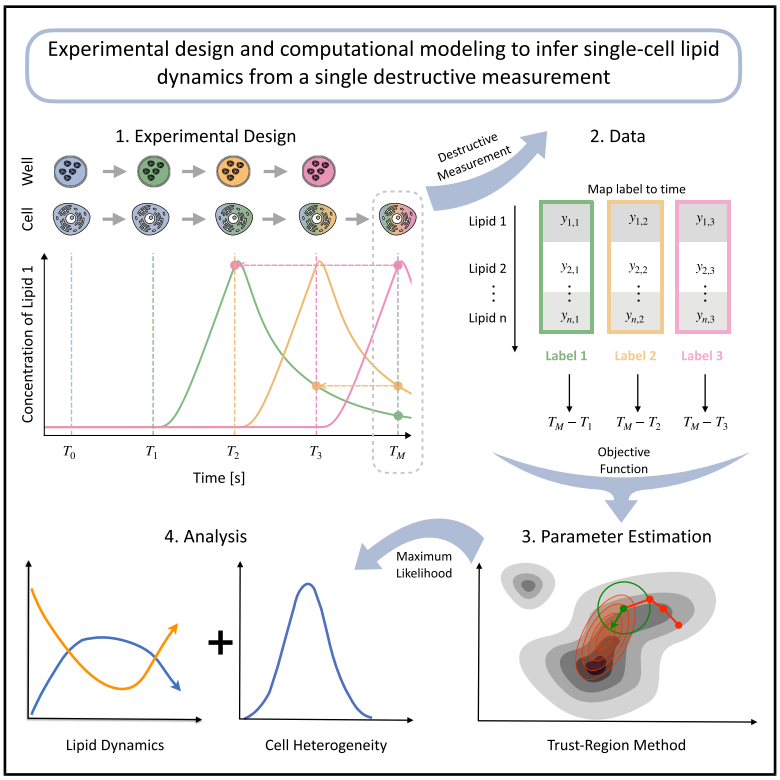
iScience

Article

Suggested experimental design and computational modeling to infer single-cell lipid dynamics from a single destructive measurement

Graphical abstract



Highlights

- Multi-label pulse-chase enables trajectory reconstruction from a single measurement
- Label-aware mechanistic ODEs capture dynamics while sharing kinetics across labels
- Maximum likelihood estimation enables quantification of single-cell parameters
- An increase in the number of labels improves parameter identifiability

Authors

Paul Jonas Jost, Daniel Weindl, Klaus Wunderling, Christoph Thiele, Jan Hasenauer

Correspondence

jan.hasenauer@uni-bonn.de

In brief

Biocomputational method; Biological sciences; Biological sciences research methodologies; Computing methodology; In silico biology; Lipidomics; Omics

Jost et al., 2026, iScience *29*, 115474
April 17, 2026
https://doi.org/10.1016/j.isci.2026.115474

CellPress

iScience

CellPress
OPEN ACCESS

Article

Suggested experimental design and computational modeling to infer single-cell lipid dynamics from a single destructive measurement

Paul Jonas Jost,[1,2] Daniel Weindl,[1,2] Klaus Wunderling,[2] Christoph Thiele,[2] and Jan Hasenauer[1,2,3,*]
[1]Bonn Center for Mathematical Life Sciences, University of Bonn, Bonn, Germany
[2]LIMES Life and Medical Sciences Institute, University of Bonn, Bonn, Germany
[3]Lead contact
*Correspondence: jan.hasenauer@uni-bonn.de
https://doi.org/10.1016/j.isci.2026.115474

SUMMARY

Cellular heterogeneity is a fundamental facet of cell biology, influencing cellular signaling, metabolism, and gene regulation. Its accurate quantification requires measurements at the single-cell level. Most high-throughput single-cell technologies provide only a snapshot of cellular heterogeneity at a specific time point because the measurement is destructive. This limits our current ability to understand the dynamics of cellular behavior and quantify cell-specific parameters. We propose a potential experimental setup combined with a model-based analysis framework, enabling the extraction of longitudinal data from a single destructive measurement. Although broadly applicable, we focus on lipid metabolism, a domain where obtaining longitudinal single-cell data have remained elusive due to technical constraints. Our method leverages multiple labels whose measurements are linked to a shared dynamic. This allows the estimation of cell-specific parameters and the quantification of heterogeneity. This framework establishes a foundation for future investigations, providing a roadmap toward a deeper understanding of dynamic cellular processes.

INTRODUCTION

Cell heterogeneity, the inherent variability among individual cells within a seemingly homogeneous population (Figure 1A), has emerged as a focal point in recent scientific research.[1–3] It can be analyzed based on measurements of a large number of biochemical species through techniques such as single-cell transcriptomics,[4] single-cell proteomics,[5] and single-cell lipidomics.[6] Complementary to this, single-cell imaging technologies have been developed for the longitudinal analysis of a few selected biochemical species.[7] These technologies had a profound impact across various domains, including cancer research,[8,9] immunology,[10] microbiology,[11] and adipocyte physiology.[12]

Most single-cell studies use inherently destructive measurement techniques, which provide information about the state of cells at the time of the measurement.[5,6,8] However, dynamic processes within cells and differences in signal processing, gene regulation, and metabolism mostly remain elusive (Figure 1B). This is problematic as most cellular processes are inherently dynamic, and small changes in rates of reactions can result in large differences, including different fate decisions.

In the field of single-cell transcriptomics, dynamic processes are starting to be addressed by distinguishing mRNAs at different processing stages[13] or metabolic labeling of mRNAs.[14] These approaches provide information about the time point of transcription, enabling velocity calculations and pseudo-time analysis.[14] Although metabolic labeling is well established in bulk proteomics[15,16] and lipidomics,[17,18] on the single-cell level, they currently have no velocity or pseudo-time analysis methods available.

The spectrum of available metabolic labeling techniques is broad. Two important labeling strategies for lipids are stable isotope labeling[19] and alkyne-fatty acid (FA) click labeling.[17] Click-labeling allows reporter molecules to be linked to their targets so that the different labeled species can be easily distinguished, for example, by fluorescence microscopy or mass spectrometry. Different reporters can be clicked on different samples, which are then combined into one measurement to reduce experimental variability. The technique has been successfully utilized with mass spectrometry-based lipidomics to assess population averages.[6,20] Moreover, the study of single cells has been demonstrated,[6] but the destructive nature of the measurement hinders the assessment of single-cell trajectories.

This study is conducted entirely *in silico* and proposes an experimental protocol paired with a model-based analysis framework to assess dynamic lipid processing within individual cells based on a single destructive measurement. The protocol includes using multiple metabolic labels introduced at different time points prior to the measurement. Under the assumption that the metabolic labeling does not influence metabolism, we attribute the measurements for the different labels to a joint dynamic, and thus, provide what can be interpreted as multiple measurements along a trajectory. The information in the measurements is then deconvoluted using a knowledge-based

