

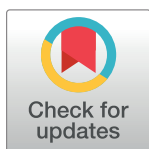
RESEARCH ARTICLE

More or less—On the influence of labelling strategies to infer cell population dynamics

Michael Gabel^{1*}, Roland R. Regoes², Frederik Graw^{1*}

1 Center for Modelling and Simulation in the Biosciences, BioQuant-Center, Heidelberg University, 69120 Heidelberg, Germany, **2** Institute for Integrative Biology, ETH Zurich, CH-8092 Zurich, Switzerland

* michael.gabel@bioquant.uni-heidelberg.de (MG); frederik.graw@bioquant.uni-heidelberg.de (FG)



Abstract

The adoptive transfer of labelled cell populations has been an essential tool to determine and quantify cellular dynamics. The experimental methods to label and track cells over time range from fluorescent dyes over congenic markers towards single-cell labelling techniques, such as genetic barcodes. While these methods have been widely used to quantify cell differentiation and division dynamics, the extent to which the applied labelling strategy actually affects the quantification of the dynamics has not been determined so far. This is especially important in situations where measurements can only be obtained at a single time point, as e.g. due to organ harvest. To this end, we studied the appropriateness of various labelling strategies as characterised by the number of different labels and the initial number of cells per label to quantify cellular dynamics. We simulated adoptive transfer experiments in systems of various complexity that assumed either homeostatic cellular turnover or cell expansion dynamics involving various steps of cell differentiation and proliferation. Re-sampling cells at a single time point, we determined the ability of different labelling strategies to recover the underlying kinetics. Our results indicate that cell transition and expansion rates are differently affected by experimental shortcomings, such as loss of cells during transfer or sampling, dependent on the labelling strategy used. Furthermore, uniformly distributed labels in the transferred population generally lead to more robust and less biased results than non-equal label sizes. In addition, our analysis indicates that certain labelling approaches incorporate a systematic bias for the identification of complex cell expansion dynamics.

OPEN ACCESS

Citation: Gabel M, Regoes RR, Graw F (2017) More or less—On the influence of labelling strategies to infer cell population dynamics. *PLoS ONE* 12(10): e0185523. <https://doi.org/10.1371/journal.pone.0185523>

Editor: Andrew J. Yates, University of Glasgow, UNITED KINGDOM

Received: March 27, 2017

Accepted: September 14, 2017

Published: October 18, 2017

Copyright: © 2017 Gabel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All methods to simulate and create the used data are within the paper and its Supporting Information files.

Funding: FG and MG are supported by the Center for Modeling and Simulation in the Biosciences (BIOMS, <http://www.bioms.de/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The ability to distinguish cells and organisms by certain markers and labels has been an indispensable asset in many biological experiments addressing population dynamics and development. For example, tracking differently labelled cells not only allows the identification of lineage pathways [1], but also the observation of dynamical changes in cell populations over time [2]. The application of labels also helps to determine the migration dynamics of cells between organs [3], or the colonisation dynamics of specific tissues by bacteria [4, 5]. In

addition, the information obtained by labelling can be used to quantify cellular turnover, such as cell activation, proliferation and differentiation dynamics [6].

For cells, there exists a large variety of experimental techniques to label and track individual populations. Besides the application of markers that are taken up during cell proliferation, such as BrdU [7, 8], deuterated glucose and heavy water [9–11], this especially concerns techniques that involve the adoptive transfer of pre-labelled cell populations. Staining cells by the fluorescent dye CFSE [12, 13] has been used extensively to infer cellular turnover and proliferation dynamics (reviewed in [6]). More fine-grained approaches that involve several different markers—e.g. by transferring cell populations bearing congenic markers [14–16] or by using naturally diverse markers, such as T cell receptor sequences [17–20]—allow to distinguish the dynamics of individual subpopulations in more detail. Finally, artificially labelling cells by unique, inheritable genetic barcodes makes it possible to follow cellular dynamics on a single cell level [21]. By this, one is able to address cell heterogeneity and to identify individual cell differentiation pathways [2, 21–23].

The adoptive transfer of labelled cells is particularly useful, if the experimental conditions prevent sampling at different times. When organs or cell cultures need to be harvested, individual measurements can only be obtained at one particular time point. In these cases, the intra-individual variability in the population dynamics of each label can provide enough information to estimate cellular turnover. Interestingly, it is also possible to quantify interacting dynamics, such as entangled migration and proliferation dynamics, even if measurements are only obtained from one of the involved compartments [4]. Thus, using multiple labels can compensate for both the lack of time-resolved data and compartments that cannot be measured.

Several different labelling strategies have been used to analyse population dynamics given these experimental limitations. These approaches differed in the number of labels and the size of each label within the transferred population [2, 4, 16]. However, it has not been determined so far if these labelling strategies allow to reliably infer the assumed dynamics, and how these different approaches influence the quantification of the kinetics: does the estimation of a cell proliferation rate benefit from a high or a small number of cells per label? To what extent would parameter estimation be improved if more labels are used? And how does the time point of sampling affect parameter identification? The impact of a labelling strategy on parameter identification needs to be evaluated in order to determine the reliability of obtained parameter estimates.

To this end, we studied the appropriateness of different labelling strategies to quantify cellular dynamics. Here, we focus on labelling approaches based on inheritable and stable markers, such as congenic markers or genetic barcodes. We considered cellular systems of various complexity that assumed either homeostatic turnover, as e.g. for naïve T cells, or cell expansion dynamics involving various steps of cell differentiation and proliferation (Fig 1A). We then simulated adoptive transfer experiments varying the composition of the labelled cell population (Fig 1B). Data sampled at a single time point were used to quantify the underlying kinetics and to evaluate the impact of the labelling strategy on parameter estimation. In addition, we analysed how experimental shortcomings, such as incomplete transfer or sampling of the labelled cell population (Fig 1C), affected the results.

Our results show that labelling strategies and the experimental limitations affect parameter estimates in multiple ways: The appropriateness of a labelling approach depends on the underlying cellular system, but also on the type of parameter that is to be identified. Labelling strategies might be biased to favour the identification of certain types of cellular characteristics, e.g. proliferation rates, with some approaches being more robust than others with respect to loss of cells during transfer or sampling.

In general, our findings argue for the use of multiple labels with an intermediate number of cells per label to reliably infer cellular transition and expansion dynamics. Furthermore, they also suggest the use of simulations to determine a-priori the appropriateness and limitations of the experimentally used labelling strategy, or to later validate obtained parameter estimates.

Materials and methods

The mathematical models of cellular dynamics

We distinguish three scenarios of cellular dynamics that consider different levels of complexity (see Fig 1A). These scenarios are described as follows.

(1) Homoeostatic turnover. Under homoeostatic conditions, a cell population is considered to be in equilibrium, meaning the total number of cells is assumed to be constant over time. However, the cell population is usually not static as cells constantly die and are replaced. Examples of homoeostatic turnover among immune cells are the dynamics of naïve T cells before antigen encounter, or the pool of memory T cells that is maintained after an infection [24]. In our model, we assume that a cell population, here termed naïve cells, N , proliferates with rate ρ and dies with rate δ . The dynamics are described by the following differential equation:

