Solution to the mathematical model (Methods, eq. (M16)) at the maximum likelihood estimate (table 1).

rise to the dynamical model (fig. 2a)

$$\begin{aligned} T &\xrightarrow{\beta} S, \\ S &\xrightarrow{\lambda} E \xrightarrow{p\beta} S + F, \end{aligned} \tag{1}$$

where β [min^{-1}] is the recycling rate, λ [min^{-1}] is the internalisation rate, and p is the probability that an endocytosed antibody disassociates allowing receptor recycling.

Given that the number of receptors in each cell is relatively large, eq. (1) can be formulated as a linear ordinary differential equation with exact solution $\mathbf{x}(t) = \mathbf{x}_0 \exp(\mathbf{M}t)$, where $\mathbf{M}$ is a matrix of coefficients and $\mathbf{x}(t)$ denotes the number of molecules in each compartment (Methods M.2.2). Initially the system is in equilibrium, so

$$S(0) = \frac{\beta}{\lambda + \beta}, \tag{2}$$

where molecule counts are taken with respect to the total number of receptors on the cell, denoted R, so $S(0) + T(0) = 1$.

Inference using mean fluorescence intensity measurements

Flow cytometry measurements are typically summarised using mean fluorescent intensity (MFI); point statistics equivalent to the geometric mean of the fluorescence distribution (fig. 2b). Cy5 MFI measurements from samples that are not quenched are related to the total amount of antibody in the sample, $A(t) = S(t) + E(t) + F(t)$, and measurements from quenched samples are related to the amount of internal antibody in the sample, $I(t) = E(t) + F(t)$. Assuming incomplete quenching so the probability that a Cy5-labelled probe is quenched is $\eta \approx 0.94$ (supplementary material) and average autofluorescence E_Q, MFI measurements can be modelled

by

$$
\begin{aligned}
Q_{\mathrm{MFI}}(t) &= \alpha_1 A(t) + E_Q + \kappa_1, \\
\overline{Q}_{\mathrm{MFI}}(t) &= \alpha_1 \Big[I(t) + \eta S(t) \Big] + E_Q + \kappa_2.
\end{aligned} \tag{3}
$$

Here, we denote by $Q_{\mathrm{MFI}}(t)$ MFI measurements from the quenchable probe in the samples that are not quenched, and by $\overline{Q}_{\mathrm{MFI}}(t)$ that of quenched samples. We capture variability in MFI measurements between experiments by assuming measurement error $\kappa_1, \kappa_2 \sim \mathrm{Normal}(0, \sigma^2)$. We refer to eq. (3) as the *homogeneous model* since the dynamical parameters λ and β, and the number of receptors, R, does not vary cell-to-cell and are fixed for the population.

To assess the suitability of the dynamical model and provide a baseline to assess our model that captures biological heterogeneity, we calibrate eq. (3) to experimental data using maximum likelihood estimation. We tabulate estimates and confidence intervals approximated using the observed Fisher information in table 1, and show the model best-fit in fig. 2b.

The homogeneous model provides a fit that qualitatively matches MFI measurements from the experimental data, and all parameters are identifiable within relatively precise intervals (fig. 2b). Estimates for the internalisation and recycling rates suggest that a proportion of approximately

$$
S(0) = \frac{\beta}{\lambda + \beta} \approx 0.31 \tag{4}
$$

of transferrin receptors lie on the surface at equilibrium. Estimates for p suggest that 6.8% (95% CI (6.3%,7.2%)) of internalised antibody disassociates, allowing receptor recycling. This is also evident from simple observations of the experimental data, since the fluorescent intensity increases throughout the experiment, suggesting that a small proportion of receptors remain on the surface while antibody accumulates inside the cell fig. 2c.

Incorporating biological variability into dynamical model of internalisation

We assume biological variability arises through both physical and physiological differences between cells in the population. Specifically, we allow number of receptors, R, and dynamical parameters λ and η to vary cell-to-cell. Without loss of generality, we set $\mathbb{E}(R) = 1$ so receptor and antibody counts are taken with respect to the average receptor count in the population.

The properties of the ith cell are given by the random variable $\boldsymbol{\xi}_i = (R_i, \lambda_i, \beta_i)$. We assume unimodality and a parametric description of $\boldsymbol{\xi}_i$ that allows us to estimate the first three moments

Table 1. Parameter estimates and approximate confidence intervals for the deterministic model. Approximate confidence intervals are calculated using the observed Fisher information matrix, calculated from the Hessian of the log-likelihood function [41].

Parameter	Estimate	95% CI	Units
λ	0.106	(0.097,0.116)	min^{-1}
β	0.047	(0.043,0.051)	min^{-1}
p	0.068	(0.063,0.072)	–
α_1	7840	(7540,8140)	Fluorescence units

and dependence structure. Marginals are given by

$$\begin{aligned} R_i &\sim \text{ShiftedLogNormal}(\mu_R, \sigma_R), \\ \lambda_i &\sim \text{ShiftedGamma}(\mu_\lambda, \sigma_\lambda, \omega_\lambda), \\ \beta_i &\sim \text{ShiftedGamma}(\mu_\beta, \sigma_\beta, \omega_\beta). \end{aligned} \tag{5}$$

where λ_i and β_i are shifted Gamma variables parameterised in terms of their respective means, standard deviations and skewness (**??**). The number of receptors is assumed to be shifted log-normally distributed [42]. To ensure positivity, we truncate $\boldsymbol{\xi}_i$. We model the dependence structure of $\boldsymbol{\xi}_i$ with a Gaussian copula parameterised by the correlation matrix

$$\mathbf{P} = \begin{pmatrix} 1 & \rho_{R,\lambda} & \rho_{R,\beta}, \\ \rho_{R,\lambda} & 1 & \rho_{\lambda,\beta}, \\ \rho_{R,\beta} & \rho_{\lambda,\beta} & 1 \end{pmatrix}. \tag{6}$$

To ensure $\mathbf{P}$ remain positive definite, we infer $\rho_{R,\lambda}$, $\rho_{\lambda,\beta}$ and $\tilde{\rho}_{R,\beta}$ (all constrained to the interval $(-1, 1)$) where

$$\rho_{R,\beta} = \rho_{R,\lambda}\rho_{\lambda,\beta} + \tilde{\rho}_{R,\beta}\sqrt{(1-\rho_{R,\lambda}^2)(1-\rho_{\lambda,\beta}^2)}. \tag{7}$$

Therefore, $\rho_{R,\lambda}$ (and similar for $\rho_{R,\beta}$ and $\rho_{\beta,\lambda}$) describes the strength of the correlation between the number of receptors, R, and internalisation rate, λ.

The *heterogeneous model* is a random ODE model where $\mathbf{x}(t)$ and its constituents are random variables [33]. For example, $A(t)$ is a random variable describing the distribution of bound-antibody between cells.

Statistical model for noisy flow cytometry data

Measurement noise in flow cytometry is primarily attributable to shot noise introduced from the photomultiplier tubes (PMT noise) that convert the photon signal to an amplified, analogue electrical signal. Recent studies suggest that the square coefficient of variation of such noise is approximately constant [28], so we model shot noise with uncorrelated white noise (i.e. Gaussian), with variance proportional to the true signal. The second source of noise is cellular autofluorescence, where the laser used to excite the labelled antibody can excite other molecules in the cell, leading to a background autofluorescence where signal is present in the absence of antibody. We build an empirical distribution of autofluorescence (E_Q, E_U) using a sample where cells have not been introduced to labelled antibody (supplementary material).

We interpret measurements from the Cy5-labelled probe, which is quenchable and denoted by $Q(t)$, and the BDP-labelled probe, which is not quenchable and is denoted by $U(t)$. Therefore,

$$Q(t) = \overbrace{\alpha_1 A(t) R} \quad + \quad \overbrace{\varepsilon_1 \sqrt{\alpha_1 A(t) R}} \quad + \quad \overbrace{E_Q}, \tag{8}$$

$$U(t) = \underbrace{\alpha_2 A(t) R}_{\text{Antibody}} \quad + \quad \underbrace{\varepsilon_2 \sqrt{\alpha_2 A(t) R}}_{\text{PMT noise}} \quad + \quad \underbrace{E_U}_{\text{Autofluorescence}}, \tag{9}$$

where $\varepsilon_1, \varepsilon_2 \sim \text{Normal}(0, \sigma_k^2)$. Similarly, the measurements from quenched samples are given by

$$\overline{Q}(t) = \alpha_1 \Big[I(t) + (1-\eta) S(t) \Big] R \quad + \tag{10}$$
$$\varepsilon_1 \sqrt{\alpha_1 \big[I(t) + (1-\eta) S(t) \big] R} + E_Q,$$
$$\overline{U}(t) = \alpha_2 A(t) R + \varepsilon_2 \sqrt{\alpha_2 A(t) R} + E_U. \tag{11}$$

Calibration and uncertainty quantification

We take a Bayesian approach to parameter estimation, calibrating the noisy heterogeneous model to SHIP assay data using a novel approximate Bayesian computation (ABC) algorithm that matches the empirical distribution of flow cytometry measurements, under the assumption that measurements from each probe are linearly correlated (fig. 3a–b). Model simulations based on $n = 1000$ cells per observation time, per condition (quenched or not quenched) are compared to experimental data using the Anderson-Darling distance [43] and Pearson correlation. Before observation of data, knowledge about the parameters are encoded in prior distributions, chosen to be independent and uniform with bounds given as axis limits in fig. 3. Full details are given in Methods.

