A data-driven and probabilistic approach for data integration, calibration, and analysis in mechanistic models of cellular processes

By

Michael Warren Irvin

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Interdisciplinary Studies: Systems Biology and Machine Learning

August 13, 2021

Nashville Tennessee

Approved:

Vito Quaranta, Ph. D.

Joey Barnett, Ph. D.

Christopher Fonnesbeck, Ph.D.

Vivian Gama, Ph.D.

Dedicated to Henry Turner

Your granddaughter remembered the arithmetic lessons you gave her.

You told her rely on her smarts and to persevere, the way you did.

Enslavement did not break you.

Your granddaughter (my grandmother) kept your traditions alive, so will I.

Table of Contents

LIST OF FIGURES

Figure **Page**

LIST OF BOXES

Box **Page**

LIST OF TABLES

Table **Page**

Introduction

The twentieth century witnessed two revolutions in science[1,2]. Physicists (chemists, engineers, etc.) translated physical laws into mathematical models as part of a model-based scientific paradigm that massively accelerated technological innovation and an increasingly detailed understanding of the physical world. Biologist, meanwhile, deconstructed biological processes into their foundational functional components as part of a reductionist scientific paradigm that culminated in an exhaustive genetic parts-list for several organisms[3]. The achievements of both approaches set the stage for a fusion of methods from biology and physics to drive a new model-based investigation of biological processes: *systems biology*[4-7]. Systems biology entered the scene on a promise of new insights necessary to propel biomedical innovation and discovery, on par with what occurred in physics. Instead, systems biology discovered a frustrating combination of challenges unlike that encountered by physicists. Specifically, issues of biological complexity, multiscale organization, and heterogeneity[8] complicate access to quantitative measurements of biological systems[9]. The dearth of quantitative biological measurements constrains the modeling efforts needed for a model-based investigation of biology. This challenge has garnered considerable attention from systems biologists who have devised various experimental and modeling strategies to contend with or confront the data-problem of biology. In this work, I introduce a modeling approach that enables incorporation of nonquantitative measurements -- as more abundant substitute for quantitative measurements.

This incorporation of nonquantitative measurements entails a model of the measurement -- i.e., a model of the relationship between the measured observables and underlying biological mechanisms. Previous work in this area focused on the uncertain constraints encoded by categorical data. Their approaches complicate application Bayesian modeling methods[10, 11] and/or introduce *ad hoc* assumptions into the modeling process[12]. The limitations of previous approaches obscure estimates of model uncertainty and distort model predictions. This work shifts the focus toward the measurement; the measurement process and its partially uncharacterized connection to underlying biological mechanism. It

introduces a probabilistic and data-driven measurement model that, as a probabilistic model, enables easy application of Bayesian modeling methods and, as a data-driven model, avoids the introduction of *ad hoc* assumptions. This measurement model approach also incorporates machine learning to resolve uncertainty in link between measurement and specific properties of the underlying biological mechanism. This work focuses on the application of nonquantitative measurements to mechanistic (ordinary differential equation (ODE) based) models of cellular processes, but its concepts readily apply to other modeling scenerios.

The first chapter describes data-problem in biology. It begins with a summary of the modeling approaches in systems biology (though this work focuses on bottom-up modeling approaches). Bottom-up modeling approaches construct mechanistic models by assembling mathematical representations of first-principal factors (i.e., individual interactions, etc.) into a system of equations describing the mechanism. Empirical data drives essential steps of the bottom-up modeling process (e.g., model calibration). This chapter details how the dearth of quantitative data complicates model calibration and limits modeling capabilities. It also describes why the dearth of quantitative data persists despite ongoing efforts to improve measurement technologies in biology. Finally, this chapter describes modeling and measurement strategies that have recently emerged to address the dearth of quantitative measurements in biology.

The second chapter introduces a new probabilistic data-driven measurement model that integrates nonquantitative data into model calibration strategies. This chapter describes the model calibration step which tunes or estimates the values of the mechanistic model's free parameters to maximize it alignment with data. This alignment between model and data is defined using a likelihood or objective function; traditional formulations of this function include some function of the sum of squared errors (SSE)[13] and do not accommodate nonquantitative input. This chapter explains how a measurement model overcomes this challenge, then demonstrates it via calibration of a model of apoptosis signaling to nonquantitative measurements: ordinal values of the abundance of apoptotic signaling proteins and nominal observations of apoptosis vs survival. The calibrations used Bayesian methods, which enabled easy uncertainty estimation and thereby revealed how the size and composition of a dataset impacts the uncertainty of a

calibrated mechanistic model. This chapter also demonstrates the risk of strategies that introduce *ad hoc* assumptions about the measurement to introduce artifacts into the predictions made by a calibrated mechanistic model. Finally, it shows how machine learning can identify connections between measurement and underlying cellular process: e.g., features encoded in the mechanistic model that best predict nominal observations of apoptosis vs survival outcomes.

In the final chapter, I discuss future directions for the probabilistic and data-driven measurement model. It starts with a demonstration of the original premise of systems biology: a model-based investigation of biological processes. I employed the bottom-up modeling approach to investigate a complex network of regulatory crosstalk interactions linking apoptotic and necroptotic cell death signaling[14]. Modeling these interactions enabled prediction of new details in the regulatory crosstalk mechanism and design of experimental confirmation of those details. This demonstrates the investigative potential of bottom-up modeling and encourages continued efforts to resolve challenges stemming from a dearth of quantitative measurements in biology. The rest of the chapter focuses on the challenges and potential of extending the measurement model concept to a broader suite of cellular measurements. It focuses on measurements or observations that may lack a known link to variables encoded in the model; I explore feature selection and dimensionality reduction to learn a representation of the model variables that best confirms with the measured data. It also focuses on heterogeneity and its impact on the measurement process; there, I explore ways of encoding heterogeneity as part of the measurement model. Throughout, this work highlights the unique challenge of biology: its complexity, multiscale organization, and heterogeneity. While it offers a workaround to the dearth of quantitative data, it does not replace the need for more data. On the contrary, it encourages scientists to collect as many measurements (quantitative or nonquantitative) as possible.

Chapter 1

The demand for more data in systems biology and methods of addressing it

1.1. Summary

Systems biology applies interdisciplinary techniques to the problem of understanding how biological species engage a complex network of interactions to produce an emergent behavior. Its potential for generating useful insights for advancing drug design, bioengineering, synthetic biology, etc. has resulted in a surge in its popularity. Systems biology has shortcomings, however, that attest to its dependence on large amounts of precise biological measurements. The fact that these measurements do not yet exist prevalently, presents a significant obstacle to progress in systems biology. Scientists have addressed the demand for more data in system biology by creating innovative measurement methods, trimming models to accommodate the current dearth of useable data, and inventing ways to more efficiently use existing data. This review covers the strategies scientists use to address the demand for more data in systems biology.

1.2. Introduction

The 1950s marked the beginning of a transition toward increased application of basic sciences to guide -- and hopefully accelerate – technological innovation in several areas, including drug development[15]. The transition proved wildly successful to a myriad of industries[16]. Pharmaceuticals, however, witnessed a paradoxically steady decline in production during this period as it churned out half as many successful new drugs per billion USD every decade[17] Drug discovery methods rely on reductionist approaches to biology, i.e., approaches that try to model intricate biological behaviors as the work of only one or a few biological components[18]. Attempts to targets these biological "magic-bullet" components with drugs routinely encounter unanticipated pharmacokinetics problems, low efficacy,

toxicity, and adverse effects – and ultimately fail[19]. Scientists have criticized the overuse of the reductionist approach biology[20, 21].

These failures point to a prevailing challenge to drug design and biological science in general: biological *complexity*. Biological properties are complex in that they do not follow the action of a single component; instead they emerge from a network of dynamic interactions involving many biological components[22]. This motivates the advancement of approaches that investigate biological behaviors as properties of complex biological networks. Such approaches have coalesced into a body of methods and analyses encompassed in the term, *systems biology*[4-7].

Biological systems possess unique properties that frustrate scientific efforts to investigate them. Specifically, biological systems create a near intractable demand for high precision quantitative data[8, 23]. Systems biology addresses this challenge with new measurement technologies and modeling strategies to increase the availability and utilization of biological data. The following describes systems biology, albeit with an emphasis on the bottom-up approach. It then explores the role of data in the success and failure of systems biology efforts.

1.3. Systems Biology

Modern definitions of systems biology vary widely, but often remark that systems biology reflects a "departure from reductionist thinking" in biology. Jacques Loeb introduced the reductionist view, in 1912, that biological components act as mechanical cogs driving their assigned biological functions[2]. This view inspired more than half a century of reductionist approaches that sought to identify which components drive specific biological behaviors. Reductionism culminated in whole genome projects, and the amassing of a nearly comprehensive "parts list" of biology[3, 24]. This contribution positioned scientists to begin investigating the connections between biological components and systems-level consequences of their interconnectedness. The establishment of system biology has begun to steer biology away from the reductionist thinking that preceded it[7].

The emergence of systems biology also coincides with the advent of technologies that make its investigations feasible i.e., high-throughput screening and "omics" technologies, measurement advances (e.g., in imaging[25], mass spectrometry[26, 27] and flow cytometry[28]), microfluidics[29], data analysis advances (e.g., network modeling[30], deep learning[31, 32]) and parallel computing[33, 34]. Indeed, systems biology ebbs and flows with the pace of innovations that allow better measurement and analysis of biological data (i.e., more precise and comprehensive measurement, more efficient storage and curation, and better analysis, modeling and visualization biological data)[35]. This technological requirement for progress in systems biology presents special opportunities for physicists, mathematicians and engineers. Their influx engenders the interdisciplinary culture of systems biology but has also expanded and blurred the definition of systems biology[7]. Systems biology now refers to a wide variety of research; much of which clusters into two broad and interacting domains: top-down and bottom-up[36].

Top-down, or data-driven, approaches start with comprehensive descriptions of a biological system (i.e., observed biological behaviors and accompanying measurements of thousands of biological species) and attempt to characterize properties of its biological components (specifically, the network of interactions amongst the biological components). In contrast, bottom-up approaches start with well characterized properties of the biological components, and their interactions, and attempt to reconstruct observed biological behaviors[36]. Note, the definitions of "component" and "system" vary with the scope of the investigation. For instance, a protein structure model of Bax might define Bax oligomerization at the mitochondrial outer membrane as a system-level property[37] while a dynamical model of mitochondria dependent apoptosis[38] defines it as the property of a component.

The top-down and bottom-up approaches occupy opposite ends of a spectrum, described by Doug Laffenburger, of approaches to investigate and model biological properties[39]. Along the spectrum, there exists models that blend aspects of top-down and bottom-up approaches in order to push investigations beyond the limits of either individual approach. All systems biology approaches must negotiate a compromise between the kinds of insights they provide and the amount of data they require[40]. Therefore,

scientists must carefully consider which modeling approaches best apply to their investigation. The following section briefly describes conventional modeling approaches used in systems biology.

1.3.1. Top-down Approach to Systems Biology

Top-down approaches use statistical and machine learning frameworks to explain, and model, the patterns of abundance or activity of biological species that correspond to biological phenotypes and/or experimental treatments. These data-driven models typically uncover large networks of statistical interactions (e.g. correlative associations between the biological species). Clustering algorithms sort biological components, e.g., genes, into groups of similarly behaving gene-clusters[42-44]. These gene-clusters may include (supervised) or exclude (unsupervised) phenotype and experimental treatment information. Regression and classification models identify a set of genetic markers and/or gene expression profiles that best predict observed biological phenotypes[44]. Correlation and mutual information frameworks give more detail by inferring networks of pair-wise statistical associations between the biological components[45]. These networks explain which interactions occur but lack information about how the interactions proceed. Bayesian network models describe directed interactions, i.e. wherein one gene that exerts influence on another[46].

The models produced in top-down modeling approaches investigate large datasets for useful insights for predicting biological behaviors. Top-down network models can predict interesting structural properties (network motifs and hubs, scale-free attributes, etc.) that may prove insightful for understanding biological mechanism and observed phenotypes[47]. Scientists can further interpret top-down models considering existing knowledge and literature documented biological mechanisms, to glean mechanistic insights[48]. Top-down approaches leverage data to uncover, ab initio, the network of interactions behind a biological process, but uncovering the specific mechanistic details of these interactions requires bottom-up approaches.

1.3.2. Bottom-up Approach to Systems Biology

Bottom-up approaches to systems biology model biological phenotypes as the integrated effect of a network of interactions; wherein each modeled interaction possesses a detailed mechanistic description based on underlying biophysical first-principles[49]. The near comprehensive "parts list" of biology, that amassed through decades of reductionist work, includes functional definitions of biological species (e.g. as ligands, protein complex scaffold, enzymes). Systems biologists attach to these biological species an abstract representation[50] (e.g., as depicted in Kitano diagrams) and a mathematical description of the first-principal properties (e.g., kinetics, transport, mechanics) that govern their interactions. The bottom-up modeling approach can help predict biological behaviors and offer mechanics explanations for those behaviors. It generally accomplishes via three model building steps: model assembly, model calibration, and model selection.

In model assembly, scientists decide on a model topology: i.e. which interactions to include as relevant to the experimental question, and how to represent those interactions mathematically. This step relies heavily on molecular biology literature and, assembling an adequate set of interactions requires substantial effort and time. The interactions can take on various mathematical representations: Boolean[51], petri-net[52], and agent-based modeling expressions[53], mass-action kinetics rate laws (e.g., as ordinary differential equations (ODEs)), and spatially resolved expressions (e.g., partial differential equations and cellular automata[54]). Each mathematical framework offers a kind of trade-off between the level of detail in the model and the amount of data required to validate the model[55, 56]. For instance, Boolean models generally require less data than ODE models. In model assembly, scientists usually employ simplifying assumptions to limit the scope and size of a model so that it specifically addresses a biological question, while adhering to technical constraints. There exists, however, a continued effort to model entire cells via bottom-up modeling approaches[57, 58].

Model calibration presents a significant challenge to bottom-up modeling approaches, as these more detailed models possess notably more free parameters than do top-down models[59]. In this step, scientists try to identify a narrow range of parameter values that optimize the agreement between the

model and experimental data. Bottom-up models commonly feature non-linear dynamics and a therefore intractable non-convex model calibration objective function[59]. Therefore, model calibrations commonly recur to stochastic optimization algorithms (e.g. Approximate Bayesian Markov-chain Monte Carlo Algorithms[60], Differential Evolution[61], Particle Swarm Optimization[62]). Precise model calibration presents a significant challenge to bottom-up modeling approaches as it requires expensive data: quantitative time-course measurements (e.g., via fluorescent markers) of the abundance and activity of biological species[13, 59]. Even with perfectly suited data, model sloppiness (i.e., that model behaviors depend on only a few covariant combinations of parameters) imposes a hard limit on the parameter estimates' precision[63-64]. Qualitative modeling (e.g., Boolean) approaches allow scientists to address experimental questions while at least partially circumventing the model calibration step[51].

Model selection helps scientists derive new mechanistic insights into a biological process via the bottom-up modeling approach[65]. The bottom-up modeling process can expose unconsidered gaps in the understanding of a biological process. Scientists address these knowledge gaps with a barrage of plausible hypothetical mechanistic explanations. Model calibration then produces an ensemble of equally valid parameterized models per each hypothetical model topology. Bayesian evidence-based model selection, which uses Bayes Factor to compare competing model topologies, entails a costly integration over the space of parameterized models (or in statistical mechanics terms, calculations of the Helmholtz free energy between ensembles of parameterized models[66]). Sethna and Brown accomplished model selection using a metric of model sloppiness to approximate the energy difference between competing models[66]. Scientists extend model selection algorithms to generalized exploration of plausible biological mechanisms, but such efforts can have prohibitive computational and data requirements.

Biological processes have subtle and complicated elements that will overwhelm the scientist who tries to deconstruct them "in their head". Isaac Newton remarked that "equations are smarter than us" [67**]. As such, the equations of bottom-up modeling approaches provide a powerful tool for elucidating the nuances of biological mechanism[67]. Scientists employ bottom-up approaches to investigate biological processes when they involve of a small number of well characterized dynamical interactions and when

precise time-course measurements are available[56]. Bottom-up modeling approaches find and fill gaps in the mechanistic understanding of biological processes. These approaches have their foundations in engineering, where modeling drives design. Scientists investigate model uncertainties and error to define the nature and scope of model predictions and propose targets of new optimally designed experiments. Model based design applies insights from systems biology to support of ongoing efforts in bioengineering[68], synthetic biology[69], and drug design[70].

The potential of bottom-up modeling approaches to understand subtle details of biological processes has contributed to its growing appeal. I explore this potential further in the next two chapters. The increasing presence of bottom-up modeling approaches in biological research and drug design has set the course toward a future where modeling and experimentation occur within a standardized, completely integrated, collaborative workflow[71]. Natural language processing methods and model visualization tools will automate and streamline key aspects of model assembly[72]. Advances in parallel computing will accelerate model calibration and model selection and support the development of larger bottom-up models[73, 74].

1.3.3. Middle-Out Approach to Systems Biology

Biologists consult top-down modeling approaches, early in biological investigations, to build large network models of the interactions responsible for a set of biological phenotypes. As these biological investigations develop, biologists migrate toward bottom-up modeling approaches to resolve the mechanistic details of a subset of these potential interactions[4]. This research trajectory takes biological investigations through a handoff region where neither top-down nor bottom-up modeling approaches fully apply. Middle-out approaches occupy this region (at the center of the Lauffenburger spectrum) where methods aim to translate large data-driven network models into smaller refined mechanistic models; and/or vice versa[4]. Top-down models serve as templates for exploring biological mechanism using middle-out models that overlay their network edges with corresponding coarse-grained mechanistic representations. These models begin with simple representative operations but gradually acquire more

detailed expressions (e.g., ODEs) for increased local granularity[75]. Modelers have extended this approach to genome scale kinetic models[76]. Conversely, the mechanistic details of bottom-up models adhere to a local modeling and experimental scope, but researchers can extend this scope via by appending data-driven edges. The resulting hybrid model captures a wider scope of biological behaviors while retaining the high local resolution supplied by the mechanistic model[77]. These kinds of hybrid models have proved particularly useful in multi-scale modeling[77, 78]. Various middle-out modeling strategies have emerged to mediate the compromise between large top-down networks of low-detail interactions and technically demanding more detailed bottom-up mechanistic models. The remaining sections discuss topics that apply most to bottom-up modeling approaches but have general relevance to all systems biology modeling approaches.

1.4. Challenges to Systems Biology

Systems biology adapts physics and engineering inspired approaches to advance an understanding of the systems-level consequences of interacting biological species. Scientists expect such approaches to revolutionize biology; giving it systematic theory-based experimentation, streamlined model-based drug-design, and theoretical foundations commensurate with that of physics or engineering. Despite maintained progress in this direction, systems biology has not yet met those initial promises. Modeling efforts still struggle to produce precise, accurate, and insightful enough predictions for drug design and other applications. Sydney Brenner criticized systems biology as "low-input, high-throughput and no-output"[79]. Indeed, some models, though technically demanding, hardly outperform random guesses[80]. Unsurprisingly, there still exists a stark difference (as Woltosz noted) between the high-throughput "trial and error" used in drug design and the methodical model- based engineering utilized in designing e.g., airplanes[81]. Biology presents unique challenges that transcend the immediate capabilities of methods borrowed from physics and engineering. Issues of complexity, multiscale organization and heterogeneity combine make biological systems more difficult to study than many other systems.

Biological systems possess a level of complexity that distinguishes them from other large and complicated systems[82]. In this context, I make an important distinction between complicatedness and complexity: a complicated system involves a large number of interacting parts; a complex system can possess few parts and a simple set of governing interactions, and yet produce a vast and unpredictable range of possible outcomes. Biological systems have such an evolved network of pleotropic signaling hubs (i.e., that influence several processes), single-link nodes (that only serve a single, perhaps redundant function) and other kinds of nodes in order to confer a complex robustness against noise and injury[83].

Biological complexity also manifests as the hierarchical "multi-scale" organization of biological systems. In biological systems, atoms dictate the properties of a diverse array of macromolecules, which integrate into complexes and compartments, which then form cells, tissues, organs and, organisms. Each level of organization occupies a specific spatial-temporal scale, contains a complex network of interactions, and connects to other levels. Multi-scale organization allows specialization and diversification of biological processes by isolating them spatially and temporally within a given scale.

Biological systems exhibit nondeterministic variable behaviors that originate, in part, from random fluctuations in the activity, localization and abundance of its components. This biological noise has extrinsic sources (e.g. fluctuations in the cellular milieu, and inherited variations in cell size and content) and intrinsic sources (e.g. stochastic changes in the abundance of cellular protein count and Brownian motion of diffusing molecules. Biological systems tolerate and even utilize noise[84-86]. Feedback circuits[87, 88] and other signaling motifs[89] attenuate noise and/or collapse noisy signals into a few discrete stable states. Biological processes also amplify noise in order to promote heterogeneity, as this helps create resilient subpopulations and drives specialization of diverse cellular functions[90,].

Complexity, multiscale organization and heterogeneity conspire to complicate and impede both the measurement and modeling of biological systems. Biological complexity (detailed in Figure 1.1 and Box 1.1) increases the number of components models require, and concomitantly increases the amount of data needed to describe biological processes[56, 59]. The multi-scale organization of biological systems exacerbates this demand for more detailed models while limiting the feasibility of useful biological

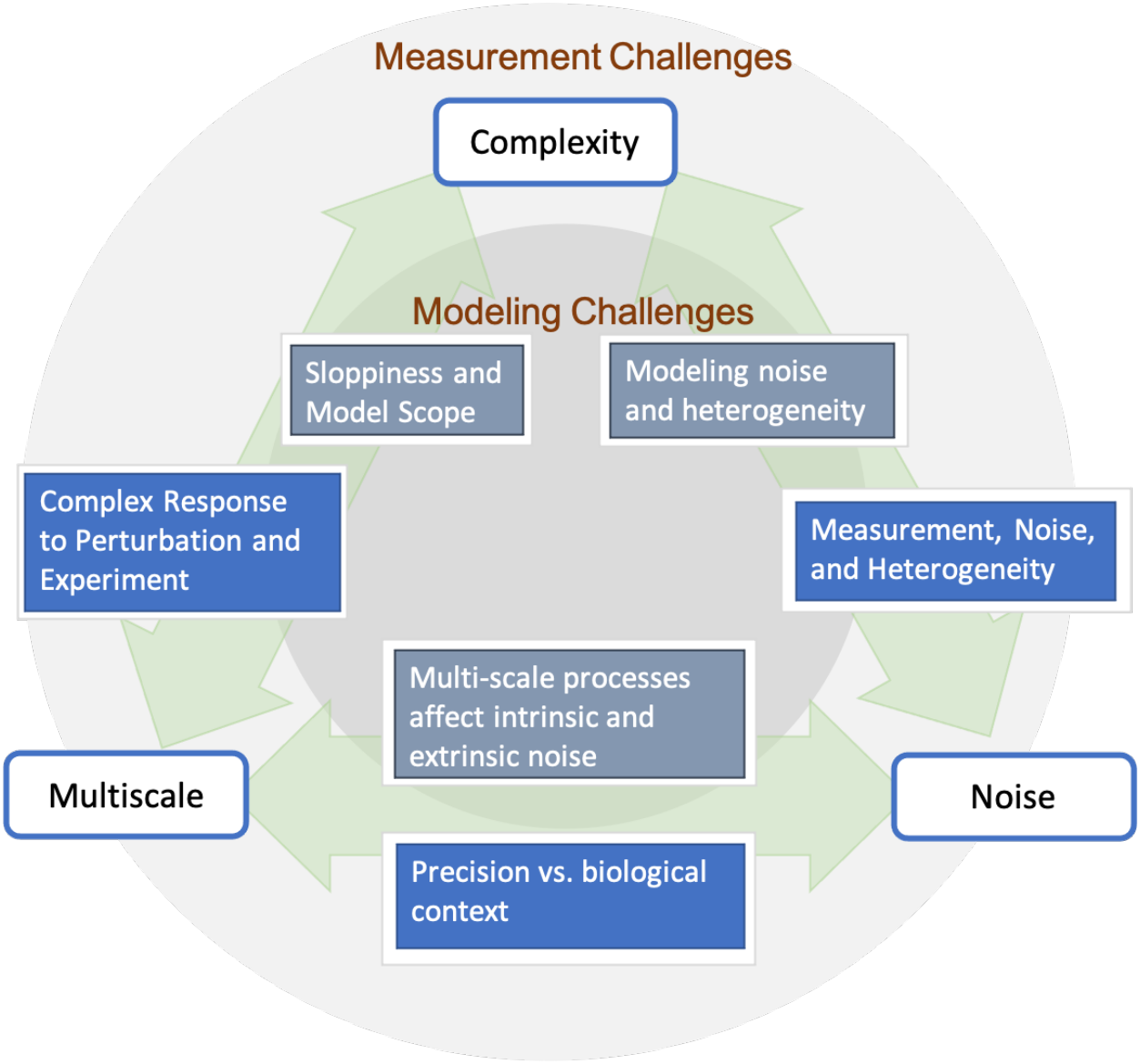


Figure 1.1. Challenging Properties of Biology: Complexity, Multiscale and Heterogeneity As detailed in Box 1.1., three properties of biology (complexity, multiscale organization and heterogeneity) interact to impede measurement and modeling of biological systems.

Model Sloppiness and Scope: Complexity gives rise to model sloppiness, a property that dictates that model behavior depends on (instead of a few parameters) a few combinations of several parameters. This property expands the set of components and interactions (i.e., the model scope or size) necessary to describe biological processes and limits the identification of which components and interactions to include in the model. Biology's hierarchy of scale compounds this effect of complexity by requiring detailed descriptions of components and interactions that span multiple temporal and spatial scales.

Multi-scale Processes Affect Intrinsic and Extrinsic Noise: Biological systems possess components whose behaviors depend on underlying complex processes that occur (usually asynchronously) on a smaller spatial and temporal scale. These processes produce variability, i.e., intrinsic noise, in how the components interact within the larger biological system. Conversely, biological systems respond to their environment, a complex network of signals that occupy larger spatial and temporal scales (i.e., biological context). These signals affect extrinsic noise. Multi-scale interactions expand and complicate models of biology.

Modeling Noise and Heterogeneity: Noise propagates through complex systems to yield a wide variety of emergent behaviors. In biological systems, this variability gives rise to heterogeneity. Some modeling efforts may require complicated frameworks that account for and/or simulate biological noise. Modeling noise expands the scope of a model and complicates its simulation and model calibration.

Noise/Heterogeneity and measurement: The measurement process commonly incorporates inherently noisy biological processes (e.g., antibody binding, fluorescently tagged protein expression) that limit the measurement's sensitivity and specificity. The inherent variability in biological systems obscures measurements, especially those which average over a heterogeneous population. Experiments and measurements that aim to capture this variability often introduce perturbations and compromise biological context.

Precision vs Biological Context: Biological systems interact with their local environment. Cells, for example, engage paracrine signals supplied by their surrounding tissue. *In vitro* experiments and measurements permit more precise cellular measurements but strip away the paracrine signals that establish aspects of the cell's biological context. Experiments and measurements contend with a tradeoff between precision and biological context.

Complex Response to Perturbation and Experiment: Complex systems spread the impact of an intended localized perturbation across several interactions in efforts to maintain some preset biological state. This lends to uncertainties in the design and implementation of experiments. Accomplishing an intended perturbation or experimental condition may disrupt other essential drivers of the biological phenotype or sacrifice the biological context of the experiment.

Box 1.1. Interaction of Biology Complexity, Multiscale and Heterogeneity with Models and Measurements

measurements. Noise obscures the details of biological processes but, when appropriately considered, it can provide important insights into the mechanisms that drive a particular phenotype[91]. Models that encode the added details of intrinsic and extrinsic noise sources, however, contain additional free parameters and require measurements and experiments that can capture biological response to noise. These measurements tend to disrupt other important drivers of the phenotype[92].

1.5. Addressing the Demand for Measured Data

The persistent challenges of systems biology attest to an inextricable link between models and measured biological data. The mechanistic insights and predictive power of systems biology models depend on the amount, kind and precision of available measured data. The most appropriate measurements often do not exist and may be awaiting the advent of new measurement technologies. Systems biology cannot advance unless supported by an adequate supply of useful biological measurements. Scientists address this challenge with a myriad of technologies and practices.

1.5.1. Simpler Models Offset the Demand for Data

Systems biology has adopted strategies to cope with the current dearth of the measurements need for modeling biological processes. Specifically, model parameterization requires a sufficient number of quantitative measurements. Modelers can negotiate this demand for data by limiting the number of model parameters[56]. As mentioned above, simplifying assumptions help limit the size of a model and constrain its scope to a particular experimental question. Common examples include assumptions that neglect intracellular diffusion and localization; approximate large chains of reactions with a single Hill equation; and neglect avidity of protein oligomers. My colleagues in the Lopez lab and I have encoded models of extrinsic apoptosis with varying degrees of model simplification[93-95]. This has allowed more amenable models and insights that pertain to our specific experimental scope(s). These simplifying assumptions, however, prevent reliable extrapolation beyond the initial experimental scope. Course-grained descriptions (e.g. Boolean) of biological mechanism limit or eliminate the model parameters. These less

detailed models explain and predict non-quantitative observations; and therefore, require little or no quantitative data, but can only provide qualitative predictions[96]. Hybrid models append ODE modules to otherwise coarse-grained descriptions of a biological process, in order to reach a compromise between the need for more reliable and detailed model predictions and the need for a simplified model that tolerates absent of data[77]. This set up adds ODE parameters that require quantitative dynamic data to estimate. The output from the ODE module inputs into a Boolean, or similar, model that predicts qualitative results of the dynamics modeled in the ODE model. For instance, J. Molina-Mora et al. parameterized a model of sphingolipid metabolism fluorescence measurements of a sphingolipid indicator. This model predicted the dynamics of sphingolipid in perturbed cells and related those prediction, via a fuzzy logical model, to cell viability[97].

Scientists can, at least partially, forgo model calibration in more detailed models (e.g., ODE models) by instead analyzing parameter-independent properties of the model. For instance, Deyan Luan et al. avoided calibration of their ODE model of human extrinsic coagulation dynamics by focusing on qualitative information that can be gleaned through analysis of robust properties over a random (uncertain) set of parameters[98]. These analyses focus on model properties (e.g., sensitivity coefficients[98], critical components, bifurcation behaviors) that rely more on the model's topology than its parameters. Steady-state and other constraint-based models estimate a wide range of ODE model parameters that correspond to end-point steady-state behaviors. Model calibration for these models does not require as much quantitative or time-course data or computational resources, but the models provide low-precision predictions of dynamic behaviors.

1.5.2. Facilitating easier access to existing data

The most obvious response to the dearth of measured data in systems biology involves practices and techniques that increase the amount of useable measured data. Modeling efforts generally include the time-consuming step of mining (typically manually) through hundreds of publications, figures and data files to extract pertinent information and measurements. With a new finding appearing in PubMed every

thirty seconds[99], the production of new data far outpaces its integration into systems biology. These data have no standardized presentation and can often possess ambiguous naming, multiple references to a single object and, incompletely or inconsistently described experimental setups, contexts, labs, etc. Advances in automated and crowd-sourced data curation help accelerate the identification, extraction and organization of these biological data into more quickly searchable, interoperable and usable databases[100].

1.5.3. Technologies generate new quantitative data

While the modeling process can include information for different kinds of measured data, the model parameterization step requires specifically quantitative measurements (which constitute only a small fraction of the measurements of biological processes). Technological innovation drives ongoing improvement in the precision, reliability and capabilities of biological assays. Mukesh Kumar et. al. recently improved the quantification capabilities of Western blots by replacing the antibodies, that mediate detection, in a conventional western blot with isotopically label protein chimeras for mass spectrometry (MS) based detection[101]. This enabled absolute quantification of proteins in *Drosophila melanogaster* tissue[102]. Other researchers have transferred the Western blot into microfluidic devices in order to measure of protein abundance single cells[103]. This technology accomplished detection and profiling of rare circulating tumor cells[104]. The application of microfluidics to MS (using isobaric tag or iTRAQ-labels) has permitted precise quantitative measurement protein abundance single cells[105]. Integrating nanoscale "nanostraws" into microfluidic devices has given researchers to non-destructive access to intracellular protein concentration[106]. This potentially allows high-precision quantitative time-course measurements of currently elusive cellular contents.

Technological advances have also improved the characterization of protein location, activity, regulation (via post-translational modification), etc. For example, new technologies allow better assaying of kinase activity. Quantitative microscopy using novel peptide-based biosensors has recently accomplished multiplex, non-destructive, spatial-temporal, measurement of intracellular kinase

activity[107]. Multiplex substrate profile via MS allows quantitative profiling of a kinase's peptide substrates and characterizes the rate of phosphorylation for those substrate[108]. Proteins possess multiple phosphorylation sites. Recent application of MS has produced a quantitative landscape of the site-specific post-translational modification of a protein *in vitro*[109, 110].

These technologies improve the quantification and characterization of intracellular proteins, but restrict their measurements, generally, to a single snapshot per sample; the most precise measurements destroy the sample. Desire to maintain the biological context of an experiment may preclude the use of these technologies. Further, these technologies require technical skill and money to implement. Researchers may have reservations about investing in such expensive quantitative measurements if modeling efforts that incorporate those data do not substantially augment their investigation[71].

1.5.4. Improved processing of measurements

The quantitative measurements, typically needed for model model calibration, constitute only a small fraction of the data in biological research. There exist methods for making the remaining non-quantitative data more useful to conventional model calibration methods. Figure 1.2 describes different non-quantitative measurements of intracellular protein concentration[111]. For instance, the relationship between the intracellular concentration of a biological species and the fluorescent intensity of its indicator can be explained (to some extent) by some function, $x_{obs} = f(x, \boldsymbol{c})$. Noise and uncharacterized influences on the measurements (e.g. sensitivity and saturation limits) confound the relationship. Hence, the term "semi-quantitative" refers to uncertainty that exist in the structure and parameter values of the function. Researchers construct a standard curve of known concentrations and accompanying fluorescence measurements in order to model their relationship. In a recent example of elaborate such standard curves can become, scientists formulated a model of fluorescence measurements of mitochondrial transmembrane potential that accounted for cell-to-cell variations in mitochondrial number, size, affinity for the fluorophore etc[112]. This added detail improved the measurement by extracting the noise related to the highly variable cellular mitochondrial content[112].

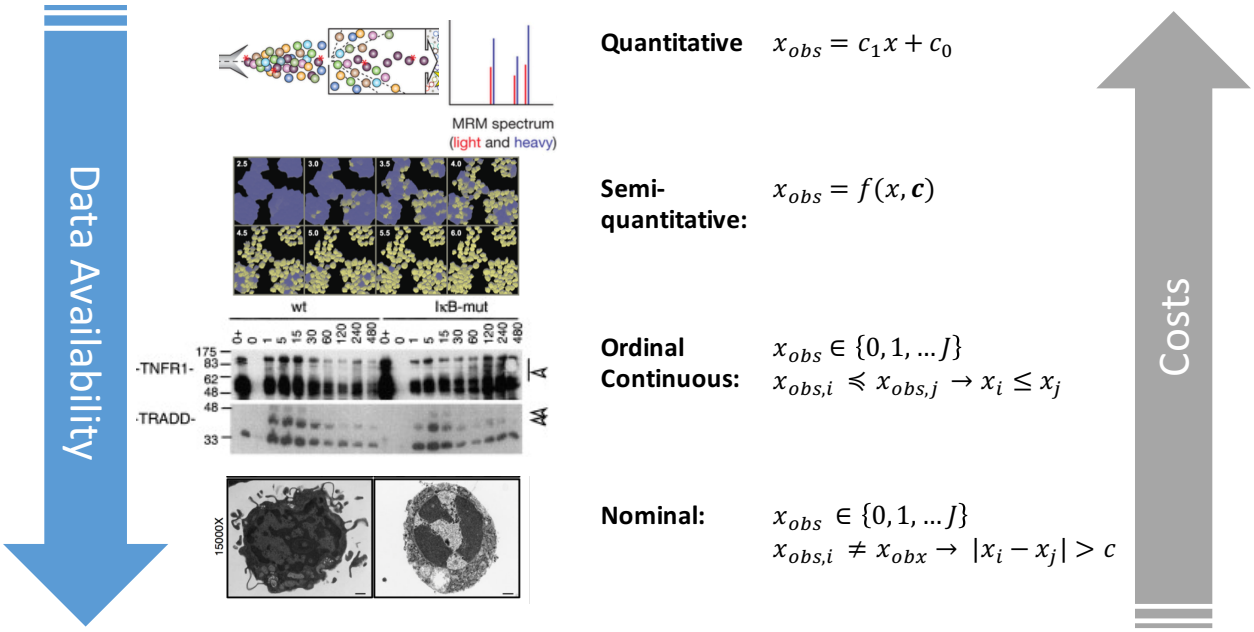


Figure 1.2: Types of Measurements in Biology Measurements assign a value, x_{obs}, to a characteristic, x, of the systems. The relation between the characteristic and its measured value depends on the kind of measurement. Quantitative measurements provide the most unambiguous description of a characteristic, and therefore are most useful. Less expensive, and less informative, measurements abound in biology. A, Quantitative measurements (e.g. mass spectrometric measurements of protein abundance) give a linear relationship between the characteristic, concentration, and its measured value. The measurements often contain a ratio or interval relation to a reference. B, Semi-quantitative measurements (e.g., fluorescence indicators or protein abundance) have a less certain nonlinear relationship between the characteristic to its measured value. C, Ordinal measurements (e.g., Western Blot of protein abundance) relate pairs of measurements by ordering them according to their magnitude. The larger measured value corresponds, generally, to an increased display of the characteristic. D, Nominal measurements (e.g., cell fate decision) further relate pairs of measurements according to whether are similar or different. Two measurements in the same nominal category represent similar displays of the characteristic. Two measurements with different nominal categories represent to dissimilar displays of the characteristic, and a value, c, represents the degree of dissimilarity.

Biological noise, and technical noise (i.e., arising in the measurement process) can prevent precise measurement of the interval between two quantities (e.g., in Western Blots[113]). Instead, the measurement only distinguishes which quantity as greater than the other. These ordinal measurements constitute the bulk of molecular biological data. Quantitative Western blots (QWBs) feature a precise fluorescent tag and standard curve to aid quantification[114]. Quantitative western blot measurements, however, contain several sources of uncertainty (e.g., nonlinearities in antibody affinity[115], unreliable loading controls[116], misleading normalization[113, 117]), that when neglected render the measurement ordinal[118]. Western blots (and QWBs), despite their abundance, go unused in model calibration efforts in systems biology. Other prominent measurements, nominal measurements, only convey binary information, (e.g. absent vs. present) or information that lacks an explicit ordering (e.g. cell type). Nominal data can provide useful information to modeling efforts, described below, when they have an established relationship to metric markers. Uncovering those relationships, however, requires additional measurements.

1.5.5. Using non-quantitative data as a substitute

The inadequate supply of quantitative measurements in biology has prompted modelers to explore model calibration strategies that leverage the far more abundantly available non-quantitative data. Perhaps the most common approach involves pseudo-arbitrarily assigning numerical values to stand in for the non-quantitative (specifically ordinal and nominal) measurements as comparators for the parameterized model[119]. For example, Sabrina Spencer used ODE models to characterize the contribution of non-genetic biological noise to heterogeneity in TRAIL induced apoptosis. This model arbitrarily defined apoptosis with a threshold at 50% maximum simulated PARP cleavage. Prior knowledge about the delayed uptick in the activity of PARP cleaving caspases that immediately precedes apoptosis supports use of this threshold[120]. Modelers extend this idea by employing signal temporal logic to construct intricate time-resolved numerical definitions of non-quantitative observations. Researchers recently modeled the dynamics endocytosis and subsequent GFP expression in newly transfected cells. This model represents

endocytosis and mRNA release as a single binary event; a Boolean trigger, or temporal threshold, that dictated an accompanying change in GFP translation rate[121].

When fluorescence measurements lack an accompanying standard curve (described above), data-driven normalization schemes scale semi-quantitative data and model predictions so that they possess consistent max and min, quartiles, means and standard deviations, etc[122]. Data-driven normalization schemes assume a linear, log-linear relationship, etc. between fluorescence and concentration. (The necessity of such normalization curves qualifies these measurements "semi-quantitative".) Arbitrarily scaling, normalization, and signal temporal logic impose assumptions about the relationship between some characteristic of the biological system (e.g., concentration) and its non-quantitative measurement (e.g., fluorescence). These assumptions potentially bias the model, rendering it prone to incorrectly predict new behaviors. The next chapter explores the risk of arbitrarily specifying the measurement model.

The incorporation of non-quantitative data into a model requires a functional definition of the non-quantitative data in terms of metrics present in the model. Experiments can help establish this definition. Andrew Paek et. al. characterized a relationship between the dynamics of p53 accumulation in cisplatin treated HCT116 cells and their commitment to cell death[123]. This produced a functional definition of cell death (specifically early and rapid accumulation of p53) that was used to simulate fractional killing of cells in a heterogeneous population[124]. This model then predicted bifurcate dynamics in p53 regulated cIAP and Bax abundance distinguish dying and surviving cells. Roux et al. similarity characterized a Heaviside function classifier that separates dead and surviving cells based on the underlying dynamics of caspase activity[125]. More elaborate methods for defining markers of cell fate exist but are not yet integrated into the mechanistic modeling paradigm[44]. This approach, though effective, does not entirely remove the requirement for quantitative time-course data since fluorescence data were required to characterize the relationship between cell signaling and cell death.

Non-quantitative measurements possess a minimal mathematical definition (listed in Figure 1.2); attaching additional details to these definitions requires the support of data to assure against potentially

biasing the model[126]. Optimal scaling accomplishes model calibration while only using these minimal mathematical definitions. For example, optimal scaling defines ordinal and nominal measurements with a range of corresponding values. The upper and lower boundaries of this representative range are positioned to maximize support of the model. The moveable boundaries in optimal scaling introduce additional free parameters to the model. Michael Pargett et. al. recently applied optimal scaling to partial differential equation models of embryonic expression of bone morphogenic protein (BMP) in drosophila[10]. Ordinal values of BMP expression were derived from colorimetric and fluorescence BMP indicators. These ordinal values were modeled as bounded intervals on the BMP concentrations. Adjusting the parameters in the partial differential equation model along with the boundaries of the BMP ordinal categories accomplished low-bias model parameter calibration. Schmiester et al. incorporated this strategy into PyPESTO, a model model calibration framework[127]. Optimal scaling uses hard boundaries to demark the edges of ordinal values. This rigorous treatment ignores uncharacterized noise sources that contribute to the nonquantitative character of measurements in biology (e.g. how noise sources can add uncertainty to the boundaries to demark the edges of ordinal values). Researchers therefore employ arbitrary gap intervals, between the boundaries of adjacent categories, to account for noise and increase the utility of the data[10, 127]. The discretization of continuous variables that occurs in optimal scaling also prevents easy implementation of Bayesian model calibration methods. Mitra et al. adapted a constraints-based method of calibrating models to certain categorical data to employ Bayesian methods[12]. Their work avoids discretizing continuous variables by instead defining a continuous probability of each categorical value. However, their formulation features *ad hoc* assumptions that may bias the model calibration.