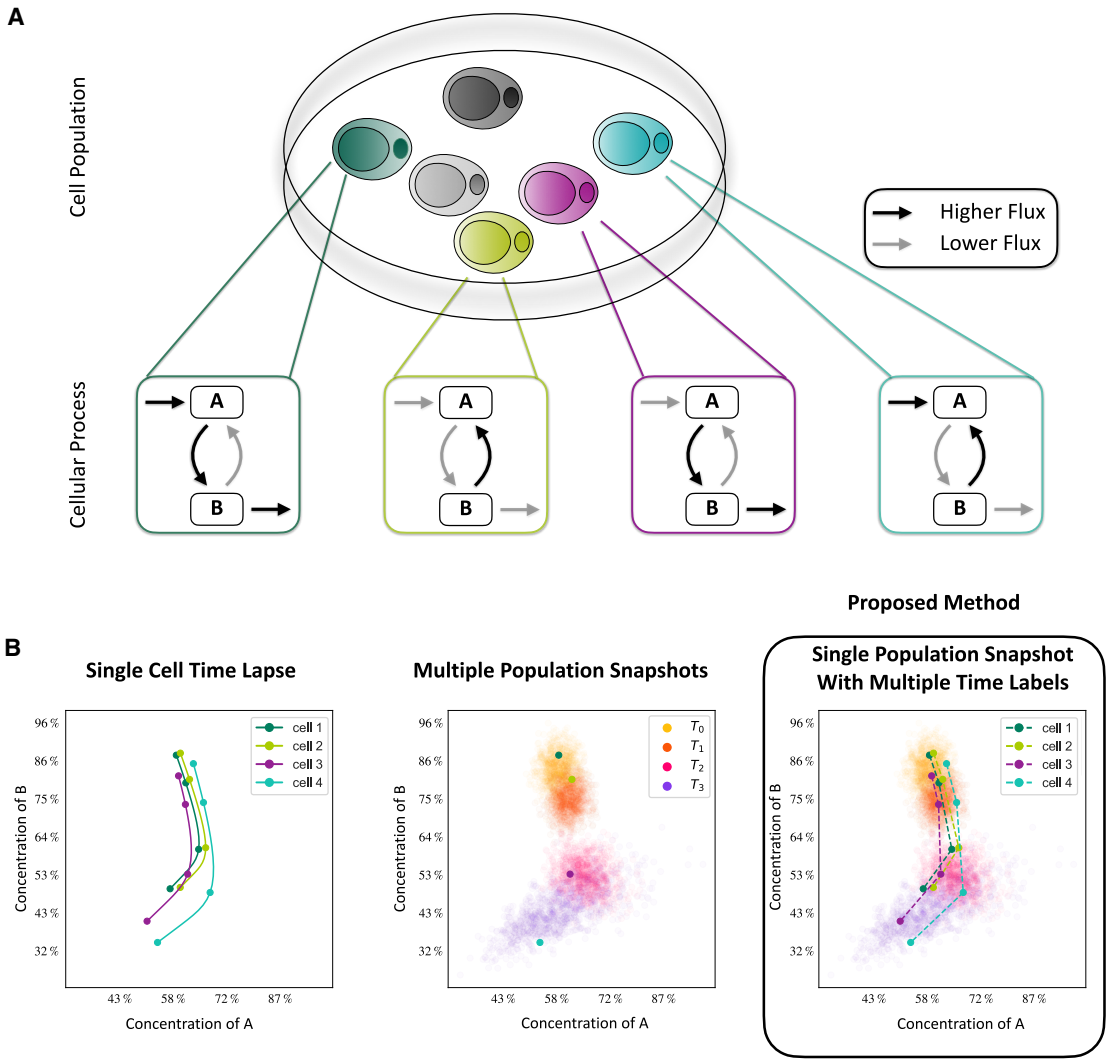


Figure 1. Cell heterogeneity and different types of measurements

(A) Different cells may differ in their cellular process activities, preventing two distinct cells from being used as measurements at two different time points.

(B) To capture the time-lapse of single cells (left), single-cell destructive measurements at different time points can only yield information on the distribution over time and no accurate dynamics (center). Our proposed method resolves this issue, allowing for accurate dynamics of a single cell as well as distributions for the whole population (right).

dynamic model of lipid metabolism. Knowledge-based models are generally preferred over machine learning when data are limited, as they require fewer parameters and allow assigning biological meaning to them. Thus, in contrast to black-box machine learning models, mechanistic models are particularly well suited for hypothesis-driven analysis in data-scarce settings. We show that using this setup, a single destructive measurement can provide information about the state of the dynamic lipid processing in individual cells. Overall, the approach shares similarities with isotopically nonstationary metabolic flux analysis[21] but allows for the analysis of cell-to-cell variability. We use exclusively *in silico* synthetic data as a proof of concept and show that we can accurately recover single-cell parameters, allowing us to quantify cell heterogeneity.

RESULTS

Model-based approach for assessing cellular dynamics based on a single measurement

To assess the cell-to-cell variability of metabolic processing, we propose a simple experimental protocol paired with a model-based analysis framework. The core idea is to provide cells with differently labeled metabolites at different time points and to measure the concentration of the metabolites at a single time point. As the labeled metabolites have remained in the system for varying amounts of time, the labels implicitly provide a time stamp. The model-based approach enables the assessment of the dynamical properties of individual cells by leveraging this implicit information.

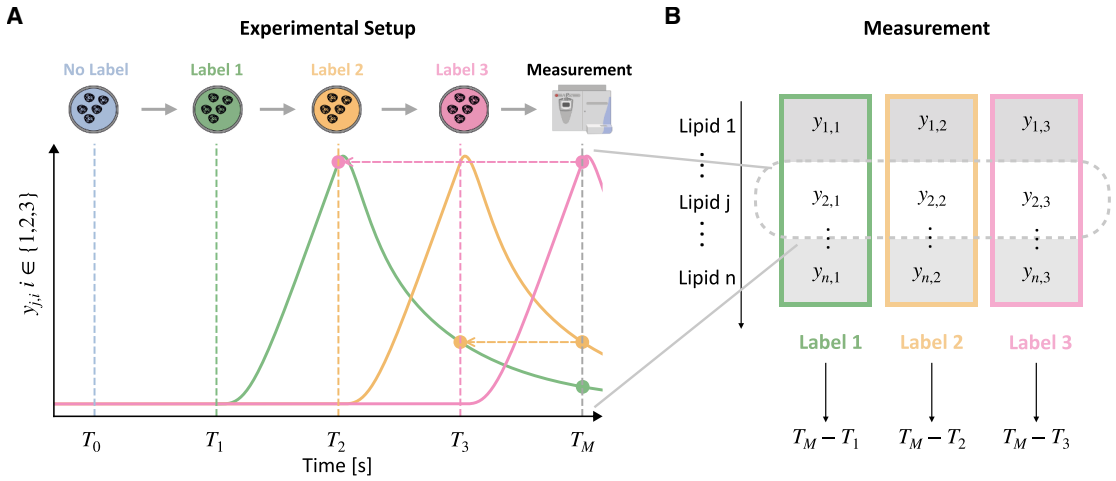


Figure 2. Illustration of proposed experimental setup and measurement

(A) Experimental setup paired with exemplary dynamics of a labeled chemical species for the case of three distinct labels. The experiment starts in the steady state for the label-free growth medium, and then growth media containing labeled species are provided after washing the cells (much like a pulse-chase experiment). The growth media only differ in the label provided and are otherwise identical. This setup results in three time-shifted trajectories of the differently labeled species.

(B) The cells are harvested, and label-sensitive measurements provide a pseudo-time trajectory.

We focus on a specific setup for illustrative purposes. We consider cells in a growth medium, which at specific time points are susceptible to changes within this medium. Specifically, the only perturbation within the medium is the introduction of different labels at different time points. We make two additional assumptions: (1) unlabeled and labeled metabolites possess the same biochemical properties—meaning that reactions taking place within cells are agnostic to labeling, and (2) the metabolite levels (disregarding the label) in the medium are considered to be constant throughout the experiment. In such a setup, the label itself, as well as the availability of labeled metabolites, does not perturb the system, and it will remain in a metabolic steady state, meaning that for any metabolite, the sum over all its unlabeled and labeled forms will stay constant.

The experimental setup and the dynamics of a metabolite are illustrated in Figure 2A. In the case of three labels, the setup is as follows:

(Step 0) cells are incubated with label-free growth medium at time point T_0 for a sufficiently long period to allow the cells to reach a steady state.

(Step 1) at time point T_1 ($\gg T_0$), cells are washed, and a growth medium containing a metabolite with label 1 is applied.

(Step 2) at time point T_2 ($> T_1$), cells are washed, and a growth medium containing the same metabolite with label 2 is applied.

(Step 3) at time point T_3 ($> T_2$), cells are washed, and a growth medium containing the same metabolite with label 3 is applied.

(Step 4) at time point T_M ($> T_3$), cells are harvested and concentrations of the differently labeled (and potentially also the label-free) metabolites are measured.

The experiment can be performed with any number of labels $N_L \geq 1$. As we will discuss, while more labels provide more information about individual cells, they also increase the complexity of the experiment and subsequent analysis.

Notably, under the experimental conditions outlined above, the concentrations of the metabolites that possess only label l are time-shifted compared to the metabolites with label l'. The time shift is equal to the difference between the time points at which the respective labels were applied to the medium, i.e., $T_l - T_{l'}$ (Figure 2B). A mathematical proof for this core result is provided in the supplemental information. The proof of time-shifting confirms that one measurement at a single time point can provide information about a time course if a cell is exposed to a sequence of growth media with different labels. However, extracting the longitudinal information encoded in such measurements requires a label-aware dynamical model, which can easily be generated from a rule-based framework.