$$\frac{dN}{dt} = (\rho - \delta)N. \tag{1}$$

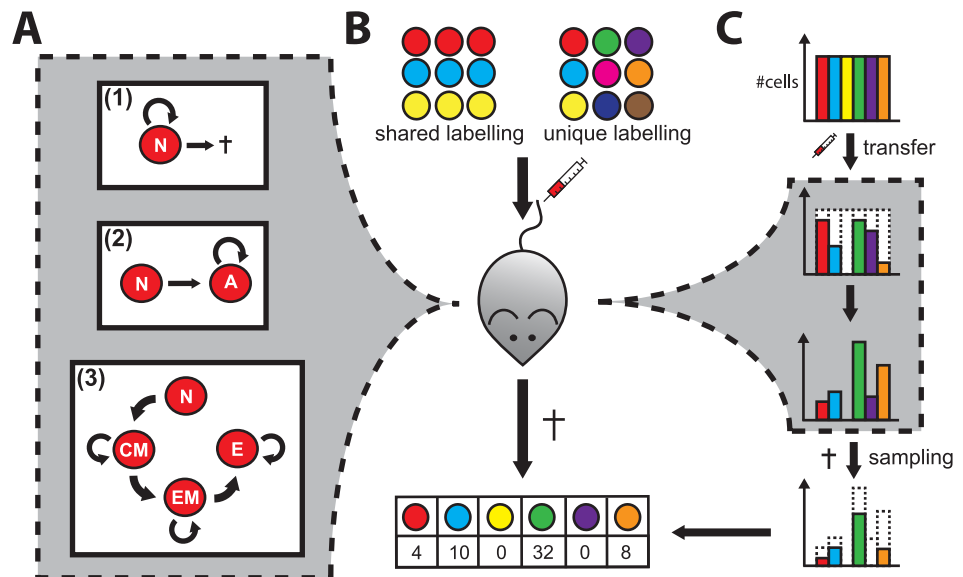


Fig 1. Population dynamics, experimental setups and technical shortcomings. (A) Schematic of different models with increasing complexity describing cellular turnover: (1) *Homoeostatic turnover*: Naïve cells proliferate only to compensate cell death, therefore maintaining a stable number of cells. (2) *Simple expansion dynamics*: By encountering their respective antigen naïve cells are activated and start to proliferate. (3) *Complex expansion dynamics*: In comparison to (2), we consider several steps of cell differentiation and proliferation. Upon activation, naïve cells differentiate into central memory precursor (CM) and subsequently into effector memory precursor cells (EM), and finally effector cells (E). For simplicity, net-proliferation rates combining cell proliferation and death are considered at this point [16]. (B) A labelling strategy using inheritable labels is defined by the number of different labels and the label size, i.e. the number of cells per label. The depicted labelling strategies show a shared and a unique labelling approach. After transfer into a host, these cells are thought to follow one of the three cellular dynamics. At a specific time, cells are sampled and used for evaluation. Data are gathered in the form of *count* data measuring the number of cells of a specific label within the sampled population. (C) Potential experimental shortcomings: Cells can be lost during transfer and/or sampling.

<https://doi.org/10.1371/journal.pone.0185523.g001>

To ensure homeostatic turnover, $\rho = \delta$. In the following, we set $\rho = \delta = 0.5 \text{ d}^{-1}$ to allow for reasonable simulation times. This is done without loss of generality as the timescale of the simulations can be rescaled to allow interpretation for much lower turnover rates, as e.g. observed for naïve T cells [6] (see also S1 Fig).

(2) Simple expansion dynamics. Another dynamics is the activation and subsequent proliferation of cells (Fig 1A). After encountering their cognate antigen, naïve T cells are activated and start to fight effectively against the invading pathogen by massively expanding in numbers and simultaneously differentiating into effective subpopulations [16]. To model a simple expansion dynamics, we distinguish between naïve, N , and activated cells, A [25]. Naïve cells are activated with rate μ , and activated cells start to proliferate with rate ρ . The dynamics of this model can be described by the following system of ordinary differential equations:

$$\begin{aligned} \frac{dN}{dt} &= -\mu N \\ \frac{dA}{dt} &= \mu N + \rho A \end{aligned} \tag{2}$$

For simplicity, cell death of both naïve and activated cells is neglected in this model, as we are mainly interested in the net-expansion rates.

(3) Complex expansion dynamics. In a third step, we extended the simple expansion model by additionally accounting for heterogeneous subpopulations among the activated cells. As for example for T cells, several functionally diverse subsets are distinguished that indicate different steps of cell differentiation [2, 26]. Each of these subsets is assumed to follow individual proliferation and differentiation dynamics. Following the study by Buchholz et al. [16], we distinguish between central memory precursor (CM), effector memory precursor (EM) and effector cells (E). The relation between these compartments is assumed to follow a linear differentiation pathway as depicted in Fig 1A and is defined by the following system of ordinary differential equations:

$$\begin{aligned} \frac{dN}{dt} &= -\mu_N N \\ \frac{dCM}{dt} &= \mu_N N + (\rho_{CM} - \mu_{CM}) CM \\ \frac{dEM}{dt} &= \mu_{CM} CM + (\rho_{EM} - \mu_{EM}) EM \\ \frac{dE}{dt} &= \mu_{EM} EM + \rho_E E, \end{aligned} \tag{3}$$

where μ_x and ρ_x describe the differentiation and proliferation rates, respectively of the corresponding compartments. We used estimates derived from Buchholz et al. [16] to parametrise the model; the respective values are given as $\mu_N = 2.2 \text{ d}^{-1}$, $\mu_{CM} = 0.2 \text{ d}^{-1}$, $\mu_{EM} = 0.04 \text{ d}^{-1}$, $\rho_{CM} = 0.85 \text{ d}^{-1}$, $\rho_{EM} = 1.42 \text{ d}^{-1}$ and $\rho_E = 1.6 \text{ d}^{-1}$.

Simulating labelling experiments

To simulate experimental data, we performed stochastic simulations of the systems defined by Eqs (1)–(3) based on the Gillespie algorithm [27]. Simulations were carried out in the R-language of statistical computing using the package *adaptivetau* [28]. Each simulation starts with a specified number of labelled naïve cells at time $t = 0$. These cells then proliferate, differentiate or die stochastically, according to the underlying model. We assume inheritable markers,

meaning that the label of each individual cell is retained during activation or differentiation, and it is passed onto every daughter cell while proliferating. At a specified sampling time $T > 0$ the system is stopped and the number of cells per label in each cellular subset is assessed.

In addition to the model parameters characterising the cellular dynamics, each simulation depends on the following experimental parameters: The sampling time, T , at which cells are sampled, and the labelling strategy, which is defined by the number of different labels, L , and the label size M , i.e. the number of cells per label in the initial cell population. Unless stated otherwise, we assume uniformly distributed labelling strategies, i.e. every label has initially the same number of cells.

To account for possible loss of cells during transfer (Fig 1C), the fraction of cells that is assumed to pass the transfer is sampled randomly from the initial cell population. This sampled transfer fraction is then used as an initial condition for the model systems. Similarly, by randomly sampling a predefined fraction of cells from the stochastically generated simulation output, we account for incomplete sampling that might occur during experiments. The sampled cell population is then used to estimate the parameters of the underlying system.

Parameter estimation

Parameter estimates for the rates describing cell activation, proliferation and differentiation are obtained by fitting the predicted summary statistics for each cell population to the sampled count data, which provide the absolute number of cells for each label. Each sample is evaluated individually. The considered summary statistics include the expected mean, the coefficients of variation (CV) and, if applicable, the correlation coefficients (CC). The predicted summary statistics are obtained by solving the corresponding master equations of the systems (Eqs (1)–(3)) (see S1 Appendix for a detailed description of the calculations). Fitting is then performed based on χ^2 -minimisation using the `optim`-function in the R-language of statistical computing [28].

Confidence intervals for parameter estimates are obtained by bootstrapping the data using the built-in R-package `boot`. These intervals are calculated based on Efron's non-parametric and accelerated bootstrap (BCa) method [29] with 999 repeats and a significance level of $\alpha = 0.05$.

To allow for comparison with the original approach by Buchholz et al. [16], the compartment of naïve cells, N , was not considered when fitting both the simple and the complex expansion model.

Evaluating the quality of parameter estimates

The appropriateness of different labelling strategies to retrieve the underlying cellular dynamics is determined by different quantities [30]. These quantities characterise the robustness of parameter estimates and their deviation from the true parameter.

Bias. The bias indicates on an absolute scale how much the average parameter estimate deviates from the true value.

In mathematical terms, if $\hat{\theta}_i$, $i = 1, \dots, m$ are estimates for the true parameter θ with $\hat{\theta} := \frac{1}{m} \sum_{i=1}^m \hat{\theta}_i$ defining the empirical mean, the bias is calculated by

$$\text{Bias} := \hat{\theta} - \theta. \quad (4)$$

Percentage bias. The percentage bias determines on a relative scale how much the average estimate differs from the true value. This allows the simultaneous comparison of estimates for several parameters of different scales.

The percentage bias is defined by

$$\text{pBias} := \frac{\hat{\theta} - \theta}{\theta}. \quad (5)$$

Mean confidence interval length (MCIL). The MCIL serves as a measure of uncertainty for the parameter estimate. If $\text{CI}_i = [a_i, b_i]$ is the estimated confidence interval for parameter θ in run i , with $l(\text{CI}_i) = b_i - a_i$ defining the length of the confidence interval, then the mean confidence interval length is calculated by

$$\text{MCIL} := \frac{1}{m} \sum_{i=1}^m l(\text{CI}_i). \quad (6)$$

Here, m denotes the total number of individual runs performed. In some cases the MCIL cannot be calculated (e.g. due to an unlimited confidence interval of at least one of the confidence intervals used for calculation). This is indicated in the corresponding plots by a grey coloured box for the respective parameter combination.