1.6. Recommendations Moving Forward

Systems biology has become an essential driver of forward progress in biology. Its shortcomings attest to the mismatch between common goals and promises of systems biology modeling efforts and the amount of useable data required to accomplish those goals. Researchers and modelers often interact only after their individual projects have reached a stubborn road block. Researchers should initiate

collaborations between experimentalists and models at the start of biological investigations. Modelers should clearly communicate what realistic modeling capabilities exist, provided available or prospective data. This will help focus the modeling efforts on attainable goals that best support a specific investigation. Top-down data-driven modeling efforts have begun to incorporate larger and more diverse datasets. This trend will, hopefully, expand into bottom-up modeling efforts.

Chapter 2

Model certainty in biological network processes with uncertain data.

Michael W. Irvin, Arvind Ramanathan, and Carlos F. Lopez

bioRxiv 2021.05.18.44474

2.1. Summary

Mathematical models are often used to study the structure and dynamics of network-driven cellular processes. In cell biology, models representing biochemical reaction networks have provided significant insights but are often plagued by a dearth of available quantitative data necessary for simulation and analysis. This has in turn led to questions about the usefulness of biochemical network models with unidentifiable parameters and high-degree of parameter sloppiness. In response, approaches to incorporate highly-available non-quantitative data and use this data to improve model certainty have been undertaken with various degrees of success. Here we employ a Bayesian inference and Machine Learning approach to first explore how quantitative and non-quantitative data can constrain a mechanistic model of apoptosis execution, in which all models can be identified. We find that two orders of magnitude more ordinal data measurements than those typically collected are necessary to achieve the same accuracy as that obtained from a quantitative dataset. We also find that ordinal and nominal non-quantitative data on their own can be combined to reduce model uncertainty and thus improve model accuracy. Further analysis demonstrates that the accuracy and certainty of model predictions strongly depends on accurate formulations of the measurement as well as the size and make-up of the nonquantitative datasets. Finally, we demonstrate the potential of a data-driven Machine Learning measurement model to identify informative mechanistic features that predict or define nonquantitative cellular phenotypes, from a systems perspective.

2.2. Introduction

The combination of systems approaches and quantitative data promised a novel understanding of cellular mechanisms that would spur science-driven innovation in biology and medicine – as happened in physics, chemistry, and engineering[81]. Despite massive research efforts and data accumulation, our understanding of cellular regulation, signaling and many other processes as biomolecular systems remains rudimentary. The systems and quantitative biology fields continue to employ strategies from physics and engineering to construct models of biological mechanism from first principles[128, 129]. However, these strategies are incompatible with the types of measurements and observations that predominate biological investigations. Observations from biological experiments investigating cell fate outcomes (apoptosis, necroptosis, etc.) are collected as categorical values, which are hard to define in terms of variables encoded in mathematical mechanistic models of biological processes[130]. Therefore, the connection of mechanistic models to corresponding biological measurements is subject to practitioner interpretation. As a result, vast amounts of existing nonquantitative data in cell biology have led to mechanistic formulations based on simple inference and informal reasoning. Noise, complexity and the hierarchical organization of biology limits how we can experimentally perturb and measure biological systems[8, 59, 131, 132]. Therefore, a relative dearth of quantitative data exists that reveals itself in mechanistic models with poor parameter constraints. Unfortunately, both non-quantitative and quantitative data, collected in an unplanned manner, results in missed opportunities to quantitatively explain complex cellular mechanisms[133].

This data-to-knowledge problem in biology has prompted researchers to incorporate nonquantitative data as a complement or substitute for quantitative data in the development of mechanistic models[10-12, 127]. The traditional workflow employed to train mechanistic models to data comprises mechanistic models and experimental measurements linked through a calibration method (Box 2.1.)[134, 135]. Such workflows have been adapted to incorporate nonquantitative data into mechanistic models and have revealed their intrinsic value in mechanistic hypothesis exploration. For example, pioneering work by Pargett and co-workers employed optimal scaling and multi-objective optimization

for training mechanistic models to large ordinal datasets[10]. Schmiester et al. incorporated this strategy into PyPESTO, a model model calibration framework[127]. Their formulation imposes discrete boundaries on the mechanistic model to reflect discrete ordinal values in the data, but this approach limits their ability to integrate multiple datatypes or use Bayesian methods for training and uncertainty estimation of mechanistic models. More recently, Mitra et al. applied predefined constraint-based models of categorical data and modified their approach to allow definition of a likelihood function within a Bayesian formalism[11, 12]. However, the ad hoc nature of their constraint models leaves room for biasing assumptions. Given the limited application of Bayesian methods and biases introduced by ad hoc assumptions, the field still has a limited understanding of the contribution of nonquantitative and quantitative data to mechanistic knowledge in biological systems.

In this work, we tackle the data-to-knowledge challenge by introducing the concept of a measurement model, a statistical construct, into systems modeling approaches, which aims to rigorously define measurements and observations in terms of an underlying mechanisms[111]. This definition entails formulation of a function that maps variables encoded in a mechanistic model to values in the nonquantitative data. Our approach departs from previous work in that it uses machine-learning based classifiers whose free parameters are estimated to accomplish data-driven identification of measurement model properties. It also uses a probabilistic formulation that lends itself to Bayesian methods and can therefore provide an unbiased evaluation of the predictive power of models trained to nonquantitative data. In what follows, we present our findings about common types of biological measurements, followed by a presentation of our methodology. In this work we use a mechanistic model of apoptosis execution to demonstrate how the amount and type of data applied to a mechanistic model can affect its predictive power. It is well established that apoptosis signaling is involved in many cellular processes in health, disease, and development[136]. Its biological importance is further underscored by available quantitative and nonquantitative empirical data[125]. We also establish how an *ad hoc* formulation of a measurement model can lead to spurious results and further show how these a priori assumptions can be examined within a Bayesian, data-driven context. Finally, we demonstrate the potential of a machine learning measurement

model formulation to identify phenomenological links between features (e.g. predictors and drivers) of a biomolecular mechanism and emergent biological phenotype. We expect our approach to improve our understanding of the data-to-knowledge relationship in biological processes, leading to a probabilistic understanding of biochemical mechanisms, and accelerated identification of systems-level interactions that drive biological network dynamics.

2.3. Results

2.3.1. Contributions and biases from different data types to mechanistic models

We first explored how experimental data measurements are used to constrain mathematical models of cellular processes. Mechanistic models typically employ physical chemistry formalisms comprised of reaction rates and chemical species concentrations to represent networks of biochemical reactions. Direct quantitative measurement of all chemical reactions and species would provide needed model parameters to carry out simulations and in silico experiments. However, these measurements are typically not available and likely untenable for real systems, thus leading to indirect measurements used to infer model parameter values using an objective function (Eq. 7) or a likelihood function (Eq. 8). When these functions are optimized, the resulting mathematical model can provide valuable new predictions and insights about the cellular process. Measurements from cell biology experiments comprise four broad types, namely, nominal, ordinal, semi-quantitative, and quantitative (Figure 2.1); each data type reveals different insights about the cellular process. In apoptosis signaling, for instance, nominal observations supported early research where it helped identify key components in the apoptosis signaling pathway[137]. Apoptosis and survival outcomes – as indicated by nominal nuclear fragmentation data (Figure 2.1. top row) – helped determine two parallel signaling arcs that proceed following initiator caspase activation: mitochondria-dependent and -independent pathways[137]. These pathways trigger apoptosis by activating effector caspases[137]. We built an abridged Extrinsic Apoptosis Reaction Model (aEARM)[95], which represents these extrinsic apoptosis execution mechanisms as biomolecular reactions (Figure 2.2. A).

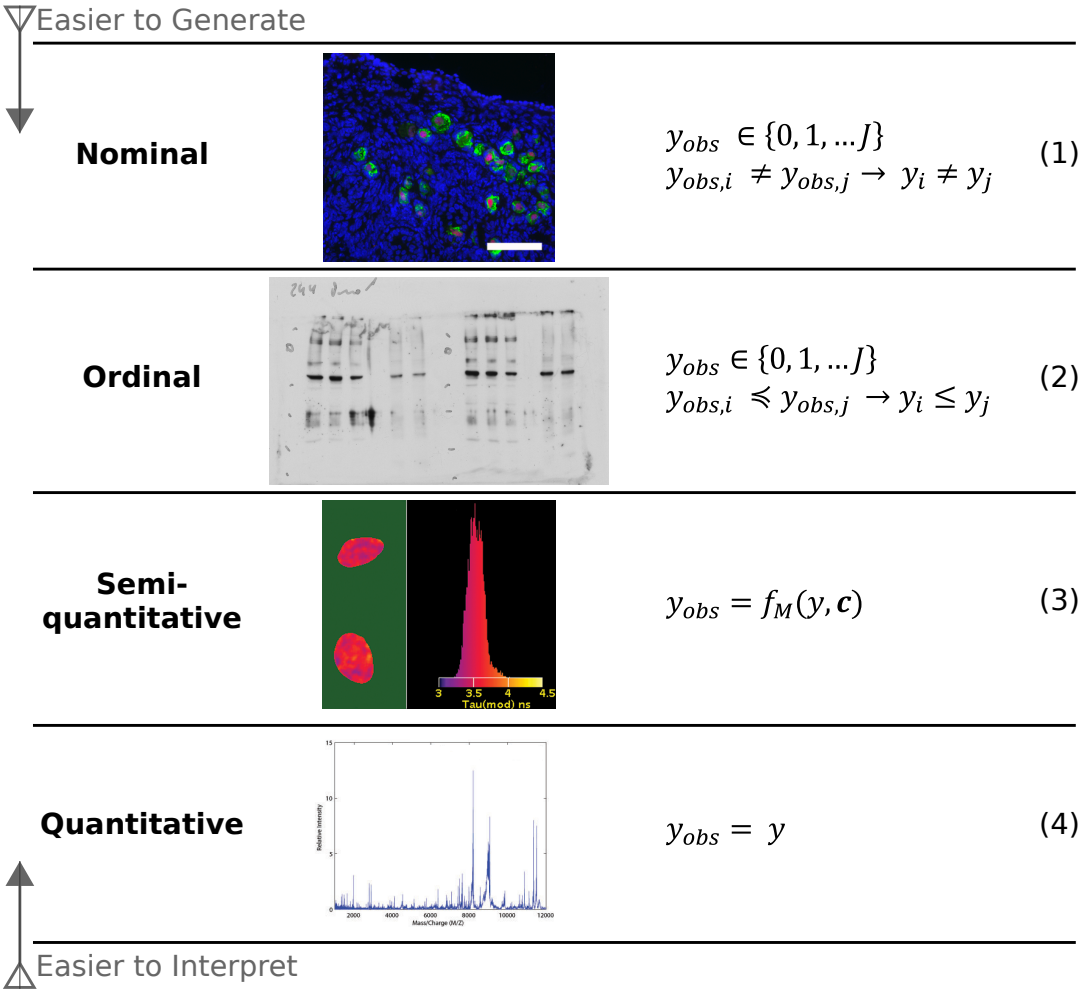
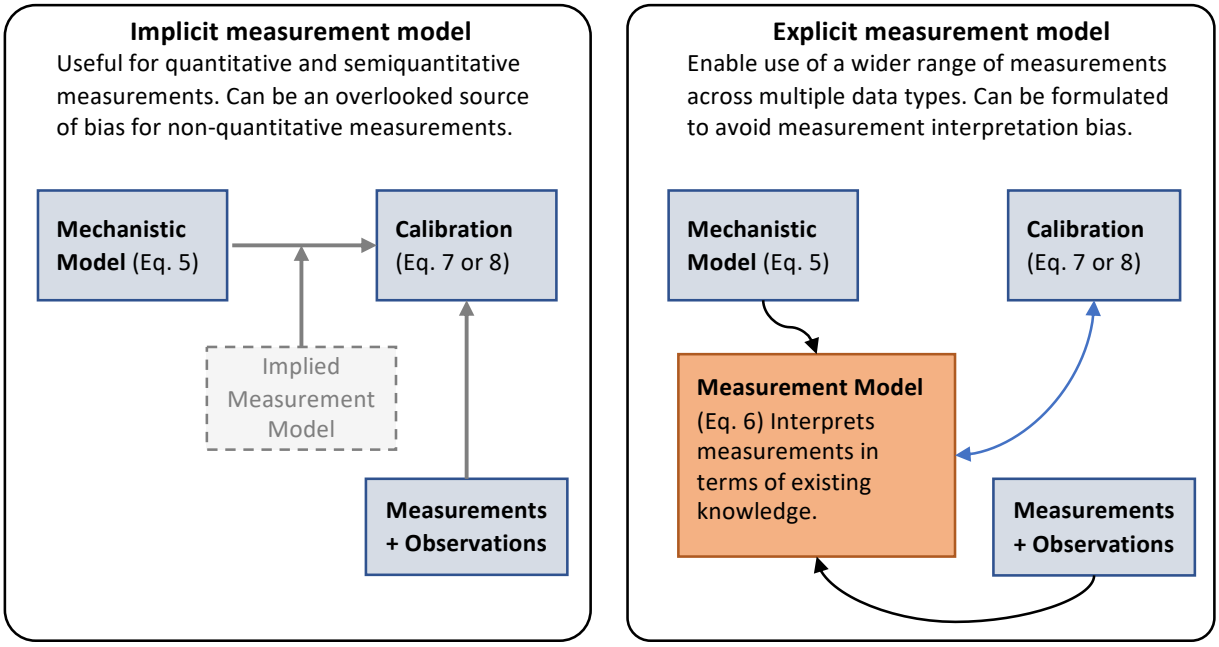


Figure 2.1. Measurements encountered in cell biology *Nominal* measurements (top) can help understand intracellular signaling activity as it relates to broader cellular and physiological behaviors. With cellular phenotype markers or drivers, we can attribute different nominal observations to distinct (intra)cellular states. This is often modeled as in Eq. 1, where each observable measurement (y_{obs}) corresponds to a given state. *Ordinal* measurements (second row) can be graded cellular phenotype observations (e.g., cell state transitions in cellular differentiation) or measurements of intracellular contents where noise can obscure intervals between values (e.g., Western Blots). Ordinal measurements imply a relative ordering of quantities along an axis but not their relative distance, i.e., we may know $y_i \leq y_j$ without knowing $y_i - y_j$ (Eq. 2). Semi-quantitative measurements (third row) typically arise when an investigation has progress toward a more quantitative understanding of the intracellular signaling. Semi-quantitative measurements (e.g., fluorescent intracellular markers) imply a quantitative relationship but a scaling function is necessary for true quantitation (Eq. 3). True quantitative measurements (bottom row) do not imply assumptions and the quantity measured can be used directly in the model (Eq. 4), such as mass-spectrometry protein concentration measurements. As shown schematically on the left triangle schematic, ordinal and nominal measurements are more abundant in biology due to their ease of production but are more difficult to interpret, whereas semi-quantitative and quantitative measurements are less common but have a more straightforward interpretation. (This figure is an updated version of Figure 1.2. and is included in this chapter for additional context and reader convenience)



Mechanistic Models of dynamic cellular processes are often *encoded* as systems of ordinary differential equations (eq. 5). Models are typically calibrated to one or many observables.

$$\boldsymbol{x}(t) = f(t, \boldsymbol{\theta}). \quad (5)$$

The **Measurement Model** translates measurements and observations into mechanistic model variables. The measurement model encodes understanding of the measurement as it relates to mechanistic knowledge. The measurement model is most often implied (left box), which can introduce significant biases for non-quantitative measurements. An explicit, adaptive measurement model (right box) can significantly alleviate this potential bias.

$$y(t_i, \boldsymbol{\theta}) = f_M\big(\boldsymbol{x}(t)\big) \quad (6)$$

Calibration minimizes the distance(d) between observed data ($\hat{\boldsymbol{y}}$) and related model predictions. Bayesian methods maximize likelihood functions $\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta})$ to optimize this distance. Calibration can include a measurement model to minimize bias between measurements/observations and the model measurands.

$$d = \sum_{i=1}^{n} w_i\big(\hat{y}(t_i) - y(t_i, \boldsymbol{\theta})\big)^2 \quad (7)$$

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}) = c + \sum_{i=1}^{n} \frac{-1}{2\sigma_i^2}\big(\hat{y}(t_i) - y(t_i, \boldsymbol{\theta})\big)^2 \quad (8)$$

Box 2.1. Objective functions and the role of a measurement model. Mechanistic models of biological processes are typically encoded as systems of (ordinary) differential equations (Eq. 5). Model calibration relies on an objective function (Eq. 7) -- or in a Bayesian setting, a likelihood function (Eq. 8) -- quantifies the degree of dissimilarity or similarity between model variables and corresponding measurements. Note, the objective or likelihood function uses measurement model (Eq. 6) which converts modeled variables $x(t)$ to a quantity $y(t_i, \theta)$ that can be compared to data $\hat{y}(t_i)$. In physics and engineering, where measurements are typically quantitative, the measurement model can be neglected. For nonquantitative measurements and observations, the measurement model takes more consideration.

Nominal observations do not provide a definitive estimation of their quantity of interest (i.e., their measurand) and instead, encode weak constraints on the measurand values (Eq. 1). They can guide mechanistic modeling by revealing salient structural elements of a cellular process but provide limited insight into the dynamics and complex regulatory cues of apoptosis signaling. Ordinal measurements have featured prominently in works investigating apoptosis signaling. They have uncovered clues about the dynamics and complex regulatory mechanisms of apoptosis. For instance, ordinal measurements of DISC (i.e., a ligand-dependent membrane bound 'death inducing signaling complex') components, initiator and effector caspases (Figure 2.1. second row), bid, etc. revealed how cells resist apoptosis by limiting (but not completely eliminating) pro-apoptotic cues[138]; the sub-maximal pro-apoptotic signaling presents as delay in the dynamics of caspase activation[139]. To better understand caspase activation dynamics and its effect on apoptosis and survival, we need mathematical models of the apoptosis signaling dynamics. Ordinal measurements, however, do not readily support a mathematical description of apoptosis signaling dynamics. Emerging work has leveraged ordinal and nominal measurements in the development of mathematical models of biological signaling but the weak constraints encoded by these measurements (Eq. 1 and Eq. 2) add uncertainty and bias to the modeling process.

Technical challenges confine our quantitative and semi-quantitative measurement to just a few apoptotic signaling proteins. Fluorescence indicators of caspase activity[125] (and by proxy, caspase substrate cleavage) enabled time course measurements of Bid and PARP cleavage dynamics (Figure 2.1. third row)[125]. They revealed pro-apoptotic activation of Bid and PARP, in TRAIL induced apoptotic HeLa cells, follows sigmoidal dynamics with delays and switch times that are sensitive to various regulatory factors. These measurements provide the details necessary for a mathematical description of apoptosis signaling dynamics and complexity. Our mathematical model aEARM captures the events from initial death ligand cue, initiator caspase activation, BID truncation (tBID), mitochondrial outer membrane permeabilization (MOMP) and eventual PARP cleavage (cPARP), as shown schematically on Figure 2.2. [140-144]. The model was calibrated to above fluorescence data, as described in Methods[145]. Semi-quantitative measurements like fluorescence, like non-quantitative measurements lack a definitive

estimation of the measurand because their interpretation requires mathematical manipulation, typically through scaling (Eq. 3), which can also add uncertainty and bias. Quantitative measurements can be used directly in a model without further modifications (Figure 2.1., fourth row) thus minimizing the uncertainty and bias introduced in the model from measurement interpretation. Therefore, the specific type of measurement and its interpretation could add significant uncertainty and bias to the mechanistic explanation of a given process.

To study the bias and uncertainty originating from different types of measurements, we introduce a concept from statistics, and social sciences: the measurement model (Box 2.1.)[146]. Briefly, a measurement model is a function (Eq. 6) that describes the relationship between the measurement and its measurand. This function maps variables from the mechanistic model $\boldsymbol{x}$ to the values expressed in the data $\hat{y}$. This function is often assumed or implied, particularly for semi-quantitative data that can more readily be applied to the model calibration. However, the application of nominal and ordinal datatypes to mechanistic models is not straightforward, because their interpretation (as we show in the following sections) can significantly bias model-derived insights. Consequently, modeling efforts have relied almost exclusively on quantitative and semi-quantitative data. By contrast, the much more abundant non-quantitative datatypes are often ignored or used inappropriately.

Early modeling efforts interpreted nonquantitative data as a series of arbitrary surrogate quantities for the ordinal or nominal values in a corresponding dataset[119, 134]. More recently, discrete boundaries on the values of the measurand were imposed along with a distance metric to describe how well the mechanistic model satisfies nominal or ordinal constraints in the non-quantitative data[10-12, 127]. These approaches reveal the value of nonquantitative data for mechanistic model calibration, but the often-ad hoc nature of these constraint-based measurement models has been an overlooked source of model bias. To minimize biases from the interpretation of non-quantitative datatypes and apply Bayesian inference methods for model calibration, we developed a data-driven probabilistic measurement model (Box 2.2.). Our measurement model is data-driven in that it possesses free parameters that are calibrated to match data; this lets us replace a priori assumptions about the measurement with a data-driven parametrization,

and thereby calibrate mechanistic models whose accuracy and precision better reflect the information contained in the data. Our measurement model is probabilistic as it replaces discrete boundary-based measurement models and distance metrics with a probability (Box 2.2., Eq. 9) of the ordinal or nominal value, which enables easy formulation of a likelihood function and application of Bayesian optimization methods that utilize MCMC sampling. In our approach, the measurement model is a mathematical construct that represents the measurand through a Machine-Learning probabilistic classifier whose free parameters are simultaneously estimated with the free parameters of the mathematical model during calibration (Box 2.2.). As a probabilistic classifier, the measurement model effectively describes the probability of the categories encoded in the non-quantitative data given values of the measurand (Eq. 9). The measurand, in our case, is encoded in the mechanistic model. For example, the measurement model (Eq. 9, Box 2.2.) can use ordinal logistic classifiers to model the probability of a categorical value as a function of variable(s) encoded by the mechanistic model. Also, the probability that a cell death or survival observations represents a specific state of the mechanistic model. In the calibration process, the measurement model is an explicit intermediate step between simulation of the mechanistic model dynamics and calculation of the likelihood (Box 2.2.). As described in the Methods section, this approach uses the Python based PySB models-as-programs framework and PyDREAM, a Python implementation of the DREAM(ZS) algorithm to sample posterior values of models' free parameters. However, other model building and parameter sampling (or optimization) algorithms could be employed by the user. In what follows, we examine the impact of different measurement modalities and interpretations on mechanistic model constraints in apoptosis execution. This work motivates an approach that could be generalized to any mathematical model to rigorously integrate quantitative and nonquantitative data types.

2.3.2. Uncertainty associated with different data types in model calibration

To date, molecular biology investigations of intracellular signaling processes and their mechanisms predominantly report nonquantitative measurements. However, it is unclear exactly how well these measurements support the development of mechanistic models. We therefore asked how various

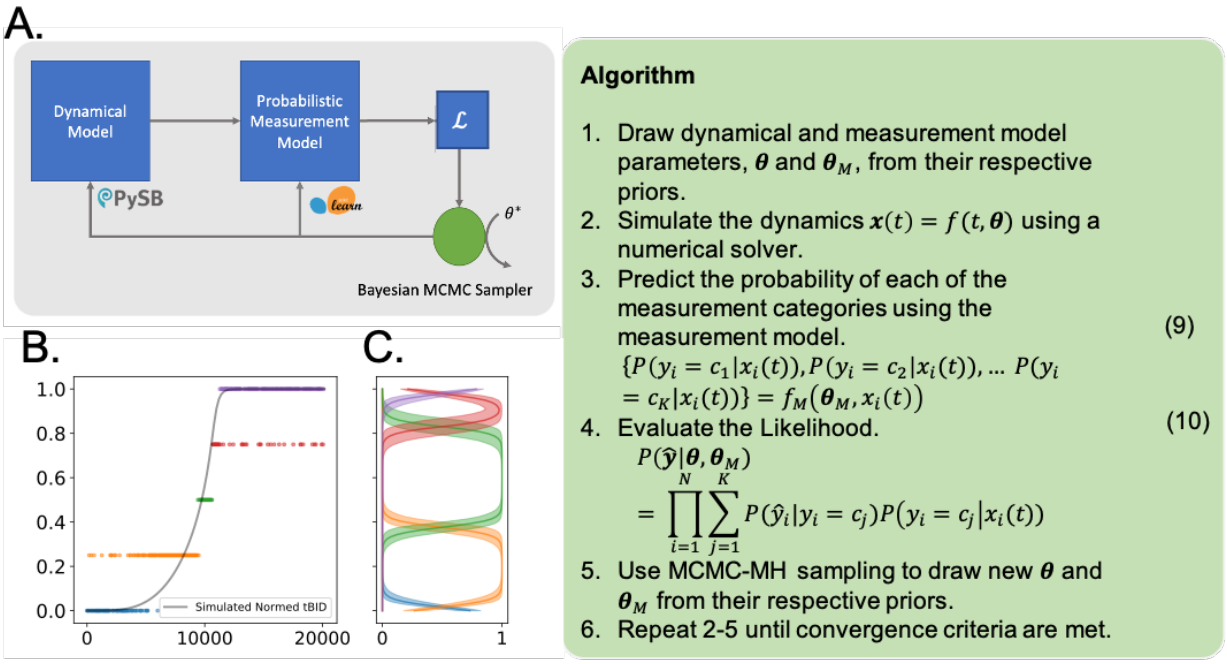
measurement datatypes impact the certainty and accuracy of model calibrations. Specifically, we explored how to adjust the size and make-up of nonquantitative datasets to better support mechanistic inferences. The resulting posterior predictive region for tBID dynamics of aEARM calibrated to (semi-quantitative) fluorescence data is shown in Figure 2.2. (B, top row). As expected, the data can effectively constrain the model and the 95% credible region of posterior predictions for tBid dynamics falls within the data uncertainty region. We then extracted a parameter vector from the fluorescence optimized data and used it as a baseline (reference) to generate ordinal datasets for tBID and other aEARM variables as described in Methods and shown Figures 2.2. (B, bottom four rows). These synthetic datasets could be considered as numerical representations of a time-course western blot dataset. We then calibrated aEARM kinetic rate and measurement model parameters to the ordinal and nominal datasets.

As shown in Figure 2.2. B, ordinal datasets accurately predicted quantitative predictions of "ground truth" dynamics for tBID. The 95% credible region of posterior predictions of tBID dynamics of aEARM trained to these ordinal datasets each contained "ground truth" dynamics for tBID. We also use the area bounded by the 95% credible region of posterior predictions of tBID as a measure of model certainty; with a smaller area indicating higher certainty. The ordinal dataset containing measurements at every 25-minute interval (i.e., typical of time-dependent western blot datasets), however, did not significantly constrain the posterior predictive regions of these dynamics (Figure 2.2. D). Increasing the number of measurements, however, increases the certainty of the posterior predictions of tBID dynamics; this certainty approaches that of the typical semi-quantitative (fluorescence) dataset, which has an area of 2.7, when then the number of ordinal measurements is increased threefold, which had an area of 6.2. The areas bounded by the 95% credible region for each ordinal time-course dataset is described in the Figure 2.2. B (Bottom two rows).

To explore the impact of nominal data on model optimization, we again extracted a parameter vector from the fluorescence optimized data and used it as a baseline (reference) to generate nominal datasets akin to an apoptosis execution observation as described in Methods. Previous work has described how features of apoptosis signaling dynamics can predict cell death vs survival18. The generated nominal

dataset describes binary cell-fate outcomes that emerge as a consequence of extrinsic apoptosis signaling dynamics. We encode this information in a nominal measurement model as described in Methods. Parameters of aEARM and the free-parameters encoded in the measurement model were jointly calibrated to a synthetically generated dataset of 400 survival vs death outcomes as shown in Figure 2.3. A (left). As shown in Figure 2.3. A (right), the binary cell-fate data minimally constrain the posterior predictive region of tBID dynamics relative to the prior constraints on the model. This is expected as the binary cell-fate data-type essentially condenses complex apoptotic signaling dynamics to a single categorical value.

In lieu of its limited ability to constrain mechanistic models, modeling efforts understandably disregard nominal data. However, we hypothesized that combining nonquantitative datatypes and covering multiple variables in the model could improve model certainty. To explore the effect of multiple data type combinations on model calibration, we again optimized the aEARM model parameters, but this time to a dataset containing nominal and ordinal measurements. As described in Methods, we added a synthetic dataset containing 61 ordinal time-course measurements for the DISC complex to the nominal dataset described above (Figure 2.3. B (left)). We modeled the likelihood of this combined dataset as the product of the likelihoods of the individual constituent datasets (see Methods for details). In Figure 2.3. A and 2.3. B (right), we see the nominal and ordinal datasets yields larger 95% credible regions for the posterior predictions of tBID dynamics. However, (in Figure 2.3. C) the combined dataset better constrained the posterior predictions of normalized tBID dynamics than either dataset alone, with a 95% credible region area of 26.5 (compared to 55.0 and 56.4 for the ordinal and nominal datasets alone). Therefore, the model uncertainty stemming from only using tBID nominal data was decreased by including more detailed upstream measurements. However, the contribution of DISC ordinal data alone was comparable to that of the tBID nominal data in isolation (Figure 2.3. B (right)). This data suggests that distributed measurements across multiple variables in a pathway yield synergistic effects on calibrated model accuracy and certainty.



Algorithm

1. Draw dynamical and measurement model parameters, $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_M$, from their respective priors.
2. Simulate the dynamics $\boldsymbol{x}(t) = f(t, \boldsymbol{\theta})$ using a numerical solver.
3. Predict the probability of each of the measurement categories using the measurement model. (9)
$$\{P(y_i = c_1|x_i(t)), P(y_i = c_2|x_i(t)), \dots P(y_i = c_K|x_i(t))\} = f_M(\boldsymbol{\theta}_M, x_i(t))$$
4. Evaluate the Likelihood. (10)
$$P(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\theta}_M) = \prod_{i=1}^{N} \sum_{j=1}^{K} P(\hat{y}_i|y_i = c_j) P(y_i = c_j|x_i(t))$$
5. Use MCMC-MH sampling to draw new $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_M$ from their respective priors.
6. Repeat 2-5 until convergence criteria are met.

Box 2.2. Model calibration with the data-driven probabilistic measurement model. A.) The measurement model is an intermediate step between the mechanistic model and likelihood function of the measurement/observations. It receives variables from the mechanistic model and transforms for use in the likelihood function. This probabilistic machine-learning measurement model estimates *probabilities* of class membership as a function of the mechanistic model variables (Eq. 9). This measurement model is data-driven in that it contains free-parameters that are evaluated via the likelihood function (Eq. 10). B.) The measurement model uses values of e.g., tBID (grey curve) to estimate the probability of membership in an ordinal category (dotted data). C.) Plots a posterior ensemble of estimates of the probability of membership into 5 ordinal categories (x-axis) as a function of normalized tBID concentration (y-axis). The plot shows the median (solid line) and 95% credible region (shaded region) of the predictions (Category colors match data plotted in B). **Algorithm)** The mechanistic model and measurement model are calibrated simultaneously using Bayesian sampling methods through stepwise operations as described in each numeral.

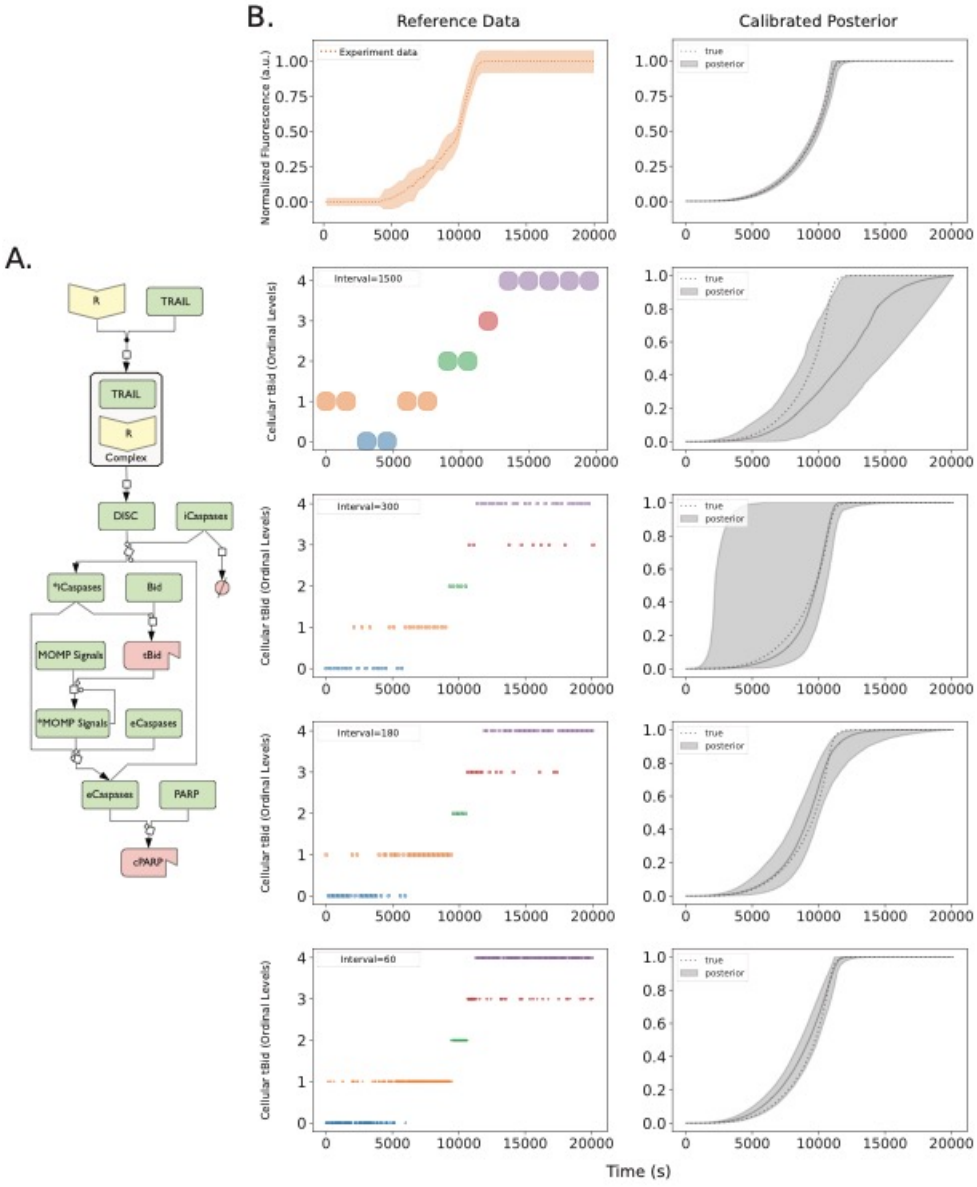


Figure 2.2. Predicted Bid truncation dynamics of aEARM trained to different sized ordinal datasets. Multiple Bayesian optimizations were run on the A.) abridged Extrinsic Apoptosis Reaction Model (aEARM) using different sized ordinal dataset to probe how dataset size influenced certainty of aEARM predictions. B.) Initiator caspase reporter (IC-RP) fluorescence time-course measurements (at 180s intervals) were measured (top left) as a proxy for truncated tBid (data from Albeck et al[28]). The plot shows the mean (dotted line) +/- 1 standard deviation (shaded region) for each time point. The 95% credible region (top right) of posterior predictions (shaded region) for tBID concentration in aEARM, calibrated to fluorescence measurements of IC-RP and EC-RP (See also Figure B.2.). The median prediction (solid line) and true (dotted line) tBID concentration trajectories are shown. In the next four rows (from top to bottom), Ordinal measurements of tBID (left) at every 1500, 300, 180 and 60s interval, respectively. The 95% credible region of predictions (shaded region), median prediction (solid line) and true (dotted line) tBID dynamics for aEARM calibrated to ordinal measurements of tBID and cPARP occurring at every 1500, 300, 180 and 60s timepoint are plotted in plots on the right. The plots for cPARP ordinal measurements and predictions are found in Figure B.2.

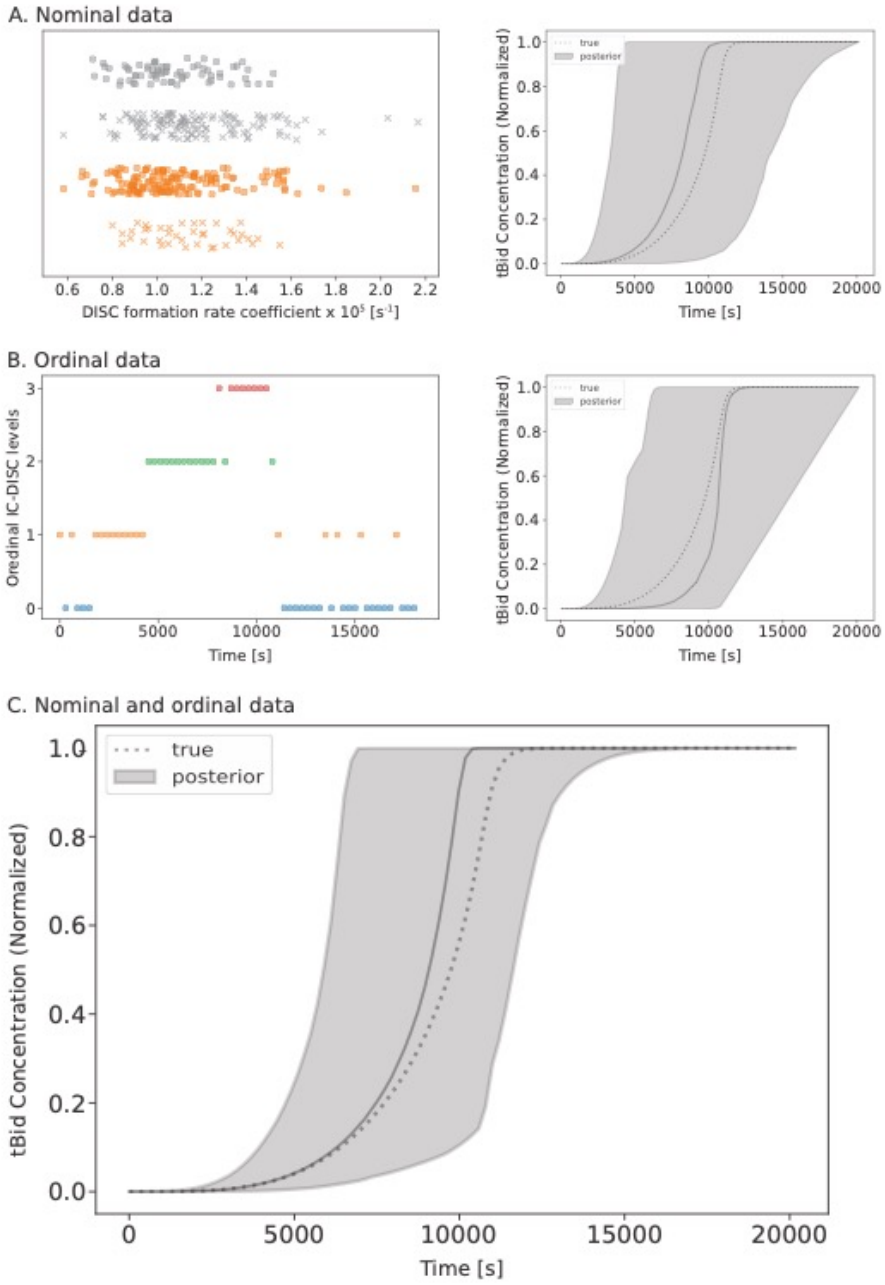


Figure 2.3. Predicted Bid truncation dynamics of aEARM trained to nominal and ordinal datasets. A.) Nominal cell death (x) vs survival (o) outcomes data for cells treated with 10ng/mL (orange) and 50ng/mL (grey) of TRAIL and with known relative values of DISC formation (x-axis). The 95% credible region (shaded region) of posterior predictions of tBID dynamics of aEARM calibrated to nominal data (right plot). The median prediction (solid line) and true (dotted line) are also plotted. B.) Ordinal measurements for initiator caspase-DISC colocalization (IC-DISC) at 300s intervals (left plot). The 95% credible region (shaded region) of posterior predictions of tBID dynamics of aEARM calibrated to ordinal IC-DISC data (right plot), and C.) of aEARM calibrated to nominal *and* ordinal IC-DISC data. The median prediction (solid line) and true (dotted line) were also plotted. The fit to IC-DISC data is shown in Figure B.8.