In the following, we will show how the proposed experimental setup can be used to assess cell-to-cell variability in metabolic processes and to determine causal differences between cells. Therefore, we assume that experimental data of the proposed setup are available for n_c cells of a heterogeneous cell population, yielding datasets $\mathcal{D}_j$ with $j = 1, \ldots, n_c$. We complement this with dynamic modeling to quantify the cell-to-cell variability. The mathematical model is a set of ordinary differential equations (ODEs) based on the biochemical reactions that are part of the considered metabolic process,

$$R_k : \sum_{i=1}^{n_X} S_{i,k}^{-} X_i \rightarrow \sum_{i=1}^{n_X} S_{i,k}^{+} X_i, \quad k = 1, \ldots, n_R,$$

with biochemical species X_i and stoichiometric coefficients $S_{i,k}^{-}, S_{i,k}^{+}$. The reaction rates are modeled using mass-action kinetics. To account for the inherent cell-to-cell variability, we allow the vector of reaction rate parameters, $\theta = (\theta_1, \ldots, \theta_{n_\theta})^T$, to differ between cells. A common assumption is that it is sampled from a probability distribution, $\theta \sim p(\theta)$,[22] e.g., a multivariate log-normal distribution.[23]

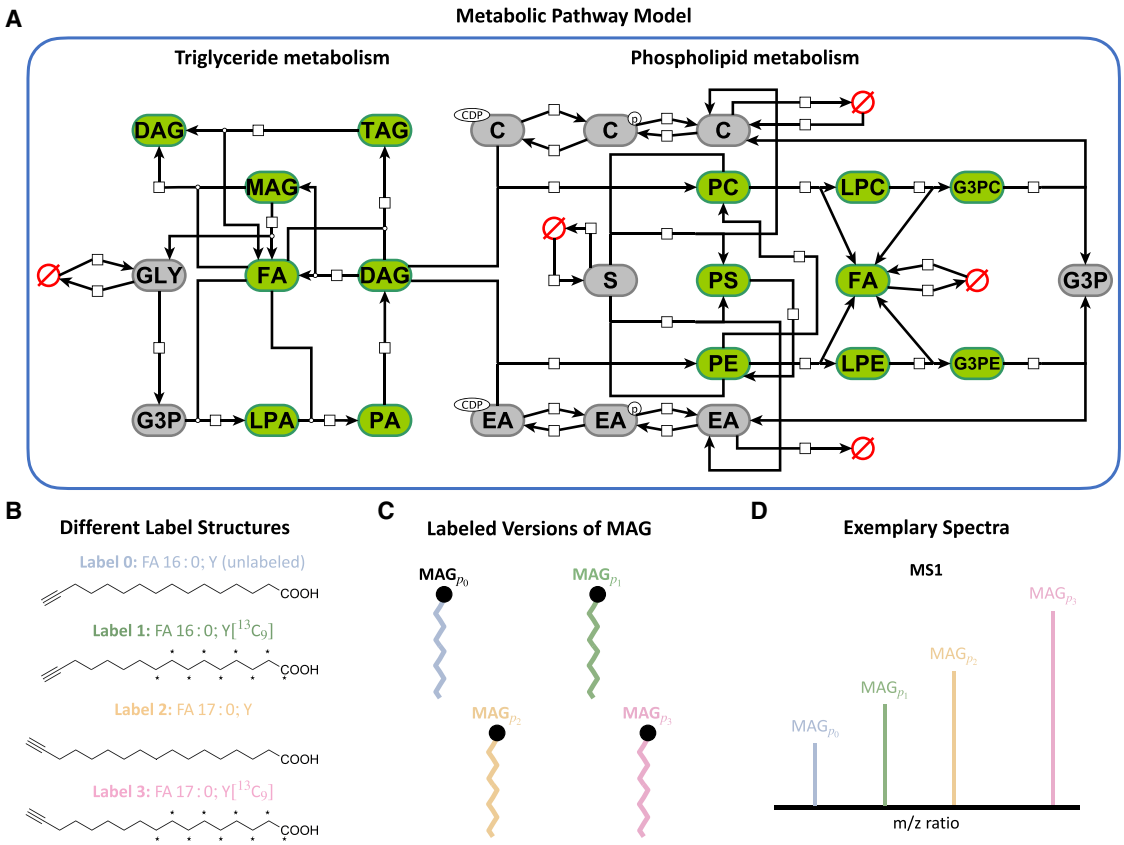


Figure 3. *In silico* mechanistic model of lipid metabolism and metabolic labeling setup

(A) Visualization of the considered model of lipid metabolism in adipocytes, considering the two main parts of triglyceride and phospholipid metabolism (visualization following SBGN[27] notation). Species with green boxes are observed, while species with gray boxes are not observed but are still considered in the model. Red empty sets symbolize the in- and outfluxes of the model. FA was used twice for representation purposes only.

(B) Chemical structures of the three alkyne-fatty acids FA 16:0;Y, [FA 16:0$^{13}C_9$];Y, and FA 17:0;Y, which serve as labels in the experimental setup.

(C) Illustration of the species MAG with different labels. A more extensive example can be found in Figure S1.

(D) Illustration of the mass spectrometry measurement. The labels are chosen so they can already be distinguished in the mass spectrometer.

Given the single-cell datasets and the mathematical model, assessing the (source of the) cell-to-cell variability is equivalent to determining the cell-specific parameters $\theta^{(j)}$ from $\mathcal{D}_j$, for $j = 1,\ldots,n_c$. In this study, we solve this problem using a maximum likelihood approach. Parameter optimization is performed using a multi-start local optimization scheme, while parameter uncertainty analysis is performed using the Fisher information matrix (see STAR Methods section and supplemental information).

Mathematical model for fatty acid metabolism in adipocytes

To assess the possibility of the reconstruction of single-cell dynamics and parameters from single-cell snapshot measurements, we performed a comprehensive *in silico* study. In our model, we consider FA metabolism in adipocytes. This model is suitable from both a technical and biological standpoint. Technically, adipocytes contain large amounts of lipids, which increases the likelihood of surpassing the detection threshold in mass spectrometry-based measurements. Biologically, adipocytes play a key role in energy storage[24] and are known to express high cell-to-cell heterogeneity,[25,26] making them a compelling cell type to assess the benefits of single-cell lipidomics.

For the *in silico* study, we constructed a model for the core components of FA metabolism (Figure 3A). The model accounts for triacylglycerol (TAG) synthesis and lipolysis,[28] accordingly, including glycerol (GLY) and glycerol-3-phosphate (G3P), lysophosphatidic and phosphatidic acids (LPAs, PAs) as well as mono-, di-, and tri-acylglycerols (MAGs, DAGs, and TAGs). Additionally, we include the phospholipids phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylserine (PS), which are key components of cell membranes and in cell signaling. Additional metabolites of the phospholipid metabolism were also incorporated. We simplify the model by assuming constant ATP and CoA levels, subsequently omitting them from the model. Thus, we use FAs directly in reactions involving acyl-CoA. Overall, the model accounts for 22 metabolites and 40 reactions. A comprehensive list is provided in Tables S1 and S2. The reaction rates are assumed to follow mass action kinetics. Hence, each reaction possesses one parameter. Those parameters are generally unknown, and we

use dynamic modeling to estimate them. Cell-to-cell variability can be quantified through their differences across cells.

As the lipid metabolism includes and processes extracellular FAs, we considered an *in silico* experimental setup in which cells are provided with alkyne- FAs functioning as labels. FAs vary in chain length, degree of unsaturation (i.e., number and position of double bonds), and other structural features such as hydroxylation, all of which can influence their biochemical behavior. One of the key factors driving FA mobilization rate is the chain length.[29] Thus we use alkyne-FAs of similar chain length, namely FA 16:0;Y, [FA 16:0$^{13}C_9$];Y, and FA 17:0;Y as labels (Figure 3B), which can be distinguished using mass spectrometry to determine the label-resolved concentrations of lipids or groups of lipids (Figure 3D). In line with this, the model accounts for multiple labels that do not alter the biochemical properties of the molecules but can be distinguished in the measurement process (Figures 3C and 3D). A lipid class carries multiple FAs in the form of esterified fatty acyl chains (two in diacylglycerols and three in triacylglycerols). Accordingly, each lipid class can be associated with as many labels as the number of its fatty acyl chains. For these lipid classes, the number of considered state variables corresponds to the number of possible label combinations. We assume that the position of the label can be disregarded, reducing the number of possible combinations. Additionally, in this study, we do not model lipid beta-oxidation explicitly (see supplemental information for details on the model formulation).

The resulting dynamical model for FA metabolism in adipocytes translates for the case of three labels to an ODE model with 101 metabolites and 263 reactions. The number of metabolites and reactions grows quickly with the number of labels. In contrast, the number of parameters remains constant as the labels are assumed not to affect the processing of metabolites. Considering every metabolite that contains one or more FAs within the model as observable results in a total of 92 observables. For the subsequent assessment, we used this model to generate quantitative noise-corrupted data for different numbers of labels.

Optimization allows for reliable fitting of dynamical models

To evaluate the feasibility of using a model-based approach to assess experimental data collected under the proposed protocol, we conducted an extensive *in silico* study (Figure 4A). Utilizing the mathematical model for FA metabolism in adipocytes, synthetic data were generated for a heterogeneous population comprising $n_c = 1,000$ cells (see STAR Methods section for parameter distribution details), providing a ground truth as a reference for evaluation.

To fit the model to single-cell data, we employed a multi-start local optimization approach. Specifically, we performed 1,000 local optimizations, with initial points sampled randomly within parameter bounds.

The analysis of the computed parameter values revealed variability in the final objective function values across different cells, with the objective function value serving as an indicator of the goodness of fit (Figure 4B, left). This can be attributed to differences in the parameter vectors and noise realizations. Nevertheless, for all cells, we observed a good agreement between the data and model fit—as illustrated for a small subset of the observables (Figure 4B, center). This is especially noteworthy, as we still only have a single measurement per observable, but we were still able to estimate the time-resolved dynamics. Notably, the local optimization exhibited robust convergence, as evidenced by the prominent plateau in the waterfall plot (Figure 4B, right). Over 80% of the local optimization runs converged to the presumed global optimum for most cells. This suggested that the objective function landscape possesses only a tiny number of local minima with a substantial region of attraction and indicated that the number of optimization runs might be reduced if computational constraints arise.