False coverage rate (FCR). The false coverage rate is defined as the fraction of simulation runs in which the estimated confidence interval does not contain the predefined rate.

Results

The influence of transfer loss on parameter estimation

During adoptive transfer of cells into a living host, it is unlikely that all cells will survive the transit. Common obstacles include experimental limitations, such as imperfect injections to the target tissue, or host-induced rejection of cells, e.g. when using congenic markers [31]. Therefore, one would expect that only a fraction of the original labelled cell population enters the system and can be recovered later. As we show below, neglecting this *transfer fraction* when evaluating sampled data can strongly impact parameter estimates of the cellular dynamics.

Assuming homeostatic cell turnover where cells proliferate and die at similar rates, we tested the ability of two different labelling strategies to infer the underlying kinetics in case of incomplete transfer (Fig 2A). Both labelling strategies involve the adoptive transfer of $N = 800$ cells that are labelled according to a unique ($L = 800$ labels with $M = 1$ cell each) or a shared labelling approach ($L = 8, M = 100$); both of which have been successfully used in experiments [2, 16]. Cells that survive the transfer and undergo stochastic homeostatic turnover are sampled at a single time point and used for parameter estimation.

Incomplete transfer results in an overestimation of both the proliferation and the death rate in either of the two labelling strategies (Fig 2B). In addition, the parameter estimates indicate an exponential decay of cells rather than a homeostatic turnover as the death rate δ is always estimated to be higher than the corresponding proliferation rate ρ . This is due to the fact that the system has to compensate for the smaller number of cells that are recovered compared to the number of labelled cells in the inoculum. A higher transfer loss also results in a larger variation of the parameter estimates whereby a unique labelling approach allows more robust estimation.

If prior knowledge on the transfer fraction can be obtained, as for example by additional experiments [16], it is possible to adjust the estimation procedure and to obtain appropriate parameter estimates for the homeostatic turnover (see Fig 2C and S1 Appendix). This works reliably for both labelling strategies even if large fractions of cells are lost during transfer.

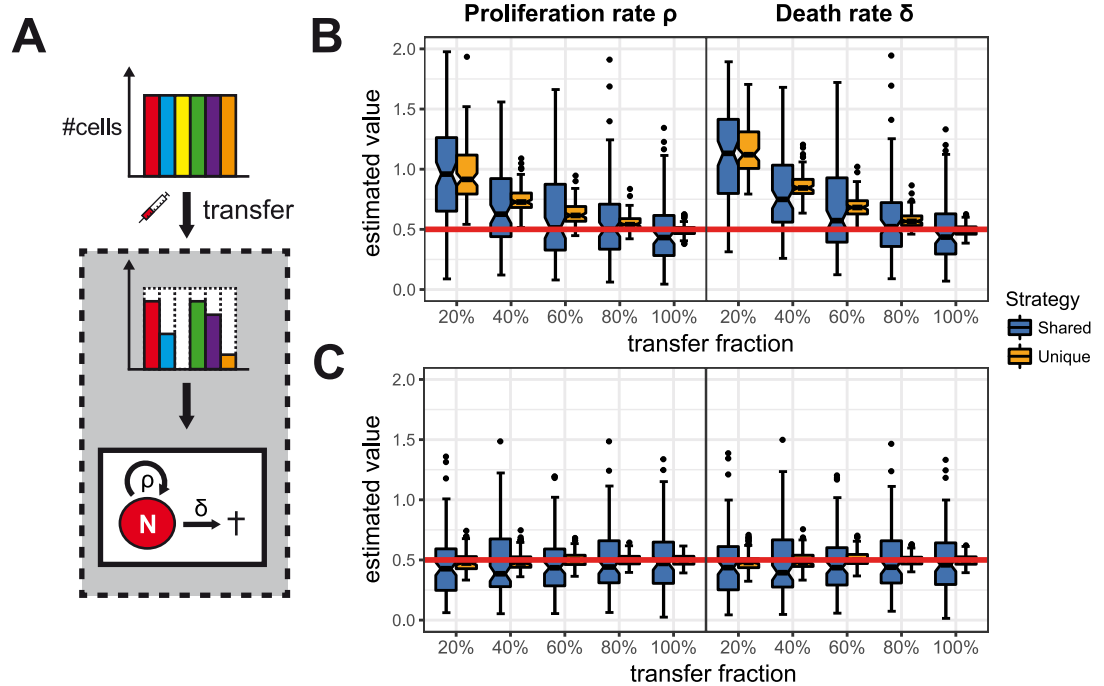


Fig 2. The influence of incomplete transfer on estimating homeostatic cell turnover. (A) Only a fraction of the initially labelled cell population might survive the transfer and follows homeostatic turnover where cells proliferate with rate ρ and die at rate δ . (B) Panels show the distribution of estimates for the proliferation rate, ρ (left), and the death rate, δ (right), for different fractions of cells surviving the transfer. Parameter estimates for two labelling strategies with $N = 800$ cells initially using either shared ($L = 8, M = 100$, blue) or unique labelling ($L = 800, M = 1$, orange) are shown. The estimation procedure did not account for the transfer loss. (C) Prior knowledge on the transfer loss improves parameter estimates even if only small fractions of cells survive the transfer. Each boxplot is based on the results of 100 individual stochastic simulations with $\rho = \delta = 0.5 \text{ d}^{-1}$ and cells being sampled 8 days after transfer. Red lines indicate the true parameter values.

<https://doi.org/10.1371/journal.pone.0185523.g002>

However, in more complex scenarios of cell expansion dynamics, even full knowledge on the transfer fraction might not be sufficient to correctly quantify the underlying dynamics. Buchholz et al. [16] studied the proliferation and differentiation dynamics of T cells and identified a linear differentiation pathway with naïve (N) cells differentiating into central memory (CM) and effector memory precursor cells (EM), and further into effector cells (E) (Fig 3A). Using a shared labelling strategy adapted from their experiment, we find that all parameter estimates besides the naïve differentiation rate μ_N seem unaffected by a loss of cells during transfer (Fig 3B). Accounting for the transfer fraction in the estimation procedure leads to more robust but not necessarily correct estimates (S2 Fig). Even if no cells are lost during transfer, the proliferation and differentiation rates associated with the naïve and central memory compartment, i.e. μ_N, μ_{CM} and ρ_{CM} , are under- and overestimated with a relative bias of 0.5 and 1.5, respectively (Fig 3B). In contrast, the unique labelling approach allows identification of the underlying kinetics for all cellular subsets if all cells survive the transfer, or if the effective transfer fraction is known and accounted for (Fig 3B, S2 Fig). If the transfer fraction is not known, this labelling approach also leads to biased estimates of the proliferation and differentiation rates μ_N, μ_{CM} and ρ_{CM} , but does not affect the estimates of the remaining parameters. As in the homeostatic scenario, the larger number of labels of the unique labelling strategy leads to less variation in the parameter estimates compared to the shared labelling approach.