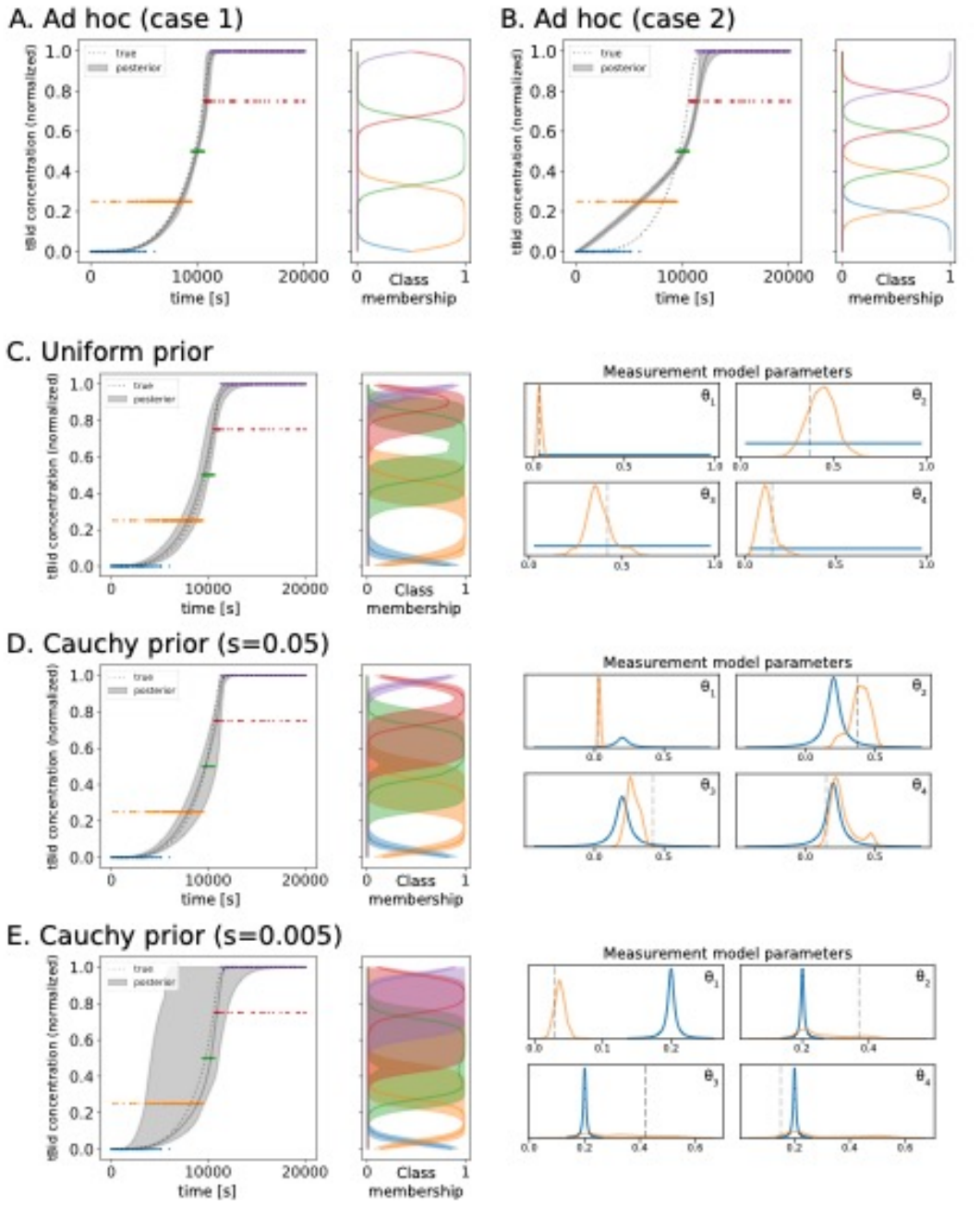


Figure 2.4. Predicted Bid truncation dynamics of aEARM trained to ordinal data using different measurement model parameterizations. A.) and B.) The 95% credible region of posterior predictions (shaded region) of tBID dynamics for aEARM calibrated to ordinal measurements two fixed parameterizations for the measurement model (see Table A.3). The adjacent panels plot the measurement models predicted probability of class membership (x-axis) as a function of normalized tBID concentration (y-axis). C.) D.) and E.) The 95% credible region of posterior predictions (shaded region) of tBID dynamics of aEARM calibrated to ordinal measurements uniform, Cauchy (scale=0.05) and Cauchy (scale=0.005) prior distributions for the parameterizations for the measurement model, respectively. In each, the median prediction (solid line) and true (dotted line) tBID dynamics are also shown. The adjacent panels give the 95% credible region of posterior predictions of the probability of class membership (x-axis) as a function of normalized tBID concentration (y-axis). Four accompanying plots show the prior (blue), posterior (orange) and true (dashed line) values of measurement model parameters.

2.3.3. Data-driven measurement model as an indicator of model bias

Traditionally, applying quantitative or semi-quantitative data to a mechanistic model has been relatively straightforward as they typically follow a well-establish and simple relationship between the measurement and the measurand. However, for non-quantitative data, measurement uncertainty can prompt researchers to make assumptions about the relationship between measurement and measurand, which may negatively impact in the resulting mechanistic model. We therefore asked how the encoding of assumptions into our models of non-quantitative measurements could impact mechanistic model calibrations. To attain this goal, we calibrated aEARM kinetic rate parameters to the ordinal dataset, but this time we replaced the free parameters in the measurement model fixed a priori parameterizations or we encoded our assumptions as priors on the measurement model's free-parameters. We tested four situations: (i) fixed parameters, a case where the measurement model is pre-parameterized by the user, presumably reflecting full confidence in their assumptions about the measurement; (ii) strong prior knowledge, a case where there is strong belief in the assumed values of the measurement model parameters; (iii) weak prior knowledge, a case where there is only weak belief in the assumed values of the measurement model parameters; and (iv) no prior knowledge, that is no constraints on the measurement model parameters.

Figures 2.4. A and 2.4. B show the ordinal class probabilities for tBID as modeled by (i) two distinct pre-parameterized measurement models. In case 1, lowest and highest categories correspond to a narrow range of tBID values, while the three internal categories each account for roughly 1/3rd of the tBID range. This parameterization might aim to account for effects of sensitivity and saturation on the measurement. In case 2, all five ordinal categories each account for 1/5th of the range of tBID values. The right panels in Figures 2.4. A. and 2.4. B. show the assumed relationship between tBID concentration and probability of each ordinal category. Figures 2.4. A. and 2.4. B. (left plots) also show posterior predictions of tBID dynamics by aEARM calibrated to the ordinal dataset using these fixed pre-parameterized measurement models. The different measurement model pre-parameterization produced markedly different posterior predictions of tBID dynamics by the resulting aEARM calibrations. This

raises potential concerns that assumptions in our interpretation of the measurement can artificially influence our interpretation of the mechanism.

In Figure 2.4. B, the 50% probability boundary between adjacent categories occurs at every 0.2 interval; essentially dividing the [0,1] range of tBID values into five equally spaced ordinal categories. Shown in Figure 2.4. C-E, we represent this as a flexible assumption by encoding it in our priors (ii – iv) i.e., Cauchy distributions centered at every 0.2 interval (as detailed in Methods). The smaller the scale – more narrowly focused – prior distributions reflect less flexibility in the free-parameter and a stronger belief in our prior assumptions. Figure 2.4. C-E shows the posterior predictions of tBID dynamics of aEARM calibrated to the ordinal dataset using increasingly more constrained priors on the measurement model parameters. The resulting posterior predictions of tBID dynamics were all less constrained than that of aEARM calibrated using fixed pre-parameterized measurement models (Figure 2.4. A and 2.4. B) but they were more accurate as they contained the "ground-truth" tBID dynamics. Strongly constrained priors on the measurement model parameters (ii) produced a less certain mechanistic model; as indicated by its wider 95% credible region of posterior predictions of tBID dynamics (Figure 2.4. E). The posterior distributions of measurement model parameters were spread out enough to give significant support of both the "ground truth" and the a priori assumed parameter values. This uncertainty in the measurement model parameter distributions translated into a less certain measurement model and less certain predictions of tBID dynamics. Weaker constraints on the measurement model parameters were encoded via larger scale prior distribution (Figure 2.4. D). In Figure 2.4. D we see these prior distributions, while centered on our a priori assumptions, includes the "ground truth" parameters. The posterior distributions of the measurement model parameters were therefore more constrained; likewise, the measurement model and posterior predictions of tBID dynamics had more certainty. This is also observed, in Figure 2.4. C, the case where no prior assumptions were applied to the measurement model parameters (iv). The accuracy of the predictions of tBID dynamics comes from the flexibility of the data-driven measurement models' parameters. This flexibility enables optimization (or prediction) of key properties of the measurement given the data. Figure 2.4. (right panels) shows the predicted probabilities of the ordinal

categories (as a function predicted cellular tBID content); these predictions are accurate in that they contain "ground truth" probabilities. Using this approach, we calibrated more accurate models of mechanism by simultaneously learning a more accurate model of the measurement. This motivates us to further explore the data-driven measurement model as a potential new avenue for insights.

2.3.4. Mechanistic insights from data-driven measurement models

We have shown thus far how a machine-learning measurement model can reduce uncertainty and increase accuracy in model calibration. Through mechanistic model calibration to categorical data, we effectively employ machine-learning classifiers to constrain mechanistic model dynamics to a corresponding categorical phenotype. We can then employ the measurement model in reverse, to better understand how properties of a biological mechanism predict, drive and define a particular phenotype. This kind of knowledge would be essential for model-driven experimental data acquisition and model-guided validation.

To demonstrate this concept, we calibrated aEARM to nominal cell survival vs death data using a measurement model that estimated the contribution of variables in aEARM to the cell survival vs. death predition. The survival vs death dataset was synthesized based on maximum log-rate of change of tBID and the time at which the rate of change maximized; these features were encoded into the measurement model, but their contribution was represented as a free parameter. In addition, the measurement model also considered the potential contribution of an unrelated variable (i.e. concentration of a reactants in reactions that occurred independently of the cell death ligand). Jointly calibrating aEARM and this measurement model to cell survival vs death data allowed data-driven predictions of how variables encoded in aEARM relate to cell survival vs death. Figure 2.5. shows posterior predictions of the values of potential predictors of cell survival vs death. The shade region marks the 95% credible interval for the line marking 50% cell survival probability. Figure 2.5. (bottom row) provides the posterior distribution of weight coefficients for each the features encoded in the measurement model. (Larger absolute values of the weight coefficient indicate greater importance of the feature.) The calibrated measurement model

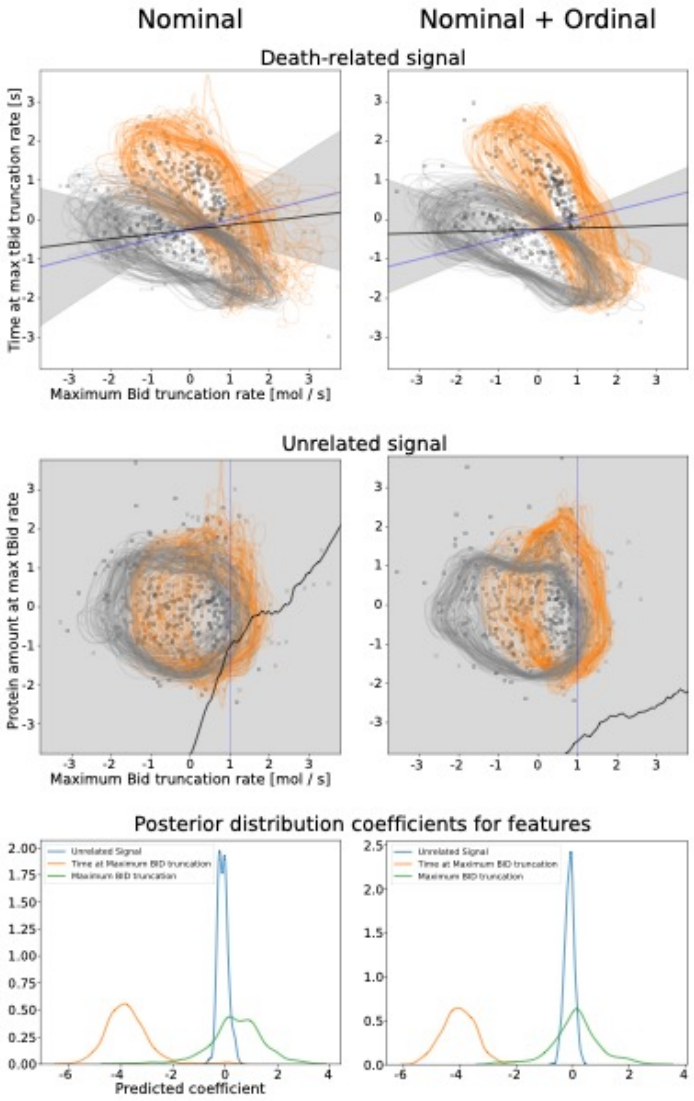


Figure 2.5. Measurement model predicts features of cell death vs. survival using aEARM calibrated to cell death datasets. Normalized predicted values of the features used in the cell death vs. survival measurement model – the x-axis is the maximum Bid truncation rate, and the y-axis is the time at maximum Bid truncation rate (top row) or an unrelated non-apoptotic signal (middle row) – for corresponding to observed cell death (x) and survival (o) outcomes. These feature values are modeled by aEARM parameterized by 100 parameter vectors randomly drawn from the posterior; for each parameterization, 5 out of the total simulated population of 400 cells were plotted. The grey and orange curves, in these plots, are 0.05 contours for the estimated density of simulated cell populations produced for each of the 100 parameter vectors – grey and orange correspond to 50 and 10ng/ml TRAIL treatments, respectively. The measurement model predicts a probability of cell death vs survival based on simulated values of the above features. The lower right region of the plots in the top row. (i.e., early maximization of Bid truncation and higher maximal Bid truncation rates) is associated with higher probability of cell death. The shaded region is the 95% credible region of the posterior prediction of the line marking 50% probability of cell death or survival. The black and blue lines are the median predicted and true 50% probability lines, respectively. The bottom row plots the posterior distributions of the weight for each feature (i.e., the product of the slope term and feature coefficient encoded in the measurement model): maximum Bid truncation rate (green), time at maximum Bid truncation (orange) and unrelated non-apoptotic signal (blue). Plots in the left column are predictions of aEARM calibrated to the cell death vs. survival dataset. Plots right column were those of aEARM calibrated to the cell death vs survival + ordinal IC-DISC combined dataset.

correctly identified time at maximum Bid truncation as the most important predictor of cell survival; and the unrelated variable as the least important predictor. Calibration of aEARM to the mixed dataset, described in the previous section, yielded a measurement model that equivalently predicted identified time at maximum Bid truncation as the most important predictor of cell survival; and the unrelated variable as the least important predictor. Calibration of a mechanistic model to categorical phenotype data, using data-driven measurement models, enabled correct identification of predictors (and potentially drivers or markers) of categorical phenotypes. The data-driven probabilistic measurement model we propose in this research was essential to this finding**.**

2.4. Discussion

We used data-driven probabilistic measurement models to calibrate, using Bayesian methods, a dynamical model of biological mechanism to quantitative and nonquantitative data. Our approach allowed us to estimate posterior predictive regions for the calibrated models and to observe how the size of a dataset, its different measurement types, and our assumptions about the measurements affect model accuracy and certainty. Our findings support results from previous studies that suggest nonquantitative data are valuable for mechanistic modeling efforts[10-12, 127]. For instance, a sufficiently large ordinal dataset can constrain the posterior predictions of a mechanistic model as much as quantitative dataset. However, we far more nonquantitative data than is typically generated would be necessary for nonquantitative assays to match the information content of quantitative assays. In Figure 2.2. B (second row), fourteen ordinal measurements of tBID – typical of common immunoblot measurements of intracellular biology – did not constrain the model around an accurate prediction of tBID dynamics. Instead, it took 24x as many ordinal measurements of tBID (336 measurements) to constrain the mechanistic model of apoptosis as well as the fluorescence dataset (112 measurements). We also found that datasets that combined categorical measurements of multiple variables in aEARM out-perform the datasets with measurements of an individual variable. These findings suggest one could overcome challenges posed by a dearth of

quantitative data by devising experiments that, while nonquantitative, produce a larger number of diverse measurements that can cover multiple variables.

We also found the posterior predictions of our mechanistic model were sensitive to the assumptions, we encode in the measurement model, about the relationship between measurement and measurand. All measurements possess uncertain (or unknown) properties, but this uncertainty has a pronounced presence in nonquantitative measurements. The limitations of nonquantitative data exist because they impose less informative constraints on models, and this leaves room for biasing assumptions and/or uncertainty. Uncertainty in nonquantitative measurements drives the, often unacknowledged and implicit, assumptions about the relationship between measurement and measurand (i.e. between data and model). With the proposed Bayesian calibration framework, we are able to observe how assumptions about measurement affected the uncertainty and accuracy of the posterior predictions, in essence providing a measurable quality of how well the model can make mechanistic predictions. We found that inaccurate ad hoc assumptions about the measurement could produce models that suggested, with a higher degree certainty, an inaccurate prediction (Figure 2.4. B). This finding suggests that ad hoc assumptions about measurements can lull practitioners into a false sense of confidence about the model and the data. This concern also motivated Schmiester and co-workers to avoid certain ad hoc assumption in their model calibration approach[127].

Having a measurement model whose attributes are determined by data creates an opportunity to learn new details about the relationship between a measurement and its measurand(s). For instance, could a model of biological mechanism plus cell phenotype observations data enable identification of cell phenotype predictors? To explored this, we encoded a small number of suspected cell-fate predictors into our measurement model and let the data (and the mechanistic model) determine, through model calibration, their respective contribution to phenotype. In doing so, model calibration using our data-driven measurement model performed feature selection to correctly identify the most important predictor of cell death. In general, this kind of measurement model, which relates mechanism to cellular phenotype,

can be used to predict phenotype outcomes and identify potentially informative experimental conditions from in silico perturbation experiments.

2.5. Conclusions

The present work presents an analysis and a proof-of-concept that can be improved upon in future work. We chose linear logistic classifiers, as they enable easy formulation of a likelihood function and application of Bayesian calibration methods, but other probabilistic classifiers could be used. We constrained our measurement representation to small number of potential features to avoid complications of high dimensionality to our machine learned measurement model. However, dimensionality reduction and feature learning (e.g. PCA) can, in theory, be integrated into the measurement model's preprocessing and/or model calibration workflow. Possibilities for integrating more complex machine learning into models of measurement will depend on dataset size, computational power, and modeling goals.

Our work introduces the concept of measurement models to the mechanistic modeling paradigm. Measurement models have their origin in social sciences and statistics[146]. They also appear in more quantitative applications; some recent examples include management[147], manufacturing[148], and computer vision[149]. These measurement models can take on more complexity than the examples we provided, depending on the unique needs of the problems in these areas. The use of measurement models in these areas is motivated by a desire to define and quantify observations of nuanced and/or subjective phenomena; and connect those observations to an underlying theory. Biology, being "harder" than social sciences, but arguably "softer" than physics will straddle the technical domains of both. As a field, we face the same challenge as these social sciences given that our mechanistic models are situated within a larger context of explaining nuanced and subjective biological phenomena (e.g. cell-fate, morphology, physiology and overall health vs. pathology). As practitioners, we never encode *everything* into our mechanistic models; instead there is always some aspect of the model (or its interpretation) that aims to connect back to these relevant biological phenomena. This fact ultimately motivates our application of data-driven probabilistic measurement models in our mechanistic models of intracellular biology.

2.6. Methods

2.6.1. Extrinsic Apoptosis Reaction Model

We built an abridged extrinsic apoptosis reaction model (aEARM) and trained it using PyDREAM to normalized fluorescence time-course data[95]. We built this abridged version of EARM to simplify convergence of Bayesian calibration algorithms and thus make feasible probability-based predictions on the model-data relationship[95]. The aEARM abstracts detailed mitochondrial reactions from the original model as two sequential mitochondrial outer membrane pore (MOMP) "signal" activation steps. In addition, apoptosome formation and effector caspase activation reactions take place in a single activation step. The aEARM does capture key dynamic characteristics, such as the snap-action delay dynamics of apoptotic effector molecules that is observed empirically[145]. For this work, three additional non-apoptotic species were encoded and linked via feedback activation and inactivation loops to test whether our data-driven measurement model could discriminate between drivers and non-drivers of apoptosis. (Supplemental Table 2). These additional species and reactions do not interact with any species or reaction in the aEARM model. The aEARM was encoded using rule-based modeling python package PySB[150].

The aEARM parameters – initial conditions and rate coefficients – were adapted from the previously developed EARM and/or calibrated to fit available fluorescence data. Initial conditions parameters were lifted from the previously developed EARM (Supplemental Table 1). Previous work characterized extrinsic heterogeneity in the expression of proteins and its effect on apoptosis. To model extrinsic heterogeneity in apoptosis signaling, initial values of certain species (marked in table 1) were sampled from a log-normal distribution such that its mean equaled that in Supplemental Table 1 and coefficient of variation was 0.20. Rate coefficients were calibrated (described below) to fit normalized fluorescence time-course measurements of initiator and effector caspase reporter proteins (IC-RP and EC-RP respectively).

2.6.2. Integrating aEARM Dynamics

Snap-action delay dynamics present challenges for Ordinary Differential Equation (ODE)-based models, as they feature rapid non-stiff to stiff transitions during integration. For this work we employed the LSODA integrator (from scipy, via the PySB solver suite), suitable for non-stiff/stiff systems[151]. However, we found that particularly poorly behaved parameter vectors could prolong integration evaluations in LSODA. Integrator settings were adjusted for efficiency and accuracy of integration as follows: `mxstep` (2^20), `atol` (1e-6 default), `rtol` (1e-3 default). The aEARM was integrated over a linear space of 100 time-points spanning 0 to 20160 seconds, in direct correspondence with the fluorescence time-course data[145]. Additional time-points in the data were obtained via linear interpolation.

2.6.3. Measurement Models and Likelihood Functions

Likelihood formulations incorporated a measurement model and resulting distance metric for each datatype in the study: fluorescence time-course data, synthetic ordinal time-course data, and synthetic survival vs death binary data for a sample of 400 initial conditions. These likelihood functions were used to calibrate the models to each dataset. In addition to their use in the likelihood formulation, the measurement models, were also used to generate synthetic non-quantitative datasets.

We first trained the aEARM to normalized fluorescence time-course data for IC-RP and EC-RP, i.e. fluorescent proxies for substrates of initiator and effector caspase, respectively (i.e. Bid and PARP, respectively). Consistent with previous work, we defined a likelihood that assume an i.i.d. Gaussian-noise component $\epsilon \sim N(0, \sigma^2)$ on normalized tBID and cPARP predictions of the aEARM; where σ^2assumedly equals the variance of the data[38,,95]. This yields a log-likelihood function (Eq. 11) where data the, $\hat{y}$, and normalized aEARM predictions, y, are compared for each time-point, t, and observable, i (i.e. tBID/IC-RP and cPARP/EC-RP). The aEARM trained to these fluorescence data served as the starting point in the synthesis of ordinal, nominal, mixed, etc. datasets, below.

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}) = \sum_1^N \sum_t^T -1/2\sigma_i(t)^2 \times \left(\hat{y}_i(t) - y_i(t, \boldsymbol{\theta})\right)^2 \quad (11)$$

To train the aEARM to synthetic ordinal time-course data, a measurement model (i.e. that models the probability of each ordinal category as a function of an aEARM variable) was defined and applied in the formulation of a likelihood function[152]. The ordinal logistic regression python package, MORD, applies empirical ordering constraints to Scikit-Learn's logistic regression class; this class then calculates a probability for each ordinal category[162]. The ordinal logistic model, encoded in MORD, defines ordinal constraints as a linear function of predicted values of an aEARM variable (e.g. $p\left(y_{tBID} \geq c_j \middle| x_{tBID}\right) = \varphi\left(\alpha x_{tBID} + \beta_j\right)$ for aEARM variable, x_{tBID}) where each ordinal constraint, j, is a logistic function $\varphi(z)$ with a different offset coefficient, β_j, but shared slope coefficient, α, for each of the ordinal categories. Each ordinal constraint function is combined, using the *sequential model* (i.e. the product of the logistic functions), to give a probability of each ordinal category, $P\left(y_i(t) = c_j \middle| x_i(t, \boldsymbol{\theta}), \alpha_i, \beta_{i,j}\right)$[153, 154]. These offset and slope coefficients are additional free parameters to be inferred in the model calibration. For example, a measurement model with K categories can be defined using $K - 1$ ordinal constraints and will therefore add a total of K free parameters (i.e. $K - 1$ offset coefficients and 1 shared slope coefficient) to the model. We also encoded error in our synthetic ordinal data by defining a 5% misclassification probability; i.e. we assume 95% probability the reported ordinal category, $c_j = \hat{y}$, and 2.5% probability of adjacent categories, $c_j = \hat{y} \pm 1$, (5% for adjacent terminal categories). We model this by the marginal probability that the observation classified into the category predicted by the model: $\sum_j^K P\left(\hat{y}_i(t) \middle| y_i(t) = c_j\right)$[154]. Together, this yields a log-likelihood function (Eq. 12) where the probability of each category c_j is calculated for each time-point, t, and observable, i; and applied toward a likelihood of the data $\hat{y}$ given the model. Where noted, we also trained the aEARM using measurement models with preset fixed parameters (Table A.3.).

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i}^{N} \sum_{t}^{T} \log \sum_{j}^{K} P\big(\hat{y}_i(t)\big|y_i(t) = c_j\big) P\big(y_i(t) = c_j\big|x_i(t, \boldsymbol{\theta}), \alpha_i, \beta_{i,j}\big) \quad (12)$$

We trained aEARM to synthetic binary (survival vs death) data by incorporating a measurement model (i.e. logistic model of the probability of each categorical outcome) similar to that used for the ordinal data. We used the Scikit-Learn logistic regression class to model the probability of a cell-death outcome, $y = c_1$, as a linear function of features, x_l, derived from the aEARM simulation: $p(y = c_1|\boldsymbol{x}) = \varphi\left(\alpha\left(\beta + \sum_l^L \beta_l x_l\right)\right)$, where α is a slope term, β is an intercept and β_l are weight coefficients for each of the L features[155]. Previous studies used *a priori* knowledge and assumptions about which features of a cell-fate marker's dynamics to associate with the binary outcome. For instance, recent work delineates necrotic and survival cell fate outcomes using a threshold in the concentration of a known necroptosis marker (this assumption enabled models of necroptosis in the absence of an established relationship between the dynamics of the marker and commitment to necroptosis). Roux et al, investigated an empirical relationship between initiator caspase reporter protein (IC-RP), a fluorescent indicator of caspase activity or proxy for caspase substrate cleavage, and apoptosis in TRAIL stimulated HeLa cells[125]. They found, instead of concentration, the maximum rate of change in IC-RP and the time when that rate of change maximized better predicted the apoptosis-survival decision[125]. The features we use in our study are based on findings by Roux et al[125]. The features are derived from aEARM simulated tBID dynamics, $x_{tBID}(t, \boldsymbol{\theta})$: time at maximum rate of change, and log-maximum rate of change. To test the measurement model's ability to discriminate between predictors and non-predictors of cell death, we encoded an additional feature: the concentration of an unrelated non-apoptotic species (USM2 in Table A.2.) when bid truncation maximizes. Together this totals three features. We interpret each observation in the dataset as an independent Bernoulli random variable. Each cell death vs survival observation is compared with these three features, $x_{l,m}$, extracted from an aEARM trajectory that was simulated from a unique vector of initial conditions. There were 400 observations; 2 sets of 200 observations corresponding to 10 and 50ng/mL initial ligand concentration. Together, this yields a log-likelihood function (Eq. 13)

where each, m, of the M aEARM simulated trajectories corresponds to an observation $\hat{y}_m$. Given the definitiveness of observed surviving vs dead outcomes, we considered the chance of misclassification to be zero (i.e. $P(\hat{y}_m|y_m = c_1) = 0$ when $\hat{y}_m \neq c_1$).

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\boldsymbol{m}}^{\boldsymbol{M}} P(\hat{y}_m|y_m = c_1) \log \varphi \left(\alpha\left(\beta + \sum_l^L \beta_l x_{l,m}\right)\right) + \sum_{\boldsymbol{m}}^{\boldsymbol{M}} \left(1 - P(\hat{y}_m|y_m = c_1)\right) \log \left[1 - \varphi \left(\alpha\left(\beta + \sum_l^L \beta_l x_{l,m}\right)\right)\right] \quad (13)$$

2.6.4. Generating Synthetic Datasets

The calibration of aEARM to IC-RP and EC-RP fluorescence time-course data provided an optimally fit vector of rate coefficient parameters, which served as the "ground truth" parameter vector in the synthesis of the nonquantitative datasets (Table A.5.). These parameters were applied to aEARM, and the resulting aEARM was used simulate time-courses for variables to be indicated in the nonquantitative data: truncated BID (tBID), initiator caspase localization to the death inducing signaling complex (IC-DISC), and cleaved PARP (cPARP).

These time-courses were converted to ordinal time-course datasets. The effective bit resolution of a measurement technology dictates how many unique values it can distinguish[156]. The total number of ordinal categories, K, was set such that resulting dataset had less than 70% of the effective bit resolution, EBR, (Eq. 14) of the IC-RP of EC-RP data. The signal to noise ratio, SNR, (Eq. 15) assumes the data, d, were subject to Gaussian noise and a 0.10 misclassification rate between adjacent values; modeled as the 0.95 quantile of a unit normal distribution[40]. Therefore, the number of ordinal categories were 5 and 4 for tBID and cPARP, respectively. The number of ordinal categories for IC-DISC were arbitrarily set to 4. Arbitrary values of slope and offset coefficients (Table A.3.) were designated "ground truth" and applied to ordinal measurement models (described above). The resulting measurement models map the values in the aEARM simulated time-courses to probabilities of each ordinal category. These probabilities were used to simulate random class assignments for synthetic ordinal datasets (see Figure 2.2.). The aEARM

was trained to time-course ordinal values of tBID and cPARP or time-course ordinal values of IC-DISC and nominal data described below.

$$K \leq 0.7 \times 2^{EBR}, \qquad EBR = -(SNR + 1.76)/6.02 \tag{14}$$

$$SNR = 20 \log_{10} q_{0.95} \mathrm{rms}(d)/(\max d - \min d) \tag{15}$$

To generate synthetic nominal (binary cell survival vs death) data, two heterogeneous populations of 200 aEARM tBID (and an unrelated non-apoptotic species, USM2) trajectories were simulated from ground truth parameters. The populations had distinct initial ligand concentrations (10 or 50 ng/mL). Heterogeneity was modeled by a log-normal random sample of certain initial conditions (described above). These time-courses were preprocessed to yield values of the features encoded in nominal measurement model, above. This measurement model (which was encoded with preset "ground truth" values of slope, intercept and weight coefficients – See Table A.4.) maps these features to probabilities of the binary outcomes. These probabilities were used to simulate random class assignments for synthetic nominal datasets (Figure 2.3. B).

To generate a synthetic distribution of times at which Bid truncation was half-maximal, two heterogenous populations of 200 aEARM tBID time-courses, corresponding to 10 and 50ng/mL initial ligand concentrations, were simulated from ground truth parameters (as above). Time at half-maximal tBID was calculated via linear interpolation and rounded to the nearest 3-minute time-point (i.e. to reflect temporal resolution of common time-series intracellular experiments) (Figure 2.3. A.).

2.6.5. Model Calibration via Bayesian Inference

The aEARM was calibrated using DREAM(ZS) algorithm for all datasets[61]. Rate parameters in aEARM were given independent log-normal distribution prior probability functions with a location equal to the ground-truth parameter vector and a scale term of 1.5. The nominal (cell death vs survival) dataset

features a heterogeneous population of values. We modeled this heterogeneity with a random sample of initial conditions (described above). This random sample was shifted and scaled according to inferred values of the model mean and variance. The mean (if estimated) was given a log-normal distribution prior probability function with a location equal to ground-truth and a scale term of 1.5. The extrinsic noise (or variance) was given inverse gamma distribution with a and b terms such that the resulting coefficient of variation had a prior mean and standard deviation of 0.20 and 0.015 respectively.

Prior probability functions were also applied to the measurement models' free-parameters. To in encode empirical ordering constraints on the ordinal measurement model, the slope terms, α, were greater than zero; they were given independent exponential distribution prior probability functions (with location of 0.0 and scale of 100.0). To insure monotonically increasing offset terms, each offset, β_j, was defined by the distance, θ_j, from its preceding offset term; $\beta_j = \beta_{j-1} + \theta_j$. The first offset, $\beta_0 = \theta_0$, and subsequent distance, θ_j, terms were given independent exponential distribution prior probability functions (with location of 0.0 and scale of 0.25). We explored the effect of increasingly biased priors on the ordinal measurement model parameters. Where noted, the slope terms, α, were given increasingly constrained independent prior probability functions: uniform (0.0 - 100.0 bounds), Cauchy (50.0 location and 10.0 scale) and Cauchy (50.0 location and 1.0 scale). The offset, β_0, and distance, θ_j, terms were similarly given independent uniform (0.0 - 1.0 bounds), Cauchy (0.2 location and 0.05 scale) and Cauchy (0.2 location and 0.005 scale) distribution prior probability functions. Parameters for the measurement model were given independent Laplace distribution prior probability functions with a location of 0.0 and scale of 1.0 for the slope, α, and 0.10 for the intercept and weighting coefficients, $\beta$and β_l.

The likelihood functions were described above. Additional settings applied to the DREAM(ZS) algorithm were: number of chains (4) number of crossover points (nCR=25), adaptive gamma (TRUE), probability of gamma=1 (p_gamma_unity=0.10), gamma term resolution (gamma_levels=8). A burn-in period wherein crossover weights are adapted was set to 50,000-step burn-in for ordinal datasets and 100,000+ step burn-in. The calibration algorithm continued until it reached the stopping criterion: when

the Gelman-Rubin metric (calculated on the latter 50% of the traces) was less than or equal to 1.2 for all free-parameters in the model; at which point the parameter traces were considered converged[61]. Gelman-Rubin metrics for each calibration are listed in Table A.6. Model calibrations were run on a x64 Intel with 32 total CPU threads (256GB RAM) and x64 AMD with 256 threads (1024GB RAM). Run times varied widely given the stochastic nature of the optimization algorithm but were typically one to seven days for simple model calibrations. Random samples of 1000 parameter were taken from the latter 50% of the resulting parameter-traces were used in subsequent analyses. Source code for the model calibrations as well as code for downloading the resulting parameter-traces is found at https://github.com/LoLab-VU/Opt2Q.

2.6.6. Model Predictions

We simulated the equal-tailed 95% credible region of the posterior predictions of aEARM via samples of the model parameters posterior distribution. This was done by randomly generating 1000 parameter sets sub-sampled from the posterior sample of parameters generated via PyDREAM. For each parameter set, tBID time-courses (and/or cPARP, IC-DISC) were simulated from aEARM. The 95% credible region of the predictions was then determined via 0.025 and 0.975 quantile bounds on the tBID (or other variables) values for each time-point in the simulated time-course. The area bounded in the 95% posterior credible interval was determined by summing the difference between the 0.025 and 0.975 quantile bounds across 100 equally spaced time points on the trajectory. The 95% posterior credible intervals on the measurement model predictions were similarly the described by calculating 0.025 and 0.975 quantile boundaries on the predictions of the measurement model parameterized via 1000 parameter set samples from a posterior. This includes the posterior probability distributions of the feature coefficients encoded in the nominal measurement model. To model predictions of the nominal dataset, however, we randomly generated 100 parameter sets via sub-sampling of the posterior parameter distribution. For each parameter set, we simulate tBID dynamics from the set of 400 initial conditions as described above; from that we compute maximum BID truncation rate and time at maximum BID truncation rate for each of the 400 trajectories. The 0.05 contour of the KDE of the resulting 400 values of

maximum BID truncation rate and time at maximum BID truncation rate was plotted for each of the 100 parameter sets.

Chapter 3

Challenges and Future Directions: Considering Complexity, Multiscale Organization, and Heterogeneity in System Biology Measurement Model Applications

3.1. Introduction

Despite ongoing experiments, emerging measurement technologies, and increasing computational abilities, progress in systems biology remains encumbered by an unmet demand for quantitative data[157]. Specifically, there exists the need for quantitative time-course measurements of intracellular dynamics (biomacromolecules, complexes, post translational modifications, etc.) needed to calibrate mechanistic models of cellular processes. This work reveals and addresses some challenges posed by the dearth of quantitative data. The previous chapter introduced an analytical tool that leverages nonquantitative data (as a more available substitute) to calibrate mechanistic models of biological processes. As a data-driven approach this innovation lets us calibrate mechanistic models while avoiding unchecked *ad hoc* assumptions about the supporting measurements. As a probabilistic approach it lets us define a representation of the relationship between a model and its supporting measurements that includes information about the uncertainty cellular measurements or observations. The measurement model opens new possibilities in that it will make available to model calibration a wide range of measurements and observations previously neglected by systems biologists. Extending the measurement model concept to these measurements and observations will still face undiscussed challenges.

The dearth of quantitative measurements of intracellular biology is a challenge that stems from and reflects more fundamental challenges in biology: complexity, multiscale organization, and heterogeneity[8]. *Complexity* gives rise to unanticipated cellular behaviors that emerge from a network of interactions and cannot easily be attributed to the action of a single factor. Systems biology began with the goal of understanding complexity in biology. The *Multiscale* organization of biological processes requires scientist consider of influence of events that occur in between individual proteins and last for few

seconds along with events that involve whole organs and last days or weeks in their investigation of biological mechanism[78]. The multiscale organization of biology is an evolved consequence (or example) of biological complexity that can drastically change the implications of experiments that focus too narrowly on a temporal and/or spatial scale. *Heterogeneity* adds an element of variability and stochasticity to biological processes and properties. Heterogeneity of a system (e.g., a cell population) can drive system-level behaviors as well as conceal distinct phenomena among the system components[84]. Future efforts to expand the use of measurement models to common measurements of intracellular biology must confront these fundamental challenges.

In this chapter, I take preliminary steps to calibrate model of cell death signaling to indirect and/or nonquantitative measurements of intracellular components (e.g., nominal cell-fate decisions, immunoblot measurements of cellular protein content, cell viability and cell viability rate). I detail the challenges of modeling these measurements and propose potential avenues for addressing them. I begin with the challenge of biological complexity. Systems biology aims to understand biological mechanisms as situated in a broader context of complexity. The biology of cell death exemplifies the challenge of complexity as multiple programmed cell death modalities engage a shared signaling molecules in a web of biomolecular regulatory *crosstalk*[158]. The interdependency between distinct cell death modalities engenders unexpected behaviors and complicates investigations of cell death. I demonstrate the potential of systems biology (despite the limitations of data) to investigate complexity through a model of apoptotic-necroptotic signaling crosstalk. I implement the three main steps of mechanistic modeling – model assembly, model calibration and model selection – to produce and analyze models of apoptosis and necroptosis signaling. The resulting modeled provided new insights into the mechanism by which apoptotic drivers regulate necroptosis initiation. This model demonstrates potential of system biology to tackle challenges of complexity, and warrants continued pursuit of solutions to the data-problem in systems biology.

The next sections focus on generalizing the measurement concept to common measurements of biology. I address challenges encountered in my work toward extending the measurement model to cell-

fate decision. This endeavor prompted me to consider how definitions of phenotype relate to underlying intracellular drivers and/or markers; much like how behaviors at one temporal spatial scale relate to properties at a smaller temporal spatial scale. This fundamental property of biology, its multiscale organization, reveals itself in our treatment of phenotype. Empirical descriptions of a phenotype (along with its broader biological impact) use a few distinguishing markers taken from a much larger set of observed mechanistic characteristics[44]. For instance, necroptosis follows nonlinear changes in hundreds of signaling proteins, but was defined (in recent work) as a necrotic morphology accompanied by MLKL phosphorylation and release of DAMPs[14]. This simplified definition let us translate cellular necroptosis to multiscale understanding of effects of necroptosis in tissue[14]. The empirical description of phenotype precedes the model of phenotype, and often the model of phenotype requires additional specification. A measurement model abstracts the complex dynamics of intracellular cell-fate signaling mechanisms to a cell-fate phenotype that is defined by a few markers, then uses this condensed model in a broader computational context. I explore potential avenues for defining models of necroptotic phenotype. The approaches I consider apply dimensionality reduction and feature methods in ways that mirror the models (in multiscale modeling approaches) of behaviors that emerge from one temporal-spatial scale and act on another temporal spatial scale[159].

The final section discusses the challenge of heterogeneity in models of biological measurements. Biological processes are inherently variable and stochastic. This makes it difficult to summarize biological processes using a single representative model of dynamics. Instead, experiments must conduct more replicate measurements and models must simulate the repeated draws from an underlying probability distribution. I discuss model calibration in the case where both the model and supporting experimental data capture heterogeneity. I then explore ways in which bulk measurement of heterogeneous populations (e.g., immunoblot measurements) obscure and distort the dynamics of the individuals. I also discuss potential modeling approaches to address the interaction between measurement and heterogeneity by modeling the measurement *process* (and how it transforms underlying heterogeneity).

The preliminary work outlined in this chapter suggests that addressing the challenges of complexity, multiscale organization, and heterogeneity, in future work will require additional integration of machine-learning into mechanistic models of biology (and their measurements)[159, 160]. While machine learning measurement models will enable adaptation to an expanding set of modeling scenarios, the necessity of data and data-driven strategies remain. Therefore, this chapter highlights challenges and possibilities of leveraging more types of measurements in order to meet (but not replace) the demand for data.

3.2. Systems Biology Addresses the Challenge of Complexity but is Limited by a Dearth Data

Biological complexity represents a significant barrier to understanding and solving problems in biology. This fundamental property of biology prevents us from translating the details of a biological process to a systems-level understanding. For instance, the last decade of cell death research has identified a myriad of new cell death modalities, each carried out by an ever-expanding parts list of biomolecular drivers[161]. Despite this wealth of detail, the association between cell death and human disease continues to baffle scientists. The current consensus view of cell death describes a morass of interdependent regulatory mechanisms, which clearly attests to the problem of biological complexity.