First, we employ a particle filter based on sequential Monte Carlo (SMC) [44] to identify the region of the 16-dimensional parameter space that produces model realisations that lie close to the data and establish a threshold below which model simulations are deemed sufficiently close to the experimental data. We then apply a Markov chain Monte Carlo (MCMC) algorithm to explore the support of the posterior distribution, interpreted as the region of the parameter space that produces model simulations that match experimental data, quantifying uncertainty in parameter estimates. In fig. 3 we plot posterior samples from four independent tuned MCMC chains thinned to a total of 400,000 samples, providing an effective sample size of at least 1,000 per parameter. Further MCMC diagnostics are provided as supplementary material. To visualise model predictions, we compute a point estimate by further thinning the chains to a total of 400 samples, and identifying the parameter set that produces the lowest average discrepancy from 100 model realisations. Model predictions at the point estimate are shown alongside experimental data in fig. 4.

Heterogenous model captures biological variability

The heterogeneous model produces realisations with excellent agreement to flow cytometry measurements, matching both marginal and jointly distributed data from both probes (fig. 4). Minor discrepancies in univariate distributions highlight the main sources of unaccounted error; for example, error relating to the precise time at which internalisation is ceased by fixation and error relating to flow cytometry gating.

Samples relating to the skewness of the internalisation and recycling rate distributions, ω_λ and ω_β, respectively, (fig. 3c,f) show that information in the experimental data is insufficient to identify the shape of the internalisation and recycling rate distributions. While the precision to which we can identify the variance of each rate, σ_λ and σ_β, (fig. 3b,e) is limited, it is clear that $\sigma_\lambda > 0.057$ (lower bound on a 95% CrI), providing evidence to suggest heterogeneity

in the internalisation rate. While p has the same interpretation between the heterogeneous and deterministic models, the estimates from the heterogeneous model, $p = 4.7\%$ (95% CrI (3.0%,6.5%)) are lower with a greater amount of uncertainty than in the homogeneous model.

In fig. 5 we plot the inferred distributions of R, λ and β. To visualise uncertainty in estimates of these distributions, we show a 95% credible internal (CrI) for the univariate probability density functions by resampling from the posterior distribution. Compared to distributions of the dynamical parameters λ and β, the distribution of the relative receptor count, R, is identified with much greater precision (fig. 5a). R does not feature in the dynamical model and is, therefore, less sensitive to issues relating to model misspecification. While results in fig. 3j–l show relatively large uncertainty in the correlations between parameters, it appears likely that the receptor count and recycling rate are negatively correlated (83% of posterior samples have $\rho_{R\beta} < 0$). Parameters identified in the homogeneous model based on MFI measurements are contained within high density regions of the inferred distributions in the heterogeneous model. This is also the case when estimates are compared to bivariate distributions in fig. 5d–f; however, the interpretation of the deterministic model parameters in the context of data with significant heterogeneity is unclear, highlighting the importance of modelling biological variability when interpreting flow cytometry data of dynamical processes like internalisation.

The data appears insufficient to distinguish between PMT noise from each probe. Initial examination of estimates for the relative magnitude of the noise from the quenchable (Cy5) and unquenchable (BDP) probe signals in fig. 3o–p suggests that the no-noise model may be appropriate, lending the study to analysis of models that assume negligible noise [32]. However, the joint distribution of σ_1 and σ_2 (fig. 3s) reveals an elliptical region, suggesting that the model requires PMT noise in the signal from *at least one* probe. Similar phenomena are observed in error-in-variables or total-least-squares problems, where errors are introduced in both independent and dependent variables, and only the ratio of the error variances is identifiable [45].

Model predicts unobservable measurements

A primary goal of flow cytometry analysis is to quantify the amount of fluorescent material present in a sample. In the context of an internalisation assay, we are interested in the *proportion* of material internalised through time. By accounting for variability introduced through receptor count, PMT noise and autofluorescence, we are able to better quantify the amount, or proportion internalised, of antibody compared with standard approaches.

Since it is not possible to collect noise-free data relating to the joint distribution of $I(t)$, provided from quenched samples, and $A(t)$, provided from sampled that are not quenched, statistics such as the proportion of antibody internalised by each cell cannot be directly measured. Rather, such statistics are typically estimated as

$$I_{\text{frac}}(t) = \frac{I_{\text{MFI}}(t)}{A_{\text{MFI}}(t)}, \tag{12}$$

where $I_{\text{MFI}}(t)$ and $A_{\text{MFI}}(t)$ are scalar estimates of the average proportion of internal and total antibody estimated using MFI [7]. Using our calibrated heterogeneous model, we can predict the *distribution* of material internalised through time by simulating the model with sources of

noise removed. In fig. 6a, we show the time-evolution of the distribution of $I(t)/A(t)$ at the model-best-fit, along with the equivalent prediction from the homogeneous model. In fig. 6b, we repeat this exercise for the total relative amount of antibody internalised, $I(t)$. To understand uncertainty in these distributions, in fig. 6c we show the time evolution of the distribution of $I(t)/A(t)$ alongside credible intervals formed by resampling parameters from the posterior distribution.

While our analysis revealed that several parameters are non-identifiable, or cannot be constrained to a relative precise interval, we are still able to produce relatively precise predictions of statistics such as the proportion of material internalised. Results in fig. 6c show a discrepancy between predictions from the homogeneous and heterogeneous models. Aside from very early time, when the distribution of material internalised is relatively wide, the deterministic model predictions lie within the lower tail of the predicted distribution. This is consistent with the discrepancy we observe in estimates of p between models: the heterogeneous model predicts that antibody disassociation and receptor recycling is rarer that what is predicted by the deterministic model. This results in a smaller proportion of surface-bound antibody at late time.

Using the inferred joint distribution of λ and β we can build a picture of the proportion of receptors present on the cell surface at equilibrium, $S(0)$ (eq. (4)). In fig. 6d, we show the inferred distribution of $S(0)$ at the model-best-fit, along with the uncertainty associated with the estimate and that predicted by the deterministic model. While we have not precisely estimated this distribution, it is clear that, on average, a smaller number of receptors are present on the cell surface than not, in agreement with the prediction of the deterministic model of 31%. We also see that the inferred distribution is highly variable; at the best-fit, for example, non-zero density at zero suggests that some cells have a very small proportion of receptors on the surface, perhaps due to inhibition of recycling. In fig. 6e, we show the inferred relationship between $S(0)$ and R at the best-fit. This result suggests that cells with a larger number of receptors—which may correlate to cells in latter stages of the cell cycle—have fewer surface-bound receptors.

3 Discussion

Heterogeneity in critical cell processes, such as the internalisation of material, is ubiquitous yet poorly understood. Paired with experimental protocols that probe these processes, flow cytometry is capable of generating vast quantities of single-cell-snapshot data that captures cell-to-cell variability. Often, such data are summarised with point statistics that provide information about the transient behaviour to the detriment of acknowledging variability between otherwise isogenic cells. In this study, we develop a mathematical model of internalisation that captures dynamical behaviour, biological variability, and measurement noise of arbitrary magnitude. We apply our model to identify key sources of biological variability in the internalisation of antibody by C1R cells.

While computationally costly, our distribution-matching ABC approach to inference carries several advantages over likelihood-based approaches; for example, those based on Bayesian hierarchical models or those that model cell properties as a finite mixture [36]. First, ABC is

robust to model error, incorporating uncertainty due to factors that are not explicitly modelled [46]. This might include the relatively small discrepancies we observe in fig. 4 that highlight potential model-misspecification as well as error introduced experimentally, such as the precise measurement time, i.e., the time at which internalisation ceases due to fixation.

Secondly, the distribution-matching approach allows the interpretation of pre-processed or summarised data. Automatic clustering algorithms [47–49] are fast replacing manual gating, providing an opportunity to analyse the parametric mixture distributions identified algorithmically, rather than relying on accurate classification of individual data points to perform analysis on the underlying data. Lastly, our approach is agnostic to both the size of the dataset and the complexity of the underlying measurement model. While the signal-to-noise ratio in our data is high (demonstrated by the amount of variability we identify as biological in origin), this is not always the case. In particular, flow cytometry measurements are often corrupted by autofluorescence and bleed-through from overlapping emission spectra. In our framework, both sources of extrinsic variability can be built into the probabilistic observation process or accounted for using pre-processing software where the compensated distributions are analysed rather than the underlying data.