In summary, the proposed parameter optimization pipeline yielded reliable fits for the rule-based model. This achievement is particularly noteworthy given that the model corresponds to an ODE with 101 state variables and 40 estimated parameters, underscoring its robustness and potential applicability.

Dynamic modeling allows for assessment of cell heterogeneity

Given that the proposed parameter optimization pipeline allows for reliable fitting of data, we assessed how well the parameters of individual cells and the parameter distribution on the population level can be reconstructed (Figure 5A). Here, we utilized the fact that the ground truth parameters are known for the considered synthetic data and examined the previously described scenario with three labels.

To evaluate how well the parameters for individual cells were recovered, we compared the correlation between the maximum likelihood estimates and the ground truth values for each cell (Figure 5B). We found Pearson correlation coefficient ranging from -0.049 to 0.998. For nine out of 40 parameters, a correlation above 0.8 was observed, while for 25 out of 40 parameters, the correlation was below 0.2. Interestingly, for the parameters with a low correlation, the estimated values spanned a much wider range, as indicated by the comparison of the empirical distributions of the maximum likelihood estimates and the ground truth distribution (Figure 5C). This indicates practical non-identifiability, which might be related to limited data or structural problems. Indeed, the uncertainties in the estimated parameters can be attributed to unobserved state variables, such as choline (C) and phosphocholine (C_{PH}) (Figure S3A). These unobserved species are interconverted, and the parameters for the conversion rates (k_6, k_{32}) are not well determined (Figure 5C). Notably, the inspection of the multi-start optimization results confirms that there are multiple different optimization endpoints with similar objective function values. Thus, these endpoints appear to be part of a high-dimensional shape (Figures S3B–S3D).

In conclusion, the proposed experimental protocol with three labels in combination with the model-based analysis pipeline allows for the reliable assessment of a substantial fraction of the single-cell parameters. For those identifiable parameters, we were able to approximate the distribution across cells with a slight overestimation of the variance within the distribution.

Quantifying the impact of the number of labels on the assessment of cellular dynamics

As the number of data points influences parameter identifiability and the quality of parameter estimates, we assessed the impact

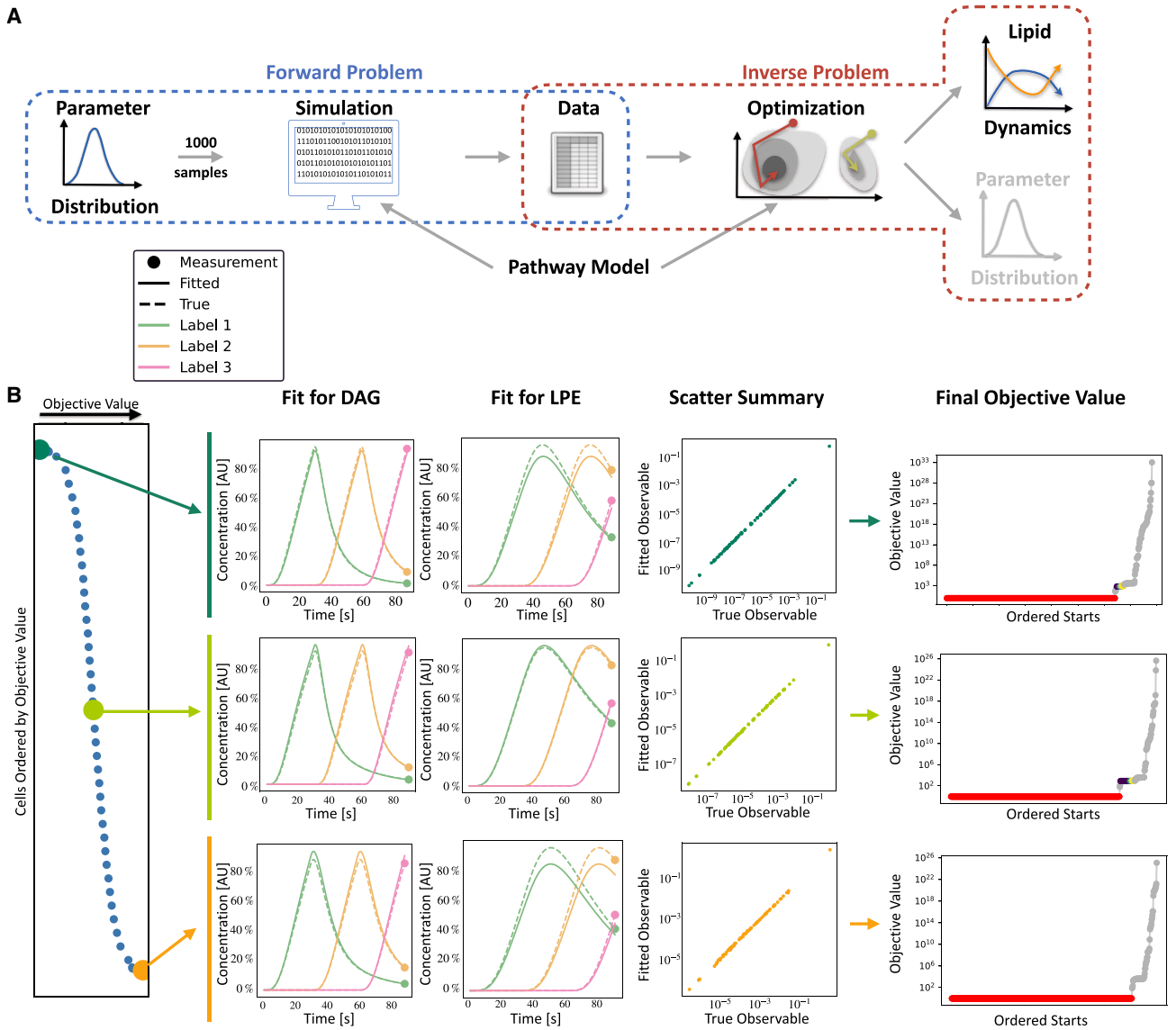


Figure 4. Evaluation setup and assessment of the parameter optimization pipeline

(A) Schematic overview of the workflow: In the forward problem, parameter vectors for individual cells are sampled independently from a multivariate log-normal distribution. For each parameter vector, the model is simulated. The resulting simulated values for each observable are corrupted by adding normally distributed noise. In the inverse problem, for each synthetic dataset, a multi-start local optimization is performed, and the results are analyzed in terms of goodness of fit. The pathway model is used in the forward and inverse problems.

(B) Illustration of fitting results for synthetic datasets: (left) ordering of 1,000 cells by objective function value, with smaller values indicating a better fit. (Middle) best fits for the two selected species, DAG and LPE, for best, median, and worst fit cells. Each color corresponds to a label; solid lines are the fit, and dashed lines are the actual dynamics. Scatterplot summary for three cells with true vs. fitted observable reveals excellent agreement. (Right) waterfall pot for the assessment of the convergence characteristics of the multi-start optimization for three cells.

of the number of employed labels. We considered the use of zero to four labels, assuming that the unlabeled species are in steady state at T_0 and that the labeling time points T_i are uniformly spaced between T_0 and the measurement time T_m (Figure 6A). For each labeling setup, we generated synthetic data for 200 cells, using the same set of 200 sampled parameter vectors across all scenarios. We performed parameter estimation on each of the datasets as described above and quantified the impact of the number of labels on model complexity, computation time, parameter estimation, and parameter identifiability.

The study of the model complexity for different labels revealed that the number of state variables and observables grows rapidly with the number of labels (Figure 6B), with a 10-fold increase in observables from 0 to 4 labels. This increase is driven by the species with the highest number of label configurations, in this case, the TAG. The increased model complexity impacted the