The most prominent form of cell death, apoptosis, is a regulated and immunologically silent cellular demise that features clear morphological characteristics: cytoplasmic and nuclear fragmentation, cellular blebbing and disintegration into small apoptotic bodies[158]. This process clears roughly 60 billion damaged, diseased, and unneeded cells from an adult human per day[**162, 163]. Proper execution of apoptosis is therefore a crucial requirement for tissue homeostasis, development, and response to pathogens. Abnormal engagement of apoptosis factors heavily in several cardiovascular, autoimmune, degenerative and neoplastic pathologies[164].Apoptosis is therefore tightly controlled by a network of regulatory interactions[164]; including interactions that elicit non-apoptotic cell death as a back-up[165]. For

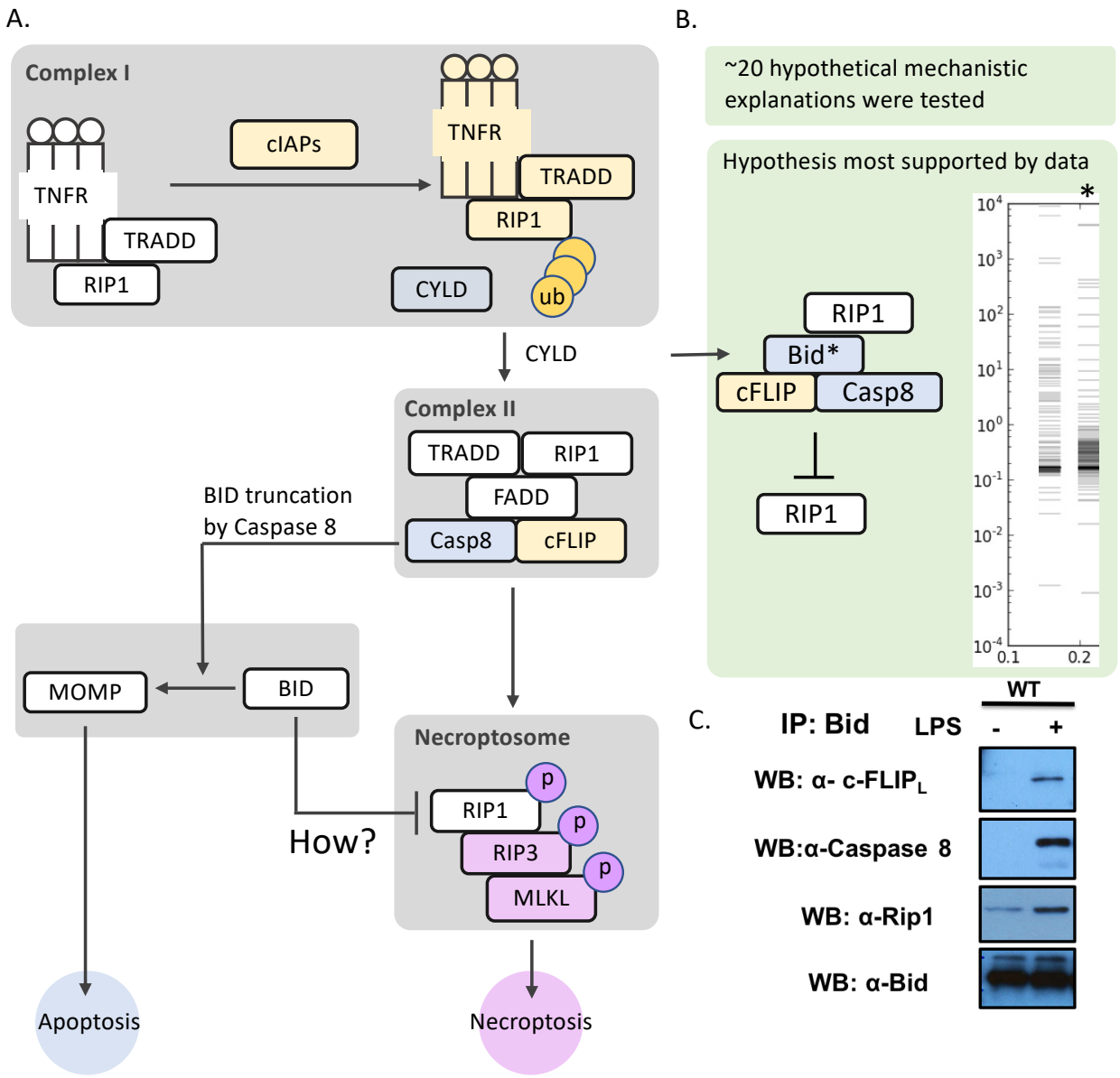


Figure 3.1. Bottom-up modeling approach investigates BID mediated apoptosis-necroptosis crosstalk The crosstalk between apoptosis and necroptosis, for example, features complexity of mechanism that makes it suitable for a bottom-up modeling approach. Specifically, to investigate the role of Bid in regulating necroptosis, we encoded a mechanistic model of TNF Complex I dynamics (TNF ligation, recruitment of scaffolding proteins, cIARP-mediated RIP1 poly-ubiquitination and CYLD mediate RIP1 de-ubiquitination); RIP1 release from Complex I and formation of cytoplasmic Complex II. This complex initiates MOMP formation and apoptosis by enabling Caspase-8 mediated BID truncation. With limited caspase activity formation of a RIP1-RIP3 dependent necroptosome and necroptosis can proceed. Recent observations, however, suggest a non-truncated BID can regulate necroptosome formation by inhibiting the accumulation of RIP1. We encoded the above mechanism as a system of ODEs and appended to the model one of several hypothetical interactions between BID and necroptosis signalers. We calibrated each model to time course data for cPARP and MLKL. By noting the calibrated models' predictions for *in silico* apoptotic and necroptotic conditions we narrowed the the set of hypotheses to those that contained a novel stable interaction between BID, cFLIP, and capsoase 8. The Hessian eigenspectral of a model with this complex featured more bands ≤ 1.0 than the model that lacked this complex (indicating more robust support for the hypothetical complex). This complex was later indicated by immunoprecipitation.

instance, necroptosis is a regulated necrotic cell death program that activates in cells facing (among other things) an interruption of normal apoptotic signaling[166]. The necrotic morphology, which appears in other programmed cell death modalities, features swelling and lysis of cellular organelles and release of intracellular contents to the extracellular space. This provokes immunogenic responses that both ameliorate and exacerbate disease. Further complicating the picture, many initiators and mediators of apoptosis also initiate and mediate necroptosis. The interactions shared by apoptosis and necroptosis permit molecular *crosstalk* between the two pathways, thereby making the problem of cell death (and its impact on health) a problem of biological complexity.

A central motivation for my dissertation research is the promise that systems biology (specifically bottom-up modeling) can provide a unique and valuable understanding of biological phenomenal that connects inextricably to complexity. I expect this understanding to help make sense of the complex network of biomolecular interactions that drive important processes like cell death. In this section, I explore this promise through an application of bottom-up modeling strategies to a question of apoptosis-necroptosis crosstalk.

Recent experiments by our collaborators, Sandra Zinkel and Patrice Wagner, suggest an unanticipated role of the pro-apoptotic BH3-only family member protein, Bid, in regulating necroptosis[14]. They constructed a mouse that blocked apoptosis in hematopoietic cells through genetic deletion of *Bak* and *Bax*. Further deletion of *Bid* (to create *BaxBakBid* triple knockout mice) resulted in increased levels of pro-necroptotic, RIP1 kinase, and thus robust activation of necroptosis. The immunogenic response to necroptosis further triggered a feedback amplification loop of the pro-necroptotic signaling; and unchecked necroptotic cell death drove myelodysplastic and bone marrow failure syndromes in the mice. Bid restricts necroptosis through suppression of Rip1 levels. Their work also suggests Bid enhances caspase 8 mediated RIP1 proteolytic degradation, while (at least partially) avoiding -- Bid is a caspase 8 substrate -- caspase 8 meditated processing. We investigated Bid mediated apoptosis-necroptosis crosstalk using a bottom-up modeling algorithm: model assembly encodes a mathematical representation of key reactions in apoptosis and necroptosis; model parameterization estimates values of the free parameters

that minimize the discrepancy between the model prediction and time-course data; model selection uses an approximation of Bayes factor to select between pairs of competing hypothetical models the model that is maximally supported by data. By integrating complex apoptotic and necroptotic signaling dynamics into as single mathematical framework, this modeling approach lets us explore, *in silico*, mechanisms by which pro-apoptotic signaling events might regulate (through Bid) necroptosis.

We modeled salient features of TNF (or LPS) induced cell death[167]. As summarized in Figure 3.1. A., TNF ligation triggers formation of a pleiotropic protein complex at the plasma membrane (Complex I). Complex I recruits pro-survival proteins (e.g. cIAP[168]), that polyubiquitinate RIP1. The complex also recruits CYLD[00], which reverses RIP1 polyubiquitination and drives the cell toward programmed cell death signaling. We encoded a distinction between RIP1 prior to ubiquitination and RIP1 after de-ubiquitination. De-ubiquitinated RIP1 initiates formation of a cytoplasmic protein complex[169] (Complex II), which recruits pro-apoptotic Caspase 8. We encoded complex II dependent dimerization of caspase 8. Caspase 8 homodimers[170] dissociate from complex II and catalyze proteolytic cleavage of RIP1, Bid, and several other substrates. Caspase/cFLIP[171] heterodimers remain in complex II and catalyze proteolytic cleavage of RIP1 and not Bid. Proteolytic cleavage of Bid by Caspase 8 activates apoptotic signaling. We modeled apoptotic reactions: Bid recruitment to the mitochondria, complexation of Bax and Bak, formation of a mitochondrial outer membrane pore (MOMP), and finally activation of apoptosis effector caspases as described in previous chapters. De-ubiquitinated RIP1 can also trigger necroptosis through recruitment of RIP3 into a complex (necrosome) that catalyzes the phosphorylative activation of pronecroptotic MLKL[172]. These interactions were encoded (using PySB rule-based modeling) as ODE representations of general mass-action descriptions of protein binding and catalysis reactions. The formation and conversion of large complexes (e.g., MOMP formation) were abbreviated to a few self-catalyzed reactions.

We used this model (which we called ANRM or apoptosis-necroptosis reaction model) to examine how hypothetical mechanisms of Bid mediated RIP1 regulation impact *in silico* predictions of apoptotic and necroptotic behaviors. We regarded effector caspase cleavage product, cleaved PARP

(cPARP), and phospho-MLKL (pMLKL) as apoptotic and necroptotic markers, respectively. ANRM predicted snap-action dynamics for cPARP and pMLKL; model simulation where cPARP reached half-maximal concentration before pMLKL did, were considered apoptotic (and vice versa). With an absence of quantitative cPARP and pMLKL time-course measurements, we generated sigmoidal temporal profiles for cPARP and MLKL that reflected comparable immunoblot data. To calibrate the model, we used a MCMC algorithm[38] as described by Eydgahi et al. along with a sum of squared error model of the log-likelihood function and log-normal priors around generic (physiological) values of reaction rate coefficient parameters. We appended to ANRM hypothetical mechanisms of Bid dependent suppression of RIP1; the model calibration process was repeated for each hypothetical model. We examined how each hypothetical model responded to *in silico* proapoptotic and pronecroptotic conditions (i.e., to varying initial levels of Caspase 8, RIP1, cFLIP, and Bid). This narrowed the set of hypothetical mechanisms to those that accurately predicted apoptosis or necroptosis in response to apoptotic or necroptotic conditions, respectively. We used ANRM to manually explore roughly 20 hypotheses, with input from Sandra Zinkel and Patrice Wagner. The remaining competing models were discriminated by comparing the range of eigenvalues of the Hessian matrix of posteriors produced by each model at their best-fit parameter locus (Figure 3.1. B.). The model with multiple smaller eigenvalues indicates it is consistent with the data over a larger volume of parameter space (thereby exhibiting greater statistical weight)[38, 173]. This analysis implicated as the most likely mechanism (Figure 3.1. B.): a ligand dependent post translational modification of Bid facilitates complexation of cFLIP and Caspase 8; this Bid-cFLIP-Caspase 8 complex then catalyzes the proteolytic degradation of RIP1. We confirmed the existence of this ligand dependent complex via immunoprecipitation with Bid (Figure 3.1. C.).

This exercise demonstrates the power of bottom-up modeling to identify, in a complex biological system, new insight into the mechanisms that engender certain nuanced biological phenomena. However, the full benefit of this analysis is lost in how we adapted to the lack quantitative time-course data. We synthesized quantitative time-course interpretations of available nonquantitative data: ordinal immunoblot measurements and nominal cell-fate (apoptosis vs necroptosis) observations. The *ad hoc* nature of these

interpretations makes it unclear to what extent our results reflect trends in the data and not artifacts of our interpretations of the measured data. Nonetheless, the successful identification of a complex that facilitates Bid dependent suppression of necroptosis demonstrates the potential of bottom-up modeling approaches to help understand biological complexity. It also warrants future work to address the challenges posed by the dearth of available measurements needed for bottom-up modeling in systems biology. The previous chapter introduced a data-driven measurement model that enables incorporation of nominal and ordinal measurements into the model calibration process. In the next sections, I consider extending the measurement model concept to common measurements of intracellular biology; I discuss potential challenges and propose possible avenues for addressing these challenges.

3.3. Generalizing the Measurement Model: Addressing a Challenge of Multiscale Biology

Biological systems and processes occupy several spatial and temporal scales. What transpires at one spatial and temporal scale has drastic impacts on other scales[78]. This fundamental property of biological systems -- its multiscale organization -- can drive unpredictable behaviors and therefore poses a significant challenge to the study of biological systems. For instance, the multiscale biology of hematopoiesis dictates that genetic knockdown of cell death effectors (Bid, Bak, and Bax) produces an unexpected *excess* cell death that drives myelodyspastic syndrome (MDS) and bone marrow failure. Hematopoeitic cells in *BaxBakBid* triple knockout mice undergo necroptosis (which emerges from short-lived intracellular interactions that depend on extracellular pro-necrotic cues)[14]. Necroptosis also triggers longer-lived tissue-level interactions that in turn affect the generation of extracellular pro-necroptotic cues. This model of MDS shows an interdependence across small (intracellular) and large (extracellular) scales with cell death at its nexus. Strategies that enable investigation and modeling of multiscale biological systems have therefore garnered considerable attention.

Models that capture multiscale properties of biology must integrate information across multiple spatial and temporal domains. It is infeasible to model the dozens or hundreds of intracellular reactions that dictate cellular behavior of each of the several thousand individual cells in a tissue. Instead, multiscale modeling abstracts the complex dynamics of the intracellular scale to a simplified model that

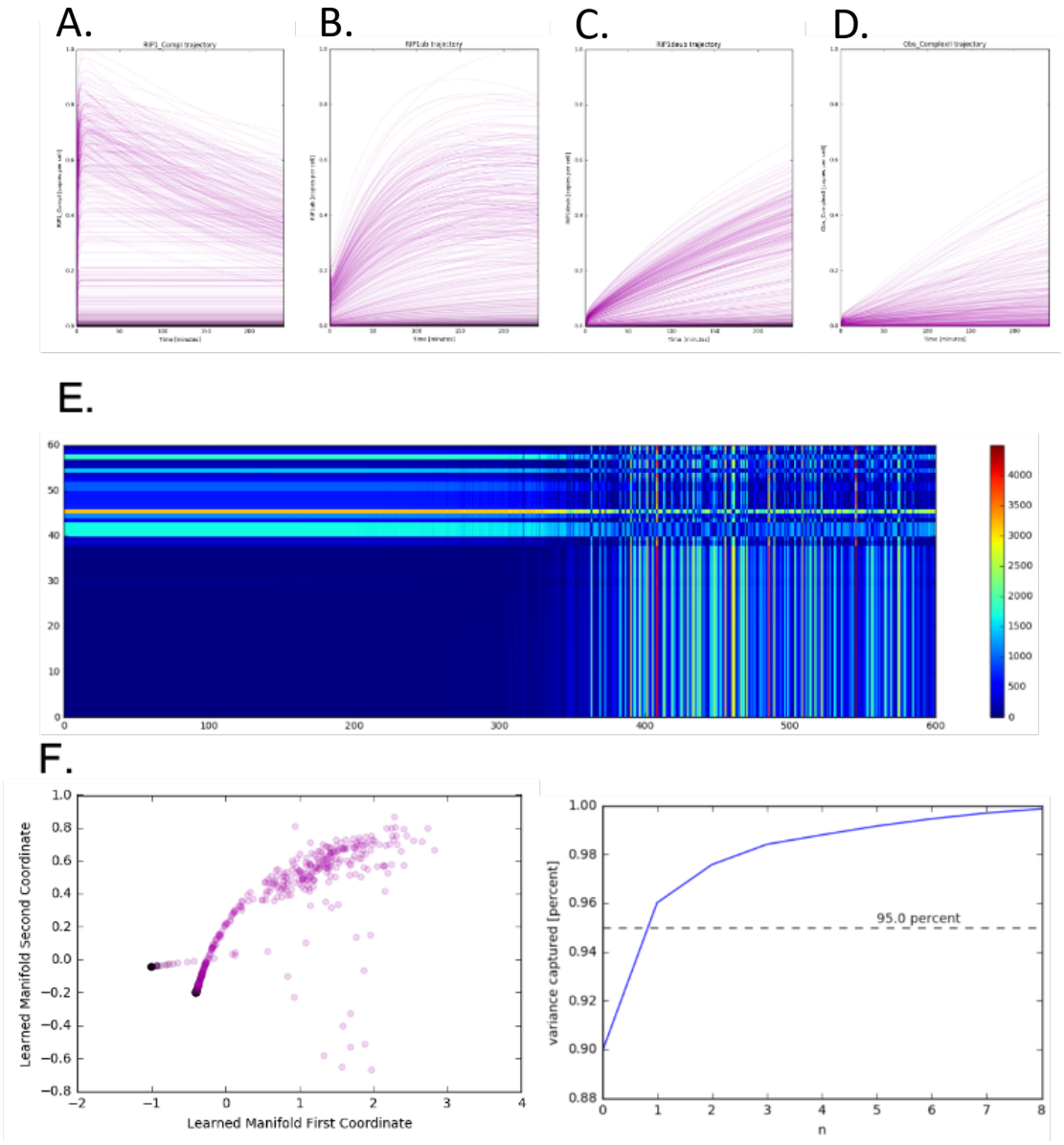


Figure 3.2. Modeling phenotype as an emergent property of a complex system through dimensionality reduction (feature extraction). The multi-scale organization of biological systems prompts us to think of phenotypes at one temporal-spatial scale as emerging from the dynamics at a smaller scale. For instance, we modeled a cellular phenotype (death vs. survival) as an emergent property of RIP1 dynamics: A.) the dynamics of RIP1 localization to the plasma membrane bound complex (Complex 1) following TNFa ligation, B.) RIP1 ubiquitination and C.) subsequent deubiquitination, and D.) RIP1 mediated formation of cytoplasmic complex (Complex 2) were simulated using an ODE model of TNF complex 1 dynamics for an *in silico* heterogeneous set (n=600) of initial conditions. E. I used dynamics time warping (DTW) to compress these time-courses to a scalar distance metric relative to 60 landmark trajectories. This use of DTW reduced the dimensionality of the model simulations to 60. The 60 DTW features can be applied directly to a probabilistic classifier that models the death vs. survival outcome. Alternatively, we further reduced the dimensionality of the modeled dynamics by via isomap multi-dimensional scaling.

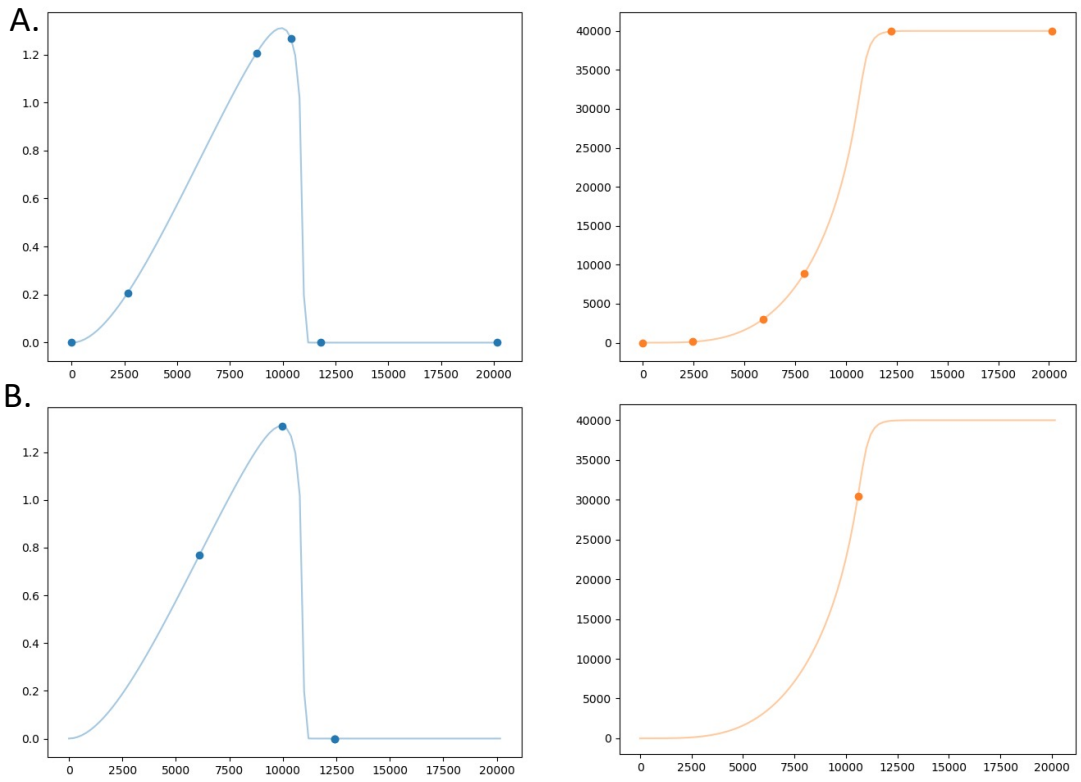


Figure 3.3. Modeling phenotype as an emergent property of a complex system through dimensionality reduction via perceptually important points or critical points. As an alternative to DTW, we compressed high-dimensional dynamics of by identifying important points in the models simulated time-courses. Dynamics of caspase 8 in the TRAIL dependent death inducing signaling complex (DISC) (left column) and dynamics of BID truncation (right column) encoded by aEARM. A. Six perceptually important points (PIP) for each trajectory are show. B. Critical points (i.e., relative max and min and points of inflection) were estimated for each trajectory. perceptually important point.

conserves enough information to enable its inclusion in a tissue-level model. The goal is a low-dimensional representation of the high-dimensional intracellular dynamics. Empirical markers of cell-fate (e.g., necroptosis markers) similarly condense high-dimensional intracellular states to a few variables (e.g., presence of pMLKL and sensitivity to RIP1 or RIP3 kinase inhibitors). The previous chapter introduced a measurement model that defined nominal observations (apoptotic cell death vs. survival outcomes) in terms of features of Bid truncation dynamics. This definition was based on experiments that identified these features as predictors of apoptosis and survival. An empirical description of cell-fate in terms of temporal dynamics of intracellular signaling is rare. To expand the measurement model concept from Opt2Q to a wider range of cell-fate observations, researchers would therefore need to consider dimensionality reduction strategies to translate high-dimensional intracellular dynamics to a manageable set of features that discriminate distinct cell fates[44, 174].

The measurement models introduced in the previous chapter (in Opt2Q) intentionally avoided dimensionality reduction in order to focus solely on idea of a measurement model. A future improvement to the measurement model would incorporate dimensionality reduction strategies. This would enable mechanistic model calibration using observations that lack a definitive link to the variables encoded in the mechanistic model. In Figure 3.2., I applied a feature extraction method simulated intracellular RIP1 dynamics. I used an ODE model of Complex I and Complex II dynamics (as schematized in Figure 3.1.) to simulate time-courses of RIP1 recruitment to Complex I (A.), polyubiquitination (B.), de-ubiquitination (C.), and recruitment to Complex II (D.). Dynamic time warping[175] (DTW) non-linearly shifts and stretches the time axis of a pair of temporal sequences in order to maximize their similarity. The distance between these temporally warped sequences enables classification of the time-courses. I used DTW to determine the distance between 600 time-courses and 60 landmark time-courses. This reduced the dimensionality of the simulated temporal profiles to 60. A measurement model could apply a probabilistic classifier to these 60 features[176] (e.g., to classify them into necroptotic vs survival outcomes). Alternatively, to further reduce the dimensionality, I applied Isomap multidimensional scaling[177] which finds a low-dimensional embedding of the 60-feature space that conserves geodesic distances between

pairs of time-courses. The first two Isomap dimensions captured over 95% of the variability of the simulated dynamics. This dimensionality reduction strategy requires a large dataset (containing hundreds of measurements) in order to provide reliable classification of the cell-fates. Further, the features extracted do not have an easy interpretation in terms of variables and temporal features of the simulated dynamics. This can complicate out-of-sample predictions of cell-fate using a measurement model that incorporates this approach. The challenges of this approach are like those that would arise using PCA or similar feature extraction methods.

The measurement model that relates a mechanistic model of intracellular dynamics (i.e., ODE simulated time-courses of various modeled species) to an observed phenotype (e.g., cell-death outcome) can exploit the relative simplicity of the curves produced by ODE models of biology. We can effectively reduce the dimensionality of such time-courses by identifying perceptually important points[178] (PIP) or critical points (e.g., relative maximum and minimum, and points of inflection). The perceptually important points algorithm identifies the subset of points along a curve necessary to recover the general shape of the curve. PIP can be applied to a wide range of curve shapes (or time-courses). Figure 3.3. A. shows the points identified by the PIP algorithm on simulated time-courses of Caspase 8 localization to the DISC and Bid truncation dynamics – as modeled by aEARM (see Chapter 2). Alternatively, in Figure 3.3. B, critical points are identified for the same time-courses. This dimensionality reduction strategy could be integrated into a measurement model of cell-fate without drastically increasing the demand for data. It is, however, limited to continuous models (e.g., ODE simulations) of intracellular dynamics. Application of this approach to stochastic models would require preprocessing (e.g., moving average calculations) to smoothen the predicted time-courses.

The measurement model, we introduced in the previous chapter, enables mechanistic model calibration using nominal cell fate outcomes. The data-driven property of this Opt2Q measurement model opens opportunities to glean insights into how intracellular mechanism connects (as predictors or drivers) to a cellular phenotype. There exist uncertainties in the definition or classification of phenotypes. For instance, our collaborators used pMLKL as a sole marker of necroptosis but, pMLKL can also indicate

pyroptosis, as it accompanies both cellular phenotypes[179]. A data-driven measurement model can also help resolve these uncertainties by identifying intracellular predictors and drivers of a phenotype. Roux et al. used fluorescence measurements of Bid truncation dynamics to identify features of Bid truncation dynamics that predict apoptosis[125]. They constructed a phenotype *phase map* of the apoptosis and survival that allowed them to reconstruct apoptotic responses to experimental and therapeutic interventions. The Opt2Q measurement model leveraged knowledge of underlying apoptotic mechanisms along with a dataset of apoptosis vs survival outcomes to also generate this phase map. By integrating dimensionality reduction methods into the Opt2Q measurement model we can narrow a large set of features to a few predictors of cell fate. This method moves the characterization of cellular phenotypes away from the *ad hoc* hypothesis driven approach and toward a data-driven approach (which, aligns well with current direction of systems biology[180]).

3.4. Generalizing the Measurement Model: Addressing a Challenge of Heterogeneity

Biological systems feature random variations in the abundance and activity of its components. These variations have two sources: intrinsic and extrinsic. Intrinsic variations, or intrinsic *noise*, arise from stochastic diffusion, binding and chemical reaction events that happen between the proteins within a cell. Intrinsic noise most notably impacts the assembly of protein complexes since these complexes have a low copy number within the cell. Extrinsic noise is the heterogeneity, between cells, in protein content. This heterogeneity results from random fluctuations in protein expression, etc. Organisms have evolved a strong tolerance for and dependence on noise[84-86]. The signaling mechanism that commits cells to apoptosis may tolerate extrinsic and intrinsic noise while the timing of apoptosis may be sensitive to noise. For instance, Albeck et al. noticed execution of apopotosis in cells subjected to low concentrations of TRAIL occurred over a larger span of time than cells subjected to higher concentrations[145]. In other instances, heterogeneity propagates into the cell death vs survival decision.

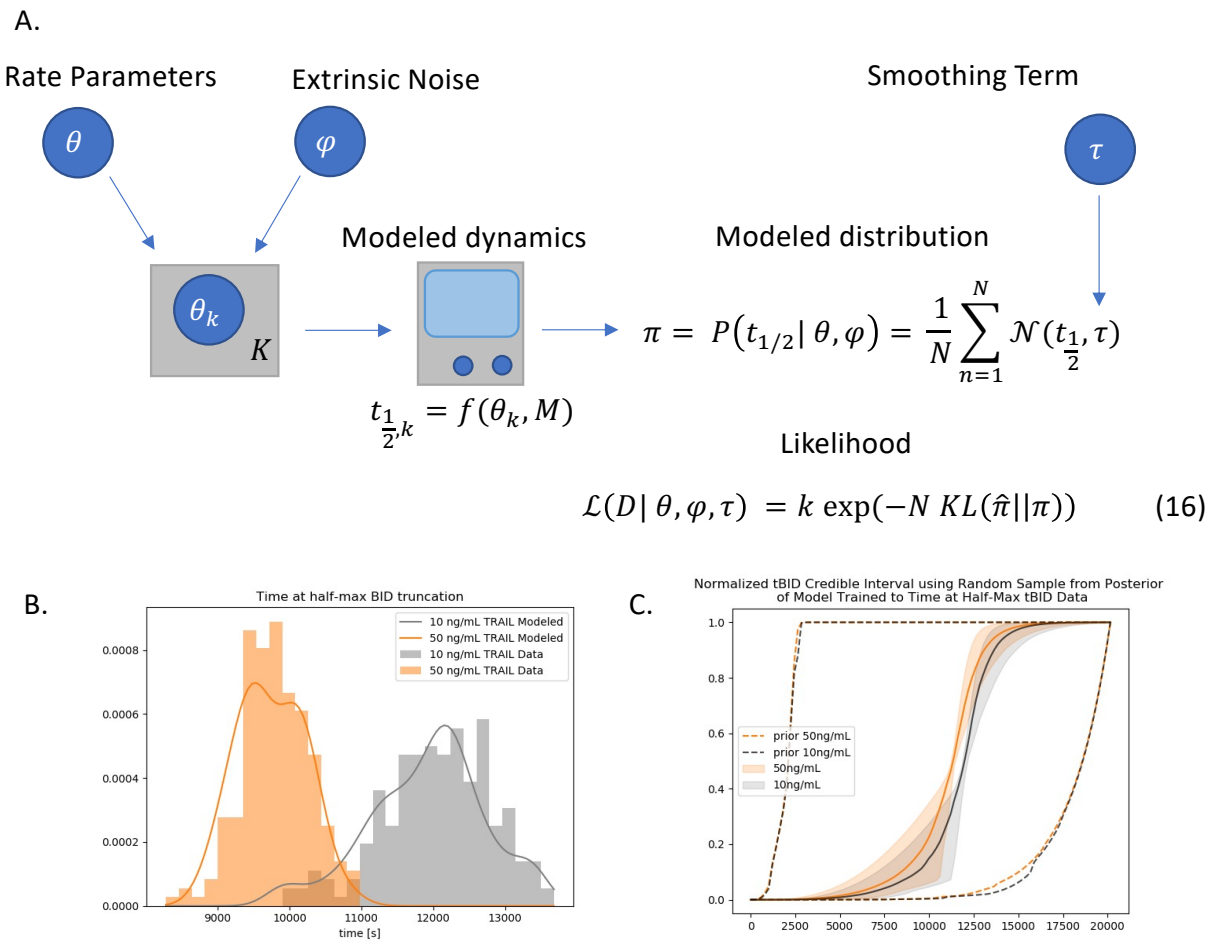


Figure 3.4. Modeling calibration and extrinsic heterogeneity. To address the affect of extrinsic heterogeneity on model calibration, we modeled dynamics of tBID in a heterogeneous population of cells; specifically, the time at half-maximal tBID concentration (B.) solid histogram). These data were expressed as a histogram, $\hat{\pi} = p(\hat{y}|\phi) = \prod_{n=1}^{N} p(\hat{y}_n|\phi)$, (B.) shaded area). The schematic A.) describes how the apoptotic cell death model (aEARM) was simulated for a set (size = K) of heterogeneous initial conditions. The distribution of the resulting *in-silico* values of time at half maximal tBID were modeled using a Gaussian mixture model (B.) solid lines). The empirical and modeled distributions were compared using a likelihood function that is proportional to their Kullback-Leibler divergence (Eq. 16)[181]. The rate parameters, extrinsic noise and smoothing terms (θ, φ, and τ, respectively) were additional free parameters. The posterior distributions (C.) of the prediction of the mean tBID dynamics are shown.

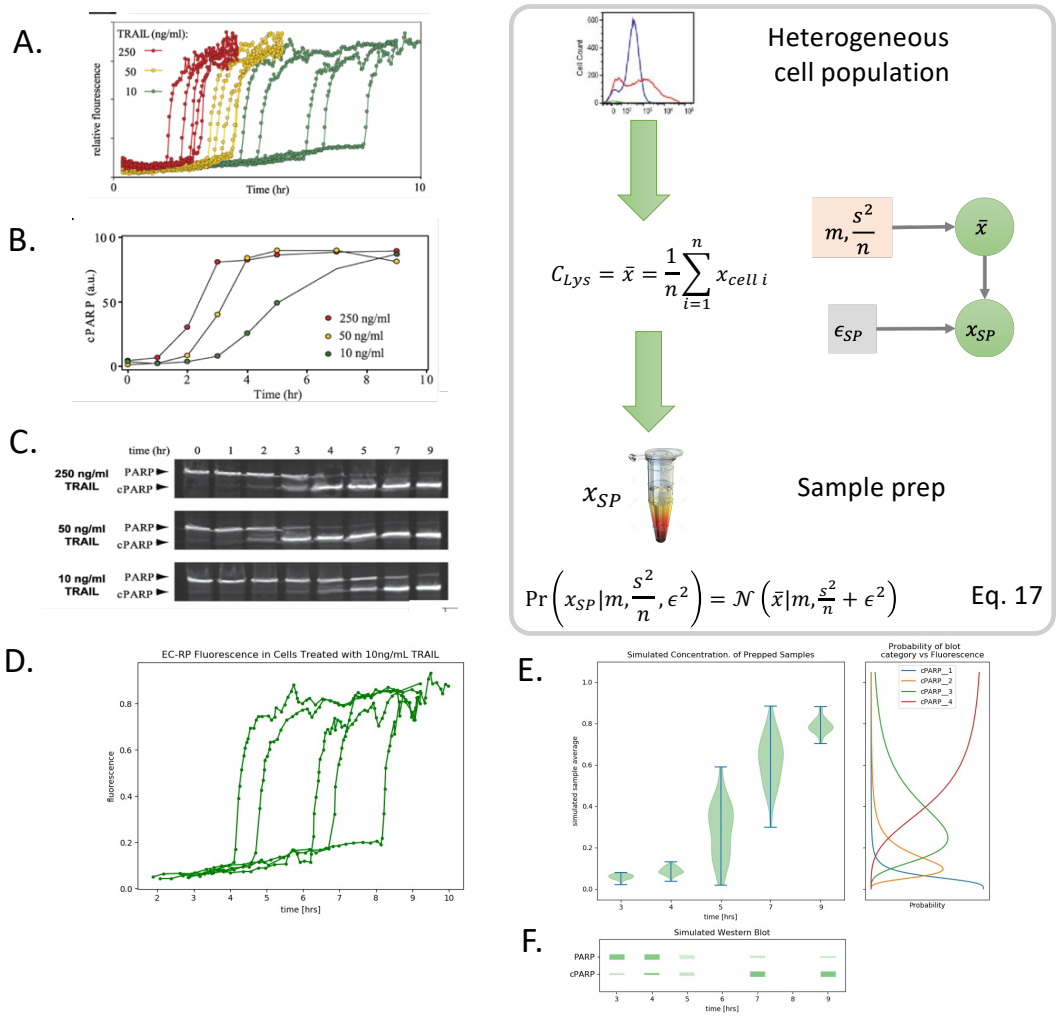


Figure 3.5. Modeling extrinsic heterogeneity in sample prep in ordinal measurements. Model calibration methods that ignore heterogeneity may miss the influence of heterogeneity on the cellular behaviors. Similarly, experimental methods that obscure heterogeneity also obscure details of the cellular behaviors. A. Cellular fluorescent intensity of effector caspase activity indicator (EC-RP) – a proxy for PARP cleavage – reveals heterogeneity in snap-switch dynamics of PARP cleavage. The snap-switch dynamics is obscured when, in B., the fluorescent measurement is averaged over the heterogeneous population of cells. Immunoblot measurements (C.) (wherein sample prep) effectively averages intracellular contents of a cellular population) mirror the averaged fluorescence measurements in B. We model the collapse of heterogeneous dynamics via sample prep by defining (Eq 17) concentration of the prepped sample x_{SP} in terms of the average of cellular concentrations $\bar{x}$ (s^2 and n are the concentration variance and cellular population size, respectively) the variability contributed by technical error ϵ^2. We apply this model of sample prep to a measurement model of ordinal immunoblot (E.) measurements. The measurement model's ordinal classifier therefore determines the marginal (i.e., marginalized over the probability density function of sample prep concentrations) probability of each category. F. is an a *in silico* representation of the predicted immunoblot intensity values. The width and shading of each rectangle are proportional to their most probable ordinal value and probability of the that ordinal value (respectively). A., B., and C. are taken from Albeck et al.[145]

Figure 3.4. A. shows an example of extrinsic noise applied to the abridged model of apoptosis (aEARM see previous chapter). Log-normal distributed noise with a mean θ and coefficient of variation φ was applied to select parameters of aEARM. By evaluating aEARM for several simulations (sample size of $K = 50$) of the model parameters we generated an *in silico* sample of the apoptosis dynamics; specifically time at half-maximal Bid truncation ($t_{1/2}$). The density of *in silico* sample of $t_{1/2}$ was modeled using a mixture of Gaussians having a scale of τ . The model parameters -- aEARM rate parameters, θ, φ, and τ -- were calibrated using a likelihood function[181] that compares empirical and *in silico* distributions of $t_{1/2}$ using a function of the Kullback-Leibler divergence. Shown in Figure 3.4. B., the resulting 95% credible region for predictions of the mean of the Bid truncation dynamics were tightly constrained around ground truth dynamics of aEARM. Conversely, calibrating aEARM to the mean and standard deviation of the $t_{1/2}$ dataset provides a far less certain model, as indicated by the broader 95% credible region of the predicted Bid truncation dynamics.

Empirical measurements of biological systems are subject to biological and technical noise sources whether the resulting measured dataset reflect the noise or not. Therefore, models of biological measurements should consider noise. The measurement model introduced in the previous chapter used a logistic model of categorical values. This modeling framework assumes the categorical measurement is a random Bernoulli distributed variable. Alternatively, a probit model of the categorical outcome could be used to model the probability of a normally distributed latent variable crossing some threshold between distinct categorical values[182]. More work remains before we can to account for heterogeneity and its effect on common measurements of intracellular biology.

Several measurements of intracellular biology entail sample preparations that collapse heterogeneous cellular properties into a single value. For instance, the cell lysis step in immunoblot measurements can artificially flatten the dynamics in individual cells. Experiments, by Albeck et al., investing the relationship between sigmoidal PARP cleavage dynamics (Figure 3.5. A.) and cPARP immunoblot measurements (Figure 3.5. C.), found that snap switch dynamics are obscured in the

immunoblot similarly to how averaging cPARP values over a heterogeneous population obscures the dynamics[145]. The most heterogeneous population (shown in green in these figures) has the flattened representation of the dynamics (shown in Figures 3.5. B., and 3.5. C.). Cell lysis steps average the intracellular content in a population of cells. This average can be expressed as having a normal probability density with mean and variance described by Eq. 17. With this formulation, we reimagine the ordinal measurement model of immunoblot measurements. A heterogeneous population of measured cPARP values, Figure 3.5. D., (or *in silico* cPARP predictions) is used to model a density function for the population average cPARP value (or the amount of cPARP in the cell lysate). The measurement model's ordinal classifier then determines the marginal (i.e., marginalized over the probability density function of sample prep concentrations) probability of each category.

A large dataset with individual cell-fate outcomes, each corresponding to known experimental treatments, is rare. Roux et al. created a dataset with apoptosis and survival outcomes and corresponding temporal profiles of a fluorescence indicator of BID truncation[125]. Such datasets do not yet exist for non-apoptotic cell death modalities. Microfluidic devices enable precise manipulation of pL cell cultures (i.e., enabling experimentation on individual cells). For instance, I subjected individual L929 cells to pro-necroptotic conditions in a nanophysiometer[183] and monitored them as they underwent necrotic cell death (See Figure C.1.). The most common measurements linking cell death to intracellular mechanism are cell viability measurements. These measurements average individual cell death outcomes to provide a value for the *proportion* of cells undergoing cell death. A measurement model of cell viability would include a similar averaging step that marginalizes the probability of cell death (i.e., as predicted using the nominal measurement model in the previous chapter) over an *in silico* population of cells (Figure 3.6.). Alternatively, cell viability could be modeled as a non-linear function of intracellular dynamics (without inclusion of a classifier to distinguish individual cellular outcomes). This approach was not explored in our work. Cell viability studies provide too few measurements to constrain the predictions of a dynamical model of cell death. Our collaborators, Maria Metzig and Alexander Hoffman, captured necroptotic cell death in a population at every 90s interval for 24 hours (as shown in Figure 3.7. A.)[184]. They used this to

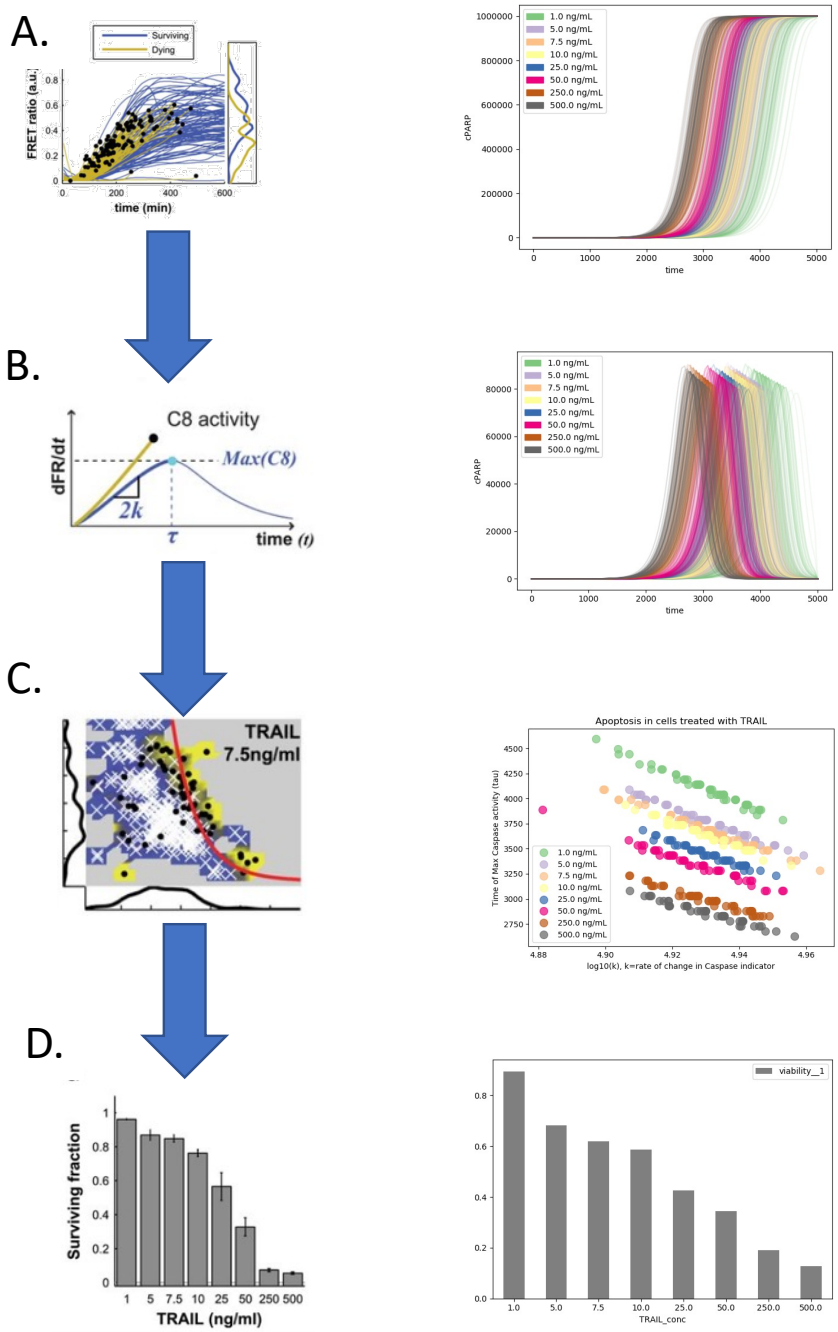


Figure 3.6. Modeling extrinsic heterogeneity in sample prep in nominal measurements. Models of nominal measurements, and observations, must also consider heterogeneity and the impact of sample prep on model calibration. Cell viability studies effectively average the individual cell death vs. survival fate outcomes for a heterogeneous population of cells. This averaging effect is appended to a model of cell fate to produce a model of cell viability. A. Initiator caspase activity (a proxy for Bid truncation) dynamics were simulated across and *in silico* population of cells (each having a distinct vector of initial conditions). B. The dynamics preprocessed to define features of apoptotic cell death. C. The resulting features, maximal initiator caspase activity (a proxy for Bid truncation) and the time when initiator caspase activity maximizes, were plotted. Empirical measurements of these features (left column) are taken from Roux et al[125]. D. By averaging probability of cell death vs survival for each in the population of cells we estimate cell viability.