Working with single-cell snapshot data provides limited information compared to single-cell trajectory data, potentially explaining why inferences relating to heterogeneity in dynamical parameters are relatively imprecise (fig. 5). In particular, the data provides no information about the joint distribution of antibody concentration between observation times. Additional results in fig. 7 illustrate the predicted dependence in internalised antibody concentration between early (10 min) and late (120 min) observation times, denoted by $Q(10)$ and $Q(120)$, respectively. An interpretation of model predictions with higher fitted correlations between $Q(10)$ and $Q(120)$ is that single-cell trajectories remain ordered: cells with a relatively lower proportion of antibody internalised at $t = 10$ min retain a relatively lower proportion at $t = 120$ min. Therefore, it is unclear from the available data (fig. 7b) whether cell trajectories remain ordered or whether cells can "catch up"; that is, whether cells that are initially slow to internalise material end up with a large amount internalised at later time points. Intuitively, assuming that such a correlation is strong (i.e., trajectories remain ordered) strongly impacts inferences. Results in fig. 7c–d show that making such an assumption narrows uncertainty in the distribution of recycling rates to distributions where cells that do not recycle (i.e., $\beta_i = 0$) are rare.

Aside from stochastic variations between isogenic cells—due to gene expression [50], for example—variability in internalisation is at least partially driven by the cell cycle [51, 52]. Therefore, we might expect lower internalisation and recycling rates in cells preparing to mitose which, therefore, have a larger number of receptors. This is also suggested by results relating to the best-fit in fig. 5, which show that the internalisation and recycling rates decrease with the number of receptors. These results raise the possibility of a non-Gaussian dependence between dynamical rates that our model cannot capture. For example, the dependence between R and λ may not be monotonic: internalisation by cells in very late stages of the cell cycle might be inhibited, whereas in general, larger cells may internalise material more quickly [53]. Distribution-free approaches [36] might better capture the dependence structures in these cases. However, given that our model is already able to match the experimental data, adding com-

plexity will exacerbate parameter non-identifiability. Therefore, further work should focus on experimental design [54]; by inhibiting recycling, or pre-sorting cells to remove variability in R, for example.

Our analysis demonstrates that inferences drawn using approaches that neglect heterogeneity can be misleading. In particular, the interpretation of predictions and parameter estimates from the homogeneous model is mathematical unclear. Generally, realisations of the homogeneous model do not represent the mean of realisations of the heterogeneous model, nor do they represent realisations where parameters in the heterogeneous model are first averaged [33]. While, in our case, parameters identified by the homogeneous model are contained within the distribution identified by the heterogeneous model, the homogeneous model produces biased predictions that are not representative of the entire population (fig. 6). These findings highlight a need to co-develop mathematical tools that account for biological variability in analysis of single-cell data.

A better understanding of heterogeneity in internalisation has important implications for drug delivery [5, 19], in addition to our understanding of pathological processes, such as the internalisation of viruses [55, 56]. In this study, we develop a novel quantitative model that captures biological variability in internalisation using arbitrarily noisy flow cytometry data. In contrast to typical approaches, we can produce predictions that give insight into the variability in material internalised while accounting for inferential uncertainty. Applying quantitative techniques that capture biological variability allows practitioners to get the most out of the vast amounts of single-cell data generated by flow cytometry and other modern experimental tools.

Data availability

Code and data are available on GitHub at `github.com/ap-browning/internalisation`.

Funding

A.P.B. is supported by the the ARC Centre of Excellence for Mathematical and Statistical Frontiers (CE140100049) research SPRINT scheme. M.J.S. is supported by the Australian Research Council (DP200100177).

Author Contributions

All authors interpreted the results and conceived the study. A.P.B. performed the mathematical analysis and implemented computational algorithms. N.A. performed the experiments. A.P.B. and N.A. drafted the manuscript. All authors provided comments and gave approval for publication.

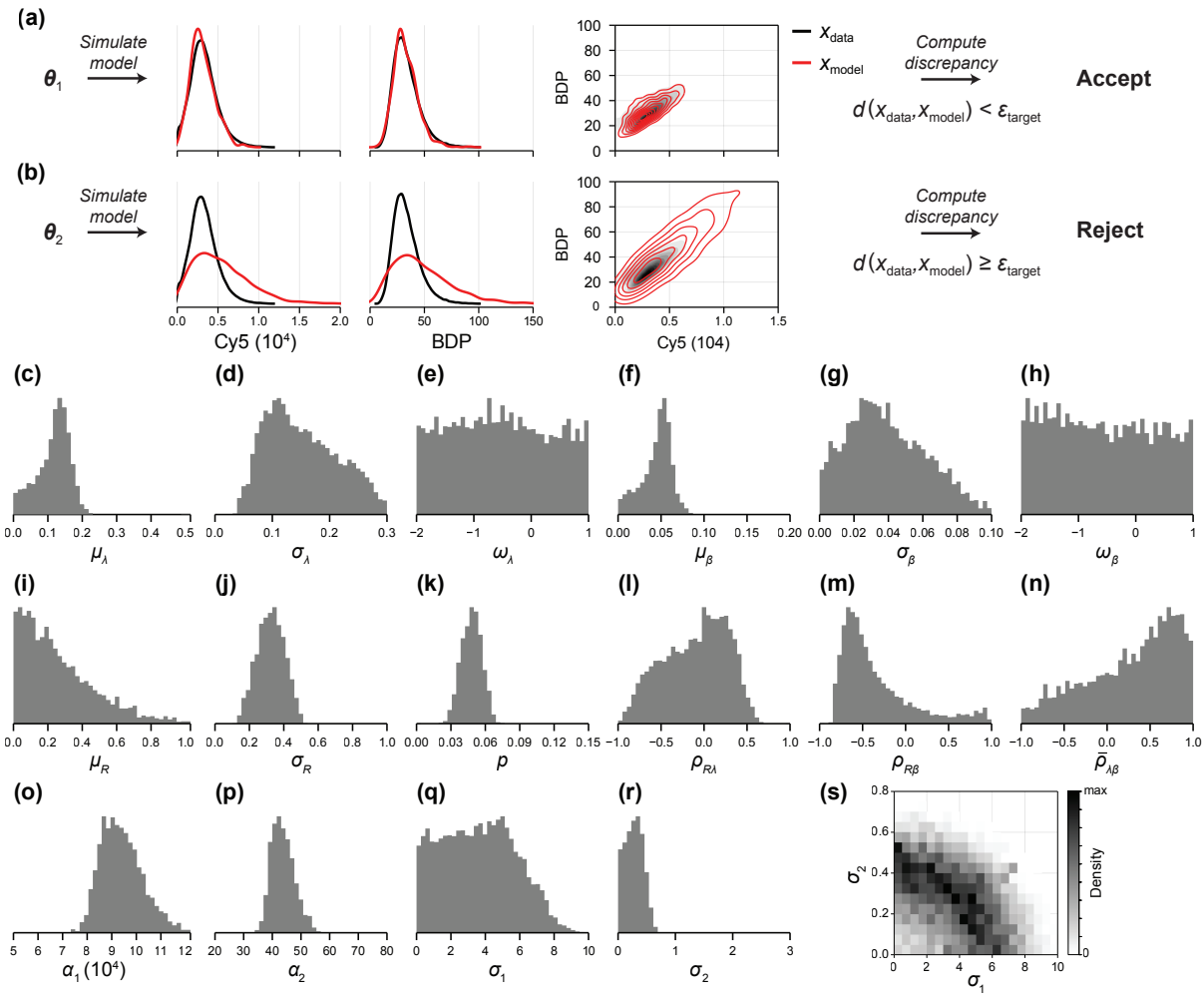

Figure 3. Model calibration and uncertainty quantification using ABC MCMC. (a–b) In ABC, data are compared to model simulations using a weighted sum of Anderson Darling distances and discrepancy in the correlations. (a) Parameter combinations that produce model realisations sufficiently similar to the experimental data, i.e. $\boldsymbol{\theta}_1$, are accepted as posterior samples. (b) Parameter combinations that do not, i.e. $\boldsymbol{\theta}_2$ are rejected. (c–s) *Posterior samples* obtained using ABC MCMC represent parameter combinations that produce realisations of the model that are similar to experimental observations. Chains are initiated at the global minimum identified using ABC SMC, and every 100th sample is retained. All axis limits correspond to the prior support (uniform priors are used). Parameter descriptions and MCMC diagnostics are given as supplementary material.