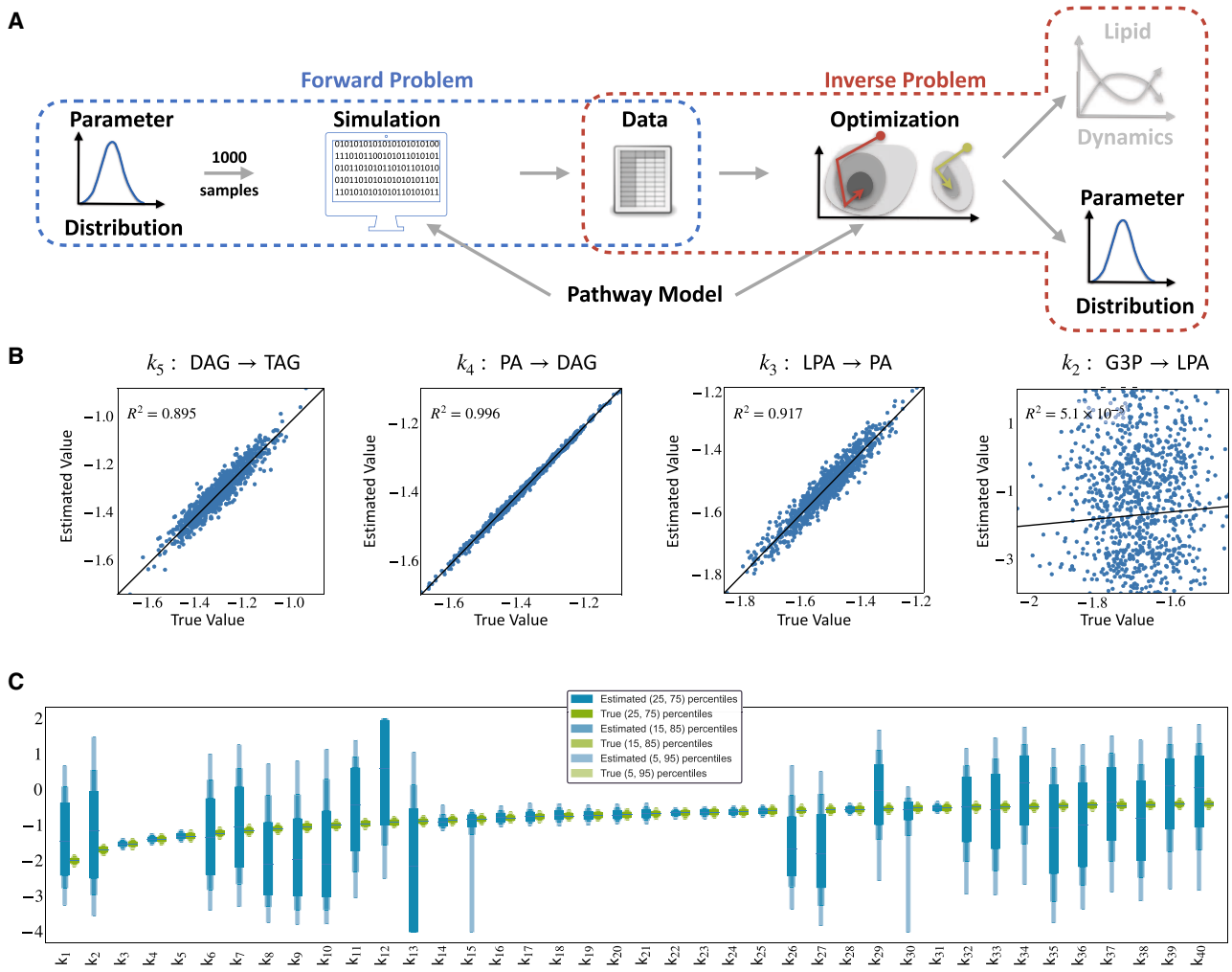


Figure 5. Analysis of single-cell parameters recovered from best local optimization

(A) Inverse problem with the focus on the recovery of parameter distribution.

(B) Scatterplot for parameters of triglyceride synthesis (G3P → TAG). Each point ($n = 1,000$) corresponds to one cell, R^2 has been computed using Pearson correlation of estimated with true parameter value. The black line indicates perfect fit, i.e., estimated = true. G3P is an unobserved state variable, while the other four metabolites are measured.

(C) Boxplots of true (green) and estimated (blue) parameter distributions for each parameter. Boxes are calculated using subsequently increasing percentile ranges around the median (black line). The estimated ones are the best-found parameters for each cell ($n = 1,000$) after optimization. We see that the estimates of half the parameters have a good agreement (with some outliers on the 95 percentiles), while the others span multiple orders of magnitude, see also Figure S3.

computation times required for optimization (Figure 6C), increasing by nearly two orders of magnitude (from 2 min to 170 min per local optimization). Importantly, for the rule-based implementation, which considers the equivalence of the labeling positions, the number of state variables and the computation time grow more slowly than the standard alternative.

To measure the influence of the number of labels on the identifiability of the parameter of individual cells, we used the Fisher information matrix. We computed the Fisher information matrix and its eigenvalue spectrum for the maximum likelihood estimates of the single-cell datasets for each number of labels. Subsequently, we compared the resulting spectra between the number of labels to assess how this number impacts identifiability. The eigenvalue spectrum of the Fisher information matrix provides information about the uncertainty of parameters[30] and hence about the information content of different experimental setups. The analysis of the eigenvalue spectrum showed that many eigenvalues are smaller than 10^{-9}, indicating directions of uncertainty in the parameter space. The setup with zero labels had the most eigenvalues in this range and smaller eigenvalues overall (Figure 6D). Incorporating additional labels shifted the eigenvalue distribution to larger values, signifying reduced uncertainty. The most significant shift occurred from 0 labels to 1 label. The increase from 1 label to 2 labels and from 2 labels to 3 labels resulted in an increase in the median by approximately one order of magnitude. Surprisingly, the increase from 3 labels to 4 labels did not add information, which is likely due to the equidistant spacing of time points. On a

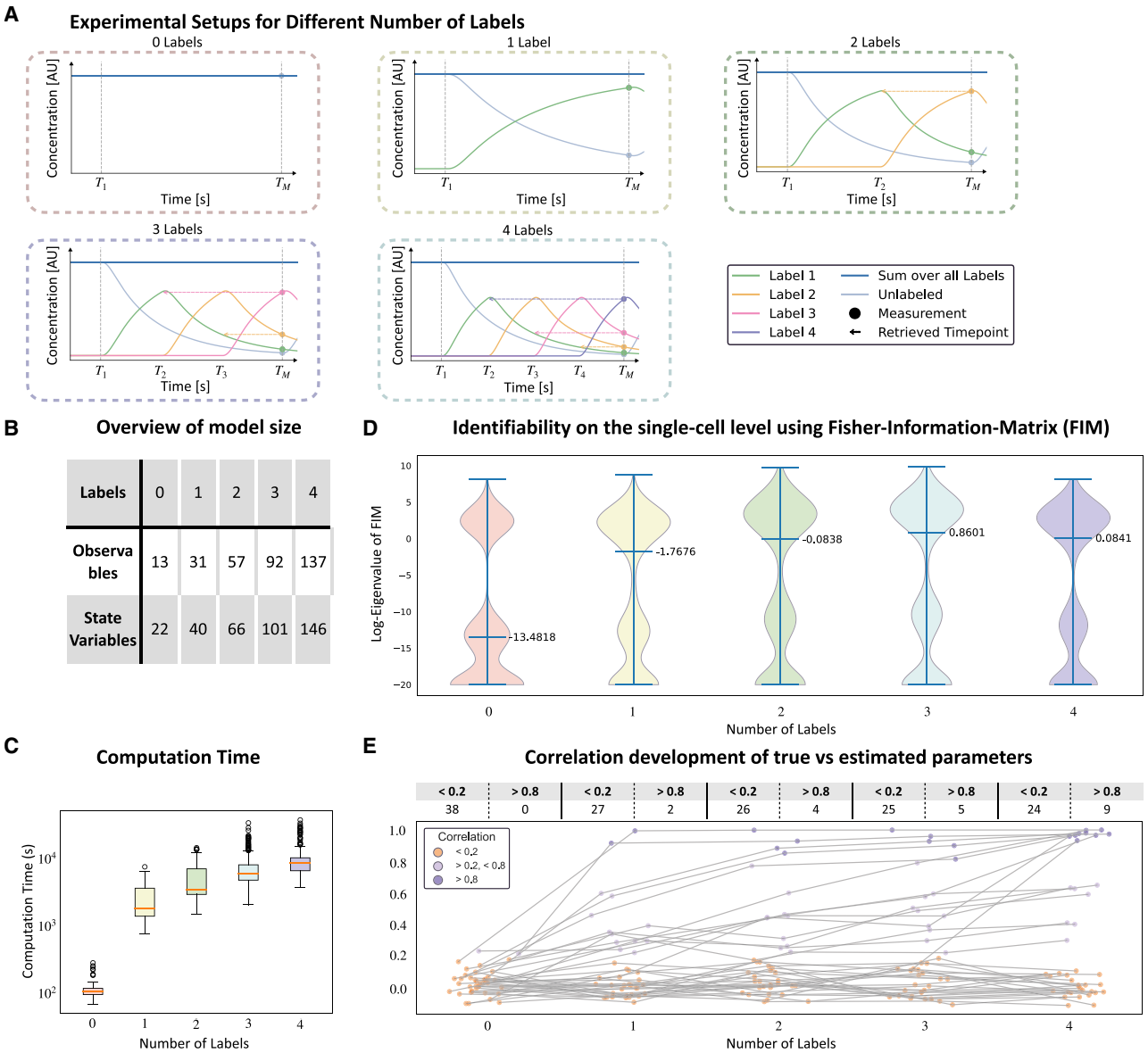


Figure 6. Impact of different numbers of labels on model size, computation time, and parameter identifiability

(A) Exemplary dynamics for each number of labels. The general experimental setup follows the same principle as Figure 2A.

(B) Overview of the model size depending on the number of labels.