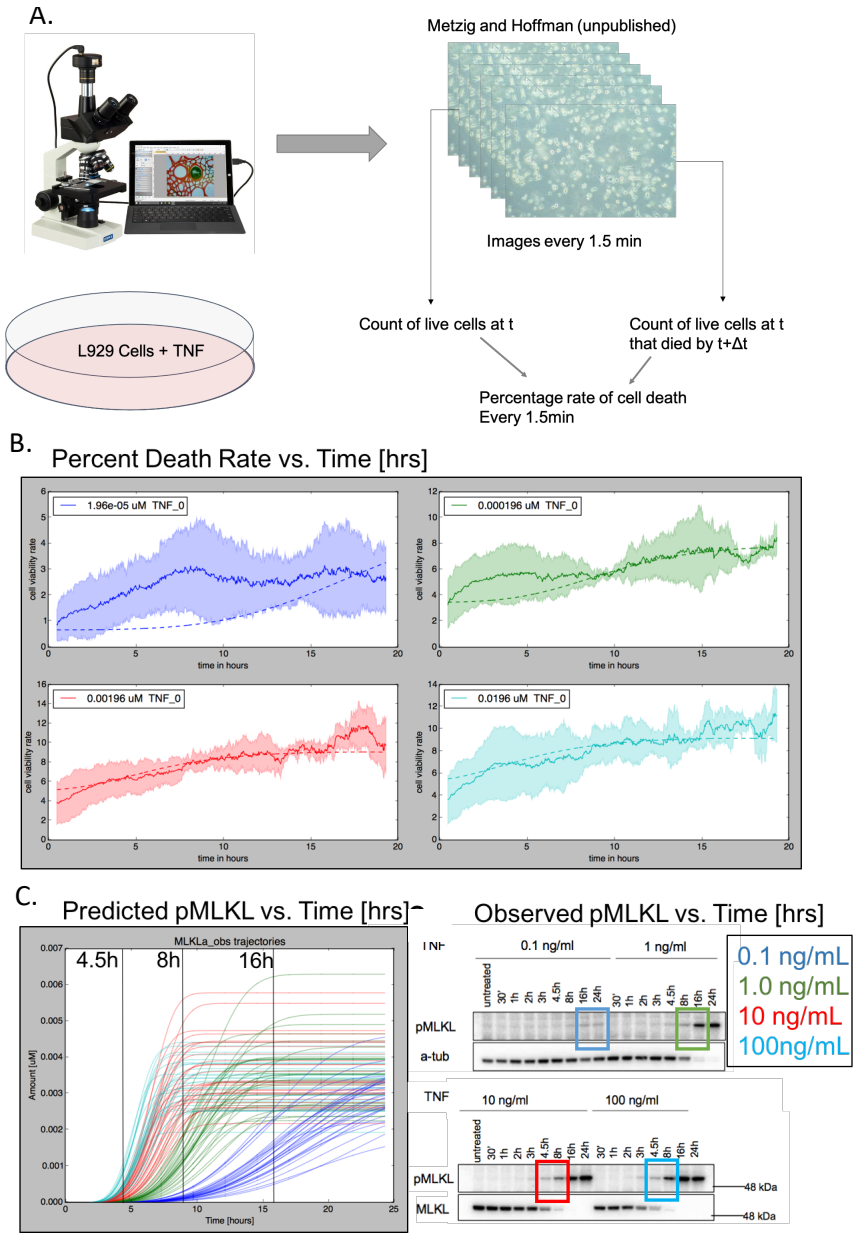


Figure 3.7. Modeling extrinsic heterogeneity in sample prep in nominal measurements. Recent work by Metzig et al.[184] provides vastly more measurements of cell viability than is typical of experiments measuring cell viability. A. They measured the rate of change in cell viability by monitoring the percentage of cells that died within a sliding 3hr window following treatment with TNF. The resulting dataset (B.) contained measurements of cell viability *rate* at every 1.5 min interval for 20hr of TNF treatment. We trained a model of necroptosis signaling to this data by applying a measurement model that modeled cell viability by averaging probability of cell death vs survival for each in an *in silico* population of cells; and cell viability *rate* by finding the difference in cell viability across a sliding 3hr window. The measurement model classified cell death vs survival using a logistic classifier on values of accumulated MLKL. C. The trained model predicted points of inflection in the amount of MLKL to occur in the 16-24hrs, 8-16hrs, 4.5-8hr and 4.5-8hrs intervals for (0.1, 1, 10 and 100ng/mL TNF, respectively). These predictions are supported by immunoblot measurements showing a transition toward full saturation occurring in the same intervals.

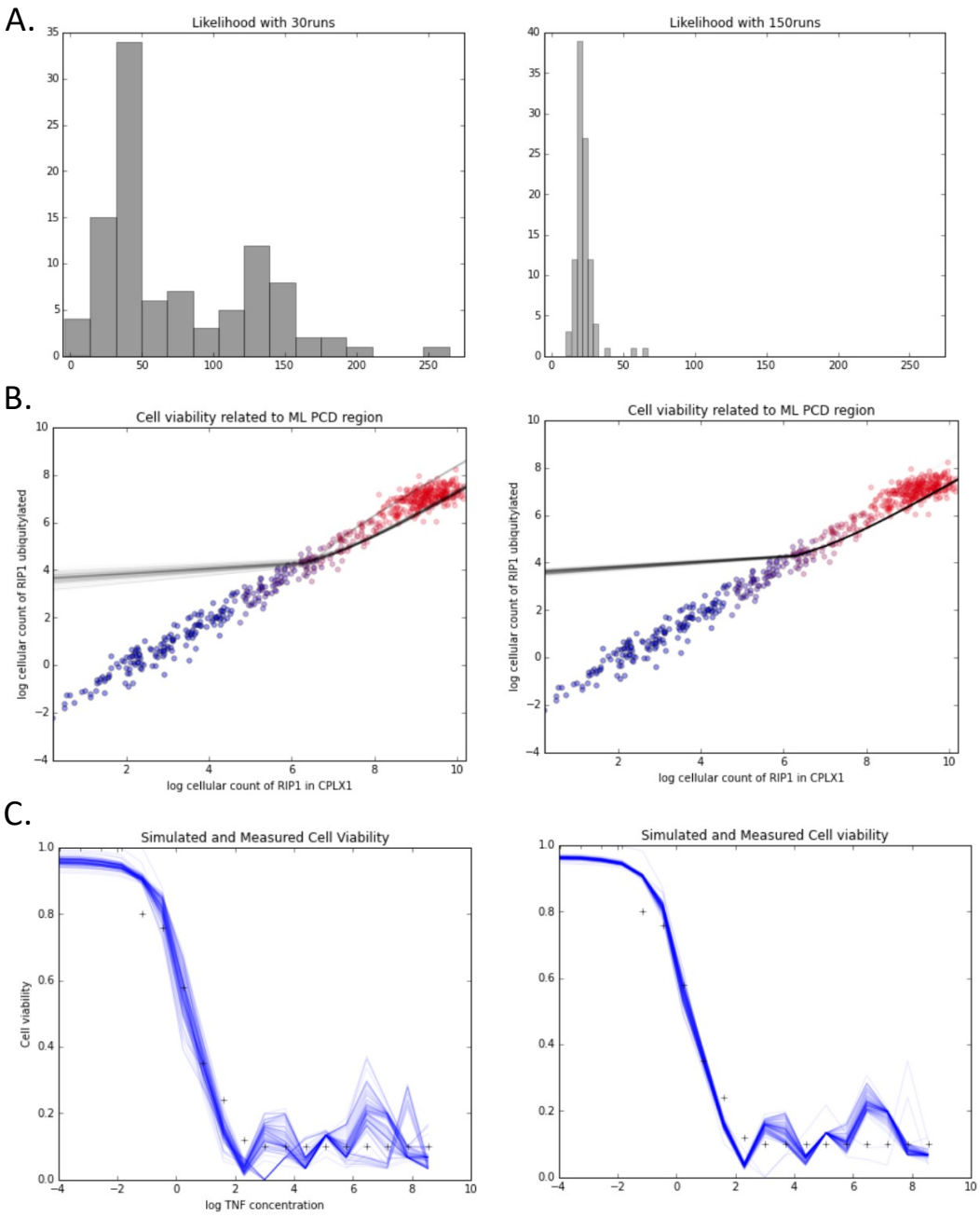


Figure 3.8. Effects of sample size in model calibrations of feature models of extrinsic heterogeneity. Models of biological mechanism that address heterogeneity typically require evaluation of an intractable likelihood. The likelihood is intractable in that it is impossible to evaluate the mechanistic model over the domain of the probability density function that describes the heterogeneity (i.e., the domain is uncountably infinite). Instead, the likelihood is approximated with evaluations over a finite sample from probability density function that describes the heterogeneity (i.e., the size of the *in silico* population). The size of the *in silico* population affects the reliability of the likelihood approximation. A. a histogram of repeated evaluations of the likelihood approximated using an *in silico* population size of 30 (left) and 150 (right). The approximated likelihood took a range of values (x-axis in both plots). A model of TNF Complex 1 dynamics of was trained to cell viability data. B. The maximal likelihood predictions of 50% cell-death probability curve for an *in silico* population of 30 (left) and 150 (right). C. The maximal likelihood predictions of cell viability for an *in silico* population of 30 (left) and 150 (right).

determine the cell viability *rate* – the change in proportion of necroptotic cells in a moving 3hr window (Figure 3.8. B.). We applied a measurement model that simulated cell viability (and thereby changes in cell viability) that uses probabilistic classifier to model the probability of cell death as a function of accumulated pMLKL. A model of necroptosis (i.e., Complex I, Complex II and Necroptosome signaling, shown in Figure 3.1. A.) calibrated to cell viability *rate* correctly predicted rapid increases in MLKL phosphorylation to occur in temporal windows that were later reflected in pMLKL immunoblot measurements.

Heterogeneity of a biological process requires more far more measurements than is typical of common biological assays to capture and model. This should direct scientist to collect more measurements to support biological research. Models of heterogeneity engender an intractable likelihood[00] that features a continuous random process or probability distribution that would require an infinite number of simulations (or draws) to evaluate. The likelihood is therefore approximated using a finite number of simulations. The size of the simulated sample set dictates the accuracy and reliability of the likelihood evaluation. To show this, I simulated a model of Complex I and Complex II dynamics (as shown in Figure 3.1.); used *in silico* intracellular amount of RIP1 in Complex I and polyubiquitinated RIP1 as predictors in a measurement model of cell death vs survival (Figure 3.8. B.); and calibrated the model to cell viability data Figure 3.8. C.). This likelihood function marginalizes the probability of cell death outcome over a population of cells; size =30 (Figure 3.8. left) or size = 150 (Figure 3.8. right) per experimental condition. Figure 3.8. A. shows the variability of the likelihood estimate for a single vector of model parameters. The smaller sample of simulations had a highly variable estimate of the likelihood. Approximate Bayesian computation (ABC-MCMC) addresses the intractable likelihood through modifications to its rejection algorithm[61]. Variability in the likelihood function, however, can produce *sticking* of a stochastic optimizer (i.e., when the optimizer receives an overestimate of the likelihood function and rejects new parameters that do not match that overestimated value) and drastically slow convergence of the Bayesian calibration. Further, the variability in the likelihood function decreases the certainty of the model calibration. Figure 3.8. B. show predictions of the line marking 50% probability of

cell death and Figure 3.8. C. shows predicted cell viabilities for a measurement model using simulated populations of size =30 (left) or size = 150 (right). Accurate approximation of a likelihood function of a model of biological heterogeneity requires a sufficiently large sample of simulations of the random process or probability distribution that models the heterogeneity. This requires many times more evaluations of the dynamical model (i.e., numerical integration of the model's ODEs) to accomplish. We used a GPU parallelized LSODA numerical integration package (CUPSODA)[74] to run hundreds of model evaluations. LSODA enables efficient and accurate ODE solutions but can stall in integration of poorly behaved models (i.e., models that possess the rapid non-stiff to stiff transitions typical of snap-action delay processes). DASSL is an alternative to the LSODA algorithm that better handles stiff-non-stiff transitions[186]. Future work would integrate high performance implementations of DASSL (e.g., GPU parallelization[187]) into PySB modeling package.

3.5. Conclusions

Systems biology has the unrealized potential to drive science-driven innovation in biology and medicine the way physics, chemistry and engineering has for several other industries. Systems biology cannot realize the potential without first confront its need for large amounts of the high-precision and quantitative measurements enjoyed in physics, chemistry and engineering. Fundamental properties of biology -- complexity, multiscale organization and heterogeneity – place tight constraints on what or how measurements occur. This leaves modeling, i.e., new modeling approaches that reconcile mechanism to measurement, as the essential response to the data problem in biology. In this work, I introduce a data-driven and probabilistic measurement model (Opt2Q) as a solution to the challenge posed by the dearth of quantitative data in systems biology. By integrating more of the measurements produced in typical investigations of cellular biology into an unbiased probabilistic modeling framework we can produce models that more accurately reflect the information in our experiments. But more work remains. With advancing computational capabilities and machine-learning methods, the Opt2Q measurement model can, in theory, adapt any *kind* of measurement to mechanistic models. This will provide new access to several

biological processes that continue to evade current modeling efforts. However, the demand, in systems biology, for *more* of data remains. Scientists should therefore devise experiments that collect more measurements (regardless of the kind(s) of measurement), if they intend to implement modeling approaches of systems biology.

References

1. Hesse MB. Models in physics. The British Journal for the Philosophy of Science. 1953 Nov 1;4(15):198-214.

2. Loeb J. The mechanistic conception of life: biological essays/by Jacques Loeb.

3. Chain PS, Grafham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR. Genome project standards in a new era of sequencing. Science. 2009 Oct 9;326(5950):236-7.

4. Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. Annual review of genomics and human genetics. 2001 Sep;2(1):343-72.

5. Kitano H. Systems biology: a brief overview. science. 2002 Mar 1;295(5560):1662-4.

6. Trewavas A. A Brief History of Systems Biology:"Every object that biology studies is a system of systems." Francois Jacob (1974). The Plant Cell. 2006 Oct 1;18(10):2420-30.

7. Schneider MV. Defining systems biology: a brief overview of the term and field. In silico systems biology. 2013:1-1.

8. Aderem A. Systems biology: its practice and challenges. Cell. 2005 May 20;121(4):511-3.

9. Faruqui N, Kummrow A, Fu B, Divieto C, Rojas F, Kisulu F, Cavalcante JJ, Wang J, Campbell J, Martins JL, Choi JH. Cellular metrology: Scoping for a value proposition in extra-and intracellular measurements. Frontiers in bioengineering and biotechnology. 2020 Jan 14;7:456.

10. Pargett M, Rundell AE, Buzzard GT, Umulis DM. Model-based analysis for qualitative data: an application in drosophila germline stem cell regulation. PLoS Comput Biol. 2014 Mar 13;10(3):e1003498.

11. Mitra ED, Dias R, Posner RG, Hlavacek WS. Using both qualitative and quantitative data in parameter identification for systems biology models. Nature communications. 2018 Sep 25;9(1):1-8.

12. Mitra ED, Hlavacek WS. Bayesian inference using qualitative observations of underlying continuous variables. Bioinformatics. 2020 May 1;36(10):3177-84.

13. Gupta HV, Beven KJ, Wagener T. Model calibration and uncertainty estimation. Encyclopedia of hydrological sciences. 2006 Apr 15.

14. Wagner PN, Shi Q, Salisbury-Ruf CT, Zou J, Savona MR, Fedoriw Y, Zinkel SS. Increased Ripk1-mediated bone marrow necroptosis leads to myelodysplasia and bone marrow failure in mice. Blood, The Journal of the American Society of Hematology. 2019 Jan 10;133(2):107-20.

15. Henderson R, Pisano GP, Orsenigo L. The pharmaceutical industry and the revolution in molecular biology: interactions among scientific, institutional, and organizational change.

16. Cardinal LB, Alessandri TM, Turner SF. Knowledge codifiability, resources, and science-based innovation. Journal of knowledge management. 2001 Jun 1.

17. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nature reviews Drug discovery. 2012 Mar;11(3):191-200.

18. Brigandt I. Explanation in biology: Reduction, pluralism, and explanatory aims. Science & Education. 2013 Jan;22(1):69-91.

19. Kubinyi H. Drug research: myths, hype and reality. Nature Reviews Drug Discovery. 2003 Aug;2(8):665-8.

20. King MR. Commentary: basic research in HIV vaccinology is hampered by reductionist thinking. Frontiers in immunology. 2016 Feb 10;7:42.

21. Van Regenmortel MH. Commentary: Basic research in HIV Vaccinology Is Hampered by reductionist thinking. Frontiers in immunology. 2016 Jul 5;7:266.

22. Weng G, Bhalla US, Iyengar R. Complexity in biological signaling systems. Science. 1999 Apr 2;284(5411):92-6.

23. Gross F, MacLeod M. Prospects and problems for standardizing model validation in systems biology. Progress in biophysics and molecular biology. 2017 Oct 1;129:3-12.

24. Sawicki MP, Samara G, Hurwitz M, Passaro Jr E. Human genome project. The American journal of surgery. 1993 Feb 1;165(2):258-64.

25. Megason SG, Fraser SE. Imaging in systems biology. Cell. 2007 Sep 7;130(5):784-95.

26. Feng, Xiaojun, et al. "Mass spectrometry in systems biology: an overview." *Mass spectrometry reviews* 27.6 (2008): 635-660.

27. Aretz I, Meierhofer D. Advantages and pitfalls of mass spectrometry based metabolome profiling in systems biology. International journal of molecular sciences. 2016 May;17(5):632.

28. McKinnon KM. Flow cytometry: an overview. Current protocols in immunology. 2018 Jan;120(1):5-1.

29. Wikswo JP, Prokop A, Baudenbacher F, Cliffel D, Csukas B, Velkovsky M. Engineering challenges of BioNEMS: the integration of microfluidics, micro-and nanodevices, models and external control for systems biology. InIee Proceedings-Nanobiotechnology 2006 Aug 1 (Vol. 153, No. 4, pp. 81-101). IET Digital Library.

30. Ingalls BP. Mathematical modeling in systems biology: an introduction. MIT press; 2013 Jul 5.

31. Gazestani VH, Lewis NE. From genotype to phenotype: Augmenting deep learning with networks and systems biology. Current opinion in systems biology. 2019 Jun 1;15:68-73.

32. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nature Reviews Genetics. 2015 Jun;16(6):321-32.

33. Nobile MS, Cazzaniga P, Tangherloni A, Besozzi D. Graphics processing units in bioinformatics, computational biology and systems biology. Briefings in bioinformatics. 2017 Sep 1;18(5):870-85.

34. Srivastava R, Varner J. Emerging technologies: systems biology. Biotechnology progress. 2007;23(1):24-7.

35. Chuang HY, Hofree M, Ideker T. A decade of systems biology. Annual review of cell and developmental biology. 2010 Nov 10;26:721-44

36. Bruggeman FJ, Westerhoff HV. The nature of systems biology. TRENDS in Microbiology. 2007 Jan 1;15(1):45-50.

37. Bleicken S, Jeschke G, Stegmueller C, Salvador-Gallego R, García-Sáez AJ, Bordignon E. Structural model of active Bax at the membrane. Molecular cell. 2014 Nov 20;56(4):496-505.

38. Eydgahi H, Chen WW, Muhlich JL, Vitkup D, Tsitsiklis JN, Sorger PK. Properties of cell death models calibrated and compared using Bayesian approaches. Molecular systems biology. 2013;9(1):644.

39. Ideker T, Lauffenburger D. Building with a scaffold: emerging strategies for high-to low-level cellular modeling. TRENDS in Biotechnology. 2003 Jun 1;21(6):255-62.

40. MacLeod M. Heuristic approaches to models and modeling in systems biology. Biology & philosophy. 2016 May 1;31(3):353-72.

41. Goldberg AV. Finding a maximum density subgraph. Berkeley: University of California; 1984 May.

42. Ng A, Jordan M, Weiss Y. On spectral clustering: Analysis and an algorithm. InAdvances in Neural Information Processing Systems (pp. 849856).

43. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC bioinformatics. 2003 Dec;4(1):1-27

44. Swan AL, Mobasheri A, Allaway D, Liddell S, Bacardit J. Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. Omics: a journal of integrative biology. 2013 Dec 1;17(12):595-610.

45. Upstill-Goddard R, Eccles D, Fliege J, Collins A. Machine learning approaches for the discovery of gene–gene interactions in disease data. Briefings in bioinformatics. 2013 Mar 1;14(2):251-60.

46. Friedman N, Linial M, Nachman I, Pe'er D. Using Bayesian networks to analyze expression data. Journal of computational biology. 2000 Aug 1;7(3-4):601-20.

47. Pritykin Y, Singh M. Simple topological features reflect dynamics and modularity in protein interaction networks. PLoS Comput Biol. 2013 Oct 10;9(10):e1003243.

48. Pino JC, Lubbock AL, Harris LA, Gutierrez DB, Farrow MA, Muszynski N, Tsui T, Norris JL, Caprioli RM, Wikswo JP, Lopez CF. A computational framework to explore cellular response mechanisms from multi-omics datasets. bioRxiv. 2020 Jan 1.

49. Edelstein-Keshet L. Mathematical models in biology. Society for Industrial and Applied Mathematics; 2005 Jan 1.

50. Rougny A, Touré V, Moodie S, Balaur I, Czauderna T, Borlinghaus H, Dogrusoz U, Mazein A, Dräger A, Blinov ML, Villéger A. Systems biology graphical notation: process description language level 1 version 2.0. Journal of integrative bioinformatics. 2019 Jun 1;16(2).

51. Wang RS, Saadatpour A, Albert R. Boolean modeling in systems biology: an overview of methodology and applications. Physical biology. 2012 Sep 25;9(5):055001.

52. Koch I. Petri nets in systems biology. Software & Systems Modeling. 2015 May 1;14(2):703-10.

53. Abar S, Theodoropoulos GK, Lemarinier P, O'Hare GM. Agent Based Modelling and Simulation tools: A review of the state-of-art software. Computer Science Review. 2017 May 1;24:13-33.

54. Ermentrout GB, Edelstein-Keshet L. Cellular automata approaches to biological modeling. Journal of theoretical Biology. 1993 Jan 7;160(1):97-133.

55. Wieland FG, Hauber AL, Rosenblatt M, Tönsing C, Timmer J. On structural and practical identifiability. Current Opinion in Systems Biology. 2021 Mar 22.

56. Szallasi Z. 10 Biological Data Acquisition for System Level Modeling—An Exercise in the Art of Compromise. System modeling in cellular biology. 2006:201.

57. Carrera J, Covert MW. Why build whole-cell models?. Trends in cell biology. 2015 Dec 1;25(12):719-22.

58. Goldberg AP, Szigeti B, Chew YH, Sekar JA, Roth YD, Karr JR. Emerging whole-cell modeling principles and methods. Current opinion in biotechnology. 2018 Jun 1;51:97-102.

59. Erguler K, Stumpf MP. Practical limits for reverse engineering of dynamical systems: a statistical analysis of sensitivity and parameter inferability in systems biology models. Molecular BioSystems. 2011;7(5):1593-602.

60. Csilléry K, Blum MG, Gaggiotti OE, François O. Approximate Bayesian computation (ABC) in practice. Trends in ecology & evolution. 2010 Jul 1;25(7):410-8.

61. Shockley EM, Vrugt JA, Lopez CF. PyDREAM: high-dimensional parameter inference for biological models in python. Bioinformatics. 2018 Feb 15;34(4):695-7.

62. Shi Y. Particle swarm optimization: developments, applications and resources. InProceedings of the 2001 congress on evolutionary computation (IEEE Cat. No. 01TH8546) 2001 May 27 (Vol. 1, pp. 81-86). IEEE.

63. Myers CR, Sethna JP. Universally sloppy parameter sensitivities in systems biology models. PLoS Comput Biol. 2007 Oct 5;3(10):e189.

64. Transtrum MK, Machta BB, Brown KS, Daniels BC, Myers CR, Sethna JP. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. The Journal of chemical physics. 2015 Jul 7;143(1):07B201_1.

65. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statistics surveys. 2010;4:40-79.

66. Brown KS, Sethna JP. Statistical mechanical approaches to models with many poorly known parameters. Physical review E. 2003 Aug 12;68(2):021904.

67. Mogilner A, Wollman R, Marshall WF. Quantitative modeling in cell biology: what is it good for?. Developmental cell. 2006 Sep 1;11(3):279-87.
** This article attributes a qouate “Equations are smarter than us” to Newton

68. Marcucio RS, Qin L, Alsberg E, Boerckel JD. Reverse engineering development: crosstalk opportunities between developmental biology and tissue engineering. Journal of Orthopaedic Research. 2017 Nov;35(11):2356-68.

69. Smanski MJ, Zhou H, Claesen J, Shen B, Fischbach MA, Voigt CA. Synthetic biology to access and expand nature's chemical diversity. Nature Reviews Microbiology. 2016 Mar;14(3):135.

70. Clancy CE, An G, Cannon WR, Liu Y, May EE, Ortoleva P, Popel AS, Sluka JP, Su J, Vicini P, Zhou X. Multiscale modeling in the clinic: drug design and development. Annals of biomedical engineering. 2016 Sep;44(9):2591-610.

71. MacLeod M, Nersessian NJ. Coupling simulation and experiment: The bimodal strategy in integrative systems biology. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences. 2013 Dec 1;44(4):572-84.

72. Waltemath D, Wolkenhauer O. How modeling standards, software, and initiatives support reproducibility in systems biology and systems medicine. IEEE Transactions on Biomedical Engineering. 2016 Jun 2;63(10):1999-2006.

73. Liepe J, Barnes C, Cule E, Erguler K, Kirk P, Toni T, Stumpf MP. ABC-SysBio—approximate Bayesian computation in Python with GPU support. Bioinformatics. 2010 Jul 15;26(14):1797-9.

74. Harris LA, Nobile MS, Pino JC, Lubbock AL, Besozzi D, Mauri G, Cazzaniga P, Lopez CF. GPU-powered model analysis with PySB/cupSODA. Bioinformatics. 2017 Nov 1;33(21):3492-4.

75. Ronen M, Rosenberg R, Shraiman BI, Alon U. Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. Proceedings of the national academy of sciences. 2002 Aug 6;99(16):10555-60.

76. Jamshidi N, Palsson BØ. Formulating genome-scale kinetic models in the post-genome era. Molecular systems biology. 2008;4(1):171.

77. Dobbe R, Tomlin CJ. Hybrid Systems Modeling for (Cancer) Systems Biology. BioRxiv. 2015 Jan 1:035022.

78. Walpole J, Papin JA, Peirce SM. Multiscale computational models of complex biological systems. Annual review of biomedical engineering. 2013 Jul 11;15:137-54.

79. Brenner S. Sequences and consequences. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010 Jan 12;365(1537):207-12.

80. Prill RJ, Marbach D, Saez-Rodriguez J, Sorger PK, Alexopoulos LG, Xue X, Clarke ND, Altan-Bonnet G, Stolovitzky G. Towards a rigorous assessment of systems biology models: the DREAM3 challenges. PloS one. 2010 Feb 23;5(2):e9202.

81. Woltosz WS. If we designed airplanes like we design drugs…. Journal of computer-aided molecular design. 2012 Jan;26(1):159-63.

82. Cotsaftis M. What makes a system complex?-an approach to self organization and emergence. InFrom System Complexity to Emergent Properties 2009 (pp. 49-99). Springer, Berlin, Heidelberg.

83. Kitano H. Towards a theory of biological robustness.

84. Rao CV, Wolf DM, Arkin AP. Control, exploitation and tolerance of intracellular noise. Nature. 2002 Nov;420(6912):231-7.

85. Hänggi P. Stochastic resonance in biology how noise can enhance detection of weak signals and help improve biological information processing. ChemPhysChem. 2002 Mar 15;3(3):285-90.

86. Levchenko A, Nemenman I. Cellular noise and information transmission. Current opinion in biotechnology. 2014 Aug 1;28:156-64.

87. Yi TM, Huang Y, Simon MI, Doyle J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. Proceedings of the National Academy of Sciences. 2000 Apr 25;97(9):4649-53.

88. Wang Z. Fluctuation Resonance of Different Genetic Feed Forward Loops. In2010 4th International Conference on Bioinformatics and Biomedical Engineering 2010 Jun 18 (pp. 1-4). IEEE.

89. Cloutier M, Wang E. Dynamic modeling and analysis of cancer cellular network motifs. Integrative Biology. 2011 Jul 1;3(7):724-32.

90. Richard M, Yvert G. How does evolution tune biological noise?. Frontiers in genetics. 2014 Oct 28;5:374.

91. Nagarajan R, Scutari M. Impact of noise on molecular network inference. PloS one. 2013 Dec 5;8(12):e80735.

92. Arriaga EA. Determining biological noise via single cell analysis. Analytical and bioanalytical chemistry. 2009 Jan;393(1):73-80.

93. Pino JC, Shockley E, Harris L, Lopez CF. Intrinsic stochasticity and variability in programmed cell death execution in a single cell.

94. Kochen MA, Lopez CF. A probabilistic approach to explore signal execution mechanisms with limited experimental data. Frontiers in genetics. 2020 Jul 10;11:686.

95. Ortega OO, Wilson BA, Pino JC, Irvin MW, Ildefonso GV, Garbett SP, Lopez CF. Probability-based mechanisms in biological networks with parameter uncertainty. bioRxiv. 2021 Jan 1.

96. Chaves M, Albert R, Sontag ED. Robustness and fragility of Boolean models for genetic regulatory networks. Journal of theoretical biology. 2005 Aug 7;235(3):431-49.

97. Molina-Mora JA, Kop-Montero M, Quiros-Fernandez I, Quirós S, Crespo-Mariño JL, Mora-Rodríguez RA. A hybrid mathematical modeling approach of the metabolic fate of a fluorescent

sphingolipid analogue to predict cancer chemosensitivity. Computers in biology and medicine. 2018 Jun 1;97:8-20.

98. Luan D, Zai M, Varner JD. Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. PLoS Comput Biol. 2007 Jul 20;3(7):e142.

99. Kambhampati SB, Vaishya R. Publication trends of PCL in the last 40 years on PubMed. Journal of Clinical Orthopaedics and Trauma. 2020 May 1;11:S354-61.

100. Hirschman L, Fort K, Boué S, Kyrpides N, Islamaj Doğan R, Cohen KB. Crowdsourcing and curation: perspectives from biology and natural language processing. Database. 2016 Jan 1;2016.

101. Kumar M, Joseph SR, Augsburg M, Bogdanova A, Drechsel D, Vastenhouw NL, Buchholz F, Gentzel M, Shevchenko A. MS Western, a method of multiplexed absolute protein quantification is a practical alternative to western blotting. Molecular & Cellular Proteomics. 2018 Feb 1;17(2):384-96.

102. Raghuraman BK, Hebbar S, Kumar M, Moon H, Henry I, Knust E, Shevchenko A. Absolute Quantification of Proteins in the Eye of Drosophila melanogaster. Proteomics. 2020 Dec;20(23):1900049.

103. Hughes AJ, Spelke DP, Xu Z, Kang CC, Schaffer DV, Herr AE. Single-cell western blotting. Nature methods. 2014 Jul;11(7):749-55.

104. Sinkala E, Sollier-Christen E, Renier C, Rosas-Canyelles E, Che J, Heirich K, Duncombe TA, Vlassakis J, Yamauchi KA, Huang H, Jeffrey SS. Profiling protein expression in circulating tumour cells using microfluidic western blotting. Nature communications. 2017 Mar 23;8(1):1-2.

105. Yang M, Cruz Villarreal J, Ariyasinghe N, Kruithoff R, Ros R, Ros A. Quantitative Approach for Protein Analysis in Small Cell Ensembles by an Integrated Microfluidic Chip with MALDI Mass Spectrometry. Analytical chemistry. 2021 Apr 5;93(15):6053-61.

106. Tay A. The Benefits of Going Small: Nanostructures for Mammalian Cell Transfection. ACS nano. 2020 Jul 6;14(7):7714-21.

107. Jena S, Damayanti NP, Tan J, Pratt ED, Irudayaraj JM, Parker LL. Multiplexable fluorescence lifetime imaging (FLIM) probes for Abl and Src-family kinases. Chemical Communications. 2020;56(87):13409-12.

108. Meyer NO, O'Donoghue AJ, Schulze-Gahmen U, Ravalin M, Moss SM, Winter MB, Knudsen GM, Craik CS. Multiplex substrate profiling by mass spectrometry for kinases as a method for revealing quantitative substrate motifs. Analytical chemistry. 2017 Apr 18;89(8):4550-8.

109. Mair W, Muntel J, Tepper K, Tang S, Biernat J, Seeley WW, Kosik KS, Mandelkow E, Steen H, Steen JA. FLEXITau: quantifying post-translational modifications of tau protein in vitro and in human disease. Analytical chemistry. 2016 Apr 5;88(7):3704-14.

110. Schlaffner CN, Kahnert K, Muntel J, Chauhan R, Renard BY, Steen JA, Steen H. FLEXIQuant-LF to quantify protein modification extent in label-free proteomics data. Elife. 2020 Dec 7;9:e58783.

111. Houle D, Pélabon C, Wagner GP, Hansen TF. Measurement and meaning in biology. The quarterly review of biology. 2011 Mar 1;86(1):3-4

112. Gerencser AA, Mookerjee SA, Jastroch M, Brand MD. Measurement of the absolute magnitude and time courses of mitochondrial membrane potential in primary and clonal pancreatic beta-cells. PLoS One. 2016 Jul 12;11(7):e0159199.

113. Degasperi A, Birtwistle MR, Volinsky N, Rauch J, Kolch W, Kholodenko BN. Evaluating strategies to normalise biological replicates of Western blot data. PloS one. 2014 Jan 27;9(1):e87293.

114. Taylor SC, Posch A. The design of a quantitative western blot experiment. BioMed research international. 2014 Oct;2014.

115. Stevens FJ, Bobrovnik SA. Deconvolution of antibody affinities and concentrations by non-linear regression analysis of competitive ELISA data. Journal of immunological methods. 2007 Dec 1;328(1-2):53-8.

116. Goasdoue K, Awabdy D, Bjorkman ST, Miller S. Standard loading controls are not reliable for Western blot quantification across brain development or in pathological conditions. Electrophoresis. 2016 Feb;37(4):630-4.

117. Thomaseth C, Radde N. Normalization of western blot data affects the statistics of estimators. IFAC-PapersOnLine. 2016 Jan 1;49(26):56-62.

118. Butler TA, Paul JW, Chan EC, Smith R, Tolosa JM. Misleading westerns: common quantification mistakes in western blot densitometry and proposed corrective measures. BioMed research international. 2019 Jan 21;2019.

119. Toni T, Jovanovic G, Huvet M, Buck M, Stumpf MP. From qualitative data to quantitative models: analysis of the phage shock protein stress response in Escherichia coli. BMC systems biology. 2011 Dec;5(1):1-6.

120. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature. 2009 May;459(7245):428-32.

121. Fröhlich F, Theis FJ, Rädler JO, Hasenauer J. Parameter estimation for dynamical systems with discrete events and logical operations. Bioinformatics. 2017 Apr 1;33(7):1049-56.

122. Degasperi A, Fey D, Kholodenko BN. Performance of objective functions and optimisation procedures for parameter estimation in system biology models. NPJ systems biology and applications. 2017 Aug 8;3(1):1-9.

123. Paek AL, Liu JC, Loewer A, Forrester WC, Lahav G. Cell-to-cell variation in p53 dynamics leads to fractional killing. Cell. 2016 Apr 21;165(3):631-42

124. Ballweg R, Paek AL, Zhang T. A dynamical framework for complex fractional killing. Scientific reports. 2017 Aug 14;7(1):1-2.

125. Roux J, Hafner M, Bandara S, Sims JJ, Hudson H, Chai D, Sorger PK. Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. Molecular systems biology. 2015 May;11(5):803.

126. Young FW. Quantitative analysis of qualitative data. Psychometrika. 1981 Dec;46(4):357-88.

127. Schmiester L, Weindl D, Hasenauer J. Parameterization of mechanistic models from qualitative data using an efficient optimal scaling approach. Journal of Mathematical Biology. 2020 Aug;81(2):603-23.

128. Kitano H. Foundations of systems biology. The MIT Press Cambridge, Massachusetts London, England; 2001 Jan 1.

129. Kitano H. Computational systems biology. Nature. 2002 Nov;420(6912):206-10.

130. Green DR. The coming decade of cell death research: five riddles. Cell. 2019 May 16;177(5):1094-107.

131. Villaverde AF, Banga JR. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. Journal of the Royal Society Interface. 2014 Feb 6;11(91):20130505.

132. Villaverde AF. Observability and structural identifiability of nonlinear biological systems. Complexity. 2019 Jan 1;2019.

133. Macklin P. Key challenges facing data-driven multicellular systems biology. Gigascience. 2019 Oct;8(10):giz127.

134. Chen KC, Csikasz-Nagy A, Gyorffy B, Val J, Novak B, Tyson JJ. Kinetic analysis of a molecular model of the budding yeast cell cycle. Molecular biology of the cell. 2000 Jan 1;11(1):369-91.

135. Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, Tyson JJ. Integrative analysis of cell cycle control in budding yeast. Molecular biology of the cell. 2004 Aug;15(8):3841-62.

136. Solano-Gálvez SG, Abadi-Chiriti J, Gutiérrez-Velez L, Rodríguez-Puente E, Konstat-Korzenny E, Álvarez-Hernández DA, Franyuti-Kelly G, Gutiérrez-Kobeh L, Vázquez-López R. Apoptosis: activation and inhibition in health and disease. Medical Sciences. 2018 Sep;6(3):54.

137. Kuwana T, Smith JJ, Muzio M, Dixit V, Newmeyer DD, Kornbluth S. Apoptosis induction by caspase-8 is amplified through the mitochondrial release of cytochrome c. Journal of Biological Chemistry. 1998 Jun 26;273(26):16589-94.

138. Flusberg DA, Roux J, Spencer SL, Sorger PK. Cells surviving fractional killing by TRAIL exhibit transient but sustainable resistance and inflammatory phenotypes. Molecular biology of the cell. 2013 Jul 15;24(14):2186-200.

139. Laussmann MA, Passante E, Hellwig CT, Tomiczek B, Flanagan L, Prehn JH, Huber HJ, Rehm M. Proteasome inhibition can impair caspase-8 activation upon submaximal stimulation of apoptotic tumor necrosis factor-related apoptosis inducing ligand (TRAIL) signaling. Journal of Biological Chemistry. 2012 Apr 27;287(18):14402-11.

140. Corazza N, Kassahn D, Jakob S, Badmann A, Brunner T. TRAIL-induced apoptosis. Annals of the New York Academy of Sciences. 2009 Aug 1;1171(1):50.

141. Elmore S. Apoptosis: a review of programmed cell death. Toxicologic pathology. 2007 Jun;35(4):495-516.

142. Kantari C, Walczak H. Caspase-8 and bid: caught in the act between death receptors and mitochondria. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 2011 Apr 1;1813(4):558-63.

143. Özören N, El-Deiry WS. Defining characteristics of Types I and II apoptotic cells in response to TRAIL. Neoplasia. 2002 Jan 1;4(6):551-7.

144. Kaufmann SH, Desnoyers S, Ottaviano Y, Davidson NE, Poirier GG. Specific proteolytic cleavage of poly (ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. Cancer research. 1993 Sep 1;53(17):3976-85.

145. Albeck JG, Burke JM, Aldridge BB, Zhang M, Lauffenburger DA, Sorger PK. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. Molecular cell. 2008 Apr 11;30(1):11-25.

146. Sajtos L, Magyar B. Auxiliary theories as translation mechanisms for measurement model specification. Journal of Business Research. 2016 Aug 1;69(8):3186-91.

147. Zubair MU, Zhang X. Hybrid Performance-Measurement Model of Elevators. Journal of Performance of Constructed Facilities. 2020 Apr 1;34(2):04020013.

148. Kamble SS, Gunasekaran A, Ghadge A, Raut R. A performance measurement system for industry 4.0 enabled smart manufacturing system in SMMEs-A review and empirical investigation. International Journal of Production Economics. 2020 Nov 1;229:107853.