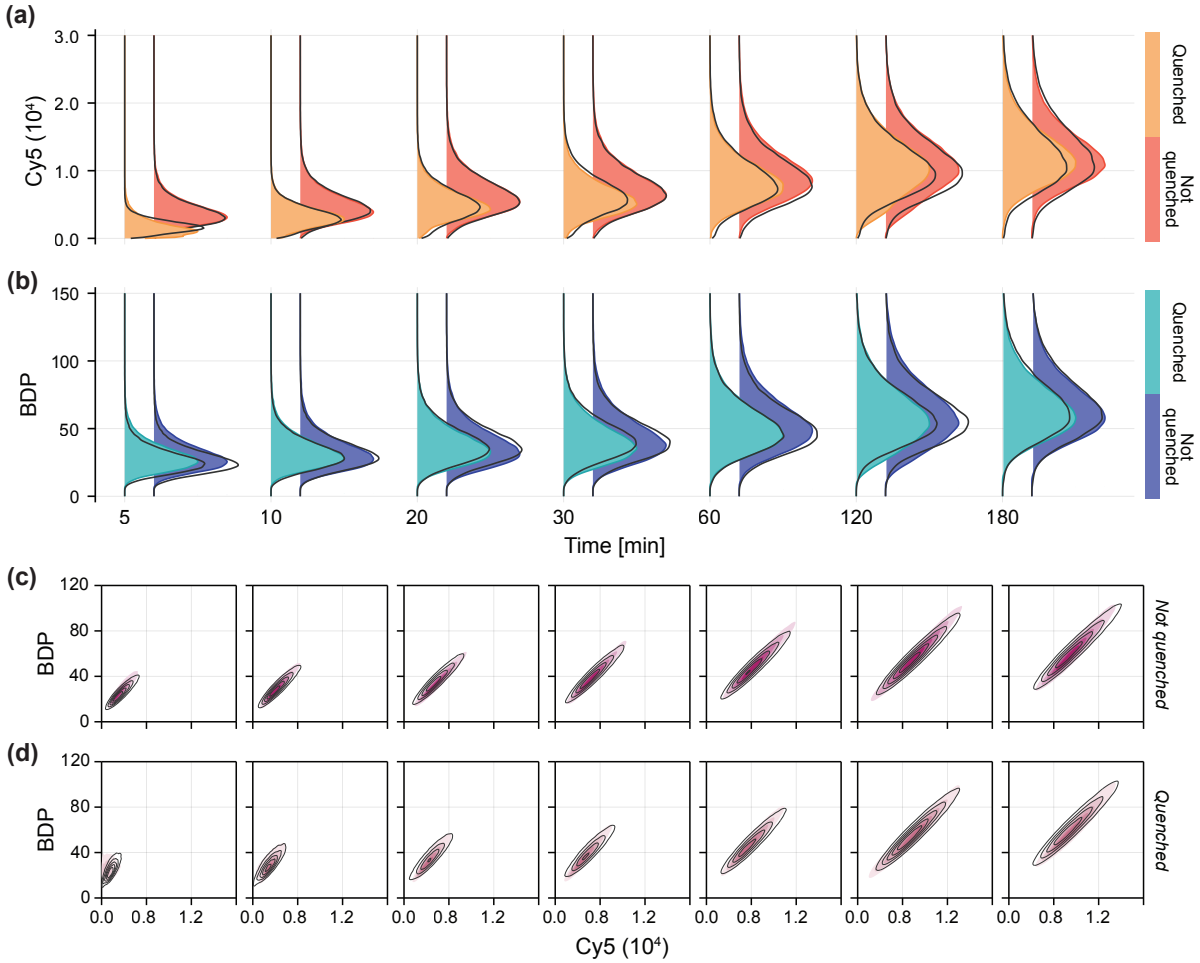

Figure 4. Mathematical model captures variability in experimental data. (a–b) Univariate kernel density estimate of the fluorescent intensity distribution from (a) the Cy5 probe, which is susceptible to the quencher dye, and (b) the BDP probe, which is not. In each case, the distribution from the quenchable experiment is shown to the left in the lighter colour. (c–d) Bivariate kernel density estimates of the joint fluorescent intensity distributions for the quenchable and unquenchable probes. The model prediction from a synthetic data set of 100, 000 cells per observation time, per condition, is overlaid in black.

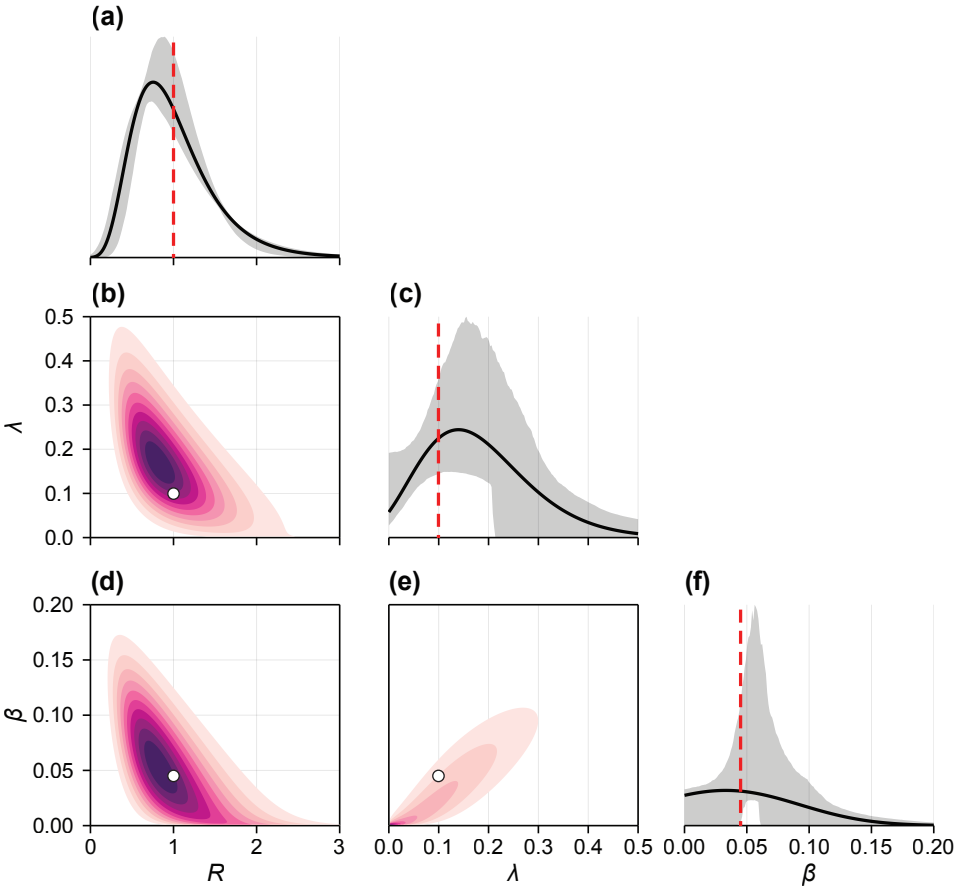

Figure 5. Inferred parameter distributions and associated uncertainty. Inferred distribution of (a) the relative number of receptors, R; (c) the internalisation rate, λ; and, (f) the recycling rate, β. Shown are the distributions at the best-fit (black), a 95% credible interval of the respective probability density functions constructed from re-sampled MCMC samples (grey), and estimates from the homogeneous model (red). (b,d,e) Bivariate distributions at the best-fit. Estimates from the homogeneous model are shown in white.