(C) Average ($n = 200$) execution time of a complete multi-start local optimization for different numbers of labels. Boxplots span first to third quartile, median is depicted in orange, whiskers extend to 1.5·IQR, and extreme points are visualized separately.

(D) Violin plots for eigenvalue spectra of Fisher information matrix. Each violin plot corresponds to a different number of labels used and plots the whole range of values, with the median highlighted separately. Eigenvalues smaller than 1×10^{-20} were set to 1×10^{-20} for ease of visualization.

(E) Pearson correlation of true vs. estimated parameters for each label. The table tracks the number of highly correlated (>0.8) and uncorrelated (<0.2) parameters.

population scale, we investigated how increasing the number of labels affected the correlation between the estimated and true parameters (Figure 6E). Here again, we saw a substantial shift between 0 and 1 labels. Additional labels resulted in a steady increase of highly correlated and a steady decrease in uncorrelated parameters, indicating higher overall accuracy of estimated parameters. In contrast to the individual scale, we observed continuous improvement of the correlation, even going from three to four labels.

The significance of the shift in both analyses from 0 to 1 label can be attributed to the qualitative transition from steady-state measurement to dynamic observation. As more labels are added, the impact becomes more quantitative rather than fundamentally changing the nature of the observed dynamics.

In summary, our investigation underscores the crucial role of labeling strategies in the study of cellular dynamics. The number of labels impacts computational demands as well as parameter estimation quality and uncertainty. An increase in label numbers decreased uncertainty at both individual and population levels. However, the benefit of additional labels tends to plateau after a certain point while the computational demands continue to rise. Therefore, determining the optimal number of labels requires balancing improved accuracy against available computational resources.

DISCUSSION

High-throughput single-cell experiments are primarily analyzed using statistical and bioinformatic approaches that provide limited information on cellular dynamics. In this study, we illustrated how cellular dynamics can be retrieved from high-throughput single-cell experiments when paired with our proposed experimental setup and analysis. We developed an integrated framework and outlined how the parameterization of cellular dynamics can be used to quantify the cell-to-cell variability of lipid metabolism.

Using synthetic data that resemble the key aspects of single-cell lipidomics measurements, we demonstrated that the proposed computational pipeline facilitates robust data fitting. In addition, a substantial fraction of the single-cell parameters was rendered identifiable. However, some parameters remained practically non-identifiable. As these parameters were mostly linked to unobserved states, this could be further investigated using structural identifiability analysis.[31] Complementary, experimental design could be used to optimize the information content about the identifiable parameters. Although deriving precise minimum requirements (e.g., number of labels or lipids) for experimental design would require a dedicated follow-up study involving information-theoretic or optimization-based methods, our current results already hint at diminishing returns beyond three to four labels and a qualitative difference between no and one or more labels. We believe these findings provide preliminary guidance for experimental planning. In addition an optimization of the timing and the use of non-equidistant labeling might be beneficial.

As some measurements share a common dynamic, it is theoretically possible to simplify the model to one that is label agnostic and incorporates the measurements as different time points instead of different label combinations. This would drastically reduce the model size, but we would also lose many mixed labeled measurements. In addition, more flexible noise distribution and distributions for the single-cell parameters could be used. Furthermore, the approach allows for the handling of missing and censored values. Accounting for this would primarily require a user-defined likelihood function in pyPESTO and modified sampling routines for the generation of the single-cell data. An investigation of the trade-offs of these kinds of models could give further insights and guide the way for more efficient dynamic modeling.

The experimental setup, as we designed it, is intended for lipid metabolism. While our long-term goal is to enable applications in single-cell lipidomics, we note that the experimental workflow and the proposed algorithms can also be used to study cell extracts, cell lines, tissue samples, or whole-organism systems. These sample types are more accessible to current lipidomics workflows and allow the generation of robust quantitative data, making them valuable for method refinement before moving toward the single-cell level. To assess the proposed sequential labeling strategy, we already studied (bulk) lipid levels in Hepatocyte (Figures S4 and S5). While the resulting datasets mask cell heterogeneity, they provide a proof-of-principle for extracting longitudinal information via a single destructive measurement. Furthermore, in principle, this approach can be generalized to other reaction networks if a set of labels can be found that satisfy the key assumptions of this setup. These key assumptions are as follows: the reactions must be agnostic to the labels, and the experiment must be able to start in a steady state. Without these conditions, tracing the dynamics of different labels back to a single one would be impossible. Additionally, we need an accommodating mathematical model that describes the process dynamics in the presence of multiple labels and accounts for the passing of a label from one metabolite to another. A prime example that could satisfy these conditions is the well-established 13C metabolic flux analysis (13C-MFA).[32] 13C-MFA, particularly nonstationary 13C-MFA (INST MFA), employs a similar concept by subdividing the dynamics. This allows for the measurement of dynamical changes within the labeled metabolites while the overall system remains in a steady state.[21] Applying our experimental setup to INST MFA is possible as it satisfies the necessary conditions. However, there are two issues: Firstly, while our study showed that even just one label could significantly improve the analysis, there currently is no other stable isotope for carbon aside from 13C. This could be tackled by using stable isotopes other than just carbon, but it would require a slightly different setup than described here since the dynamics would not be time-shifted versions of one another. Secondly, the positioning of the carbon atoms and other isotopes would need to be modeled in detail. This means that the overall number of labeling sites will, in general, be vast and thus significantly increase model complexity, a well-known problem for INST MFA.[33]

In conclusion, our new proposed experimental setup in combination with the appropriate mathematical modeling opens up new possibilities in regard to destructive measurement techniques, generating longitudinal data from a single measurement. As it is agnostic to the kind of label, it has the potential to be applied in a wide variety of fields in addition to single-cell lipidomics.

Limitations of the study

The most limiting factor to this setup and why we resorted to synthetic data is the sensitivity of measurement devices. While the feasibility of single-cell lipidomics has been demonstrated,[6] it is not a standard approach yet and requires cells with high lipid content. We would like to note that our proposed experimental setup is currently theoretical in nature and has only been assessed using synthetic single-cell data.

Another factor to consider is the assumption that the alkyne FAs are taken up by the cell at similar rates. This is necessary to get the shared dynamics of lipids across labels. However, if

other reaction parameters are still shared, the modeling approach should still yield comparable results. More detailed investigations will reveal the viability of this approach.

A possible consideration to alleviate the sensitivity problem is targeted single-cell lipidomics.[6] This could be accompanied by experimental design analyses to pinpoint both crucial and possible lipid classes to measure. Despite this constraint, recent technological advancements[34,35] suggest that such experiments will become practically viable in the foreseeable future. When transitioning into a viable experiment, we expect that Thiele et al.[6] will serve as a good basis for an experimental protocol. We, therefore, remain optimistic that the present theoretical setup will soon transition into experimental application, thereby providing deeper insights into cellular dynamics at single-cell resolution.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Jan Hasenauer (jan.hasenauer@uni-bonn.de).

Materials availability

This study did not generate new materials.

Data and code availability

- All data reported in this study will be shared by the lead contact upon request.
- All original code has been deposited at Zenodo under the DOI https://doi.org/10.5281/zenodo.12654307 as well as on GitHub at https://github.com/PaulJonasJost/Pseudo-time-trajectory-of-single-cell-lipidomics and is publicly available as of the date of publication.[36]
- Any additional information required to reanalyze the data reported in this study is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (project IDs 390685813 - EXC 2047 and 390873048 - EXC 2151) through Metaflammation, project ID 432325352 – SFB 1454, through BATenergy, project ID 450149205 - TRR 333, and AMICI, project ID 443187771, the German Federal Ministry of Education and Research (BMBF) within the e:Med funding scheme (junior research alliance PeriNAA, grant no. 01ZX1916A), and by the University of Bonn via the Schlegel professorship to J.H. The funders had no role in the experimental protocol, data collection and analysis, decision to publish, or manuscript preparation. The icon in Figure 2 was created with BioRender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, J.H. and C.T.; methodology, P.J.J., D.W., and K.W.; software, P.J.J.; formal analysis, P.J.J.; writing – original draft, P.J.J., D.W., and J.H.; writing – review and editing, all authors; funding acquisition, J.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - Mathematical models
  - Synthetic data generation
  - Parameter estimation
  - Uncertainty analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2026.115474.