149. Xia Y, Wang P, Berntorp K, Boufounos P, Orlik P, Svensson L, Granström K. Extended Object Tracking with Automotive Radar Using Learned Structural Measurement Model. In2020 IEEE Radar Conference (RadarConf20) 2020 Sep 21 (pp. 1-6). IEEE.

150. Lopez CF, Muhlich JL, Bachman JA, Sorger PK. Programming biological models in Python using PySB. Molecular systems biology. 2013;9(1):646.

151. Hindmarsh AC, Petzold LR. LSODA, ordinary differential equation solver for stiff or non-stiff system.

152. Antoniuk K, Franc V, Hlaváč V. Mord: Multi-class classifier for ordinal regression. InJoint European Conference on Machine Learning and Knowledge Discovery in Databases 2013 Sep 22 (pp. 96-111). Springer, Berlin, Heidelberg.

153. Bürkner PC, Vuorre M. Ordinal regression models in psychology: A tutorial. Advances in Methods and Practices in Psychological Science. 2019 Mar;2(1):77-101.

154. Poon WY, Wang HB. Analysis of ordinal categorical data with misclassification. British Journal of Mathematical and Statistical Psychology. 2010 Feb;63(1):17-42.

155. Hackeling G. Mastering Machine Learning with scikit-learn. Packt Publishing Ltd; 2017 Jul 24.

156. Ziemer RE, Tranter WH, Fannin DR. Signals and systems: continous and discrete.

157. Hagan T, Pulendran B. Will systems biology deliver its promise and contribute to the development of new or improved vaccines? From data to understanding through systems biology. Cold Spring Harbor perspectives in biology. 2018 Aug 1;10(8):a028894.

158. Kim R, Emi M, Tanabe K, Murakami S, Uchida Y, Arihiro K. Regulation and interplay of apoptotic and non-apoptotic cell death. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2006 Feb;208(3):319-26.

159. Peng GC, Alber M, Tepole AB, Cannon WR, De S, Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, Petzold L. Multiscale modeling meets machine learning: What can we learn?. Archives of Computational Methods in Engineering. 2021 May;28(3):1017-37.

160. Meyer P, Saez-Rodriguez J. Advances in systems biology modeling: 10 years of crowdsourcing DREAM challenges. Cell Systems. 2021 Jun 16;12(6):636-53.

161. Yan G, Elbadawi M, Efferth T. Multiple cell death modalities and their key features. World Academy of Sciences Journal. 2020 Mar 1;2(2):39-48.

162. Raj D, Brash DE, Grossman D. Keratinocyte apoptosis in epidermal development and disease. Journal of Investigative Dermatology. 2006 Feb 1;126(2):243-57.
** cited by [163]

163. Kakarla R, Hur J, Kim YJ, Kim J, Chwae YJ. Apoptotic cell-derived exosomes: Messages from dying cells. Experimental & molecular medicine. 2020 Jan;52(1):1-6.

164. D'Arcy MS. Cell death: a review of the major forms of apoptosis, necrosis and autophagy. Cell biology international. 2019 Jun;43(6):582-92.

165. Remijsen Q, Goossens V, Grootjans S, Van Den Haute C, Vanlangenakker N, Dondelinger Y, Roelandt R, Bruggeman I, Goncalves A, Bertrand MJ, Baekelandt V. Depletion of RIPK3 or MLKL blocks TNF-driven necroptosis and switches towards a delayed RIPK1 kinase-dependent apoptosis. Cell death & disease. 2014 Jan;5(1):e1004-.

166. Vanden Berghe T, Kaiser WJ, Bertrand MJ, Vandenabeele P. Molecular crosstalk between apoptosis, necroptosis, and survival signaling. Molecular & cellular oncology. 2015 Oct 2;2(4):e975093.

167. Seo J, Nam YW, Kim S, Oh DB, Song J. Necroptosis molecular mechanisms: Recent findings regarding novel necroptosis regulators. Experimental & Molecular Medicine. 2021 Jun 1:1-1.

168. Guicciardi ME, Mott JL, Bronk SF, Kurita S, Fingas CD, Gores GJ. Cellular inhibitor of apoptosis 1 (cIAP-1) degradation by caspase 8 during TNF-related apoptosis-inducing ligand (TRAIL)-induced apoptosis. Experimental cell research. 2011 Jan 1;317(1):107-16.

169. Micheau O, Tschopp J. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. Cell. 2003 Jul 25;114(2):181-90.

170. Tummers B, Green DR. Caspase-8: regulating life and death. Immunological reviews. 2017 May;277(1):76-89.

171. Oberst A, Dillon CP, Weinlich R, McCormick LL, Fitzgerald P, Pop C, Hakem R, Salvesen GS, Green DR. Catalytic activity of the caspase-8–FLIP L complex inhibits RIPK3-dependent necrosis. Nature. 2011 Mar;471(7338):363-7.

172. Liu Y, Liu T, Lei T, Zhang D, Du S, Girani L, Qi D, Lin C, Tong R, Wang Y. RIP1/RIP3-regulated necroptosis as a target for multifaceted disease therapy. International journal of molecular medicine. 2019 Sep 1;44(3):771-86.

173. Eydgahi H. *A quantitative framework For large-scale model estimation and discrimination In systems biology* (Doctoral dissertation, Massachusetts Institute of Technology).

174. Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, Petzold L. Multiscale modeling meets machine learning: What can we learn?. Archives of Computational Methods in Engineering. 2021 May;28(3):1017-37.

175. Müller M. Dynamic time warping. Information retrieval for music and motion. 2007:69-84.

176. Genkin A, Lewis DD, Madigan D. Large-scale Bayesian logistic regression for text categorization. technometrics. 2007 Aug 1;49(3):291-304.

177. Saxena A, Gupta A, Mukerjee A. Non-linear dimensionality reduction by locally linear isomaps. InInternational Conference on Neural Information Processing 2004 Nov 22 (pp. 1038-1043). Springer, Berlin, Heidelberg.

178. Yan L, Liu Y, Liu Y. Interval Feature Transformation for Time Series Classification Using Perceptually Important Points. Applied Sciences. 2020 Jan;10(16):5428.

179. Croker BA, Kelliher MA. BID-ding on necroptosis in MDS. Blood, The Journal of the American Society of Hematology. 2019 Jan 10;133(2):103-4.

180. National Institutes of Health. Bridge to Artificial Intellegence (Bridge2AI) [Internet]. Bethesda MD, National Institutes of Health; 2021[Updated 2021 June 25, cited 2021 June 29]. Available from: https://commonfund.nih.gov/bridge2ai

181. Bronstein L, Zechner C, Koeppl H. Bayesian inference of reaction kinetics from single-cell recordings across a heterogeneous cell population. Methods. 2015 Sep 1;85:22-35.

182. Liao TF. Interpreting probability models: Logit, probit, and other generalized linear models. Sage; 1994 Jun 30.

183. Faley S, Seale K, Hughey J, Schaffer DK, VanCompernolle S, McKinney B, Baudenbacher F, Unutmaz D, Wikswo JP. Microfluidic platform for real-time signaling analysis of multiple single T cells in parallel. Lab on a Chip. 2008;8(10):1700-12.

184. Metzig MO, Mitchell S, Taylor B, Hoffmann A. Time-resolved studies of necroptosis reveal distinct regulatory functions for NFκB.

185. Drovandi CC, Pettitt AN. Bayesian experimental design for models with intractable likelihoods. Biometrics. 2013 Dec;69(4):937-48.

186. Petzold LR. Description of dassl: a differential/algebraic system solver. Sandia National Labs., Livermore, CA (USA); 1982 Sep 1.

187. Stavåker K, Rolls D, Guo J, Fritzson P, Scholz SB. Compilation of Modelica array computations into single assignment C for efficient execution on CUDA-enabled GPUs. In3rd International Workshop on Equation-Based Object-Oriented Modeling Languages and Tools; Oslo; Norway; October 3 2010 Sep 21 (No. 047, pp. 81-90). Linköping University Electronic Press.

Appendices

A. Supplemental Tables for Chapter 2

Table A.1. Extrinsic Apoptosis Reaction Model Initial Conditions Parameters

Species	Symbol	Value	Notes
Ligand (TRAIL)	'L_0'	3000	3000 copies per cell equates to 50ng/mL
Receptor	'R_0'	200	
DISC	'DISC_0'	0	The death inducing signaling complex (DISC) is initially absent
Initiator Caspase	'IC_0'	2.0e4	Caspase 8
Effector Caspase	'EC_0'	1.0e4	Caspase 3
MOMP signal*	'MOMP_sig_0'	1.0e5	Uses the EARM initial value for Smac
PARP	'PARP_0'	1.0e6	PARP (Caspase 3 substrate and apoptosis marker)
Three Unrelated* Signaling Molecules	'USM1_0'	1.0e3	These signaling molecules do not interact with aEARM and are linked together via a separate set of activation and inactivation reactions (Supplemental table 2)
	'USM2_0'	1.0e3	
	'USM3_0'	1.0e3	

*Indicates values that were subjected log-normal extrinsic noise. Additionally, rate parameter 'kc0' (which catalyzes the formation of DISC) was subjected to extrinsic noise using the same procedure as for those noted in the table.

Table A.2. Reactions for unrelated signaling molecules

USM1 catalyzes USM2 activation	$USM1^* + USM2 \rightleftharpoons USM1^*{:}USM2 \longrightarrow USM1^* + USM2^*$	
USM2 catalyzes USM3 activation	$USM2^* + USM3 \rightleftharpoons USM2^*{:}USM3 \longrightarrow USM2^* + USM3^*$	
USM3 catalyzes USM1 activation	$USM3^* + USM1 \rightleftharpoons USM3^*{:}USM1 \longrightarrow USM3^* + USM1^*$	
USM1 catalyzes USM3 inactivation	$USM1^* + USM3^* \rightleftharpoons USM1^*{:}USM3^* \longrightarrow USM1^* + USM3$	
USM2 catalyzes USM1 inactivation	$USM2^* + USM1^* \rightleftharpoons USM2^*{:}USM1^* \longrightarrow USM2^* + USM1$	
USM3 catalyzes USM2 inactivation	$USM3^* + USM2^* \rightleftharpoons USM3^*{:}USM2^* \longrightarrow USM3^* + USM2$	

* Indicates activated.

Note, monotonically increasing values would correlate highly with TRAIL dependent species in aEARM, as an artifact of the aEARM species all having nonequilibrium values at t=0. This activation and inactivation scheme can produce slow oscillations, which decrease their correlation with TRAIL dependent species in aEARM.

Table A.3. – Parameterizations of Ordinal Measurement Model

"Ground Truth Parameterization"	$\alpha_{tBID} = 50.0, \boldsymbol{\beta}_{tBID} = (0.03, 0.40, 0.82, 0.97)$ $\alpha_{cPARP} = 50.0, \boldsymbol{\beta}_{cPARP} = (0.03, 0.20, 0.97)$ $\alpha_{IC-DISC} = 50.0, \boldsymbol{\beta}_{IC\_DISC} = (0.05, 0.40, 0.85)$
Ad hoc parameterization (Case 1)	$\alpha_{tBID} = 50.0, \boldsymbol{\beta}_{tBID} = (0.00, 0.33, 0.67, 1.00)$ $\alpha_{cPARP} = 50.0, \boldsymbol{\beta}_{cPARP} = (0.00, 0.50, 1.00)$
Ad hoc parameterization (Case 2)	$\alpha_{tBID} = 50.0, \boldsymbol{\beta}_{tBID} = (0.20, 0.40, 0.60, 0.80)$ $\alpha_{cPARP} = 50.0, \boldsymbol{\beta}_{cPARP} = (0.25, 0.50, 0.75)$

Table A.4. – Ground Truth Parameterization of Nominal Measurement Model

"Ground Truth Parameterization"	$\alpha = 4.0$ (slope term) $\beta = -0.25$ (intercept term) $\beta_{unrelated\ signal} = 0.0$ $\beta_{\max tBID\ rate} = 0.25$ $\beta_{time\ at \max tBID\ rate} = -1.0$

Table A.5. – Ground Truth Parameterization of aEARM

kf0	1.15594933e-07	kr2	6.87631084e-02	kc4	3.39059104e-03	kf7	1.98195485e-05
kr0	1.04056134e-05	kc2	1.09271827e-02	kf5	8.69183082e-10	kr7	1.82182416e-04
kc0	1.12028665e-05	kf3	3.82096809e-06	kr5	1.15619903e-03	kc7	3.90039620e-03
kf1	1.93083645e-06	kr3	1.39230000e-05	kc5	8.49612026e-06	kf8	1.76040074e-05
kr1	1.71446702e-03	kc3	2.95330594e-02	kf6	7.69174202e-07	kr8	1.34854804e-03
kc1	3.29355599e-01	kf4	1.61609297e-05	kr6	7.35417590e-04	kc8	1.67799299e-01
kf2	1.48867282e-08	kr4	1.71818138e-02	kc6	5.18485726e-03	kc9	8.28571278e-08

Table A.6: Gelman Rubin Values for Calibrated Models

Model Parameter	Fluorescence Data GR Values	Ordinal Data (1500s intervals) GR Values	Ordinal Data (300s intervals) GR Values
kf0	1.08514627	1.00416013	1.12613066
kr0	1.00220123	1.00731437	1.01094174
kc0	1.03226655	1.01661178	1.04884635
kf1	1.0526068	1.00787276	1.04161391
kr1	1.00504595	1.08370608	1.01497949
kc1	1.02257591	1.01803541	1.03766028
kf2	1.0130557	1.00138999	1.01795943
kr2	1.00305479	1.00047459	1.00459635
kc2	1.00393523	1.00892224	1.00669506
kf3	1.00236732	1.12747739	1.1033154
kr3	1.00518277	1.01158553	1.03838169
kc3	1.00208664	1.02737927	1.01072007
kf4	1.03636438	1.02287885	1.01291816
kr4	1.00973982	1.03539431	1.01106759
kc4	1.07074804	1.090603	1.0257023
kf5	1.06039857	1.00451871	1.05108669
kr5	1.11159688	1.03016518	1.01993528
kc5	1.06253039	1.00534456	1.19839917
kf6	1.16079145	1.00726142	1.03689339
kr6	1.01920037	1.01311067	1.00727259
kc6	1.12806401	1.03022348	1.05880415
kf7	1.1978977	1.00103472	1.05011201
kr7	1.015523	1.00363864	1.01582668
kc7	1.04973795	1.02434569	1.05180776
kf8	1.09112645	1.01930478	1.0592043
kr8	1.00130727	1.00505759	1.01064644
kc8	1.05073858	1.08118403	1.02354009
kc9	1.00629721	1.02264177	1.0120538
kf10	1.00377768		
kr10	1.00236537		
kc10	1.00127349		
kf11	1.00073078		
kr11	1.00006038		
kc11	1.01171049		
tBID_blot_c		1.03099437	1.04455791
tBID_blot_t1		1.00577635	1.04113665
tBID_blot_t2		1.00475455	1.0276283
tBID_blot_t3		1.18645775	1.01563682
tBID_blot_t4		1.03067735	1.03781179
cPARP_blot_c		1.0238484	1.0136447
cPARP_blot_t1		1.11064182	1.1217791
cPARP_blot_t2		1.09517445	1.04147523
cPARP_blot_t3		1.06658201	1.07456182
population term			
slope			
intercept			
unrelated signal coef			
max tBID rate			
time at max tBID rate			

Table A.6. (continued): Gelman Rubin Values for Calibrated Models

Model Parameter	Ordinal Data (180s intervals) GR Values	Ordinal Data (60s intervals) GR Values
kf0	1.09227853	1.19455225
kr0	1.01355927	1.00077791
kc0	1.21872921	1.19973674
kf1	1.08038175	1.09315405
kr1	1.0152864	1.00105368
kc1	1.01258181	1.02240288
kf2	1.02564596	1.00383742
kr2	1.00373621	1.00119441
kc2	1.00598853	1.00285406
kf3	1.0182979	1.05590664
kr3	1.00074794	1.00111898
kc3	1.02153065	1.01906018
kf4	1.00281514	1.0082177
kr4	1.00078925	1.0129524
kc4	1.02105192	1.19698486
kf5	1.11968392	1.0070741
kr5	1.02852199	1.00461711
kc5	1.07532827	1.01683151
kf6	1.05350227	1.01504637
kr6	1.00341359	1.00527656
kc6	1.02009389	1.01470974
kf7	1.00398446	1.06631947
kr7	1.00484732	1.00983058
kc7	1.05113986	1.02668767
kf8	1.00181596	1.00884647
kr8	1.00307851	1.00269787
kc8	1.03648993	1.05647474
kc9	1.00024721	1.00279375
kf10		
kr10		
kc10		
kf11		
kr11		
kc11		
tBID_blot_c	1.00332379	1.02397397
tBID_blot_t1	1.00534022	1.02255679
tBID_blot_t2	1.06528275	1.01778354
tBID_blot_t3	1.0136356	1.01897305
tBID_blot_t4	1.05682272	1.04328318
cPARP_blot_c	1.01365954	1.03105962
cPARP_blot_t1	1.01185214	1.02320031
cPARP_blot_t2	1.02776775	1.02698812
cPARP_blot_t3	1.0190093	1.04132186
population term		
slope		
intercept		
unrelated signal coef		
max tBID rate		
time at max tBID rate		
IC_DISC_blot_c		
IC_DISC_blot_t1		
IC_DISC_blot_t2		

Table A.6. (continued): Gelman Rubin Values for Calibrated Models

Model Parameter	Ordinal Data Uniform Priors for Measurement Model	Ordinal Data Cauchy (s=0.05) priors for Measurement Model	Ordinal Data Cauchy (s=0.005) priors for Measurement Model
kf0	1.07737054	1.03191813	1.1328619
kr0	1.00259287	1.01752463	1.00351023
kc0	1.03789017	1.04074446	1.06318023
kf1	1.07576464	1.01380773	1.0017717
kr1	1.06738237	1.00712663	1.00813843
kc1	1.10226386	1.04113557	1.01000593
kf2	1.00068557	1.00418444	1.00362501
kr2	1.0011541	1.00432116	1.00103303
kc2	1.00030986	1.01747953	1.00041386
kf3	1.03228596	1.0080699	1.03240523
kr3	1.00378028	1.00538727	1.00497349
kc3	1.08034193	1.00231253	1.01343209
kf4	1.01376285	1.06893954	1.00778619
kr4	1.00201924	1.00359284	1.00103777
kc4	1.04892641	1.03878297	1.01617983
kf5	1.04063891	1.0222424	1.02449938
kr5	1.03929407	1.01615766	1.00350635
kc5	1.02879373	1.01421718	1.00302031
kf6	1.03374728	1.05569952	1.01600363
kr6	1.01707809	1.01822194	1.01683535
kc6	1.02665254	1.02118882	1.08812762
kf7	1.04782734	1.05077333	1.00623474
kr7	1.00492933	1.01692564	1.00060767
kc7	1.06241425	1.01119753	1.13249616
kf8	1.00887196	1.01008733	1.00351389
kr8	1.00109911	1.00243843	1.0022869
kc8	1.00634403	1.13854335	1.10138409
kc9	1.0005081	1.01064269	1.00241524
kf10			
kr10			
kc10			
kf11			
kr11			
kc11			
tBID_blot_c	1.06175948	1.03161885	1.00405138
tBID_blot_t1	1.05877672	1.01005417	1.02304313
tBID_blot_t2	1.19211738	1.13115608	1.03286286
tBID_blot_t3	1.18990995	1.0739582	1.17343317
tBID_blot_t4	1.01935735	1.18526477	1.05742366
cPARP_blot_c	1.13717314	1.09934593	1.11144046
cPARP_blot_t1	1.15650209	1.10662718	1.09565958
cPARP_blot_t2	1.02989154	1.1249058	1.10627225
cPARP_blot_t3	1.0807912	1.1616861	1.15043555
population term			
slope			
intercept			
unrelated signal coef			
max tBID rate			
time at max tBID rate			
IC_DISC_blot_c			
IC_DISC_blot_t1			

Table A.6. (continued): Gelman Rubin Values for Calibrated Models

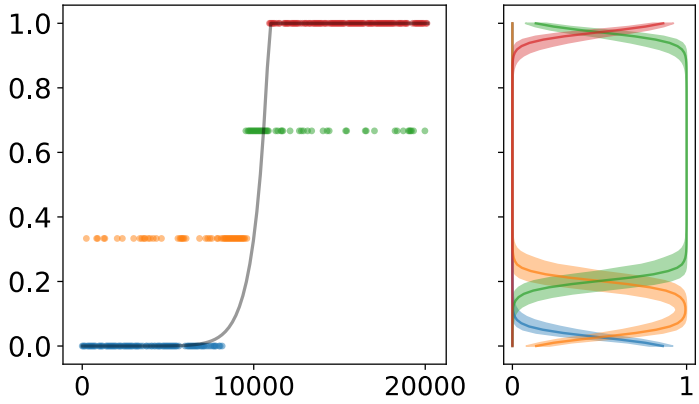
Model Parameter	Ordinal Data *ad hoc* measurement model (case 1)	Ordinal Data *ad hoc* measurement model (case 2)
kf0	1.07183296	1.00000592
kr0	1.00263502	1.00018696
kc0	1.11656015	1.00470607
kf1	1.15953028	1.11455866
kr1	1.00063062	1.05000013
kc1	1.07288909	1.16411978
kf2	1.00182892	1.19552739
kr2	1.00409273	1.11353462
kc2	1.00344124	1.18997792
kf3	1.00682301	1.17230406
kr3	1.00014235	1.00006023
kc3	1.04108872	1.06430983
kf4	1.00444363	1.33857703
kr4	1.00267552	1.01517747
kc4	1.14963217	1.14393001
kf5	1.08464278	1.02789781
kr5	1.01703471	1.01150415
kc5	1.00304983	1.02425553
kf6	1.11694745	1.01069756
kr6	1.00721167	1.00315432
kc6	1.10724666	1.072313
kf7	1.05194206	1.02347356
kr7	1.00192184	1.00212701
kc7	1.01155434	1.03646728
kf8	1.024501	1.01555239
kr8	1.00134119	1.00672554
kc8	1.16182793	1.27993175
kc9	1.00187187	1.4356259

Table A.6. (continued): Gelman Rubin Values for Calibrated Models

Model Parameter	Cell Death Dataset	Cell Death and IC-DISC Data	IC-DISC Ordinal Data
kf0	1.147082389	1.193508804	1.025888058
kr0	1.044549798	1.008386606	1.014422169
kc0	1.025309202	1.01875332	1.168214291
kf1	1.040691259	1.02818894	1.09296232
kr1	1.012969697	1.007636634	1.001200069
kc1	1.055136276	1.033107565	1.020263262
kf2	1.083912804	1.019927873	1.028984715
kr2	1.009869421	1.006757301	1.000222658
kc2	1.078414444	1.004699709	1.013178451
kf3	1.026369587	1.102132649	1.027554019
kr3	1.016666553	1.003595864	1.002519069
kc3	1.107343918	1.009143694	1.041325643
kf4	1.04447673	1.009667955	1.039020146
kr4	1.027860627	1.009341313	1.003006034
kc4	1.180435979	1.097402311	1.063978927
kf5	1.029934603	1.074718714	1.028936545
kr5	1.019712202	1.06652895	1.008325575
kc5	1.02241033	1.14509581	1.025696761
kf6	1.043800183	1.218209169	1.048466699
kr6	1.046639613	1.047536954	1.026881137
kc6	1.054575418	1.072500435	1.012896495
kf7	1.037915934	1.68575507	1.03771687
kr7	1.147434517	1.061381782	1.000948168
kc7	1.064822622	1.122795274	1.052213192
kf8	1.090951064	1.02801587	1.012671669
kr8	1.021929242	1.032207782	1.005317571
kc8	1.026358748	1.297306767	1.022151632
kc9	1.018902592	1.001625898	1.025093796
kf10	1.008415077	1.00323074	1.007755469
kr10	1.034442813	1.003569568	1.004258882
kc10	1.019863359	1.00298905	1.006463655
kf11	1.04564885	1.011607656	1.003158814
kr11	1.004836579	1.010118764	1.013272594
kc11	1.013629331	1.010063299	1.004651762
population term	1.010681473	1.012081154	
slope	1.038709572	1.010643589	
intercept	1.047805799	1.038549728	
unrelated signal coef	1.044461028	1.001983281	
max tBID rate	1.004573799	1.17125971	
time at max tBID rate	1.018539041	1.016101671	
IC_DISC_blot_c		1.034005667	1.037952935
IC_DISC_blot_t1		1.042627137	1.034646881
IC_DISC_blot_t2		1.094231403	1.115679326
IC_DISC_blot_t3		1.026081419	1.026988824

B. Supplemental Figures for Chapter 2

A.



B.

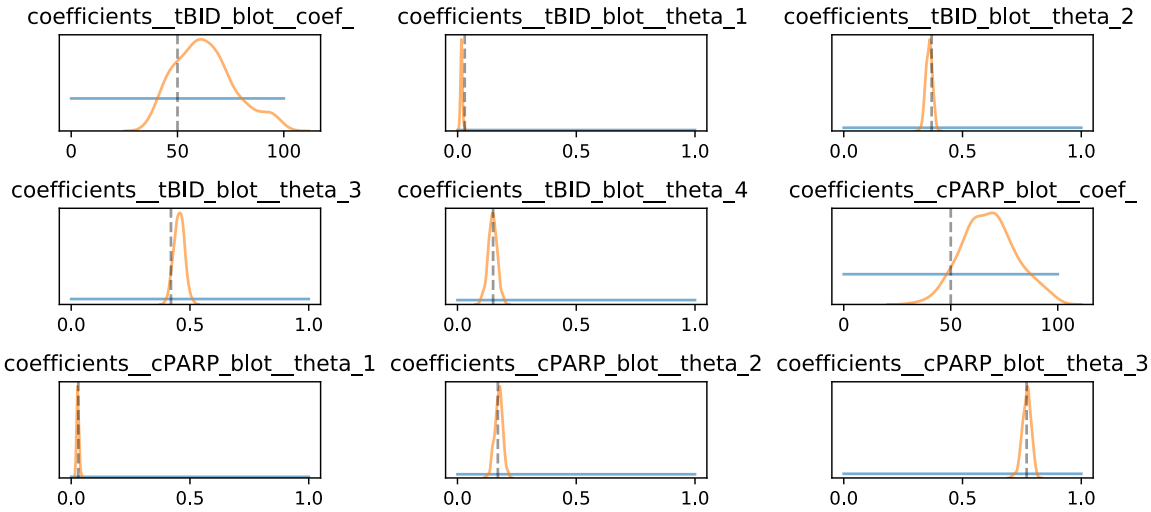
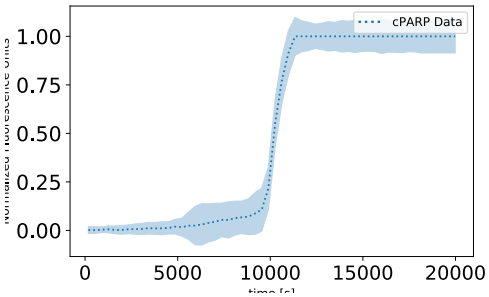
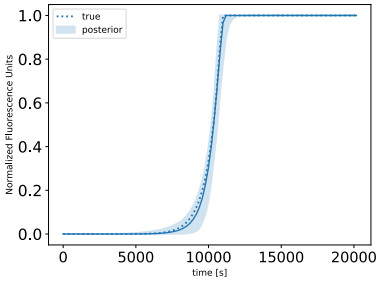
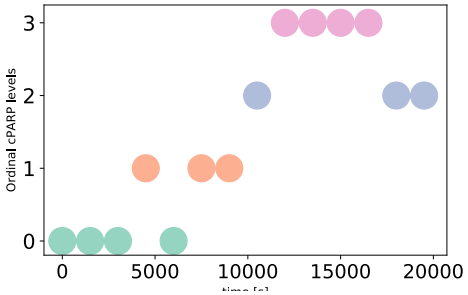
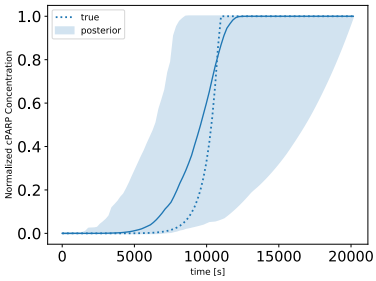


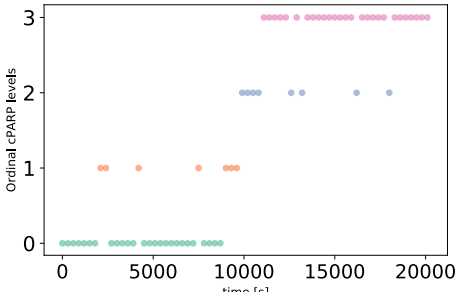
Figure B.1. Measurement model parameters calibrated to an ordinal dataset. The ordinal dataset contained a single ordinal value for tBID (Box 2, B.) and cPARP (A.) at every 60 second interval. Prior and posterior distributions for measurement model parameters are plotted in C. The 95% credible region of posterior predictions (shaded region) for the measurement model for tBID (Box 2, C.) and cPARP (B.) are plotted. The solid line in these plots is the median prediction for the measurement model. These plots give estimates of the probability of the ordinal value (x-axis) as a function of the normalized value of tBID or cPARP (y-axis). The ordinal categories are color coded and plotted in ascending order (for example, the 'blue' category is lower than the 'orange'.

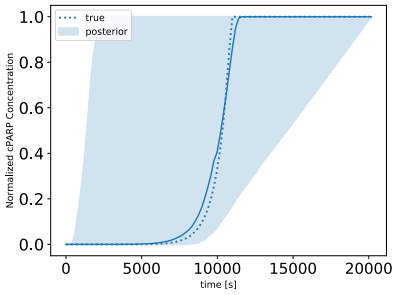

cPARP Data
1.00
0.75
0.50
0.25
0.00
0
5000
10000
15000
20000

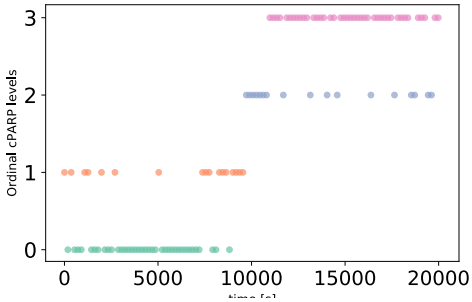

true
posterior
Normalized Fluorescence Units
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
time [s]

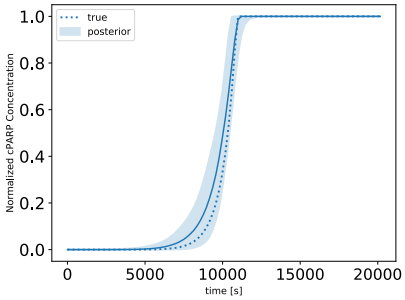

Ordinal cPARP levels
3
2
1
0
0
5000
10000
15000
20000

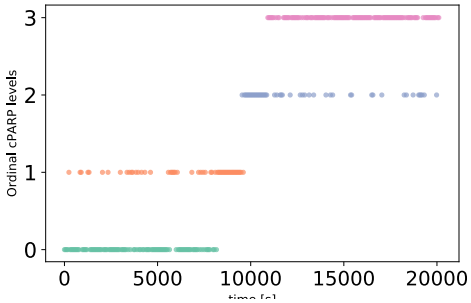

true
posterior
Normalized cPARP Concentration
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
time [s]


Ordinal cPARP levels
3
2
1
0
0
5000
10000
15000
20000


true
posterior
Normalized cPARP Concentration
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
time [s]


Ordinal cPARP levels
3
2
1
0
0
5000
10000
15000
20000


true
posterior
Normalized cPARP Concentration
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
time [s]


Ordinal cPARP levels
3
2
1
0
0
5000
10000
15000
20000

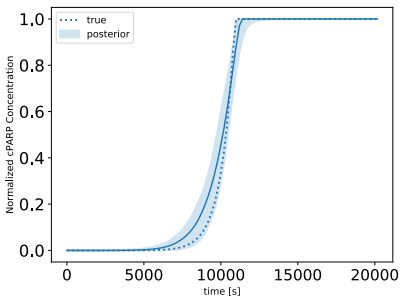

true
posterior
Normalized cPARP Concentration
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
time [s]

Figure B.2. PARP cleavage dynamics of aEARM trained to Fluorescence EC-RP data. Bayesian optimization of aEARM parameters trained to A.) Effector caspase reporter (EC-RP) fluorescence time-course measurements at 180s intervals (i.e. a proxy for PARP cleavage) and IC-RP fluorescence time-course (Fig 2A) data from [REF]). The plot shows the mean (dotted line) +/- 1 standard deviation (shaded region) for each time point. B.) The 95% credible region of posterior predictions (shaded region) for cPARP concentration in aEARM, calibrated to fluorescence measurements of IC-RP and EC-RP data. The median prediction (solid-line) and true (dotted line) cPARP concentration trajectories are shown. C., E., G., and I.) Ordinal measurements of cPARP at occurring at every 1500, 300, 180 and 60s timepoint, respectively. The 95% credible region of posterior predictions of cPARP dynamics for aEARM calibrated to ordinal measurements of tBID and cPARP occurring at every 1500, 300, 180 and 60s timepoint, respectively. The 95% credible region of predictions (shaded region), median prediction (solid line) and true (dotted line) cPARP dynamics for aEARM calibrated to ordinal measurements of tBID and cPARP occurring at every 1500, 300, 180 and 60s timepoint, are plotted in D., F., H., and J., respectively. The plots for tBID ordinal measurements and predictions are found in Figure 2.

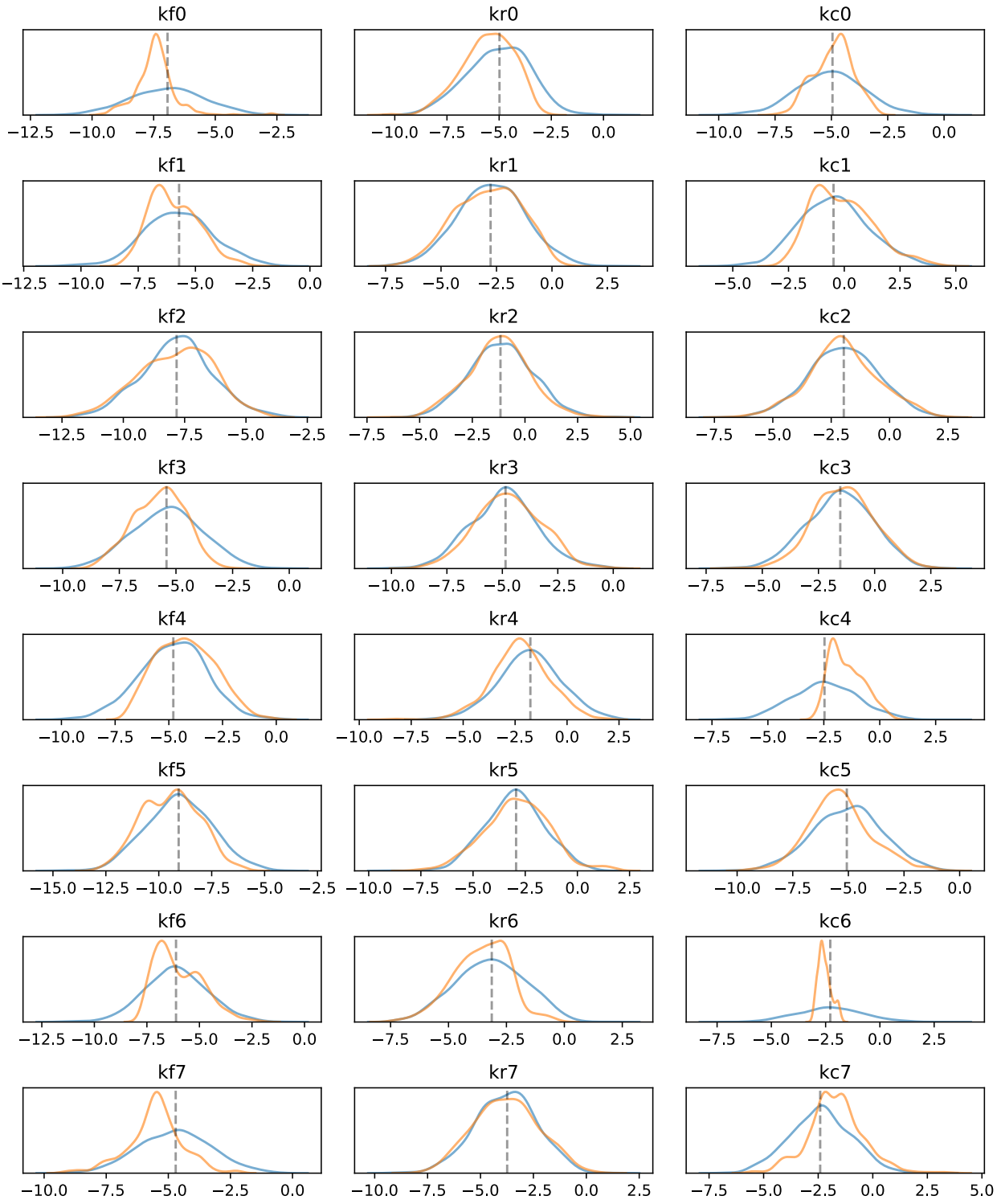


Figure B.3. Model parameters calibrated to a Fluorescence Dataset Parameters for aEARM were calibrated to fluorescence time-course measurements of EC-RP and IC-RP at every 180s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

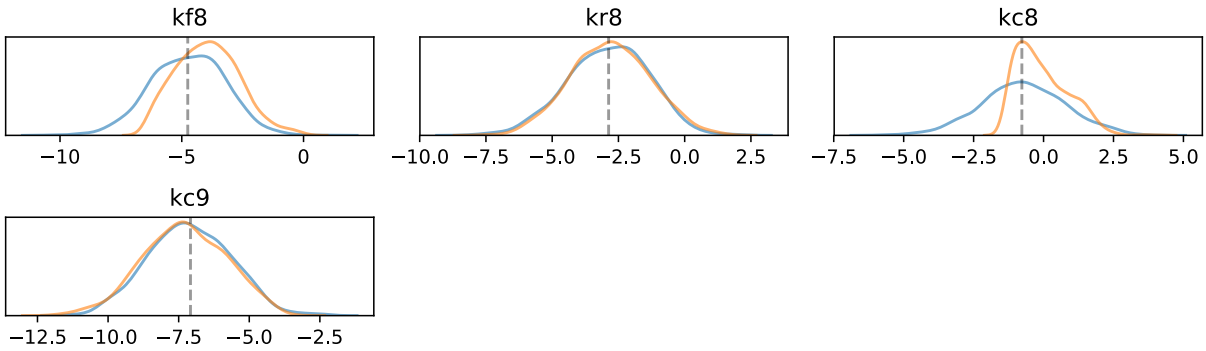


Figure B.3. (continued) Model parameters (remaining) calibrated to a Fluorescence Dataset Parameters for aEARM were calibrated to fluorescence time-course measurements of EC-RP and IC-RP at every 180s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