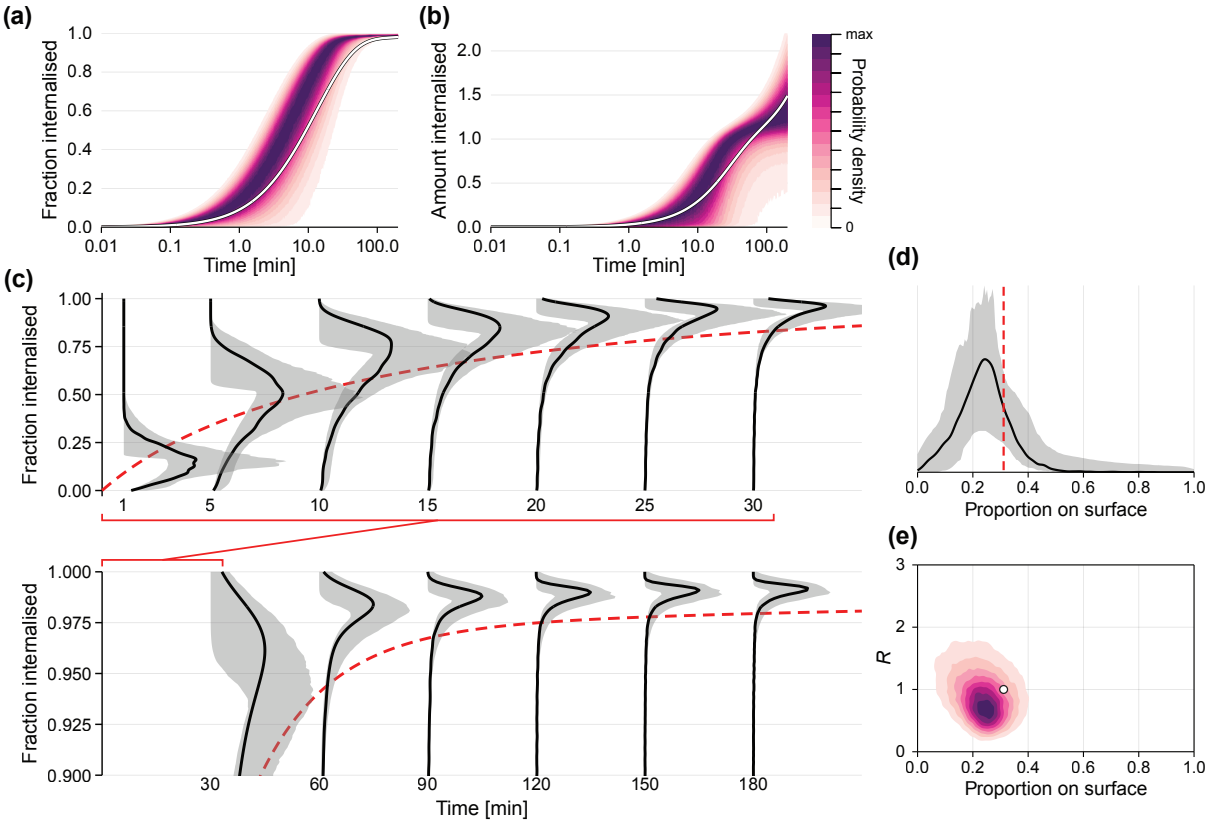

Figure 6. Model predicts unobservable measurements. Using the calibrated mathematical model, we can predict the time-evolution of the distribution of (a,c) fraction antibody inside the cell; (b) amount of antibody internalise relative to the average number of receptors on a cell, $I(t)$; and (d–e) proportion of receptors on the cell surface at equilibrium. (a,b,e) Show the distribution at the model best-fit, which is shown in black in all other plots; (c,d) additionally show 95% credible intervals constructed from MCMC samples. (c) To compare predictions with the deterministic model, we show $A(t)/I(t)$ predicted by the deterministic model (red-dashed) using the maximum likelihood estimate. In (a,b) distributions are normalised to the mode.

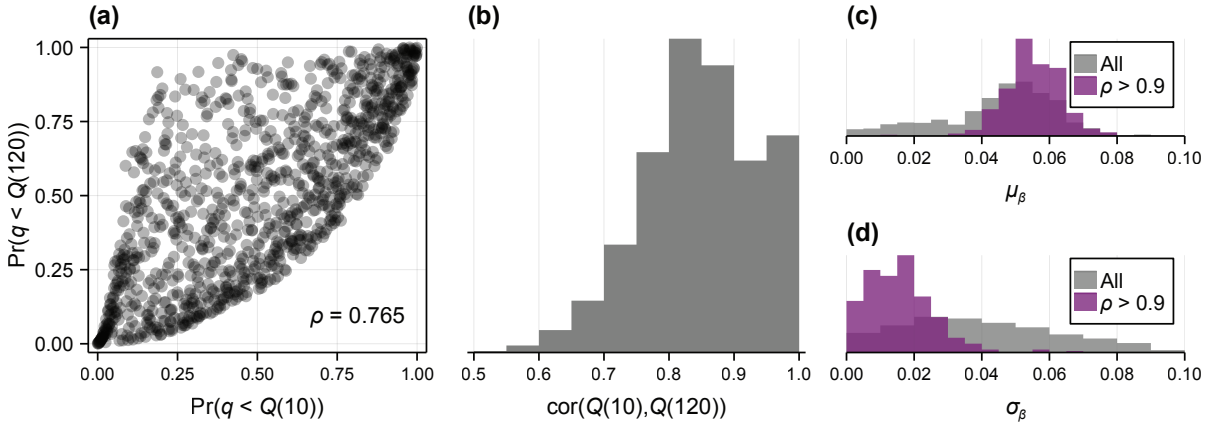

Figure 7. Dependence between the amount of material internalised at successive observation times remains uncertain. We demonstrate model predictions and uncertainty in the dependence between the noise-free quenched signal at time $t = 10$ and $t = 120$ min. (a) Quantile-quantile plot for a simulation with 1000 cells at the best-fit. Correlation of $\rho = 0.765$ is calculated based on a fitted Gaussian copula. (b) Uncertainty in the inferred correlation based on resampled posterior samples. (c,d) Assuming strong dependence between observation times affects inferences. We show the posterior distribution for the recycling rate mean and standard deviation parameters, μ_β and σ_β, respectively, if we assume a correlation of at least 0.9.

M Methods

M.1 Experimental methods

M.1.1 Cell culture

C1R cells, a human B cell lymphoblast cell line, were cultured in Dulbeccos Modified Eagle Medium (DMEM) supplemented with 10% FBS and 1% Penicillin Streptomycin, at 37 °C in a humidified 5% CO_2 atmosphere.

M.1.2 Dual-labelled fluorescent internalisation probe

Purified monoclonal IgG1 anti-human transferrin receptor antibody (OKT9) [57] was purchased from WEHI Antibody Facility.

The antibody was labelled with two fluorescent dyes; BDP-FL and FIP-Cy5. For this, OKT9 was incubated with BDP-FL-NHS ester and incubated at 4 °C overnight. BDP-FL labelled antibody was purified using a 7K MWCO Zeba spin desalting column (Thermo Scientific). The antibody was then functionalised with dibenzylcyclooctyne (DBCO)-NHS ester. Functionalised antibody was purified using a 7K MWCO Zeba spin desalting column (Thermo Scientific), and incubated with FIP-Cy5 at 4 °C overnight [58]. The dual-labelled antibody was purified using a 50k MWCO Amicon filter (Merck, Millipore), and the degree of labelling was measured by NanoDrop UV-vis spectrophotometer.

M.1.3 Internalisation assay

SHIP internalisation assays were performed by incubating the cells with dual-labelled OKT9 antibody in DMEM containing 0.1% FBS at 37 °C for different time points. After incubation, cells were washed twice with cold PBS and resuspended in propidium iodide with or without quencher (1 µM), as described previously [58]. Cells were analysed using a Stratedigm S1000EON flow cytometer and analysed using FlowJo 10.8.0.

M.2 Mathematical methods

M.2.1 Code availability

Codes are available on Github at `https://github.com/ap-browning/internalisation`.

M.2.2 Solution to dynamical model

The dynamical model is given by

$$\begin{aligned} &T \xrightarrow{\beta} S, \\ &S \xrightarrow{\lambda} E \xrightarrow{p\beta} S + F, \end{aligned} \tag{M13}$$

which we express as the system of ordinary differential equations

$$\begin{aligned}
\frac{\mathrm{d}T(t)}{\mathrm{d}t} &= -\beta T(t), \\
\frac{\mathrm{d}S(t)}{\mathrm{d}t} &= \beta T(t) - \lambda S(t) + p\beta E(t), \\
\frac{\mathrm{d}E(t)}{\mathrm{d}t} &= \lambda S(t) - p\beta E(t), \\
\frac{\mathrm{d}F(t)}{\mathrm{d}t} &= p\beta E(t),
\end{aligned} \tag{M14}$$

subject to the initial condition

$$\mathbf{x}(0) = \left(\frac{\lambda}{\lambda+\beta}, \frac{\beta}{\lambda+\beta}, 0, 0 \right), \tag{M15}$$

where $\mathbf{x}(t) = (T(t), S(t), E(t), F(t))$. The solution to eq. (M14) is given by

$$\mathbf{x}(t) = \mathbf{x}(0)\exp(\mathbf{M}t), \tag{M16}$$

where

$$\mathbf{M} = \begin{bmatrix} -\beta & 0 & 0 & 0 \\ \beta & -\lambda & p\beta & 0 \\ 0 & \lambda & -p\beta & 0 \\ 0 & 0 & p\beta & 0 \end{bmatrix}. \tag{M17}$$

Equation (M16) is solved for $A(t) = S(t) + E(t) + F(t)$ and $I(t) = E(t) + F(t)$ exactly using Mathematica. Our implementation is given in `Module/Model/deterministic.jl`.