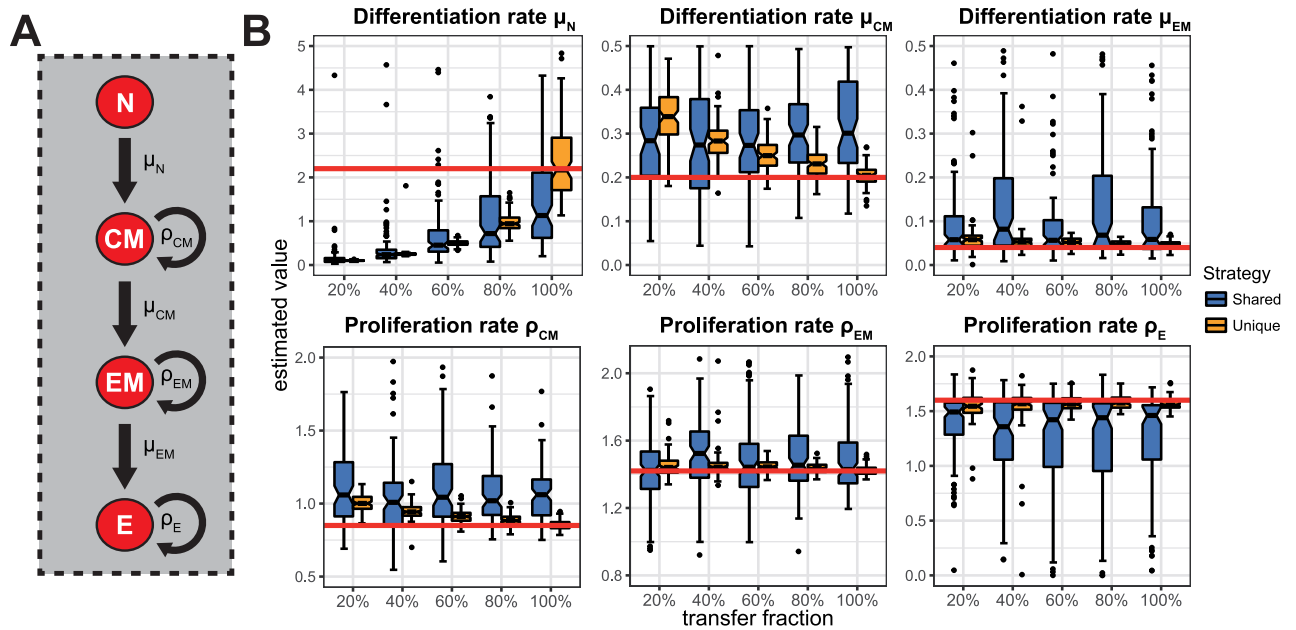


Fig 3. The influence of incomplete transfer on estimating complex cell differentiation and expansion dynamics. (A) Schematic of a linear pathway for cell differentiation and proliferation as assumed for the complex expansion model [16]. Naïve cells (N) turn into central memory precursor cells (CM), which subsequently turn into effector memory precursor (EM) and effector (E) cells. Cells differentiate and proliferate according to the corresponding rates μ and ρ , respectively. (B) Panels show the distribution of estimates for the differentiation rates (upper row) and proliferation rates (lower row) for the different cellular subsets for various fractions of cells surviving the transfer. The estimation procedure did not account for the transfer loss. Parameter estimates for two labelling strategies with $N = 800$ cells initially using either shared ($L = 8, M = 100$, blue) or unique labelling ($L = 800, M = 1$, orange) are shown. Every boxplot is based on the results of 100 individual stochastic simulations. Differentiation and proliferation rates are defined as $\mu_N = 2.2 \text{ d}^{-1}$, $\mu_{CM} = 0.2 \text{ d}^{-1}$, $\mu_{EM} = 0.04 \text{ d}^{-1}$, $\rho_{CM} = 0.85 \text{ d}^{-1}$, $\rho_{EM} = 1.42 \text{ d}^{-1}$ and $\rho_E = 1.6 \text{ d}^{-1}$ [16]. Red lines indicate the true parameter values.

<https://doi.org/10.1371/journal.pone.0185523.g003>

In summary, our results show that incomplete transfer mainly affects the quantification of cellular kinetics in early compartments while the estimation of later differentiation and expansion steps is not affected.

The influence of incomplete sampling on parameter estimation

Sampling cells from the host system represents another source of error. Most likely only a fraction of the labelled cell population can be recovered as cells migrate into different tissues or are lost during circulation [32]. In addition, pre-treatment of harvested tissue for experimental measurements can lead to additional loss of cells [33]. To determine the impact of incomplete sampling on the quantification of cellular kinetics we repeated our analysis but only considered a fraction of the cell population at the time point of sampling in the estimation procedure. For simplicity, we assumed that all cells survived the adoptive transfer.

Given homeostatic cell turnover, not accounting for incomplete sampling results in an underestimation of both the proliferation, ρ , and the death rate, δ (Fig 4B). This bias decreases with increasing sampling fractions. However, in comparison to a scenario with incomplete transfer, the relative bias of the death rate is on average substantially smaller than the relative bias of the proliferation rate (compare Fig 2B). These observations can be seen for both labelling strategies used.

In contrast, more diverse effects are observed when quantifying the cellular dynamics in the complex expansion model (Fig 4C). Incomplete sampling seems to affect the two types of

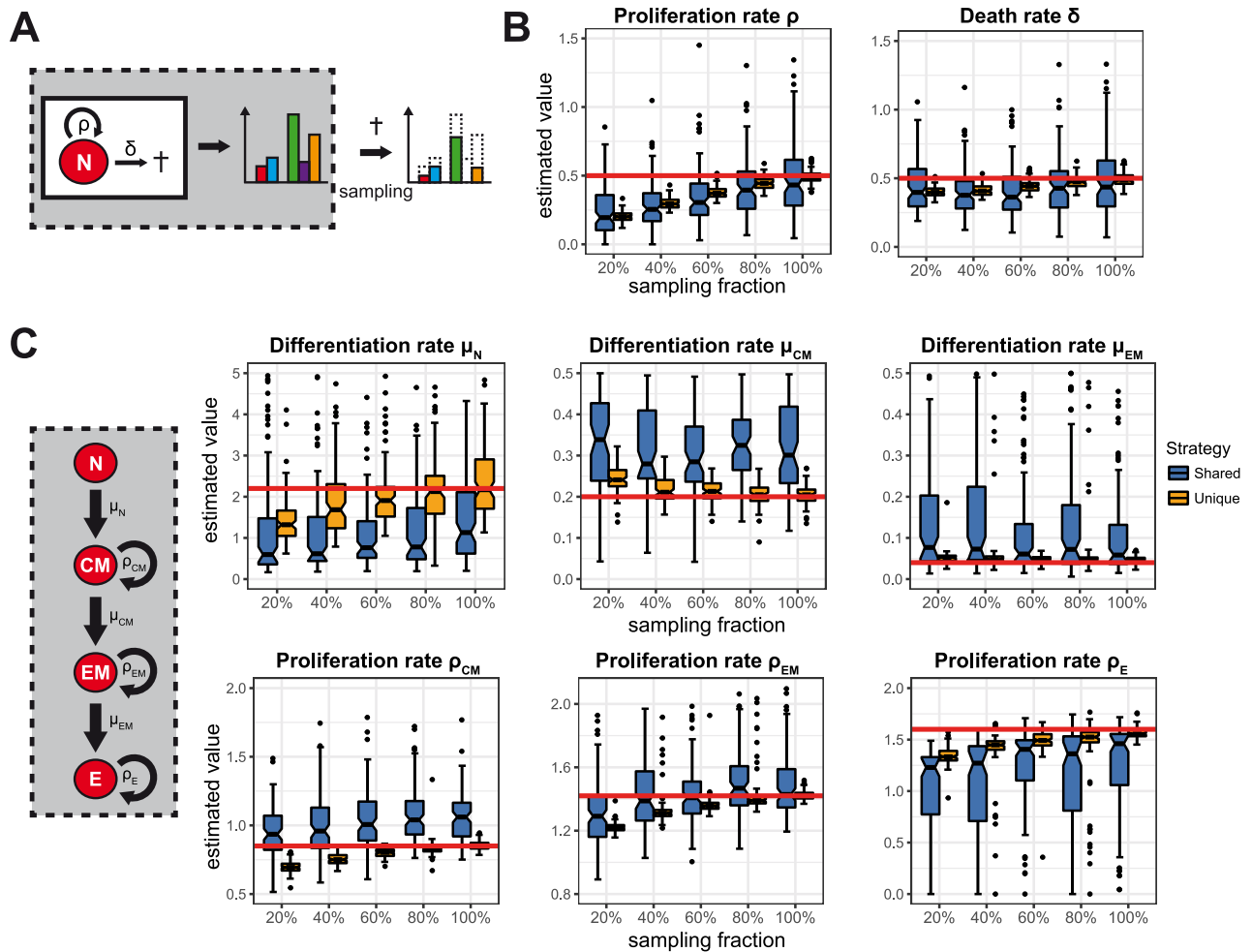


Fig 4. The influence of incomplete sampling on parameter estimation. (A) Schematic depicting the problem of incomplete sampling: Only a fraction of the labelled cells is sampled and can be used for analysis. (B) Panels show the distribution of the estimated proliferation, ρ , and death rate, δ , for the homoeostatic system shown in (A) using the shared (blue) and unique (orange) labelling approach given different sampling fractions. (C) The distribution of the estimated parameters for the complex expansion dynamics analogous to (B). Every boxplot is based on the results of 100 individual stochastic simulations. Parameters used to simulate the dynamics and the time point of sampling are defined as before. Red lines indicate the true parameter values.

<https://doi.org/10.1371/journal.pone.0185523.g004>

rates, i.e. differentiation and proliferation rates, differently: While estimates of the differentiation rates are not affected by different sampling fractions, proliferation rates are generally underestimated. This trend is especially visible for the unique labelling approach, while shared labelling is less affected by incomplete sampling (Fig 4C). However, as already seen for incomplete transfer, the shared labelling approach leads to a biased estimation of the cellular dynamics, especially for non-intermediate compartments, e.g. N, CM and E (Fig 4C).