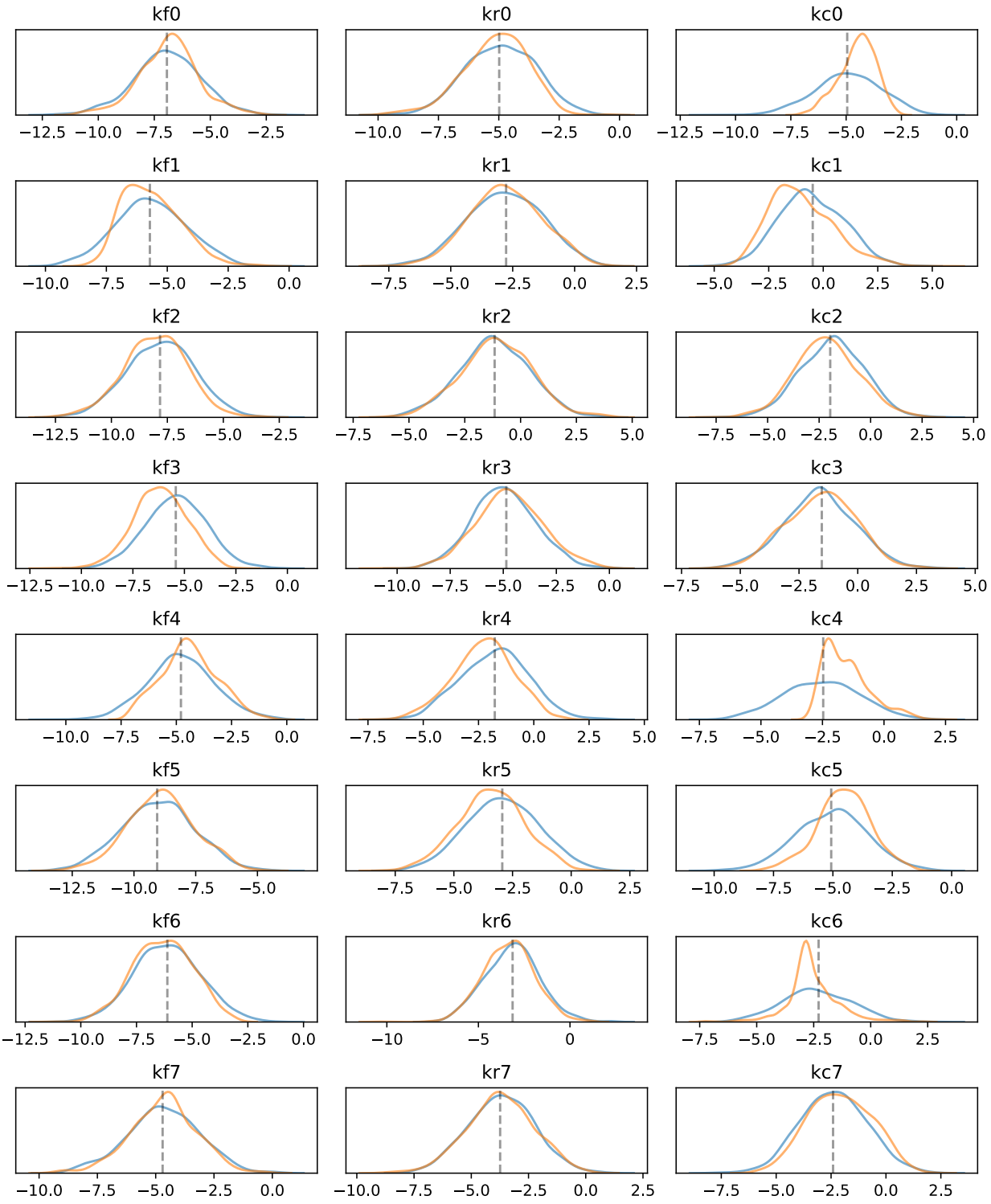


Figure B.4. Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

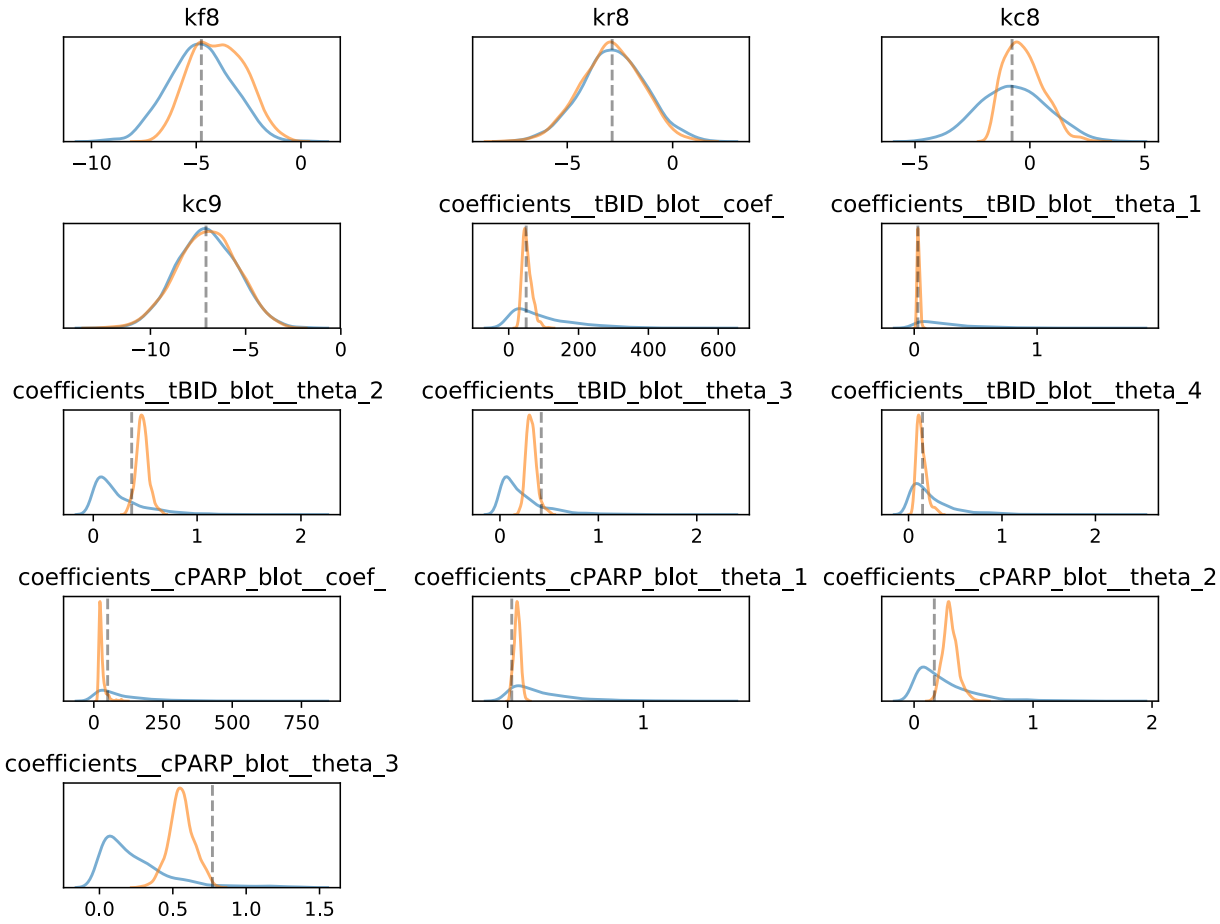


Figure B.4. (continued) Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of the aEARM parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

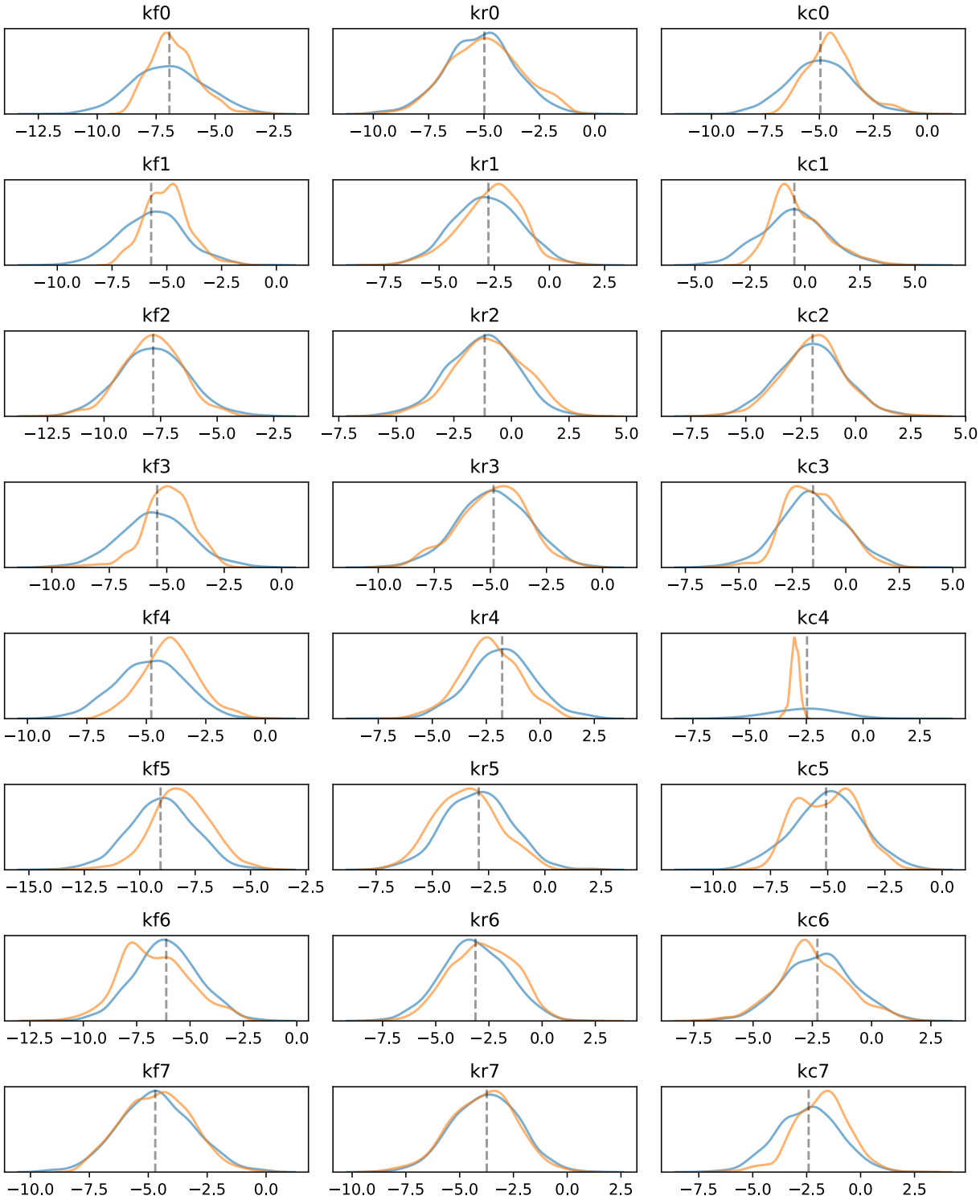


Figure B.5. Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 180s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

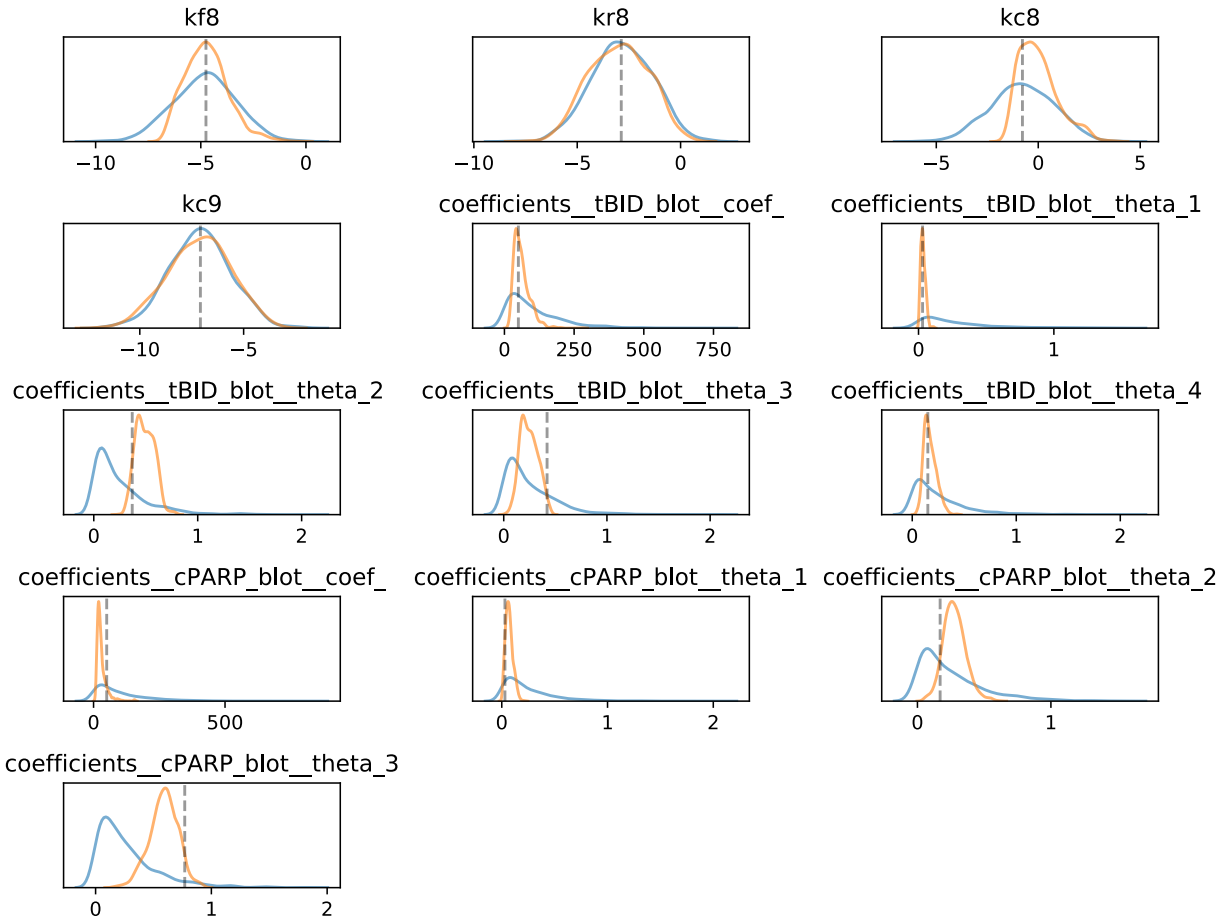


Figure B.5. (continued) Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 180s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of the aEARM parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

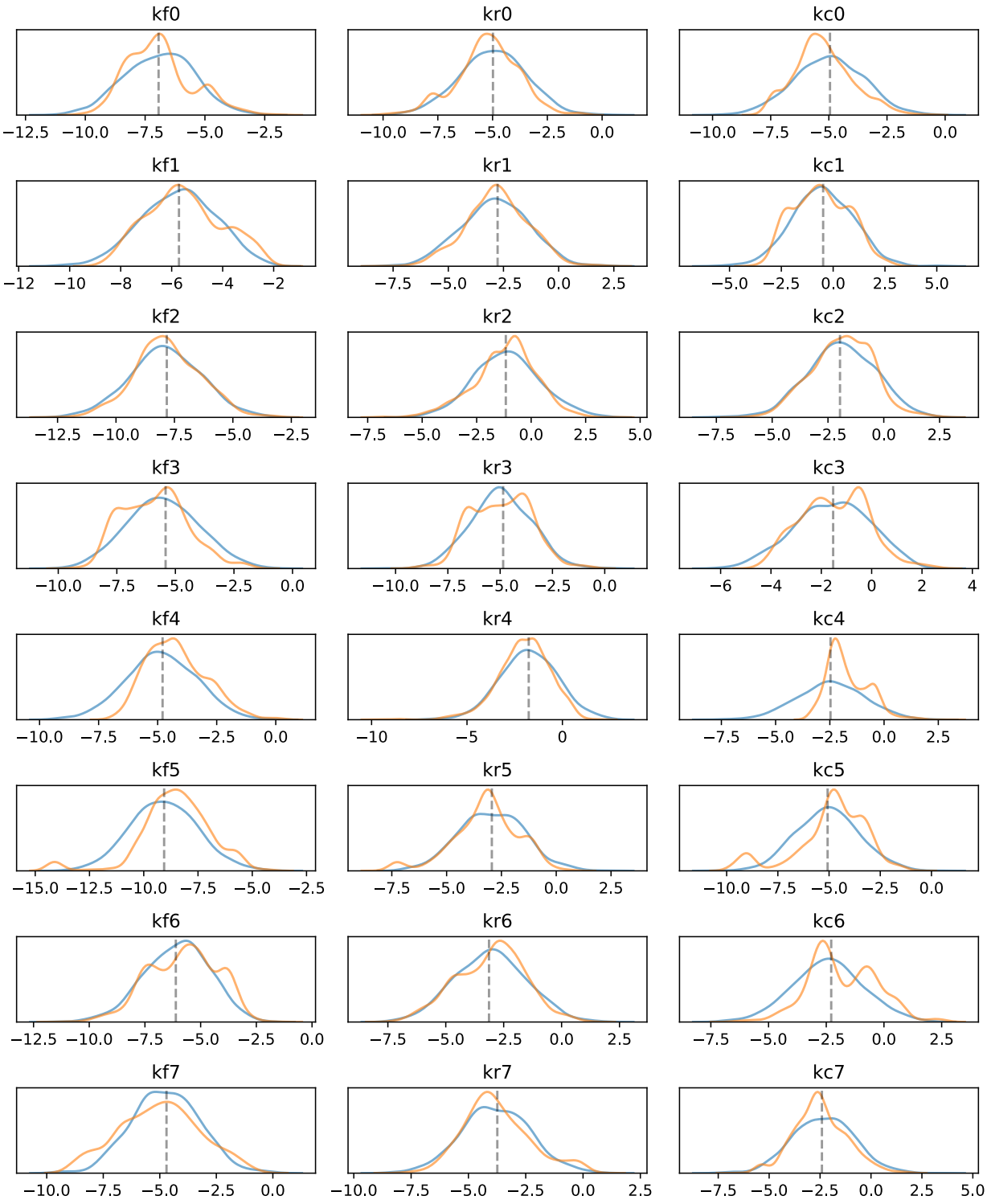


Figure B.6. Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 300s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

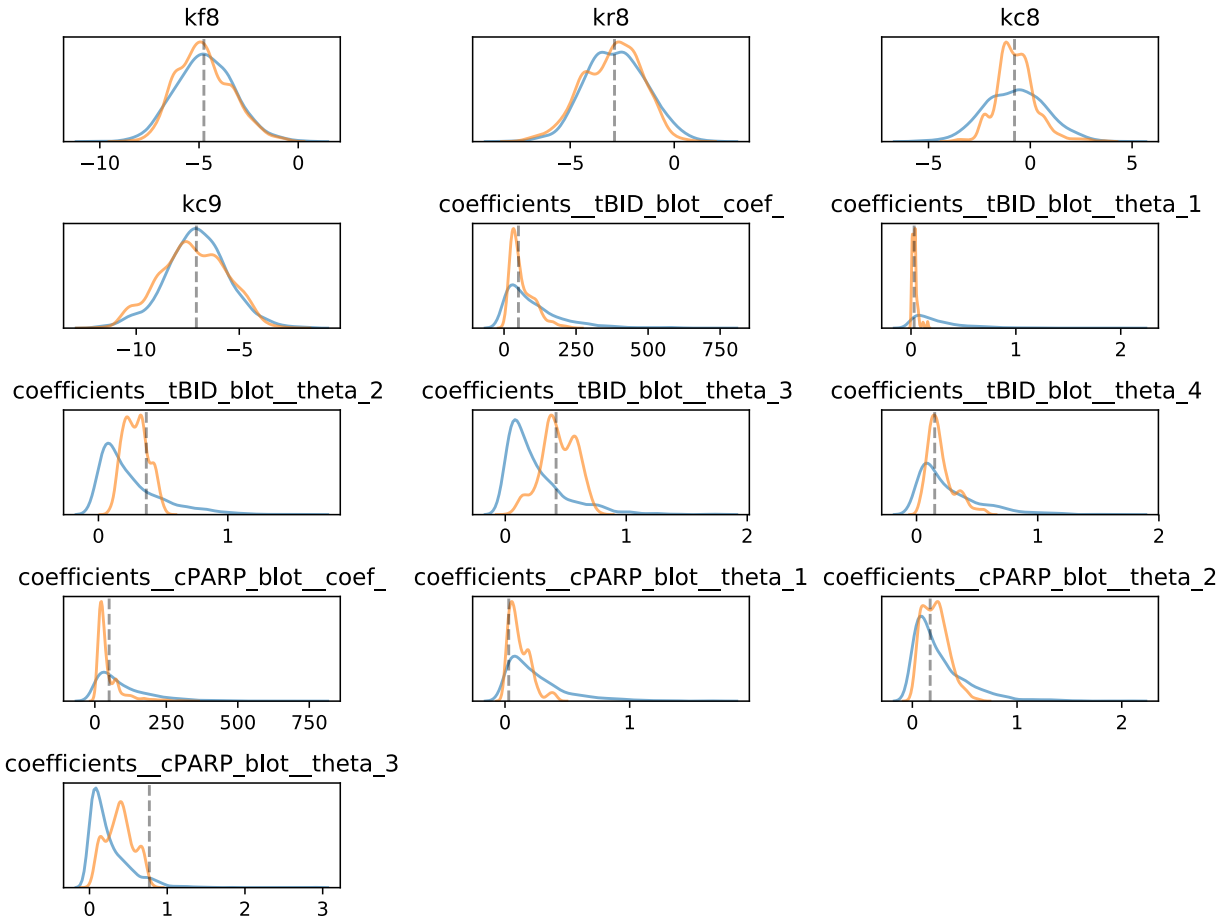


Figure B.6. (continued) Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 300s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of the aEARM parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

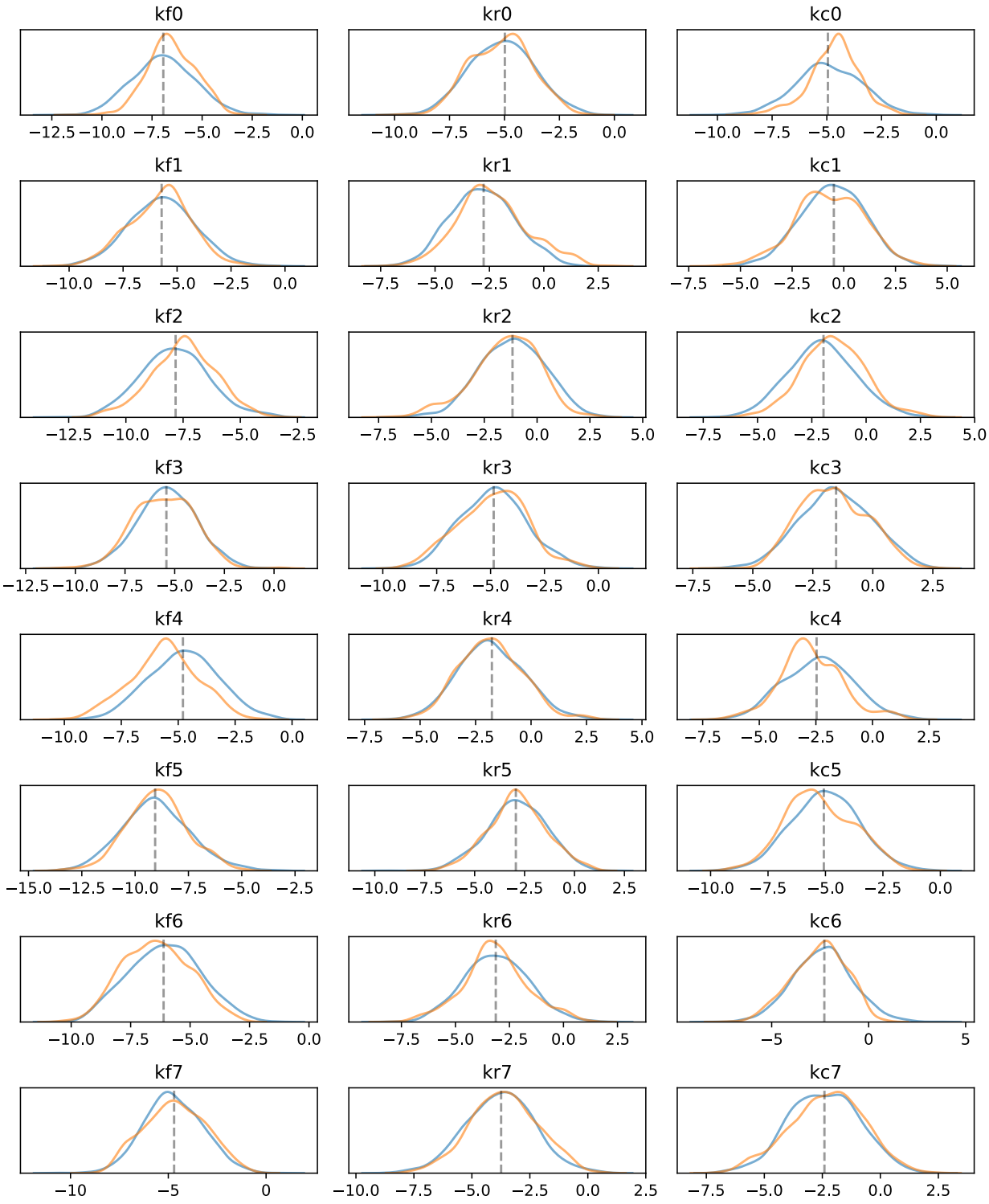


Figure B.7. Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 1500s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

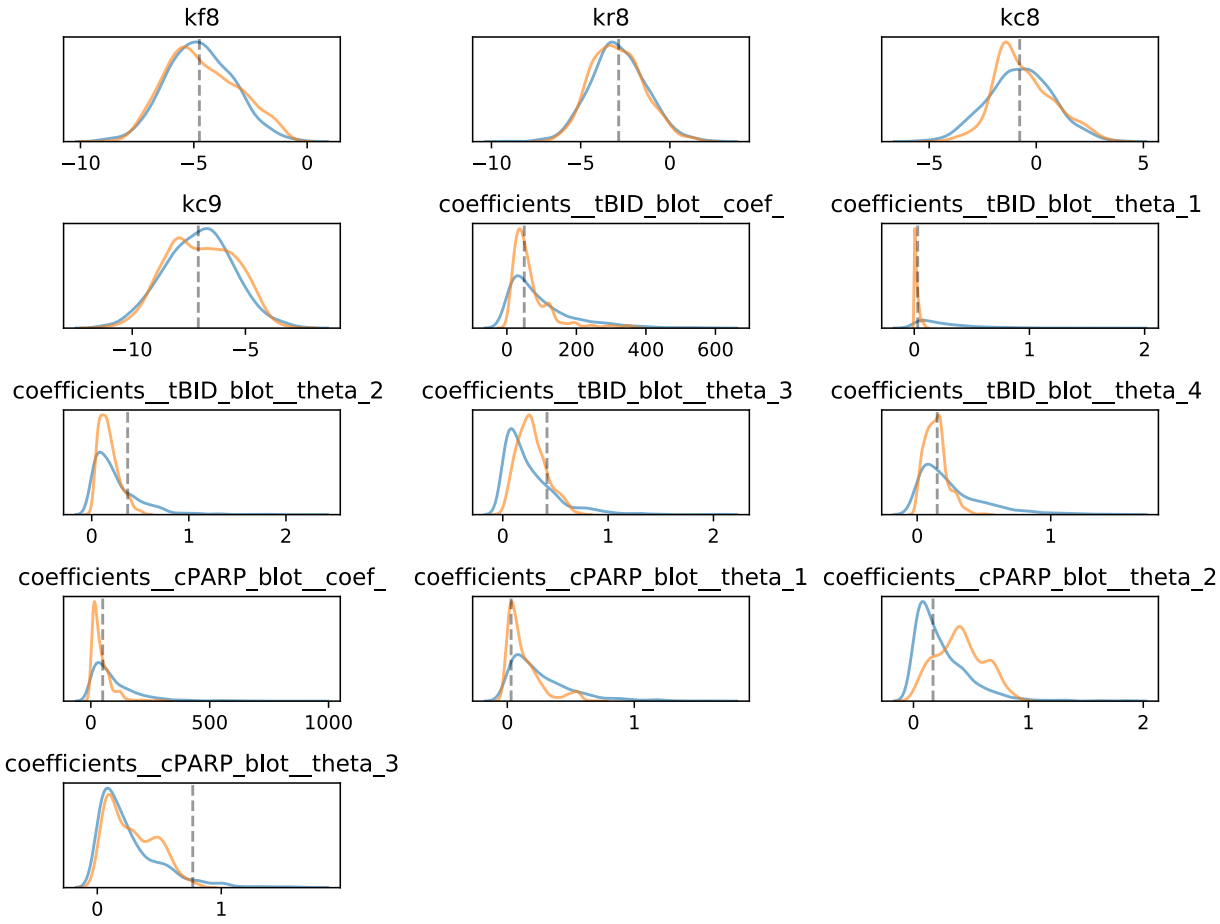


Figure B.7. (continued) Model parameters calibrated to an Ordinal Dataset Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 1500s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of the aEARM parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

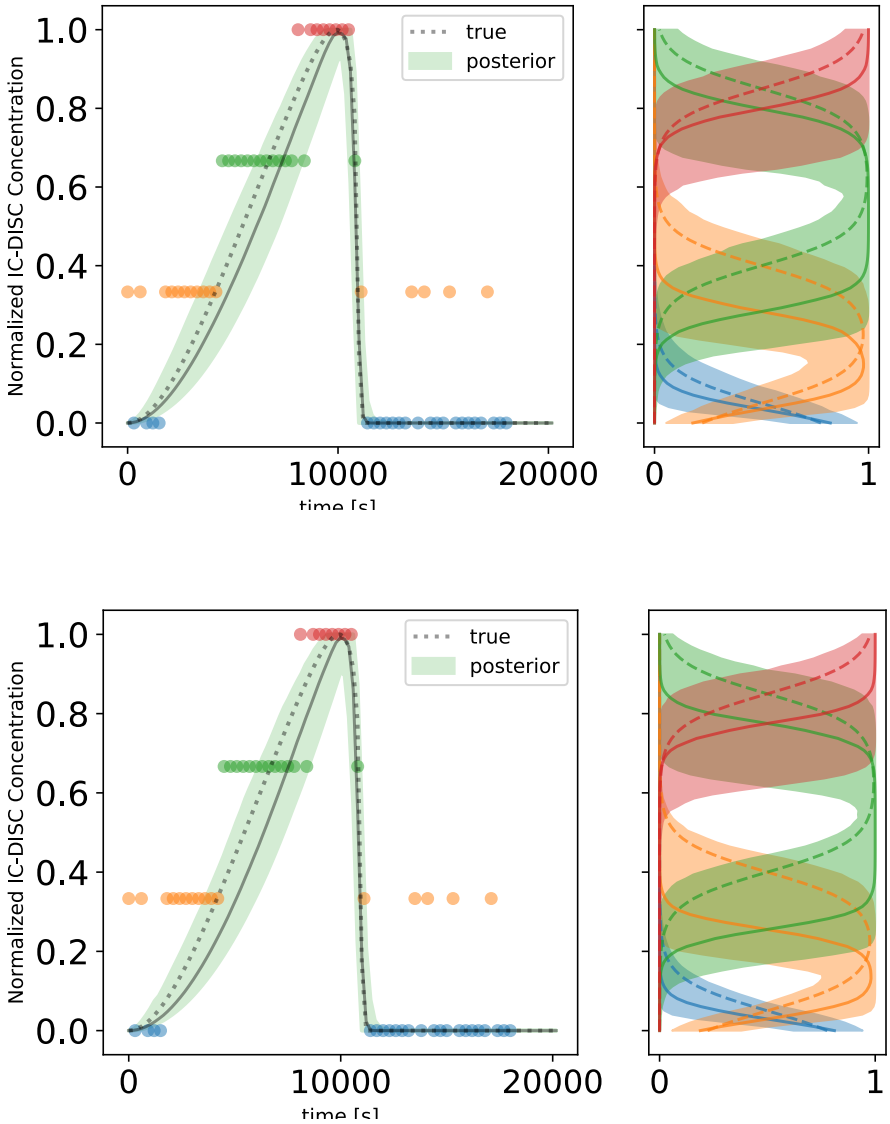


Figure B.8. Predicted Initiator caspase and DISC colocalization dynamics of aEARM trained to ordinal and mixed ordinal/nominal datasets. The 95% credible region (shaded region) of posterior predictions of IC-DISC dynamics of aEARM calibrated to ordinal IC-DISC data in Figure 4C, A.) and a mixed dataset containing the nominal data in Figure 4A the ordinal IC-DISC data, C.). The median predictions (solid-line) and true (dotted line) are also plotted. The adjacent panels B.) and D.) give the 95% credible region of posterior predictions (shaded regions) for the probability of class membership (x-axis) as a function of aEARM-simulated normalized IC-DISC concentration (y-axis). The ordinal categories are color coded and plotted in ascending order.

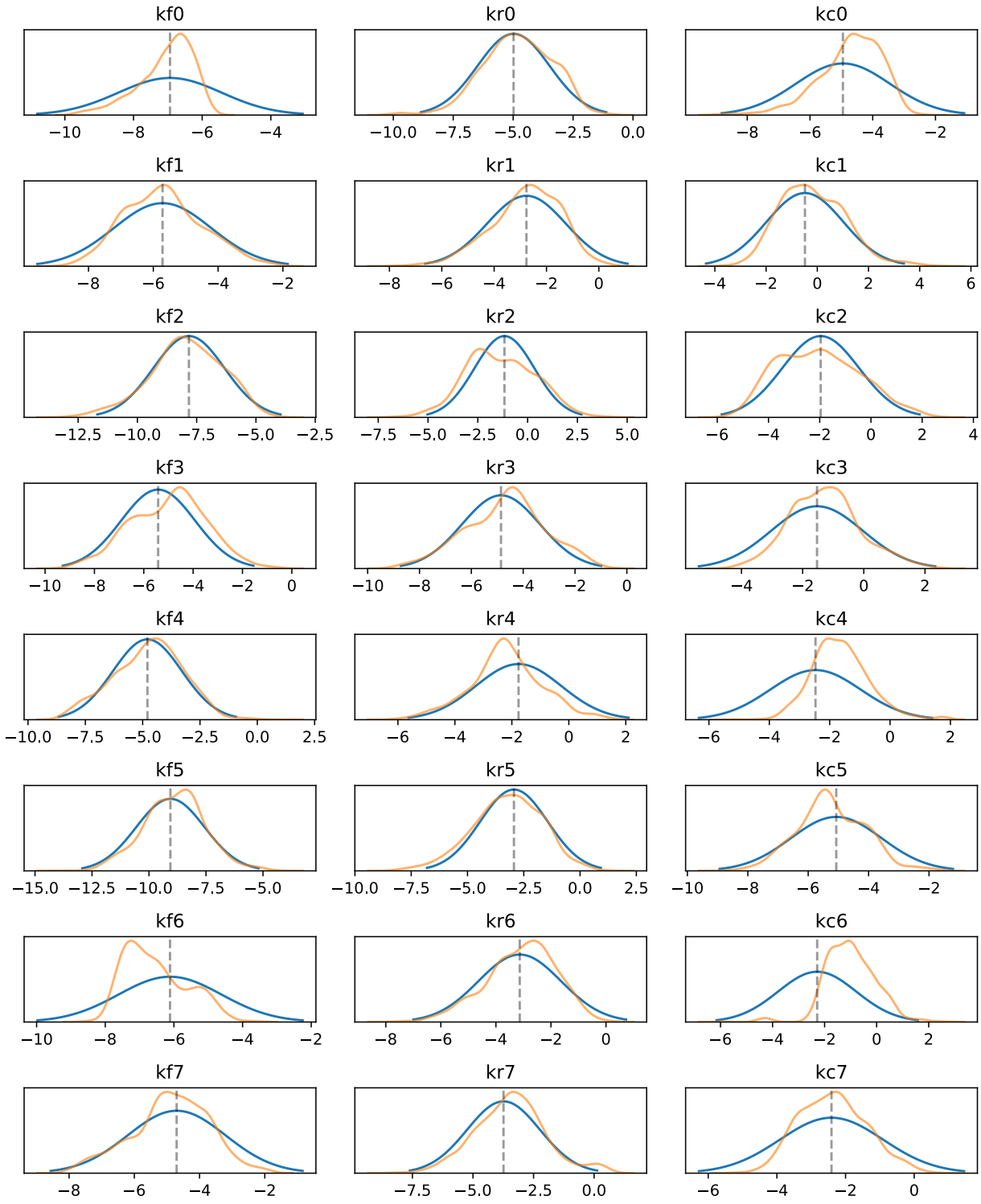


Figure B.9. Model parameters calibrated to a Cell Death Dataset Parameters for aEARM were calibrated to nominal observations of cell death vs survival. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

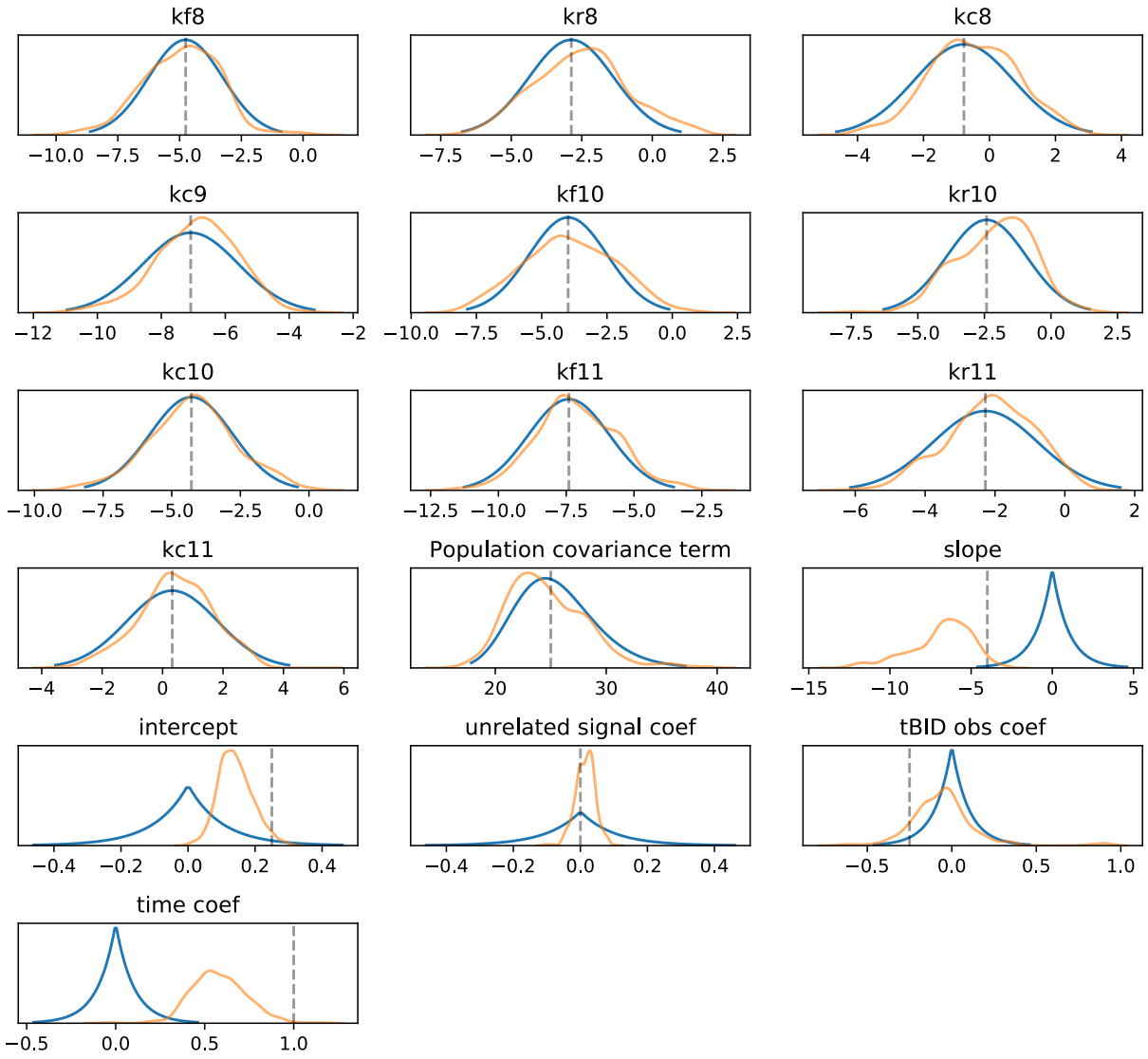


Figure B.9. (continued) Model parameters calibrated to a Cell Dataset Parameters for aEARM were calibrated to nominal observations of cell death vs survival. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of the aEARM parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

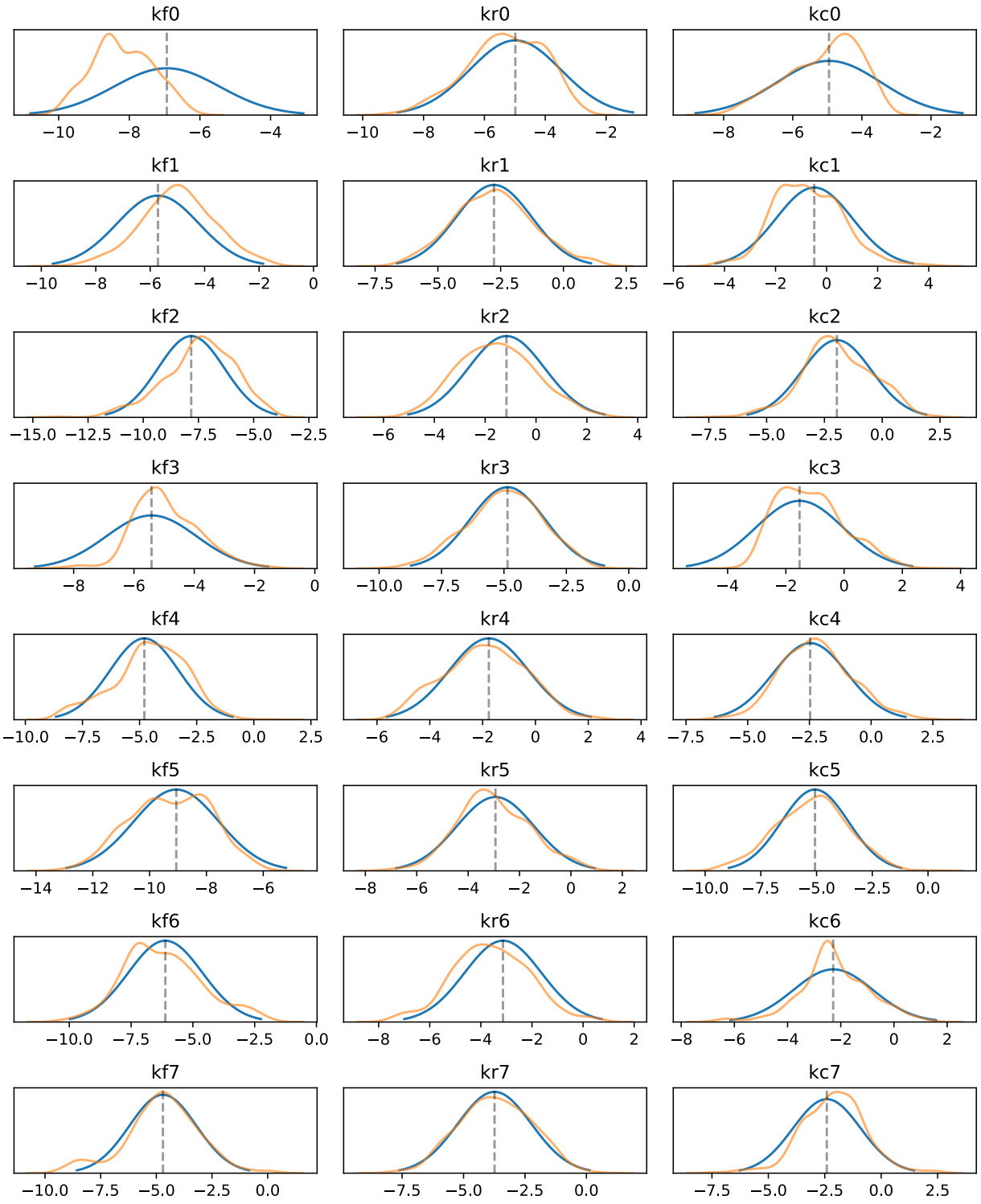


Figure B.10. Model parameters calibrated to Ordinal IC-DISC Dataset Parameters for aEARM were calibrated to ordinal values of IC-DISC abundance at every 300s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