M.2.3 Inference using approximate Bayesian computation

We perform inference using approximate Bayesian computation (ABC) [37, 39]. Given a set of experimental observations $\mathcal{X}$, we encode knowledge about the model parameters $\boldsymbol{\theta}$ in the *posterior distribution*, given by

$$\underbrace{p(\boldsymbol{\theta}|\mathcal{X})}_{\text{Posterior}} \propto \underbrace{p(\mathcal{X}|\boldsymbol{\theta})}_{\text{Likelihood}} \underbrace{p(\boldsymbol{\theta})}_{\text{Prior}}. \tag{M18}$$

Here, $p(\boldsymbol{\theta})$ denotes the *prior distribution*, which encodes prior parameter knowledge. In our work, we take a standard approach and set the prior to be uniform with independent components [59]. We choose parameter bounds to reflect either physical constraints on parameters (i.e., all correlations are bounded and rates, standard deviations, proportionality constants are positive) or realistic bounds (for example, we expect the distributions of λ and β to be negatively skewed so that support is low, but non-zero, at zero if internalisation or recycling is inhibited in a small proportion of cells).

In ABC, we approximate the posterior distribution using the ABC posterior

$$p(\boldsymbol{\theta}|\mathcal{X}) \approx p(\boldsymbol{\theta}|d(\mathcal{X}, \mathcal{Y}(\boldsymbol{\theta})) < \varepsilon). \tag{M19}$$

Here, $\mathcal{Y}(\boldsymbol{\theta})$ denotes synthetically generated observations of the model using parameters $\boldsymbol{\theta}$ and $d(\cdot,\cdot)$ is a *discrepancy measure* that measures how close synthetically generated observations lie to the experimental data. In ABC, ε is a parameter that describes the maximum discrepancy at which synthetic observations are judged to be close. In the following sections, we outline our choice of discrepancy measure followed by our algorithm to sample from eq. (M19).

Discrepancy measure

We denote by $\{X_t^Q, X_t^U\}_t$ experimental observations from samples that are not quenched at time t, where Q and U denote measurements from the quenchable and unquenchable probes, $\{\bar{X}_t^Q, \bar{X}_t^U\}_t$ those of the quenched experiments, and $\{Y_t^Q, Y_t^U\}_t$ and $\{\bar{Y}_t^Q, \bar{Y}_t^U\}_t$ as m synthetically generate observations from the mathematical model. We denote the complete set of experimental observations $\mathcal{X} = \{\{X_t^Q, X_t^U\}_t, \{\bar{X}_t^Q, \bar{X}_t^U\}_t$ and similar for $\mathcal{Y}$.

Noting that the relationship between signals from the quenchable and unquenchable probes are generally linear, we develop a discrepancy metric, $d(\mathcal{X},\mathcal{Y})$ that matches univariate marginal distributions using a modified Anderson-Darling statistic $A(X,Y)$, and matches the correlations between quenchable and unquenchable probes in the experimental and synthetic data. As we expect X_t^U and $\bar{X}_t^U$ to be similar (since $U(t) = \bar{U}(t)$), we apply a relative weighting of 1/2 to Anderson-Darling statistics involving the unquenchable probe.

Therefore,

$$\begin{aligned} d(\mathcal{X},\mathcal{Y}) = \sum_t \Bigg[& 2w_1 \left(A\left(X_t^Q, Y_t^Q\right) + A\left(\bar{X}_t^Q, \bar{Y}_t^Q\right)\right) && + \\ & w_1\Big(A\Big(X_t^U, Y_t^U\Big) + A\Big(\bar{X}_t^U, \bar{Y}_t^U\Big)\Big) && + \\ & w_2\Big\|\mathrm{Cor}(X_t^Q, X_t^U) - \mathrm{Cor}(Y_t^Q, Y_t^U)\Big\|_2 && + \\ & w_2\Big\|\mathrm{Cor}(\bar{X}_t^Q, \bar{X}_t^U) - \mathrm{Cor}(\bar{Y}_t^Q, \bar{Y}_t^U)\Big\|_2 \Bigg]. \end{aligned} \tag{M20}$$

We set $m = 1000$, $w_1 = 1$ and $w_2 = 80$.

To compare experimental observations from a marginal distribution, $X = \{x_i\}_{i=1}^n$, to a synthetically generated set of observations, $Y = \{y_i\}_{i=1}^m$, we employ a modified Anderson-Darling statistic. As the number of experimental observations, n, is large, we first compute an empirical distribution function $F_{\mathrm{obs}}(x)$ using linear interpolation and $k = 1000$ equally spaced points from $\min_i(x_i)$ to $\max_i(x_i)$. The Anderson-Darling statistic is given by [43]

$$A^2(X,Y) = m \int_{-\infty}^{\infty} \frac{\big(F_{\mathrm{sim}}(x) - F_{\mathrm{obs}}(x)\big)^2}{F_{\mathrm{obs}}(x)(1 - F_{\mathrm{obs}}(x))} f_{\mathrm{obs}}(x)\,\mathrm{d}x, \tag{M21}$$

where $F_{\mathrm{sim}}(x)$ is the empirical distribution function of the simulated data $\{y_i\}_{i=1}^m$.

Equation (M21) may be calculated by

$$A^2(X,Y) = -m - \sum_{i=1}^{n} \frac{2i-1}{m}\big(\log(F_{\mathrm{obs}}(y_i)) + \log(1 - F_{\mathrm{obs}}(y_{m+1-i}))\big), \tag{M22}$$

where $\{y_i\}_{i=1}^m$ are sorted such that $y_i \leq y_{i+1}$. To avoid blow up in cases where y_i fall outside

the range of the observed data $\{x_i\}_{i=1}^n$ (where $F_{\text{obs}}(x) = 0$ or 1) we set

$$F_{\text{obs}}(y_i) \to \min\left(1 - 10^{-9}, \max\left(10^{-9}, F_{\text{obs}}(y_i)\right)\right) \tag{M23}$$

in eq. (M22). Our implementation is given in `Module/Inference/discrepancies.jl`.

Sampling algorithm

To identify regions of the parameter space with non-neglible posterior density, we employ the sequential Monte Carlo (SMC) algorithm described in [44]. In summary, we consider an initial sample of 1000 particles, which comprise parameter combinations sampled from the prior distribution with equal weight. Initially, a model realisation is produced for each particle and compared to the experimental data using eq. (M20). Particles are then sorted by the discrepancy metric and 750 particles with the lowest discrepancy are discarded. The ABC acceptance tolerance, ε, is set to be the largest discrepancy from the remaining particles. At each successive iteration, 750 replacement particles are resampled from the 250 weighted particles that remain and perturbed using a multivariate Gaussian perturbation kernel with covariance equal to twice the empirical covariance of the remaining 250 particles. This process is repeated until 750 replacement particles are found with a discrepancy lower than ε. Particles are reweighed according to the formula given in [44]. We iterate the algorithm until the acceptance rate at the resampling step is lower than the target 0.1%.

We treat the ABC SMC algorithm as a particle-based global optimisation algorithm, the result of which is the particle, $\hat{\boldsymbol{\theta}}$, with the lowest discrepancy that is used to initiate and tune an ABC MCMC algorithm based on the Metropolis-Hastings algorithm [60] to explore the posterior. The ABC acceptance tolerance, ε, for the MCMC algorithm is chosen by taking the median discrepancy from 100 realisations of the model at $\hat{\boldsymbol{\theta}}$. All ABC MCMC chains are then initiated at $\hat{\boldsymbol{\theta}}$.

We implement ABC MCMC using `MCMCChains.jl` [61]. First, we produce four pilot chains of length 10^6 using a multivariate normal proposal with covariance equal to [62]

$$\Sigma = \frac{2.38^2}{\dim(\boldsymbol{\theta})}\hat{\Sigma}_{\text{SMC}}, \tag{M24}$$

where $\hat{\Sigma}_{\text{SMC}}$ is the empirical covariance of the 1000 SMC particles. Every 100th sample is retained, and the covariance tuned using eq. (M24) where $\hat{\Sigma}_{\text{SMC}}$ is replaced by the empirical covariance of the 40,000 samples identified using the pilot MCMC chains. Four tuned chains of length 10^7 are produced and every 100th sample retained.

Our implementation of the SMC and MCMC algorithms are given in `Module/Inference/abc.jl`. Our implementation of the ABC-SMC-MCMC sampling algorithm is given in `Results/MainResult_Inference.jl`.

M.2.4 Dependence modelling with Gaussian copula

To produce realisations of the model we require independent samples of $\boldsymbol{\xi}$, a random variable described by marginal distributions where the dependence structure is described by a Gaussian copula. In this section, we outline a computationally efficiency algorithm to sample from the

marginal distributions of $\boldsymbol{\xi}$. Additionally, we outline how we form an analytical expression for the bivariate marginal distributions used to produce visualisations in fig. 5.

Inverse transformation sampling

Sampling from the joint distribution $f(x, y, z)$ with the dependence structure described by a Gaussian copula relies on the probability inverse transformation and, therefore, becomes prohibitively expensive if the quantile function of a marginal is expensive. Therefore, we employ an approximate sampling algorithm (algorithm 1) in cases where the quantile function of the marginal is expensive (i.e., when x, y or z are Gamma-distributed) and a large number of samples is required.

Algorithm 1 Probability inverse transformation sampling of $n > 100$ samples where quantile function $F^{-1}(u)$ is prohibitively expensive.