Similar to the shortcoming of incomplete transfer, incomplete sampling can be addressed in the estimation procedure by rescaling the measured cell numbers by the sampled fraction (see S3 Fig). However, it might be experimentally difficult to obtain an estimate for this fraction.

Thus, while incomplete transfer especially affects the quantification of transition rates, incomplete sampling particularly leads to underestimation of the proliferation rates.

The influence of labelling strategies on parameter identification

Our previous analyses indicate that the composition of the labelled cell population affects parameter estimates. The unique labelling approach leads to more robust and less biased estimates than the shared labelling strategy (Figs 2–4). This increased robustness is expected, as the unique labelling strategy provides up to 800 individual measurements, i.e. one for each label. This is 100-fold the number we obtain when using the shared labelling approach. However, the latter strategy might still comprise useful aspects, because less labels are lost by stochastic effects or during sampling due to the larger number of cells per label. In addition, larger population sizes usually allow more robust experimental measurements.

In order to investigate the qualitative influence of different labelling strategies on the quantification of cellular dynamics, we studied a system of simple cell expansion in which transferred cells, N , are activated with an activation rate μ and activated cells, A , proliferate with rate ρ (Fig 1A) [25]. Here, we analysed the impact of different factors on the ability to infer the cellular kinetics. This included (i) the actual labelling strategy for the transferred cell population characterised by the number of labels, L , and the number of cells per label, M , (ii) the sampling time, T , and (iii) the activation, μ , and proliferation rate, ρ that determine the cellular dynamics. To focus on the impact of each individual factor, we always assumed complete transfer and sampling.

Influence of the labelling strategy. By varying the number of labels, L , from 2 to 50 and the number of cells per label, M , from 1 to 50, we assessed the influence of a total of 2450 different labelling strategies on their ability to quantify the cellular turnover.

We find that increasing the number of labels, L , continuously improves the estimation quality for both the activation and the proliferation rate (Fig 5). The absolute bias, as well as the false coverage rate, i.e. the probability that the actual rate is not within the calculated confidence interval, is reduced. Increasing the number of cells per label, M , only improves the robustness of parameter estimates judged by a decreasing mean confidence interval length (MCIL). Here, we observe a sharp decline between a strategy using unique labels and those relying on multiple cells per label. However, this effect quickly saturates in our scenario with increasing label sizes.

In some cases, confidence intervals for the activation rate, μ , cannot be obtained and the MCIL cannot be calculated. This is indicated by grey colour in the plots. In these cases, all activation rates above a certain threshold are equally likely to generate the observed outcome, leading to unlimited confidence intervals. This effect is mostly limited to labelling strategies with a low number of labels, L , but is also observed for unique labelling approaches having an intermediate number of labels (Fig 5C).

In summary, these results argue for the use of a large number of labels with medium numbers of cells per label as a reliable and robust labelling strategy.

Influence of the distribution of labels. Our previous analyses indicate that a large number of different labels reduces estimation bias while the use of larger label sizes generally improves the robustness of parameter estimates. As only a limited number of cells can be transferred, this leads to the question if estimation can be improved by a combination of both approaches. For example, does a strategy relying on many labels with small label sizes and few labels with more cells per label perform better than one using unique labels for all cells?

To address this question, we repeated our analyses by using a fixed total number of cells that were labelled with L different markers using either a uniformly, linearly or exponentially distributed label size (Fig 6A). The evaluation of data from non-uniformly distributed label sizes required the adaptation of our approach for the calculation of the corresponding summary statistics (S1 Appendix). To compare the performance of the different labelling strategies,

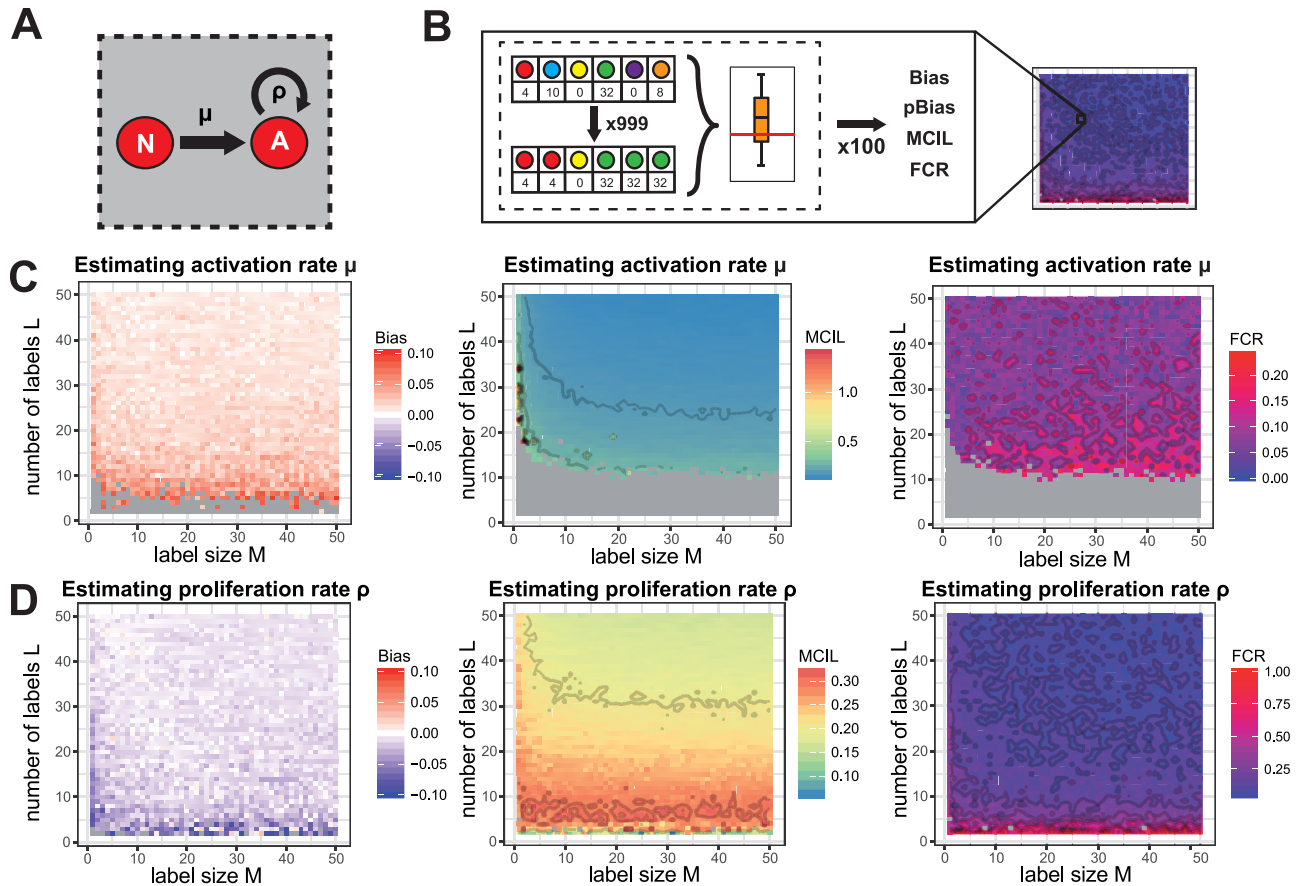


Fig 5. The influence of the labelling strategy on parameter estimates. (A) Schematic depicting the dynamics of the simple expansion model: Cells are activated with rate μ and activated cells proliferate with rate ρ . (B) Calculation of heatmaps: Each labelling strategy is used to generate 100 stochastically simulated data samples. Each data sample is then bootstrapped with 999 repeats (see [Materials and Methods](#)) to calculate the corresponding distribution of parameter estimates and the respective confidence interval. Combining these results allows the calculation of the depicted statistical quantities for the corresponding parameter combination. (C-D) The bias, the mean confidence interval length and the false coverage rate for the estimation of the activation rate, μ (C), and the proliferation rate, ρ (D), assuming a system of simple expansion dynamics. The estimation for each parameter combination is based on 100 independent stochastic simulations. Parameters not varied are fixed to $\mu = 0.3$, $\rho = 0.3$ and cells were sampled at $T = 3$. Grey colour indicates values being above or below the shown range (Bias), or that the method is not able to estimate the respective confidence interval for the corresponding parameter combination (MCIL and FCR, see [Materials and Methods](#)).