Received: April 10, 2025
Revised: August 9, 2025
Accepted: March 20, 2026
Published: March 25, 2026

REFERENCES

1. Lee, D., Jayaraman, A., and Kwon, J.S.I. (2020). Identification of cell-to-cell heterogeneity through systems engineering approaches. AIChE J. *66*, e16925. https://doi.org/10.1002/aic.169251of14.
2. Muto, Y., Wilson, P.C., Ledru, N., Wu, H., Dimke, H., Waikar, S.S., and Humphreys, B.D. (2021). Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. Nat. Commun. *12*, 2190. https://doi.org/10.1038/s41467-021-22368-w.
3. Kinker, G.S., Greenwald, A.C., Tal, R., Orlova, Z., Cuoco, M.S., McFarland, J.M., Warren, A., Rodman, C., Roth, J.A., Bender, S.A., et al. (2020). Pan-cancer single-cell rna-seq identifies recurring programs of cellular heterogeneity. Nat. Genet. *52*, 1208–1218. https://doi.org/10.1038/s41588-020-00726-6.
4. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mrna-seq whole-transcriptome analysis of a single cell. Nat. Methods *6*, 377–382. https://doi.org/10.1038/nmeth.1315.
5. Kelly, R.T. (2020). Single-cell proteomics: progress and prospects. Mol. Cell. Proteomics *19*, 1739–1748. https://doi.org/10.1074/mcp.R120.002234.
6. Thiele, C., Wunderling, K., and Leyendecker, P. (2019). Multiplexed and single cell tracing of lipid metabolism. Nat. Methods *16*, 1123–1130. https://doi.org/10.1038/s41592-019-0593-6.
7. Skylaki, S., Hilsenbeck, O., and Schroeder, T. (2016). Challenges in long-term imaging and quantification of single-cell dynamics. Nat. Biotechnol. *34*, 1137–1144. https://doi.org/10.1038/nbt.3713.
8. Hosein, A.N., Huang, H., Wang, Z., Parmar, K., Du, W., Huang, J., Maitra, A., Olson, E., Verma, U., and Brekken, R.A. (2019). Cellular heterogeneity during mouse pancreatic ductal adenocarcinoma progression at single-cell resolution. JCI Insight *5*, e129212. https://doi.org/10.1172/jci.insight.129212.
9. Yabo, Y.A., Niclou, S.P., and Golebiewska, A. (2022). Cancer cell heterogeneity and plasticity: A paradigm shift in glioblastoma. Neuro Oncol. *24*, 669–682. https://doi.org/10.1093/neuonc/noab269.
10. Schnell, A., Littman, D.R., and Kuchroo, V.K. (2023). Th17 cell heterogeneity and its role in tissue inflammation. Nat. Immunol. *24*, 19–29. https://doi.org/10.1038/s41590-022-01387-9.
11. Hatzenpichler, R., Krukenberg, V., Spietz, R.L., and Jay, Z.J. (2020). Next-generation physiology approaches to study microbiome function at single cell level. Nat. Rev. Microbiol. *18*, 241–256. https://doi.org/10.1038/s41579-020-0323-1.
12. Wang, T., Sharma, A.K., and Wolfrum, C. (2021). Novel insights into adipose tissue heterogeneity. Rev. Endocr. Metab. Disord. *23*, 5–12. https://doi.org/10.1007/s11154-021-09703-8.

13. Bergen, V., Soldatov, R.A., Kharchenko, P.V., and Theis, F.J. (2021). RNA velocity—current challenges and future perspectives. Mol. Syst. Biol. *17*, e10282. https://doi.org/10.15252/msb.202110282.

14. Erhard, F., Saliba, A.E., Lusser, A., Toussaint, C., Hennig, T., Prusty, B.K., Kirschenbaum, D., Abadie, K., Miska, E.A., Friedel, C.C., et al. (2022). Time-resolved single-cell rna-seq using metabolic RNA labelling. Nat. Rev. Methods Primers *2*, 77. https://doi.org/10.1038/s43586-022-00157-z.

15. Lindemann, C., Thomanek, N., Hundt, F., Lerari, T., Meyer, H.E., Wolters, D., and Marcus, K. (2017). Strategies in relative and absolute quantitative mass spectrometry based proteomics. Biol. Chem. *398*, 687–699. https://doi.org/10.1515/hsz-2017-0104.

16. Alabert, C., Loos, C., Voelker-Albert, M., Graziano, S., Forné, I., Reveron-Gomez, N., Schuh, L., Hasenauer, J., Marr, C., Imhof, A., and Groth, A. (2020). Domain model explains propagation dynamics and stability of histone H3K27 and H3K36 methylation landscapes. Cell Rep. *30*, 1223–1234.e8. https://doi.org/10.1016/j.celrep.2019.12.060.

17. Kuerschner, L., and Thiele, C. (2022). Tracing lipid metabolism by alkyne lipids and mass spectrometry: the state of the art. Front. Mol. Biosci. *9*, 880559. https://doi.org/10.3389/fmolb.2022.880559.

18. Triebl, A., and Wenk, M.R. (2018). Analytical considerations of stable isotope labelling in lipidomics. Biomolecules *8*, 151. https://doi.org/10.3390/biom8040151.

19. Allen, D.K., Bates, P.D., and Tjellström, H. (2015). Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: past, present and future. Prog. Lipid Res. *58*, 97–120. https://doi.org/10.1016/j.plipres.2015.02.002.

20. Wunderling, K., Zurkovic, J., Zink, F., Kuerschner, L., and Thiele, C. (2023). Triglyceride cycling enables modification of stored fatty acids. Nat. Metab. *5*, 699–709. https://doi.org/10.1038/s42255-023-00769-z.

21. Cheah, Y.E., and Young, J.D. (2018). Isotopically nonstationary metabolic flux analysis (inst-mfa): putting theory into practice. Curr. Opin. Biotechnol. *54*, 80–87. https://doi.org/10.1016/j.copbio.2018.02.013.

22. Hasenauer, J., Waldherr, S., Doszczak, M., Radde, N., Scheurich, P., and Allgöwer, F. (2011). Identification of models of heterogeneous cell populations from population snapshot data. BMC Bioinf. *12*, 125. https://doi.org/10.1186/1471-2105-12-125.

23. Fröhlich, F., Reiser, A., Fink, L., Woschée, D., Ligon, T., Theis, F.J., Rädler, J.O., and Hasenauer, J. (2018). Multi-experiment nonlinear mixed effect modeling of single-cell translation kinetics after transfection. npj Syst. Biol. Appl. *5*, 1. https://doi.org/10.1038/s41540-018-0079-7.

24. Rosen, E.D., and Spiegelman, B.M. (2014). What we talk about when we talk about fat. Cell *156*, 20–44. https://doi.org/10.1016/j.cell.2013.12.012.

25. Kahn, C.R., Wang, G., and Lee, K.Y. (2019). Altered adipose tissue and adipocyte function in the pathogenesis of metabolic syndrome. J. Clin. Investig. *129*, 3990–4000. https://doi.org/10.1172/JCI129187.

26. Luong, Q., Huang, J., and Lee, K.Y. (2019). Deciphering white adipose tissue heterogeneity. Biology *8*, 23. https://doi.org/10.3390/biology8020023.

27. Large, V., Peroni, O., Letexier, D., Ray, H., and Beylot, M. (2004). Metabolism of lipids in human white adipocyte. Diabetes Metab. *30*, 294–309. https://doi.org/10.1016/S1262-3636(07)70121-0.

28. Le Novère, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M.I., Wimalaratne, S.M., et al. (2009). The Systems Biology Graphical Notation. Nat. Biotechnol. *27*, 735–741. https://doi.org/10.1038/nbt.1558.

29. Raclot, T. (2003). Selective mobilization of fatty acids from adipose tissue triacylglycerols. Prog. Lipid Res. *42*, 257–288. https://doi.org/10.1016/S0163-7827(02)00066-8.

30. Fröhlich, F., Kessler, T., Weindl, D., Shadrin, A., Schmiester, L., Hache, H., Muradyan, A., Schütte, M., Lim, J.H., Heinig, M., et al. (2018). Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. Cell Syst. *7*, 567–579.e6. https://doi.org/10.1016/j.cels.2018.10.013.

31. Villaverde, A.F., Evans, N.D., Chappell, M.J., and Banga, J.R. (2019). Input-dependent structural identifiability of nonlinear systems. IEEE Control Syst. Lett. *3*, 272–277. https://doi.org/10.1109/LCSYS.2018.2868608.

32. Ahn, W.S., and Antoniewicz, M.R. (2011). Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry. Metab. Eng. *13*, 598–609. https://doi.org/10.1016/j.ymben.2011.07.002.

33. Wiechert, W., and Nöh, K. (2013). Isotopically non-stationary metabolic flux analysis: complex yet highly informative. Curr. Opin. Biotechnol. *24*, 979–986. https://doi.org/10.1016/j.copbio.2013.03.024.

34. Zhang, L., and Vertes, A. (2018). Single-cell mass spectrometry approaches to explore cellular heterogeneity. Angew. Chem. Int. Ed. Engl. *57*, 4466–4477. https://doi.org/10.1002/anie.201709719.

35. Li, C., Chu, S., Tan, S., Yin, X., Jiang, Y., Dai, X., Gong, X., Fang, X., and Tian, D. (2021). Towards higher sensitivity of mass spectrometry: A perspective from the mass analyzers. Front. Chem. *9*, 813359. https://doi.org/10.3389/fchem.2021.813359.

36. Jost, P.J. (2025). Code for "Experimental design and computational modeling to infer single-cell lipid dynamics from a single destructive measurement" (Zenodo). https://doi.org/10.5281/zenodo.17468379.