1: Given a vector of uniformly sampled probabilities $\mathbf{u} \in \mathbb{R}^n$, cumulative distribution function $F(x)$ and (relatively expensive) quantile function $F^{-1}(u)$. Choose interpolation grid resolution $m = 100$.
2: Generate evenly spaced grid $\mathbf{x} = (x_1, x_2, ..., x_m)$ where $x_1 = F^{-1}(\min(\mathbf{u}))$ and $x_m = F^{-1}(\max(\mathbf{u}))$.
3: Calculate associated probabilities $\mathbf{y} = F(\mathbf{x})$.
4: Create linear interpolation $x = G(y)$ where $x_i = G(y_i)$, such that $G(u)$ is an approximation of the quantile function $F^{-1}(u)$.
5: Calculate and return samples $\mathbf{z} = G(\mathbf{u})$.

Bivariate marginal distributions

Consider the joint distribution of the random variables X, Y, and Z, with respective marginal density functions $f(x)$, $f(y)$ and $f(z)$, distribution functions $u = F(x)$, $v = F(y)$ and $w = F(z)$, connected through the Gaussian copula $C(u, v, w)$. Then the marginal $f(x, y)$ is given by

$$\begin{aligned}
\int_z f(x, y, z)\,\mathrm{d}z &= \int_z c(u, v, w) f(x) f(y) f(z)\,\mathrm{d}z, \\
&= f(x) f(y) \int_0^1 c(u, v, w)\,\mathrm{d}w \quad \text{since } \frac{\mathrm{d}w}{\mathrm{d}z} = f(z), \\
&= f(x) f(y) \int_0^1 \frac{\partial^3}{\partial u \partial v \partial w} C(u, v, w)\,\mathrm{d}w, \\
&= f(x) f(y) \frac{\partial^2}{\partial u \partial v} C(u, v, 1), \\
&= f(x) f(y)\, \bar{c}(u, v).
\end{aligned}$$

Here $\bar{c}(u, v) = \frac{\partial^2}{\partial u \partial v} \bar{C}(u, v)$, where $\bar{C}(u, v) = C(u, v, 1)$ is the bivariate copula describing the marginalised dependence between x and y. Since $C(u, v, w)$ is a Gaussian copula, we have that

$$\begin{aligned}
\bar{C}(u, v) &= C(u, v, 1), \\
&= \Phi_R(\Phi^{-1}(u), \Phi^{-1}(v), \Phi^{-1}(1)),
\end{aligned}$$

where Φ_Σ is the distribution function of the multivariate Gaussian distribution with covariance matrix Σ, and $\Phi(x)$ is the distribution function of the standard univariate Gaussian distribution.

We also note that $\lim_{a\to\infty}\Phi(a) = 1$. Denoting φ_Σ the probability density function of the multivariate Gaussian distribution with covariance matrix Σ, we have that

$$\begin{aligned}
\bar{C}(u,v) &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \int_{-\infty}^{\infty} \varphi_\Sigma(a,b,c)\,\mathrm{d}c\,\mathrm{d}a\,\mathrm{d}b, \\
&= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \varphi_{\Sigma_{1:2,1:2}}(a,b)\,\mathrm{d}a\,\mathrm{d}b, \\
&= \Phi_{\Sigma_{1:2,1:2}}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right).
\end{aligned}$$

Therefore, $\bar{C}(u,v)$ is a Gaussian copula with covariance matrix $\Sigma_{1:2,1:2}$.

References

[1] Pearse BMF, Robinson MS. 1990. Clathrin, adaptors, and sorting. *Annual Review of Cell Biology* **6**:151–171. doi:10.1146/annurev.cb.06.110190.001055.

[2] Kumari S, MG S, Mayor S. 2010. Endocytosis unplugged: multiple ways to enter the cell. *Cell Research* **20**:256–275. doi:10.1038/cr.2010.19.

[3] Patra JK, Das G, Fraceto LF, Campos EVR, Rodriguez-Torres MdP, Acosta-Torres LS, Diaz-Torres LA, Grillo R, Swamy MK, Sharma S, Habtemariam S, Shin HS. 2018. Nano based drug delivery systems: recent developments and future prospects. *Journal of Nanobiotechnology* **16**:71. doi:10.1186/s12951-018-0392-8.

[4] Johnston ST, Faria M, Crampin EJ. 2021. Understanding nano-engineered particle-cell interactions: biological insights from mathematical models. *Nanoscale Advances* doi:10.1039/d0na00774a.

[5] Mitchell MJ, Billingsley MM, Haley RM, Wechsler ME, Peppas NA, Langer R. 2021. Engineering precision nanoparticles for drug delivery. *Nature Reviews Drug Discovery* **20**:101–124. doi:10.1038/s41573-020-0090-8.

[6] Burel JG, Qian Y, Arlehamn CL, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R, Gilman RH, Saito M, Silva ADd, Vijayanand P, Scheuermann RH, Sette A, Peters B. 2017. An integrated workflow to assess technical and biological variability of cell population frequencies in human peripheral blood by flow cytometry. *The Journal of Immunology* **198**:1748–1758. doi:10.4049/jimmunol.1601750.

[7] Selby LI, Aurelio L, Yuen D, Graham B, Johnston APR. 2018. Quantifying cellular internalization with a fluorescent click sensor. *ACS Sensors* **3**:1182–1189. doi:10.1021/acssensors.8b00219.

[8] Heins AL, Johanson T, Han S, Lundin L, Carlquist M, Gernaey KV, Srensen SJ, Lantz AE. 2019. Quantitative flow cytometry to understand population heterogeneity in response to changes in substrate availability in escherichia coli and saccharomyces cerevisiae chemostats. *Frontiers in Bioengineering and Biotechnology* **7**:187. doi:10.3389/fbioe.2019.00187.

[9] Dixit PD, Lyashenko E, Niepel M, Vitkup D. 2020. Maximum entropy framework for predictive inference of cell population heterogeneity and responses in signaling networks. *Cell Systems* **10**:204–212.e8. doi:10.1016/j.cels.2019.11.010.

[10] Sachdeva K, Goel M, Sundaramurthy V. 2020. Heterogeneity in the endocytic capacity of individual macrophage in a population determines its subsequent phagocytosis, infectivity and subcellular trafficking. *Traffic* **21**:522–533. doi:10.1111/tra.12752.

[11] Usaj MM, Yeung CHL, Friesen H, Boone C, Andrews BJ. 2021. Single-cell image analysis to explore cell-to-cell heterogeneity in isogenic populations. *Cell Systems* **12**:608–621. doi:10.1016/j.cels.2021.05.010.

[12] Ham L, Jackson M, Stumpf M. 2021. Pathway dynamics can delineate the sources of transcriptional noise in gene expression. *eLife* **10**:e69324. doi:10.7554/elife.69324.

[13] Elsasser WM. 1984. Outline of a theory of cellular heterogeneity. *Proceedings of the National Academy of Sciences* **81**:5126–5129. doi:10.1073/pnas.81.16.5126.

[14] Altschuler SJ, Wu LF. 2010. Cellular heterogeneity: do differences make a difference? *Cell* **141**:559–563. doi:10.1016/j.cell.2010.04.033.

[15] Li B, You L. 2013. Predictive power of cell-to-cell variability. *Quantitative Biology* **1**:131–139. doi:10.1007/s40484-013-0013-3.

[16] Gough A, Stern AM, Maier J, Lezon T, Shun TY, Chennubhotla C, Schurdak ME, Haney SA, Taylor DL. 2016. Biologically relevant heterogeneity: metrics and practical insights. *SLAS Discovery* **22**:213–237. doi:10.1177/2472555216682725.

[17] Browning AP, Sharp JA, Mapder T, Baker CM, Burrage K, Simpson MJ. 2021. Persistence as an optimal hedging strategy. *Biophysical Journal* doi:10.1016/ j.bpj.2020.11.2260.

[18] Chisholm RH, Lorenzi T, Clairambault J. 2016. Cell population heterogeneity and evolution towards drug resistance in cancer: Biological and mathematical assessment, theoretical treatment optimisation. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1860**:2627–2645. doi:10.1016/j.bbagen.2016.06.009.

[19] Inde Z, Dixon SJ. 2017. The impact of non-genetic heterogeneity on cancer cell death. *Critical Reviews in Biochemistry and Molecular Biology* **53**:1–16. doi:10.1080/10409238.2017.1412395.

[20] Dumont C, Czuba E, Chen M, Villadangos JA, Johnston AP, Mintern JD. 2017. DNA-based probes for flow cytometry analysis of endocytosis and recycling. *Traffic* **18**:242–249. doi:10.1111/tra.12466.

[21] Lee D, Ding Y, Jayaraman A, Kwon JS. 2018. Mathematical modeling and parameter estimation of intracellular signaling pathway: application to LPS-induced NFκB activation and TNFα production in macrophages. *Processes* **6**:21. doi:10.3390/pr6030021.