<https://doi.org/10.1371/journal.pone.0185523.g005>

we then calculated the difference between the bias, the FCR, and the mean-confidence interval length (e.g. $\Delta_{MCIL} = MCIL_{Uniform} - MCIL_{Linear}$).

We find that a uniformly distributed labelling strategy always performs best in terms of estimation bias and robustness of parameter estimates independent of the total number of cells transferred (Fig 6B–6E for Δ_{MCIL} , plots for pBias and FCR not shown). This observation is consistent for the activation, μ , and proliferation rate ρ . Increasing the inequality between label sizes impairs the quality of parameter estimates as an exponentially distributed labelling strategy always performs worst. Thus, a combination of several uniquely labelled cells with few labels comprising multiple cells does not improve parameter identification compared to an approach based on the same number of labels uniformly distributed among the cells.

The influence of the sampling time. In our scenario, we investigate the impact of various labelling strategies on inferring cellular dynamics if measurements can only be obtained at a

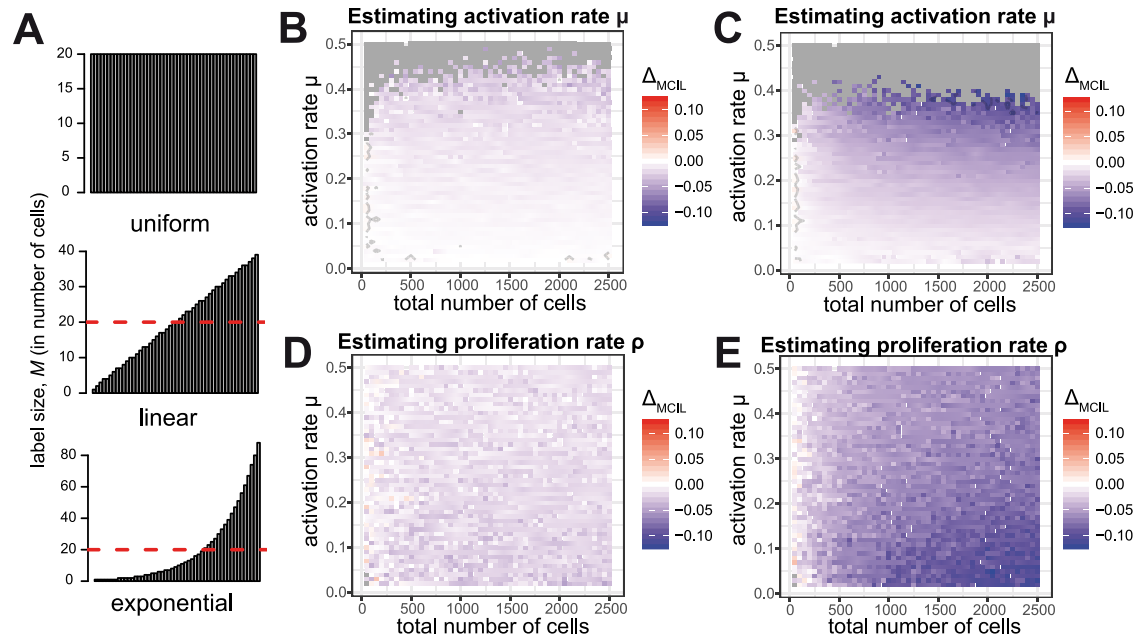


Fig 6. Influence of the label size distribution on parameter estimation. (A) Examples of the three different distributions of label sizes investigated: Uniform, linear and exponential distribution (from top to bottom). Each distribution comprises a total of 1000 cells and 50 labels. The red dotted line indicates the average label size of 20 cells. (B-C) The difference in the MCIL for the estimation of the activation rate μ between the uniformly and linearly distributed labels (B), and the difference between uniformly and exponentially distributed labels (C). (D, E) Analogous to (B, C) the difference in the MCIL for the estimated proliferation rate ρ .

<https://doi.org/10.1371/journal.pone.0185523.g006>

single time point. Thereby, the choice of this sampling time point, T , also has an influence on the ability to estimate the kinetics.

For example, if in our scenario of cell activation and subsequent proliferation the sampling time point is chosen too late, the activation rate μ cannot be reliably estimated (Fig 7A). In contrast, sampling too early leads to increased uncertainty in the estimates due to stochastic effects. Thus, sampling at an intermediate time point gives the most reliable estimates for the

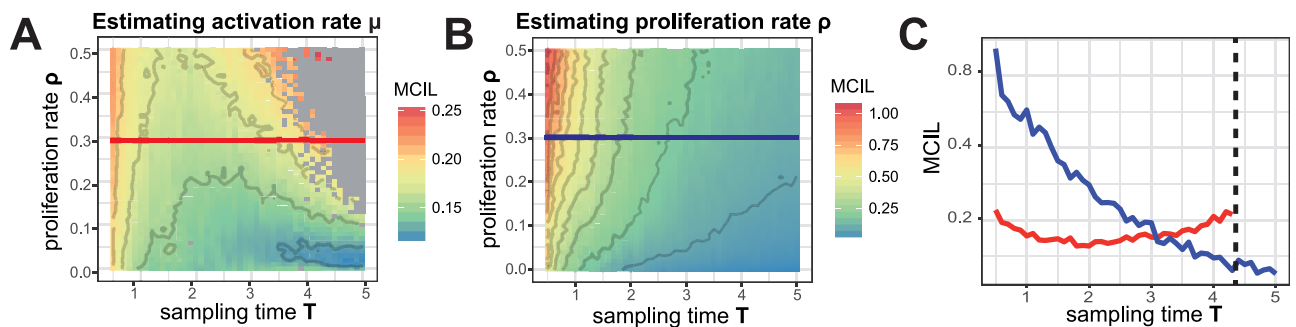


Fig 7. Influence of the sampling time on parameter estimation. MCIL of the activation rate, μ (A), and proliferation rate, ρ (B), using varying combinations of sampling times, T , and proliferation rates, ρ , in the simple system of cell activation and proliferation. Panel (C) shows the cross sections of panels (A & B) indicating the MCIL of the activation rate, μ , (red) and the proliferation rate, ρ , (blue) dependent on the sampling time for a fixed proliferation rate ($\rho = 0.3$). The black dotted line defines the time after which the estimation of the activation rate failed as all labels were sampled. The estimation for each parameter combination is based on 100 independent stochastic simulations. Parameters that were kept fixed are $\mu = 0.3$, $L = 50$ and $M = 5$. Grey colour indicates that the method is not able to estimate the respective confidence interval for the corresponding parameter combination (see Materials and Methods).

<https://doi.org/10.1371/journal.pone.0185523.g007>

activation rate. In contrast, for the proliferation rate ρ we observe that a later sampling time continuously improves robustness of the estimates (Fig 7B, MCIL) and parameter identification, i.e. leading to a reduced percentage bias (S4 Fig). Thus, there is a trade-off regarding the time point of sampling leading to more certainty in the estimates for either the activation or the proliferation rate (Fig 7C).

This trade-off is also found in the complex expansion system (see S1 Fig). Here, proliferation rates are estimated more reliably as time increases, while the transition from naïve to central memory precursor cells is captured especially well for early sampling time points. In case of the homeostatic system, no effect of the sampling time on the parameter estimation is observed, and both the proliferation and death rate are estimated reliably irrespective of the sampling time (S1 Fig).

In summary, our results show that the identification of proliferation rates benefits from later sampling times, while initially occurring transition dynamics might already be masked by then. Hence, an appropriate estimation of all involved dynamics might not be possible in many systems.

Discussion

Over the last decades, technical advances have steadily increased the possibilities to label cells by specific markers. As of today, a huge variety of labelling methods in various levels of detail exists, relying on naturally occurring or artificially induced cellular markers. These methods have been applied in adoptive transfer experiments to quantify cellular differentiation and expansion dynamics (reviewed in [6, 21]). However, to which extent the various labelling strategies actually allow the appropriate identification and quantification of the processes characterising the cellular dynamics has not been systematically studied.

To address this question, we simulated adoptive transfer experiments with various labelling strategies for different scenarios of cellular turnover. These scenarios included homeostatic cell proliferation, as well as simple and complex expansion dynamics involving several steps of cell differentiation. We particularly focused on the situation where only one single measurement can be obtained, as e.g. due to organ harvest [2, 16].

In general, we found that a larger number of labels continuously improves parameter estimation in all of the different models tested. This is not completely surprising, as each label provides an additional measurement that can be used in the analysis and, thereby, reduces estimation bias and variance.