37. Fröhlich, F., Weindl, D., Schälte, Y., Pathirana, D., Paszkowski, Ł., Lines, G.T., Stapor, P., and Hasenauer, J. (2021). AMICI: high-performance sensitivity analysis for large ordinary differential equation models. Bioinformatics *37*, 3676–3677. https://doi.org/10.1093/bioinformatics/btab227.

38. Loos, C., Krause, S., and Hasenauer, J. (2018). Hierarchical optimization for the efficient parametrization of ODE models. Bioinf *34*, 4266–4273. https://doi.org/10.1093/bioinformatics/bty514.

39. Kreutz, C., Bartolome Rodriguez, M.M., Maiwald, T., Seidl, M., Blum, H.E., Mohr, L., and Timmer, J. (2007). An error model for protein quantification. Bioinformatics *23*, 2747–2753. https://doi.org/10.1093/bioinformatics/btm397.

40. Schmiester, L., Schälte, Y., Bergmann, F.T., Camba, T., Dudkin, E., Egert, J., Fröhlich, F., Fuhrmann, L., Hauber, A.L., Kemmer, S., et al. (2021). PEtab—interoperable specification of parameter estimation problems in systems biology. PLoS Comput. Biol. *17*, 1–10. https://doi.org/10.1371/journal.pcbi.1008646.

41. Schälte, Y., Fröhlich, F., Jost, P.J., Vanhoefer, J., Pathirana, D., Stapor, P., Lakrisenko, P., Wang, D., Raimúndez, E., Merkt, S., et al. (2023). pyPESTO: a modular and scalable tool for parameter estimation for dynamic models. Bioinformatics *39*, btad711. https://doi.org/10.1093/bioinformatics/btad711.

42. Fröhlich, F., and Sorger, P.K. (2022). Fides: Reliable trust-region optimization for parameter estimation of ordinary differential equation models. PLoS Comput. Biol. *18*, e1010322. https://doi.org/10.1371/journal.pcbi.1010322.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python 3.11.1	Python Software Foundation	https://www.python.org
pyPESTO v0.3.1 (Python Package)	GitHub	https://github.com/ICB-DCM/pyPESTO
Fides v0.7.8 (Python Package)	GitHub	https://github.com/fides-dev/fides
AMICI v0.18.0 (Python Package)	GitHub	https://github.com/AMICI-dev/AMICI
Simplesbml v2.3.0 (Python Package)	GitHub	https://github.com/sys-bio/simplesbml
The code to generate in-silico data and analyze the results is available at Zenodo	–	https://doi.org/10.5281/zenodo.12654307

METHOD DETAILS

Mathematical models

To model the lipid metabolism in a single cell, we use a system of ordinary differential equations (ODEs):

$$\frac{dx}{dt} = f(x(t,\theta),\theta), \quad x(t_0,\theta) = x_0(\theta),$$

in which $x(t,\theta) \in \mathbb{R}^{n_x}$ is the vector of state variables at time t, $\theta \in \mathbb{R}^{n_\theta}$ is the vector of parameters, and f the vector field describing the time evolution of the vector of state variables. The vector of state variables is at time t_0 initialized with $x_0(\theta)$. Throughout the study, we assume that this initial state is the equilibrium point for a single cell exposed to a label-free growth medium. Accordingly, the initial condition is defined as

$$x_0(\theta) = \lim_{t\to\infty} x(t,\theta),$$

Metabolite levels are assumed to be partially observed (using, e.g., mass spectrometry). Metabolite levels are considered as sums over different lipid species (e.g., Triglycerides with different fatty acid chain lengths), only differentiating between the labeled fatty acids. The vector of observables at time t is denoted by $y(t) \in \mathbb{R}^{n_y}$,

$$y = h(x,\theta),$$

with observable mapping h. The observations are subject to measurement noise, which we consider here to be additive, normally distributed. We denote all measurements with $\mathcal{D} \in \mathbb{R}^{n_y \times n_t}$, with

$$\mathcal{D}_{i,k} = \bar{y}_{i,k} = y_i(t_k) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0,\sigma_i^2) \quad i = 1,\ldots,n_y, k = 1,\ldots,n_t,$$

in which n_y is the number of observables, and n_t is the number of measurement time points.

In our model, state variables represent concentrations of metabolites. For each different combination of labels per lipid, one state variable is used. The considered molecules, as well as their respective number of state variables, can be found in Table S1. We consider mass action kinetics for each reaction and assume that reactions including labeled lipids have the same kinetic parameter for each label combination. The reactions and their corresponding rates are listed in Table S2.

Synthetic data generation

To generate synthetic data, we sampled vectors of parameters from a pre-defined log-normal distribution. This distribution was chosen to ensure positivity in the parameter values. The means for each parameter individually were set to a value between 0.01 and 0.5. With a selected standard deviation of 0.1, we anticipated the samples of each parameter to encompass a range spanning between half and one order of magnitude. The sampled parameter vectors were used to simulate the mathematical model (Figure 4A, blue box).

For data generation, we simulated the ODE model using AMICI,[37] including a pre-equilibration step to mirror the experimental setup. To emulate real-world conditions, we introduced additive normally distributed measurement noise, with a standard deviation equal to 10 percent of the measurement value. The assumption of normality is commonly employed in systems biology studies.[38]

Indeed, the second most commonly used noise model is multiplicative log-normal noise, which, after transformation, yields on the log-transformed observable also additive normal noise.[39] The generated data served as the starting point for parameter estimation (Figure 4A, red box).

For the initial analysis, we used the mathematical model with three labels. We generated a thousand parameter sets, with each set designed to represent an artificial cell. For each cell, we created a corresponding parameter estimation problem that we encoded in the PEtab format.[40] We assume here that data and parameters belong to one singular cell, not pooled measurements of cells belonging to one cell type.

To investigate the effect of the number of labels on the parameter uncertainty (Figure 6), we used five different models, using between zero and four labels. Independent of the number of labels, the simulation time after pre-equilibration was kept constant. We generated 200 parameter sets in the same fashion as the previously generated ones. Within this set, we generated models and data for each label, resulting in five distinct datasets for each parameter set. This approach ensured consistency in parameter selection across different labels, allowing comparisons between the number of labels.

Parameter estimation

To infer the unknown model parameters from data, we apply a maximum likelihood approach. Assuming independent, normally distributed measurement noise, the likelihood function is given by

$$\mathcal{L}_D(\theta) = \prod_{i=1}^{n_y} \prod_{k=1}^{n_t} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2} \left(\frac{\overline{y}_{i,k} - y_i(t_k, \theta)}{\sigma_i} \right)^2 \right).$$

Other noise models can also be assumed and mainly change the likelihood function. In the absence of a known error structure, additive normally distributed noise is commonly used as a first approximation, justified by the central limit theorem. The maximum likelihood (ML) estimate is obtained by maximizing the likelihood function. For numerical stability and improved convergence, we minimize the negative log-likelihood:

$$\begin{aligned} \mathcal{J}(\theta) &= -\log \mathcal{L}_D(\theta) \\ &= -\frac{1}{2} \sum_{i=1}^{n_y} \sum_{k=1}^{n_t} \log(2\pi\sigma_i^2) + \left(\frac{\overline{y}_{i,k} - y_i(t_k, \theta)}{\sigma_i} \right)^2 \end{aligned}$$

For the considered example, the parameters are constrained to the interval $[10^{-4}, 10^2]$, spanning six orders of magnitude. We used multi-start gradient-based local optimization with 1000 log-uniformly distributed starting points for minimization. Optimization was performed in pyPESTO v0.3.1[41] using a PEtab[40] representation of our model and interfacing the Fides v0.7.8[42] optimizer. The model was simulated using AMICI v0.18.0.[37] For a visualization of the workflow, we refer to Figure S2.

Uncertainty analysis

To evaluate the impact of the number of labels on the identifiability of model parameters, we used the Fisher information matrix (FIM) at the maximum likelihood estimate. We computed the FIM via AMICI[37] by summing over the product of the partial derivatives of the objective function with respect to the parameters, i.e., for any parameter $\theta \in \mathbb{R}^{n_\theta}$

$$\mathrm{FIM}(\theta) = \left(\frac{\partial}{\partial \theta_i} \mathcal{J}(\theta) \frac{\partial}{\partial \theta_j} \mathcal{J}(\theta) \right)_{i,j=1}^{n_\theta}.$$

The eigenvalues of the FIM provide information about the likelihood function landscape that is close to the maximum likelihood estimate. Eigenvectors for large (positive) eigenvalues indicate weighted sums of parameters (directions in parameter space), which are well determined, while eigenvectors for eigenvalues that are zero indicate directions in which the likelihood function landscape is locally flat. Thus, the FIM at the maximum likelihood estimate provides information about the practical identifiability and sloppiness of the estimates. Accordingly, the eigenvalue spectrum can be used for uncertainty comparisons.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed using Python 3.11. Correlations were computed using Pearson correlation. The corresponding figure legends describe any pooled data, error bars as well as number of data points analyzed.