[22] O'Neill K, Aghaeepour N, pidlen J, Brinkman R. 2013. Flow cytometry bioinformatics. *PLOS Computational Biology* **9**:e1003365. doi:10.1371/journal.pcbi.1003365.

[23] Durbin KR, Phipps C, Liao X. 2018. Mechanistic modeling of antibody-drug conjugate internalization at the cellular level reveals inefficient processing steps. *Molecular Cancer Therapeutics* **17**:molcanther.0672.2017. doi:10.1158/1535-7163.mct-17-0672.

[24] FitzGerald LI, Johnston APR. 2020. It's what's on the inside that counts: techniques for investigating the uptake and recycling of nanoparticles and proteins in cells. *Journal of Colloid and Interface Science* **587**:64–78. doi:10.1016/j.jcis.2020.11.076.

[25] Eberhardt EH. 1967. Noise in photomultiplier tubes. *IEEE Transactions on Nuclear Science* **14**:7–14. doi:10.1109/tns.1967.4324498.

[26] Kell DB, Ryder HM, Kaprelyants AS, Westerhoff HV. 1991. Quantifying heterogeneity: flow cytometry of bacterial cultures. *Antonie van Leeuwenhoek* **60**:145–158. doi:10.1007/bf00430362.

[27] Banks H, Kapraun D, Thompson WC, Peligero C, Argilaguet J, Meyerhans A. 2013. A novel statistical analysis and interpretation of flow cytometry data. *Journal of Biological Dynamics* **7**:96–132. doi:10.1080/17513758.2013.812753.

[28] Galbusera L, Bellement-Theroue G, Urchueguia A, Julou T, Nimwegen Ev. 2020. Using fluorescence flow cytometry data for single-cell gene expression analysis in bacteria. *PLOS ONE* **15**:e0240233. doi:10.1371/journal.pone.0240233.

[29] Huang Y, Liu D, Wu H. 2006. Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics* **62**:413–423. doi:10.1111/j.1541-0420.2005.00447.x.

[30] Hasenauer J, Hasenauer C, Hucho T, Theis FJ. 2014. ODE constrained mixture modelling: a method for unraveling subpopulation structures and dynamics. *PLOS Computational Biology* **10**:e1003686. doi:10.1371/journal.pcbi.1003686.

[31] Loos C, Moeller K, Frhlich F, Hucho T, Hasenauer J. 2018. A hierarchical, data-driven approach to modeling single-cell populations predicts latent causes of cell-to-cell variability. *Cell Systems* **6**:593–603.e13. doi:10.1016/j.cels.2018.04.008.

[32] Lambert B, Gavaghan DJ, Tavener SJ. 2021. A Monte Carlo method to estimate cell population heterogeneity from cell snapshot data. *Journal of Theoretical Biology* **511**:110541. doi:10.1016/j.jtbi.2020.110541.

[33] Soong T. 1973. *Random Differential Equations in Science and Engineering*. volume 103 of *Mathematics in Science and Engineering*. ISBN 9780126548501. doi:10.1016/s0076-5392(08)60159-9.

[34] Lawson BAJ, Drovandi CC, Cusimano N, Burrage P, Rodriguez B, Burrage K. 2018. Unlocking data sets by calibrating populations of models to data density: A study in atrial electrophysiology. *Science Advances* **4**:e1701676. doi:10.1126/sciadv.1701676.

[35] Waldherr S, Hasenauer J, Allgwer F. 2009. Estimation of biochemical network parameter distributions in cell populations. *IFAC Proceedings Volumes* **42**:1265–1270. doi:10.3182/20090706-3-fr-2004.00210.

[36] Hasenauer J, Waldherr S, Doszczak M, Radde N, Scheurich P, Allgwer F. 2011. Identification of models of heterogeneous cell populations from population snapshot data. *BMC Bioinformatics* **12**:125. doi:10.1186/1471-2105-12-125.

[37] Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics* **162**:2025.

[38] Sisson SA, Fan Y, Tanaka MM. 2007. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences* **104**:1760–1765. doi:10.1073/pnas.0607208104.

[39] Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface* **6**:187–202. doi:10.1098/rsif.2008.0172.

[40] Strand J. 1970. Random ordinary differential equations. *Journal of Differential Equations* **7**:538–553. doi:10.1016/0022-0396(70)90100-2.

[41] Pawitan Y. 2013. *In all likelihood: statistical modelling and inference using likelihood.* Oxford University Press, Oxford.

[42] Furusawa C, Suzuki T, Kashiwagi A, Yomo T, Kaneko K. 2005. Ubiquity of log-normal distributions in intra-cellular reaction dynamics. *Biophysics* **1**:25–31. doi:10.2142/biophysics.1.25.

[43] Anderson TW, Darling DA. 1954. A test of goodness of fit. *Journal of the American Statistical Association* **49**:765–769. doi:10.1080/01621459.1954.10501232.

[44] Vo BN, Drovandi CC, Pettitt AN, Pettet GJ. 2015. Melanoma cell colony expansion parameters revealed by approximate Bayesian computation. *PLOS Computational Biology* **11**:e1004635 22. doi:10.1371/journal.pcbi.1004635.

[45] Linnet K. 1993. Evaluation of regression procedures for methods comparison studies. *Clinical Chemistry* **39**:424–32.

[46] Wilkinson RD. 2013. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Statistical applications in genetics and molecular biology* **12**:129 141. doi:10.1515/sagmb-2013-0010.

[47] Pyne S, Hu X, Wang K, Rossin E, Lin TI, Maier LM, Baecher-Allan C, McLachlan GJ, Tamayo P, Hafler DA, Jager PLD, Mesirov JP. 2009. Automated high-dimensional flow cytometric data analysis. *Proceedings of the National Academy of Sciences* **106**:8519–8524. doi:10.1073/pnas.0903028106.

[48] Johnsson K, Wallin J, Fontes M. 2016. BayesFlow: latent modeling of flow cytometry cell populations. *BMC Bioinformatics* **17**:25. doi:10.1186/s12859-015-0862-z.

[49] Minoura K, Abe K, Maeda Y, Nishikawa H, Shimamura T. 2019. Model-based cell clustering and population tracking for time-series flow cytometry data. *BMC Bioinformatics* **20**:633. doi:10.1186/s12859-019-3294-3.

[50] Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* **297**:1183–1186. doi:10.1126/science.1070919.

[51] Fielding AB, Willox AK, Okeke E, Royle SJ. 2012. Clathrin-mediated endocytosis is inhibited during mitosis. *Proceedings of the National Academy of Sciences* **109**:6572–6577. doi:10.1073/pnas.1117401109.

[52] Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, Teichmann SA, Marioni JC, Stegle O. 2015. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nature Biotechnology* **33**:155–160. doi:10.1038/nbt.3102.

[53] Khetan J, Shahinuzzaman M, Barua S, Barua D. 2019. Quantitative analysis of the correlation between cell size and cellular uptake of particles. *Biophysical Journal* **116**:347–359. doi:10.1016/j.bpj.2018.11.3134.

[54] Kreutz C, Timmer J. 2009. Systems biology: experimental design. *FEBS Journal* **276**:923–942. doi:10.1111/j.1742-4658.2008.06843.x.

[55] Snijder B, Sacher R, Rm P, Damm EM, Liberali P, Pelkmans L. 2009. Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461**:520–523. doi:10.1038/nature08282.

[56] Kalemera M, Mincheva D, Grove J, Illingworth CJR. 2019. Building a mechanistic mathematical model of hepatitis C virus entry. *PLOS Computational Biology* **15**:e1006905. doi:10.1371/journal.pcbi.1006905.

[57] Sutherland R, Delia D, Schneider C, Newman R, Kemshead J, Greaves M. 1981. Ubiquitous cell-surface glycoprotein on tumor cells is proliferation-associated receptor for transferrin. *Proceedings of the National Academy of Sciences* **78**:4515–4519. doi:10.1073/pnas.78.7.4515.

[58] Liu H, Johnston APR. 2013. A programmable sensor to probe the internalization of proteins and nanoparticles in live cells. *Angewandte Chemie International Edition* **52**:5744–5748. doi:10.1002/anie.201301243.

[59] Hines KE, Middendorf TR, Aldrich RW. 2014. Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach. *The Journal of General Physiology* **143**:401–16. doi:10.1085/jgp.201311116. URL `https://rupress.org/jgp/article-pdf/143/3/401/649400/jgp_201311116.pdf`.

[60] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* **21**:1087–1092. doi:10.1063/1.1699114.

[61] MCMCChains.jl. https://github.com/TuringLang/MCMCChains.jl. Version 5.0.1.

[62] Roberts GO, Rosenthal JS. 2001. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science* **16**:351–367. doi:10.1214/ss/1015346320.