Testing two extreme labelling strategies involving either the transfer of 800 uniquely labelled cells [2] or using only 8 labels with 100 cells each [16], we found that the complexity of the system influences the required number of labels. Both labelling strategies showed a similar average bias for the quantification of cell proliferation and death within a homeostatic model, with the unique labelling approach leading to less variation (Figs 2 and 4). However, within a system of complex cell expansion and differentiation dynamics as considered by Buchholz et al. [16], a shared labelling approach similar to the one used in their experiment generally leads to a slight systematic bias when estimating cell differentiation and proliferation rates. Only rates associated with intermediate compartments for which measurements of the previous and subsequent differentiation steps can be obtained (i.e. EM compared to CM and E, as N was not measured) can be reliably identified (Figs 3 and 4). In addition, the relative relationship between the cell proliferation and expansion rates of the different compartments could not be recovered in the estimates. This suggests that previous estimates for the proliferation and differentiation dynamics of T cells [16] should be taken with care as the labelling approach might be insufficient to determine those rates reliably.

Unique labelling is usually preferred to infer lineage differentiation pathways, such as for immune cell differentiation [1, 2, 16, 26] or hematopoiesis [34–36]. However, our analysis indicates that unique labelling is not always the best approach when estimating cellular turnover or expansion. Estimates on the proliferation dynamics are more robust if larger label sizes are used. Such label sizes will make the labelled population less prone to stochastic effects, although this improvement quickly saturates with increasing label sizes—at least for the analysed simple expansion dynamics.

Interestingly, we found that strategies combining labels with smaller and larger label sizes perform worse than uniformly distributed labels. In general, a strategy using uniformly distributed label sizes provided the most reliable results. However, an approach combining unique and large labels can still be beneficial, as the unique labels can be used to estimate the potential fraction of cells lost during transfer [16]. Due to the varying dependency of cell differentiation and proliferation rates on population sizes, a trade-off can be observed with regard to the choice of the sampling time. Proliferation rates belonging to continuously expanding cell compartments are estimated more reliably given later sampling times. However, harvesting cells at a late time point might severely impair the estimation of activation or initial transition dynamics. Hence, a robust estimation for all parameters might not be achievable if only one sampling time point is available.

The possible loss of cells during adoptive transfer or by incomplete sampling are experimental limitations that can strongly impact the quantification of cellular dynamics. Prior-knowledge on these quantities could be used to correct parameter estimation. However, while transfer loss could be experimentally approximated by using unique labels [16], determining the actual fraction of cells sampled remains difficult. Experimental methods might bias measurements against certain cellular subsets and thereby underestimate the total cell population [32]. However, sampling only a fraction of cells particularly affected cellular expansion rates while transition rates remained mostly identifiable.

In our analyses, we generally assumed that each label is independent and stable. Furthermore, we assumed that the label itself does not interfere with the underlying cellular dynamics. While this is appropriate for artificial markers, such as genetic barcodes [2] or unrelated congeneric markers [4, 16], this assumption will most likely be violated in case of naturally occurring markers, such as α - and β -chains of T cell receptors (TCR) [20, 37, 38]. Here, the actual β -chain could affect T cell affinity and, thus, influences T cell activation [39, 40]. In addition, due to TCR β -chain rearrangements, these labels might not be considered as stable, impairing the possibility to track populations of cells [41, 42]. Novel analysis methods have to be developed to determine if such markers can still be used to infer cellular dynamics.

In summary, our results suggest that a generally suitable labelling strategy consists of a large number of shared labels, with an intermediate number of cells per label. This approach would likely lead to reliable estimates for different cellular systems, even in the case of incomplete transfer or sampling. In general, assumed model systems should always be tested in the context of the applied experimental labelling strategies in order to validate obtained parameter estimates. Performing a-priori simulations or a-posteriori testing allows to identify potential pitfalls, such as consistent bias or a susceptibility of parameter estimates to incomplete transfer or sampling. More systematic analyses of the relationship between labelling strategies and specific cellular systems are needed to infer appropriate labelling strategies in terms of actual cell numbers. Advances in single cell technologies [43, 44] and cell sorting might provide the necessary techniques to customise labelled cell populations used for adoptive transfer.

Supporting information

S1 Appendix. Mathematical derivation of estimation methods.

(PDF)

S1 Fig. Parameter estimates for the homoeostatic and complex expansion system given different sampling times.

(PDF)

S2 Fig. Parameter estimates for the complex expansion system corrected by the transfer fraction.

(PDF)

S3 Fig. Parameter estimates for the homoeostatic and complex expansion system corrected by the sampling fraction.

(PDF)

S4 Fig. Influence of the sampling time on the estimation quality given the simple expansion model.

(PDF)

Author Contributions

Conceptualization: Frederik Graw.

Formal analysis: Michael Gabel, Frederik Graw.

Investigation: Michael Gabel, Roland R. Regoes, Frederik Graw.

Methodology: Michael Gabel, Frederik Graw.

Software: Michael Gabel.

Visualization: Michael Gabel, Frederik Graw.

Writing – original draft: Michael Gabel, Frederik Graw.

Writing – review & editing: Michael Gabel, Roland R. Regoes, Frederik Graw.

References

1. Perie L, Hodgkin PD, Naik SH, Schumacher TN, de Boer RJ, Duffy KR. Determining lineage pathways from cellular barcoding experiments. *Cell Rep*. 2014; 6(4):617–624. <https://doi.org/10.1016/j.celrep.2014.01.016> PMID: 24508463
2. Gerlach C, Rohr JC, Perie L, van Rooij N, van Heijst JW, Velds A, et al. Heterogeneous differentiation patterns of individual CD8+ T cells. *Science*. 2013; 340(6132):635–639. <https://doi.org/10.1126/science.1235487> PMID: 23493421
3. Ganusov VV, Auerbach J. Mathematical modeling reveals kinetics of lymphocyte recirculation in the whole organism. *PLoS Comput Biol*. 2014; 10(5):e1003586. <https://doi.org/10.1371/journal.pcbi.1003586> PMID: 24830705
4. Kaiser P, Slack E, Grant AJ, Hardt WD, Regoes RR. Lymph node colonization dynamics after oral *Salmonella Typhimurium* infection in mice. *PLoS Pathog*. 2013; 9(9):e1003532. <https://doi.org/10.1371/journal.ppat.1003532> PMID: 24068916
5. Kaiser P, Regoes RR, Dolowschiak T, Wotzka SY, Lengefeld J, Slack E, et al. Cecum lymph node dendritic cells harbor slow-growing bacteria phenotypically tolerant to antibiotic treatment. *PLoS Biol*. 2014; 12(2):e1001793. <https://doi.org/10.1371/journal.pbio.1001793> PMID: 24558351
6. De Boer RJ, Perelson AS. Quantifying T lymphocyte turnover. *J Theor Biol*. 2013; 327:45–87. <https://doi.org/10.1016/j.jtbi.2012.12.025> PMID: 23313150
7. Tough DF, Sprent J. Turnover of naive- and memory-phenotype T cells. *J Exp Med*. 1994; 179(4):1127–1135. <https://doi.org/10.1084/jem.179.4.1127> PMID: 8145034

8. Mohri H, Bonhoeffer S, Monard S, Perelson AS, Ho DD. Rapid turnover of T lymphocytes in SIV-infected rhesus macaques. *Science*. 1998; 279(5354):1223–1227. <https://doi.org/10.1126/science.279.5354.1223> PMID: 9469816
9. Hellerstein M, Hanley MB, Cesar D, Siler S, Papageorgopoulos C, Wieder E, et al. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. *Nat Med*. 1999; 5(1):83–89. <https://doi.org/10.1038/4772> PMID: 9883844
10. Ribeiro RM, Mohri H, Ho DD, Perelson AS. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proc Natl Acad Sci USA*. 2002; 99(24):15572–15577. <https://doi.org/10.1073/pnas.242358099> PMID: 12434018
11. Mohri H, Perelson AS, Tung K, Ribeiro RM, Ramratnam B, Markowitz M, et al. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J Exp Med*. 2001; 194(9):1277–1287. <https://doi.org/10.1084/jem.194.9.1277> PMID: 11696593
12. Lyons AB. Analysing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution. *J Immunol Methods*. 2000; 243(1–2):147–154. [https://doi.org/10.1016/S0022-1759\(00\)00231-3](https://doi.org/10.1016/S0022-1759(00)00231-3) PMID: 10986412
13. Yates A, Chan C, Strid J, Moon S, Callard R, George AJ, et al. Reconstruction of cell population dynamics using CFSE. *BMC Bioinformatics*. 2007; 8:196. <https://doi.org/10.1186/1471-2105-8-196> PMID: 17565685
14. Shen FW, Saga Y, Litman G, Freeman G, Tung JS, Cantor H, et al. Cloning of Ly-5 cDNA. *Proc Natl Acad Sci USA*. 1985; 82(21):7360–7363. <https://doi.org/10.1073/pnas.82.21.7360> PMID: 3864163
15. Kearney ER, Pape KA, Loh DY, Jenkins MK. Visualization of peptide-specific T cell immunity and peripheral tolerance induction in vivo. *Immunity*. 1994; 1(4):327–339. [https://doi.org/10.1016/1074-7613\(94\)90084-1](https://doi.org/10.1016/1074-7613(94)90084-1) PMID: 7889419
16. Buchholz VR, Flossdorf M, Hensel I, Kretschmer L, Weissbrich B, Graf P, et al. Disparate individual fates compose robust CD8+ T cell immunity. *Science*. 2013; 340(6132):630–635. <https://doi.org/10.1126/science.1235454> PMID: 23493420
17. Maryanski JL, Jongeneel CV, Bucher P, Casanova JL, Walker PR. Single-cell PCR analysis of TCR repertoires selected by antigen in vivo: a high magnitude CD8 response is comprised of very few clones. *Immunity*. 1996; 4(1):47–55. [https://doi.org/10.1016/S1074-7613\(00\)80297-6](https://doi.org/10.1016/S1074-7613(00)80297-6) PMID: 8574851
18. Lin MY, Welsh RM. Stability and diversity of T cell receptor repertoire usage during lymphocytic choriomeningitis virus infection of mice. *J Exp Med*. 1998; 188(11):1993–2005. <https://doi.org/10.1084/jem.188.11.1993> PMID: 9841914
19. Turner SJ, Diaz G, Cross R, Doherty PC. Analysis of clonotype distribution and persistence for an influenza virus-specific CD8+ T cell response. *Immunity*. 2003; 18(4):549–559. [https://doi.org/10.1016/S1074-7613\(03\)00087-6](https://doi.org/10.1016/S1074-7613(03)00087-6) PMID: 12705857
20. Blattman JN, Sourdive DJ, Murali-Krishna K, Ahmed R, Altman JD. Evolution of the T cell repertoire during primary, memory, and recall responses to viral infection. *J Immunol*. 2000; 165(11):6081–6090. <https://doi.org/10.4049/jimmunol.165.11.6081> PMID: 11086040
21. Schumacher TN, Gerlach C, van Heijst JW. Mapping the life histories of T cells. *Nat Rev Immunol*. 2010; 10(9):621–631. <https://doi.org/10.1038/nri2822> PMID: 20689559
22. Schepers K, Swart E, van Heijst JW, Gerlach C, Castrucci M, Sie D, et al. Dissecting T cell lineage relationships by cellular barcoding. *J Exp Med*. 2008; 205(10):2309–2318. <https://doi.org/10.1084/jem.20072462> PMID: 18809713
23. Naik SH, Schumacher TN, Perie L. Cellular barcoding: a technical appraisal. *Exp Hematol*. 2014; 42(8):598–608. <https://doi.org/10.1016/j.exphem.2014.05.003> PMID: 24996012
24. De Boer RJ, Homann D, Perelson AS. Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *J Immunol*. 2003; 171(8):3928–3935. <https://doi.org/10.4049/jimmunol.171.8.3928> PMID: 14530309
25. De Boer RJ, Oprea M, Antia R, Murali-Krishna K, Ahmed R, Perelson AS. Recruitment times, proliferation, and apoptosis rates during the CD8(+) T-cell response to lymphocytic choriomeningitis virus. *J Virol*. 2001; 75(22):10663–10669.
26. Kaech SM, Wherry EJ. Heterogeneity and cell-fate decisions in effector and memory CD8+ T cell differentiation during viral infection. *Immunity*. 2007; 27(3):393–405. <https://doi.org/10.1016/j.immuni.2007.08.007> PMID: 17892848
27. Gillespie D, Cao Y, Petzold L. Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*. 2006; 124(4):044109–044109–11. <https://doi.org/10.1063/1.2159468> PMID: 16460151
28. R Development Core Team. R: A Language and Environment for Statistical Computing; 2006. Available from: <http://www.R-project.org>.

29. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med*. 2000; 19(9):1141–1164. [https://doi.org/10.1002/\(SICI\)1097-0258\(20000515\)19:9%3C1141::AID-SIM479%3E3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9%3C1141::AID-SIM479%3E3.0.CO;2-F) PMID: 10797513
30. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Stat Med*. 2006; 25(24):4279–4292. <https://doi.org/10.1002/sim.2673> PMID: 16947139
31. Moon JJ, Chu HH, Hataye J, Pagan AJ, Pepper M, McLachlan JB, et al. Tracking epitope-specific T cells. *Nat Protoc*. 2009; 4(4):565–581. <https://doi.org/10.1038/nprot.2009.9> PMID: 19373228
32. Steinert EM, Schenkel JM, Fraser KA, Beura LK, Manlove LS, Igyarto BZ, et al. Quantifying Memory CD8 T Cells Reveals Regionalization of Immunosurveillance. *Cell*. 2015; 161(4):737–749. <https://doi.org/10.1016/j.cell.2015.03.031> PMID: 25957682
33. Hawkins ED, Turner ML, Dowling MR, van Gend C, Hodgkin PD. A model of immune regulation as a consequence of randomized lymphocyte division and death times. *Proc Natl Acad Sci USA*. 2007; 104(12):5032–5037. <https://doi.org/10.1073/pnas.0700026104> PMID: 17360353
34. Cvejic A. Mechanisms of fate decision and lineage commitment during haematopoiesis. *Immunol Cell Biol*. 2016; 94(3):230–235. <https://doi.org/10.1038/icb.2015.96> PMID: 26526619
35. Hofer T, Busch K, Klapproth K, Rodewald HR. Fate Mapping and Quantitation of Hematopoiesis In Vivo. *Annu Rev Immunol*. 2016; 34:449–478. <https://doi.org/10.1146/annurev-immunol-032414-112019> PMID: 27168243
36. Nguyen LV, Makarem M, Carles A, Moksa M, Kannan N, Pandoh P, et al. Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells. *Cell Stem Cell*. 2014; 14(2):253–263. <https://doi.org/10.1016/j.stem.2013.12.011> PMID: 24440600
37. Wong J, Mathis D, Benoist C. TCR-based lineage tracing: no evidence for conversion of conventional into regulatory T cells in response to a natural self-antigen in pancreatic islets. *J Exp Med*. 2007; 204(9):2039–2045. <https://doi.org/10.1084/jem.20070822> PMID: 17724131
38. Zarnitsyna VI, Evavold BD, Schoettle LN, Blattman JN, Antia R. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front Immunol*. 2013; 4:485. <https://doi.org/10.3389/fimmu.2013.00485> PMID: 24421780
39. Stone JD, Chervin AS, Kranz DM. T-cell receptor binding affinities and kinetics: impact on T-cell activity and specificity. *Immunology*. 2009; 126(2):165–176. <https://doi.org/10.1111/j.1365-2567.2008.03015.x> PMID: 19125887
40. Chervin AS, Stone JD, Soto CM, Engels B, Schreiber H, Roy EJ, et al. Design of T-cell receptor libraries with diverse binding properties to examine adoptive T-cell responses. *Gene Ther*. 2013; 20(6):634–644. <https://doi.org/10.1038/gt.2012.80> PMID: 23052828
41. Penit C, Lucas B, Vasseur F. Cell expansion and growth arrest phases during the transition from precursor (CD4-8-) to immature (CD4+8+) thymocytes in normal and genetically modified mice. *J Immunol*. 1995; 154(10):5103–5113. PMID: 7730616
42. Rohr JC, Gerlach C, Kok L, Schumacher TN. Single cell behavior in T cell differentiation. *Trends Immunol*. 2014; 35(4):170–177. <https://doi.org/10.1016/j.it.2014.02.006> PMID: 24657362
43. Chattopadhyay PK, Gierahn TM, Roederer M, Love JC. Single-cell technologies for monitoring immune systems. *Nat Immunol*. 2014; 15(2):128–135. <https://doi.org/10.1038/ni.2796> PMID: 24448570
44. Proserpio V, Lonnberg T. Single-cell technologies are revolutionizing the approach to rare cells. *Immunol Cell Biol*. 2016; 94(3):225–229. <https://doi.org/10.1038/icb.2015.106> PMID: 26620630