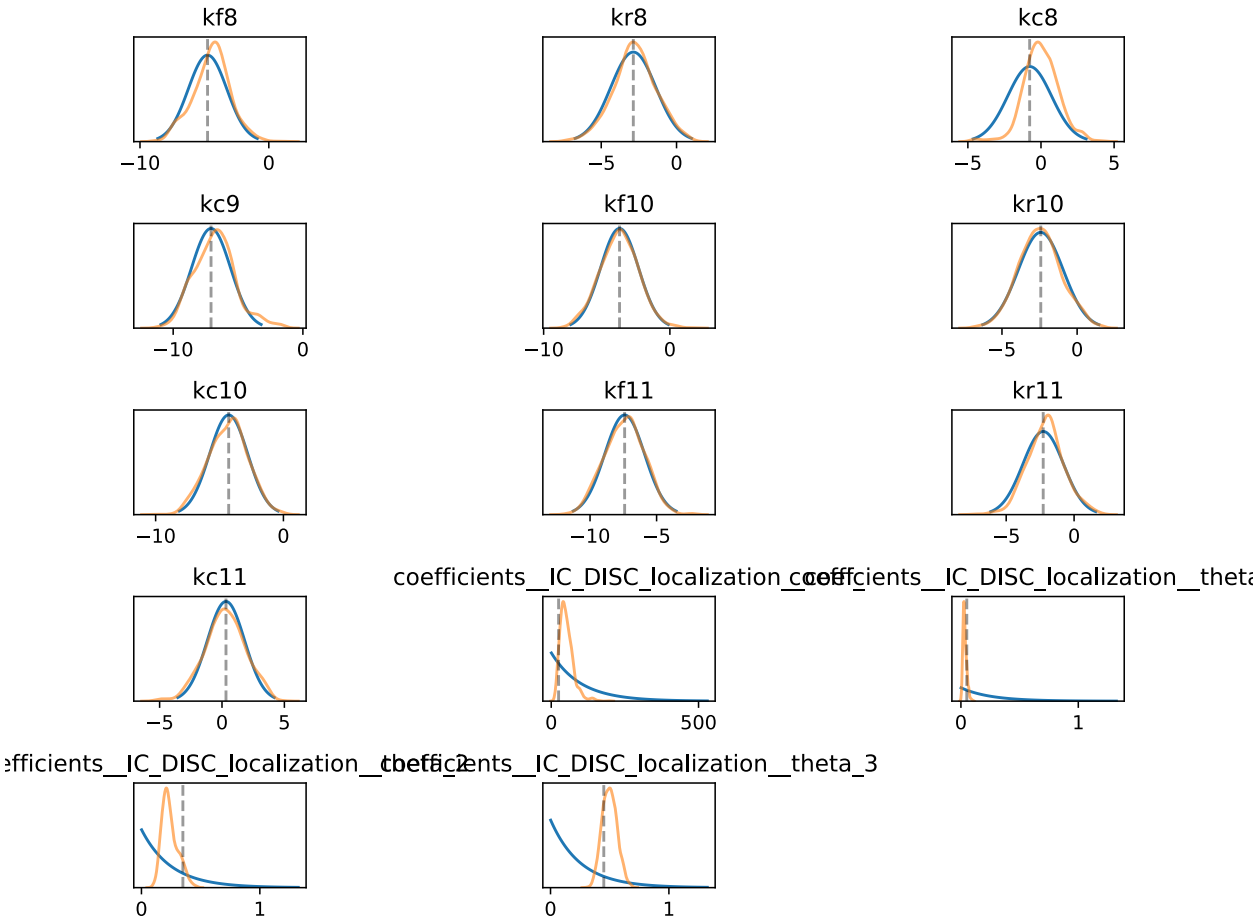


Figure B.10. (continued) Model parameters calibrated to Ordinal IC-DISC Dataset Parameters for aEARM were calibrated to ordinal values of IC-DISC abundance at every 300s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

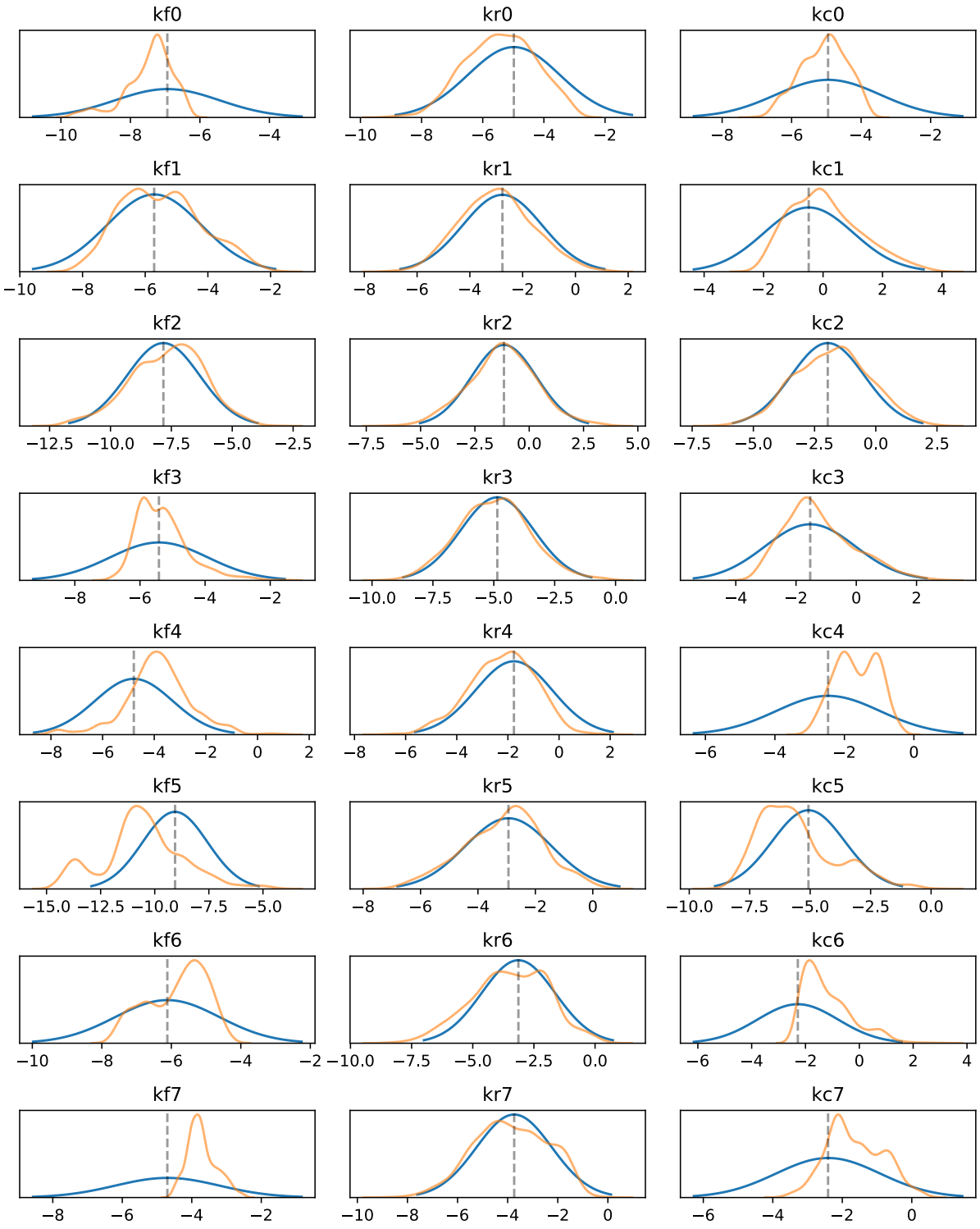


Figure B.11. Model parameters calibrated to Mixed Ordinal and Nominal Dataset Parameters for aEARM were calibrated to ordinal values of IC-DISC abundance at every 300s interval and to nominal observations of cell death vs survival. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown.

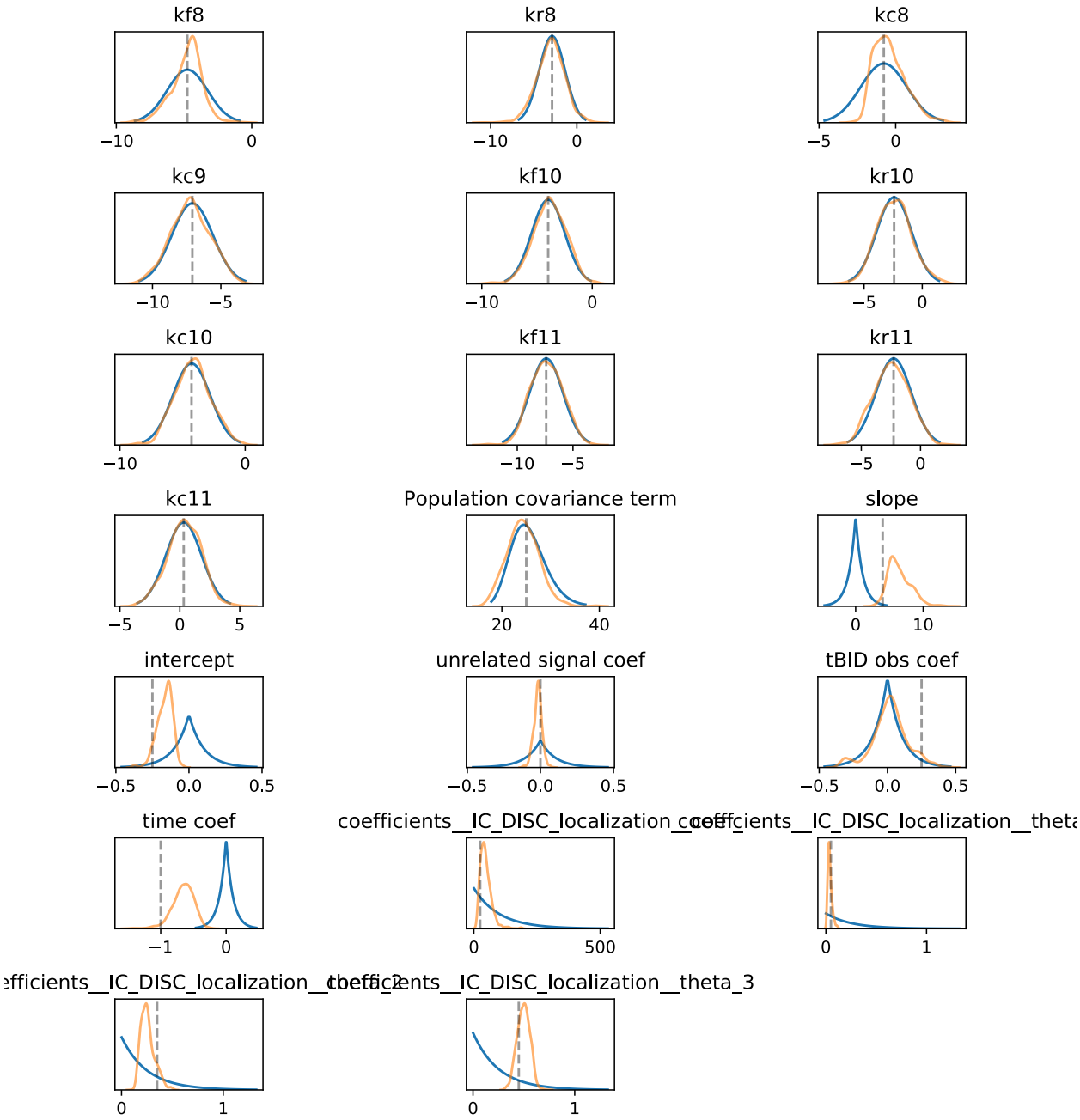


Figure B.11. (continued) Model parameters calibrated to Mixed Ordinal and Nominal Dataset Parameters for aEARM were calibrated to ordinal values of IC-DISC abundance at every 300s interval and to nominal observations of cell death vs survival. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Prior and posterior distributions of the value of measurement model coefficients are also shown (these are not log-scale).

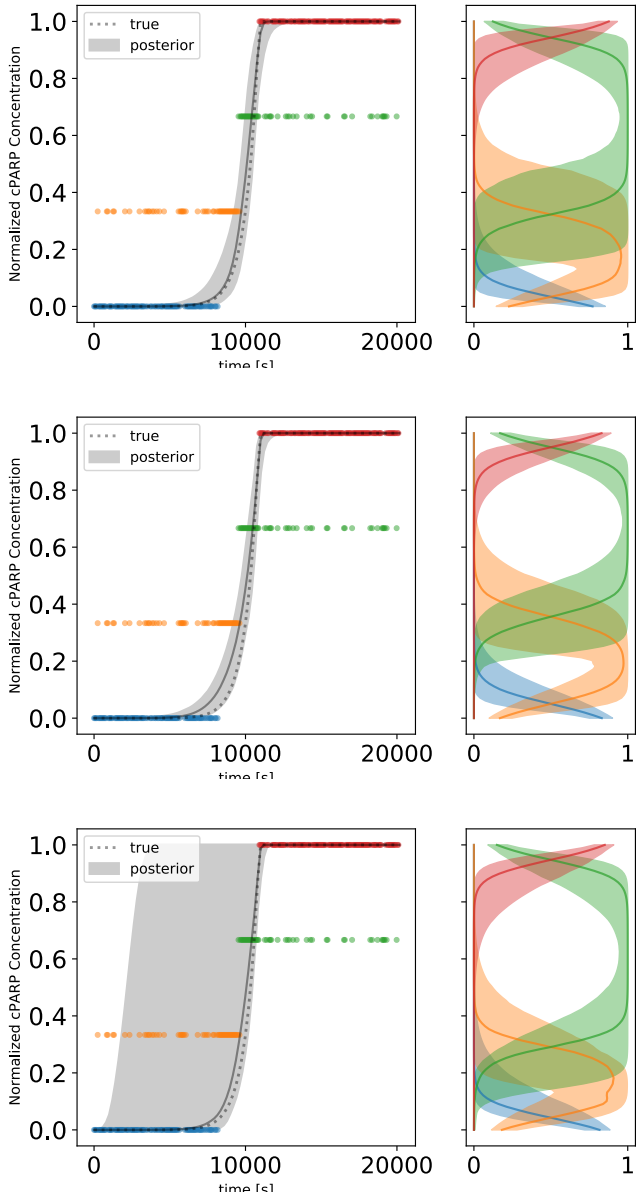


Figure B.12. Predicted PARP cleavage dynamics of aEARM trained to ordinal data using different measurement model priors.

The 95% credible region of posterior predictions (shaded region) of cPARP dynamics of aEARM calibrated to ordinal values of tBID and cPARP at every 60s interval. Uniform, Cauchy (scale=0.05) and Cauchy (scale=0.005) prior distributions for the parameterizations for the measurement model, respectively. In each, the median prediction (solid line) and true (dotted line) cPARP dynamic sare also shown. The adjacent panels give the 95% credible region of posterior predictions of the probability of class membership (x-axis) as a function of normalized cPARP concentration (y-axis).

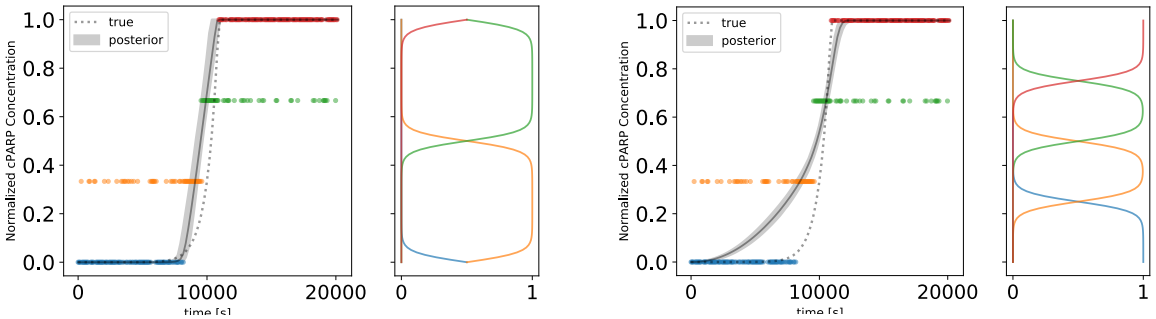


Figure B.13.Predicted PARP cleavage dynamics of aEARM trained to ordinal data using different *ad hoc* parameterizations of the measurement model. The 95% credible region of posterior predictions (shaded region) of cPARP dynamics for aEARM calibrated to ordinal measurements of tBID and cPARP at every 60s interval. Two fixed parameterizations for the measurement model. The adjacent panels plot the measurement models predicted probability of class membership (x-axis) as a function of normalized cPARP concentration (y-axis). The adjacent panels give the 95% credible region of posterior predictions of the probability of class membership (x-axis) as a function of normalized cPARP concentration (y-axis).

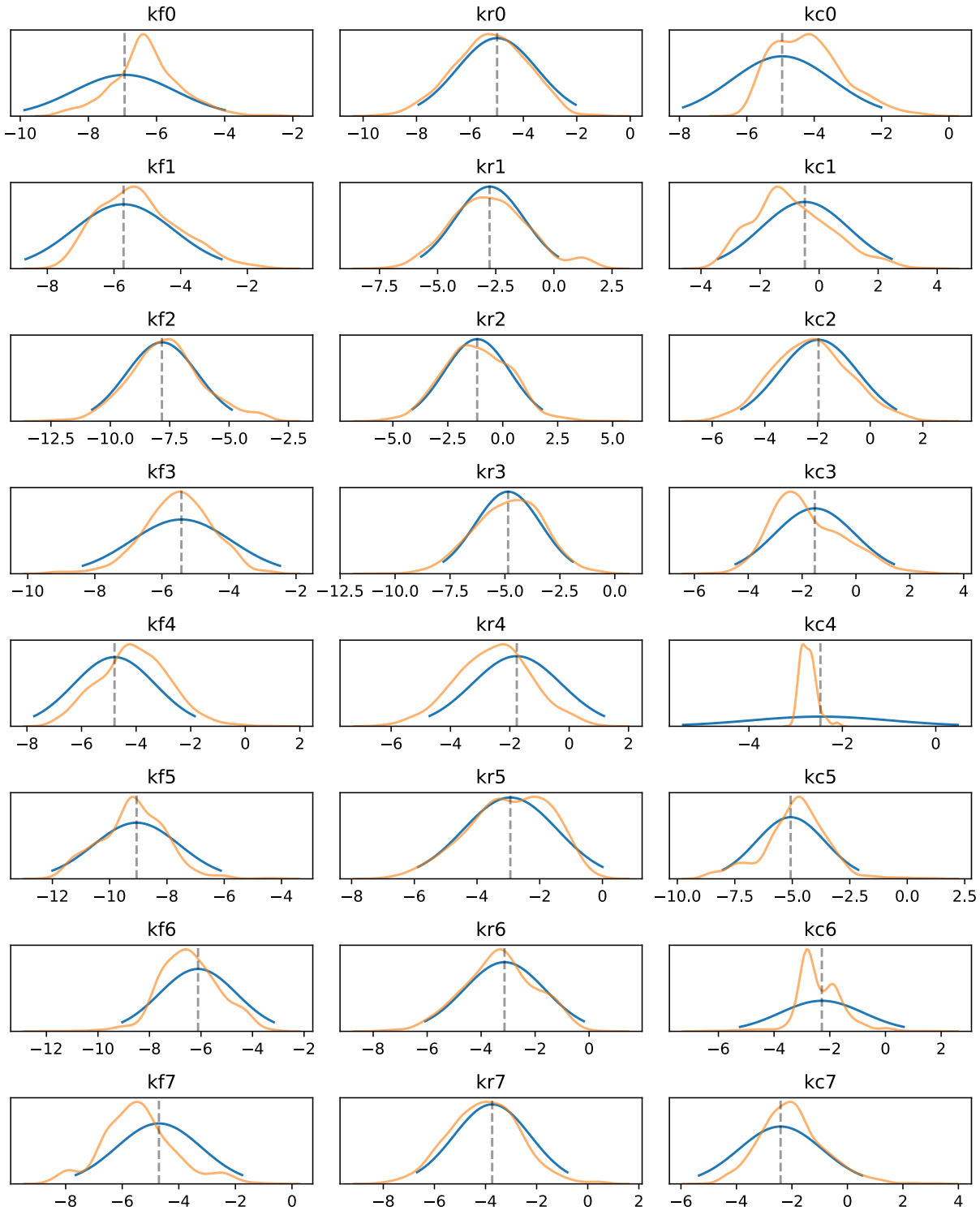


Figure B.14. Model parameters calibrated to Ordinal Dataset using Uniform priors on Measurement Model Parameters Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Uniform priors were placed on the measurement model parameters (these are not log-scale).

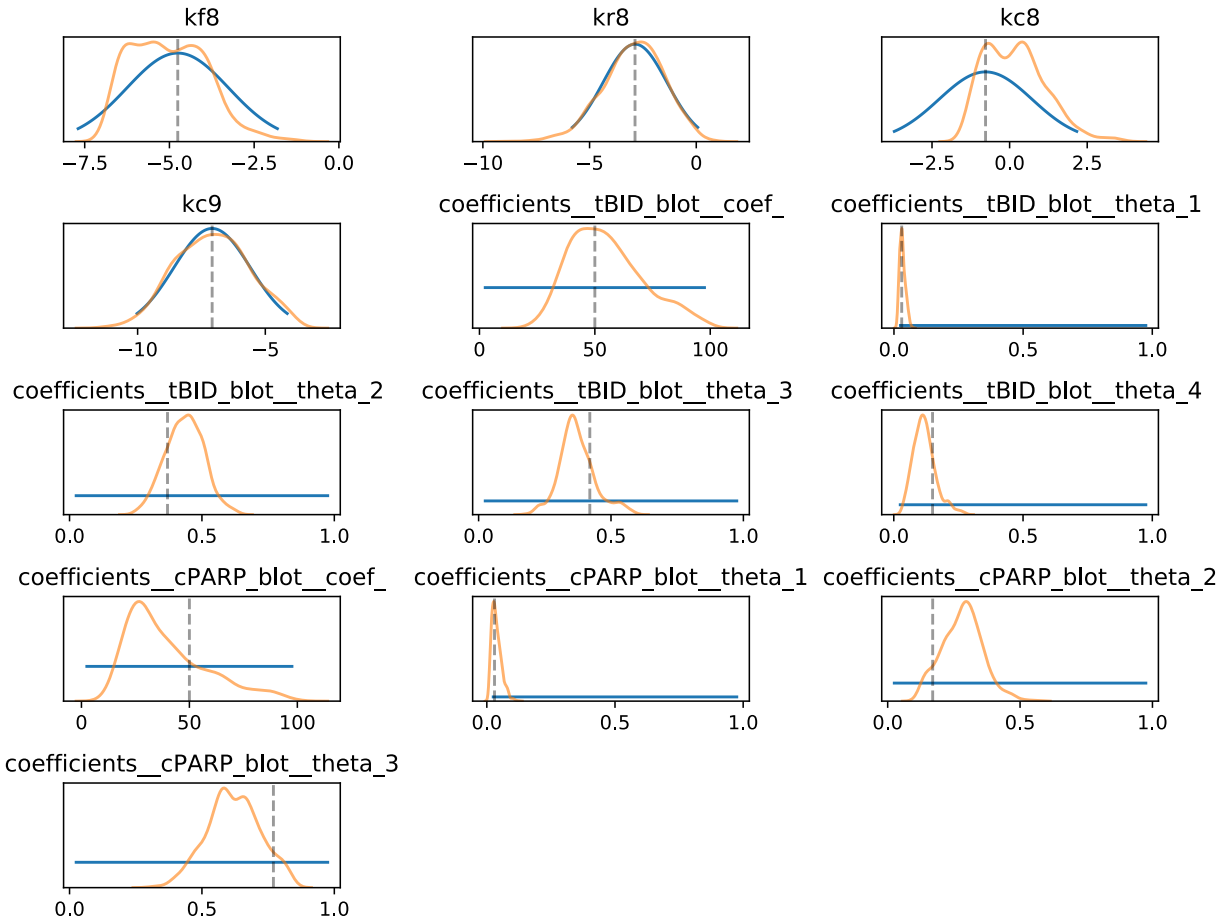


Figure B.14. (continued) Model parameters calibrated to Ordinal Dataset using Uniform priors on Measurement Model Parameters Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Uniform priors were placed on the measurement model parameters (these are not log-scale).

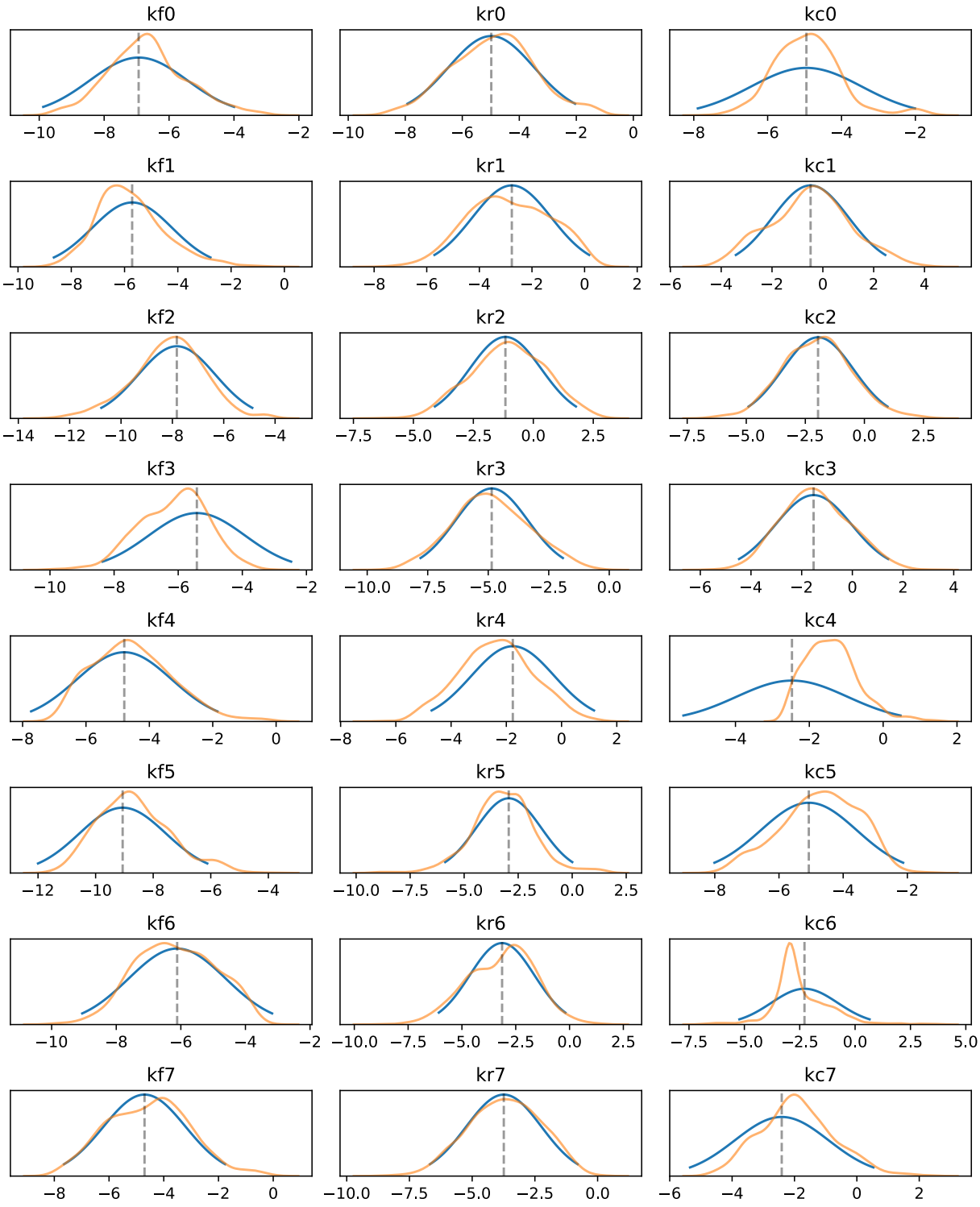


Figure B.15. Model parameters calibrated to Ordinal Dataset using Cauchy priors on Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Cauchy priors with a scale term of 0.05 were placed on the measurement model parameters (these are not log-scale).

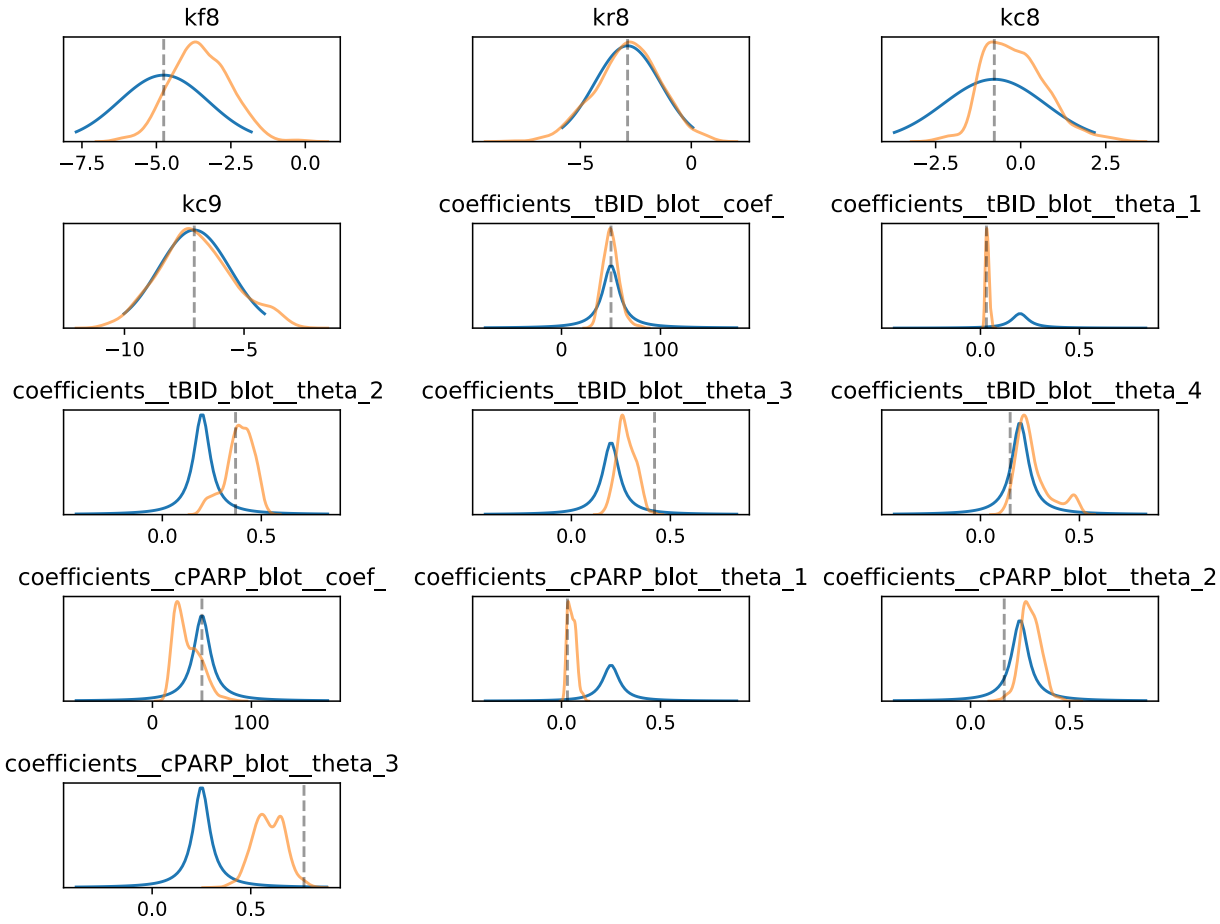


Figure B.15. (Continued) Model parameters calibrated to Ordinal Dataset using Cauchy priors on Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Cauchy priors with a scale term of 0.05 were placed on the measurement model parameters (these are not log-scale)

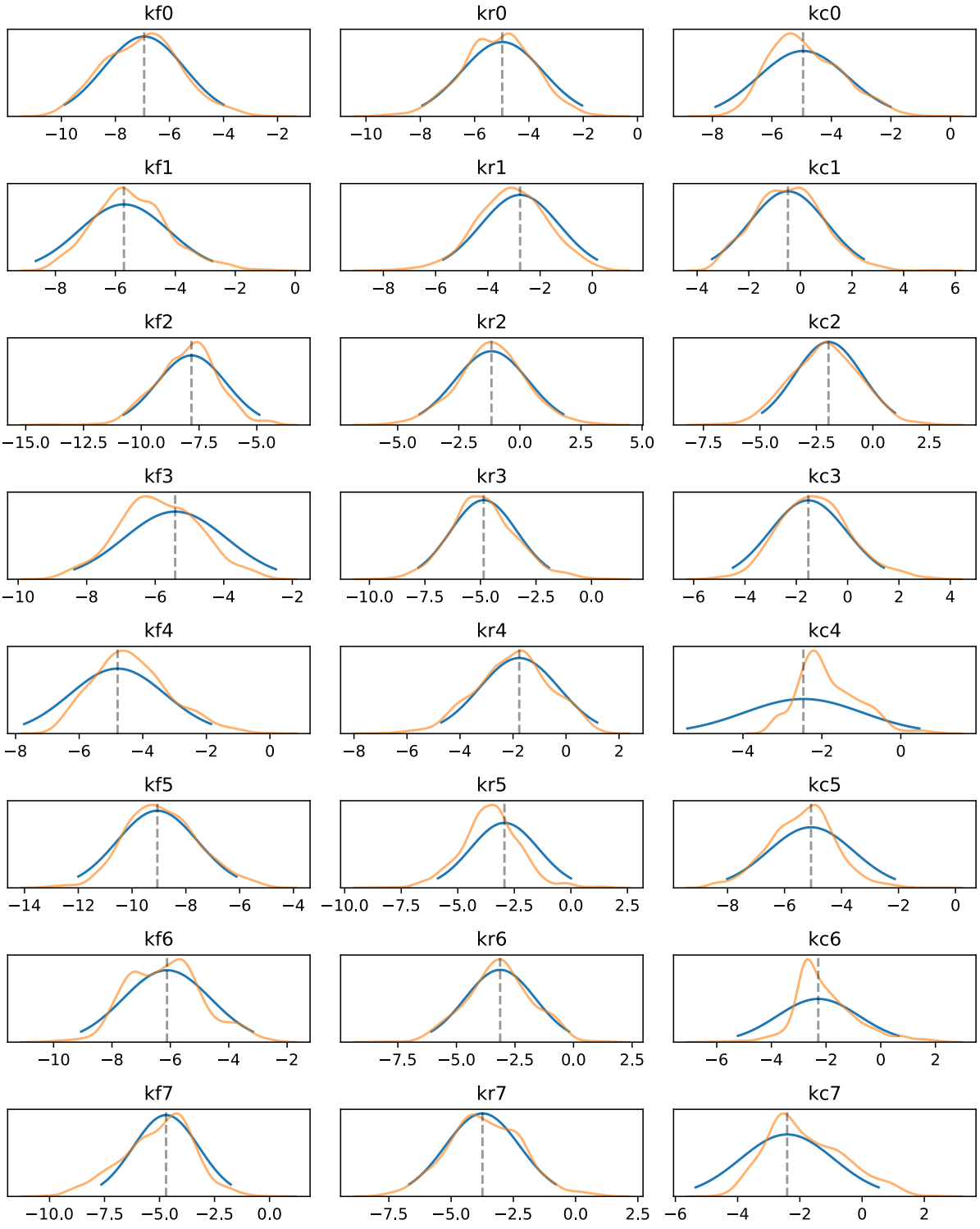


Figure B.16. Model parameters calibrated to Ordinal Dataset using Cauchy priors on Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Cauchy priors with a scale term of 0.005 were placed on the measurement model parameters (these are not log-scale).

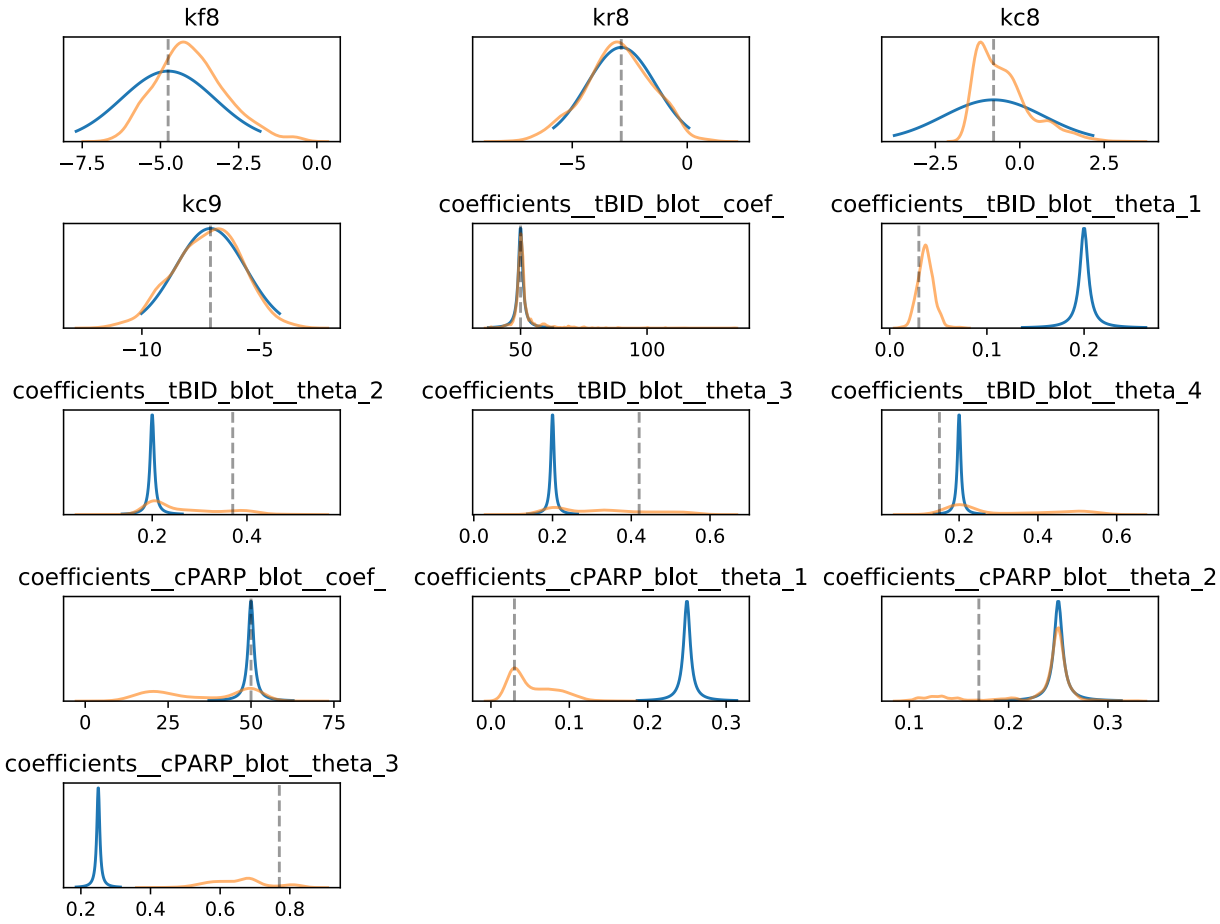


Figure B.16. (Continued) Model parameters calibrated to Ordinal Dataset using Cauchy priors on Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Cauchy priors with a scale term of 0.005 were placed on the measurement model parameters (these are not log-scale).

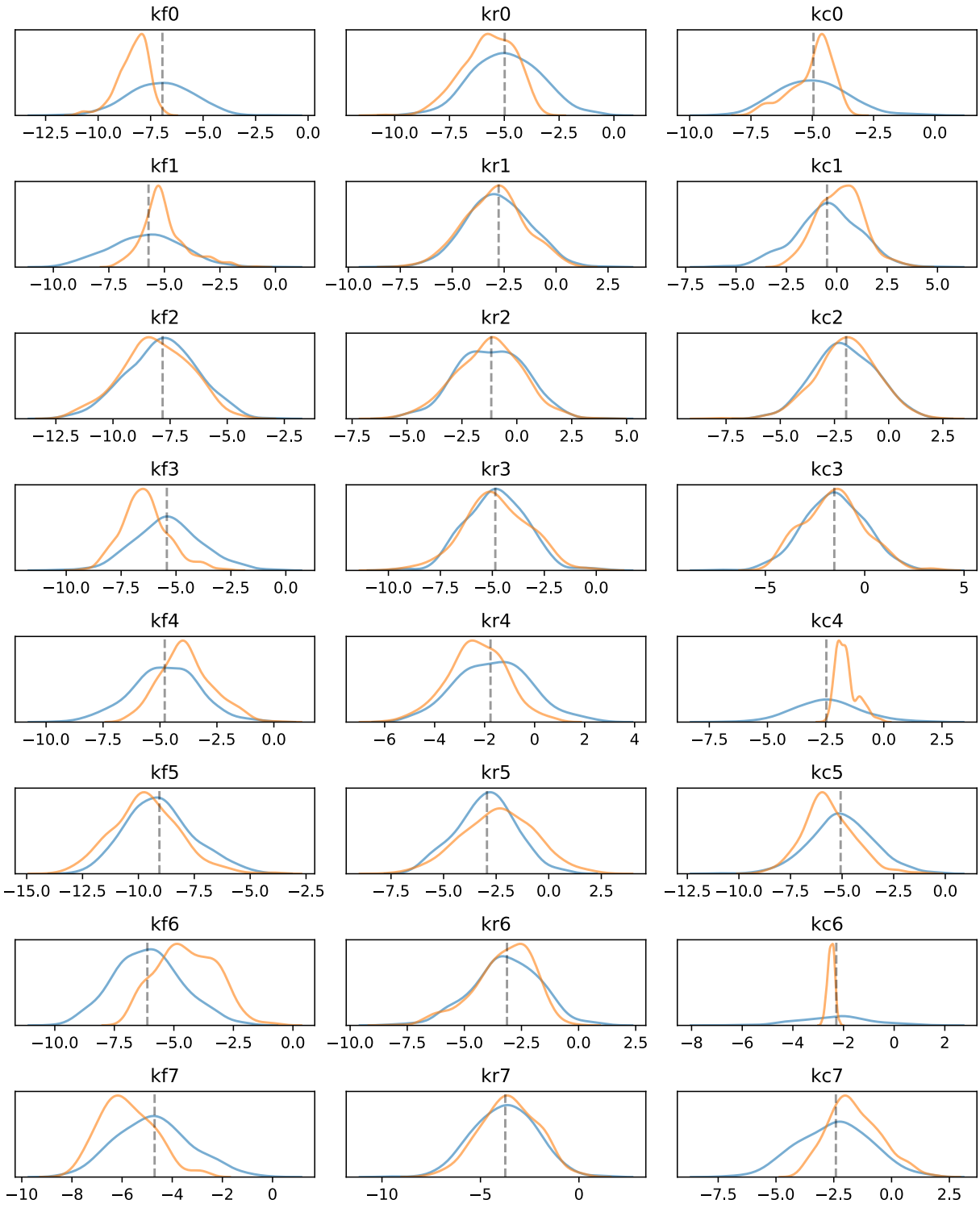


Figure B.17. Model parameters calibrated to Ordinal Dataset using Fixed *ad hoc* Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Fixed *ad hoc* values were applied to the measurement model parameters (see Table A.3. (Case 1))

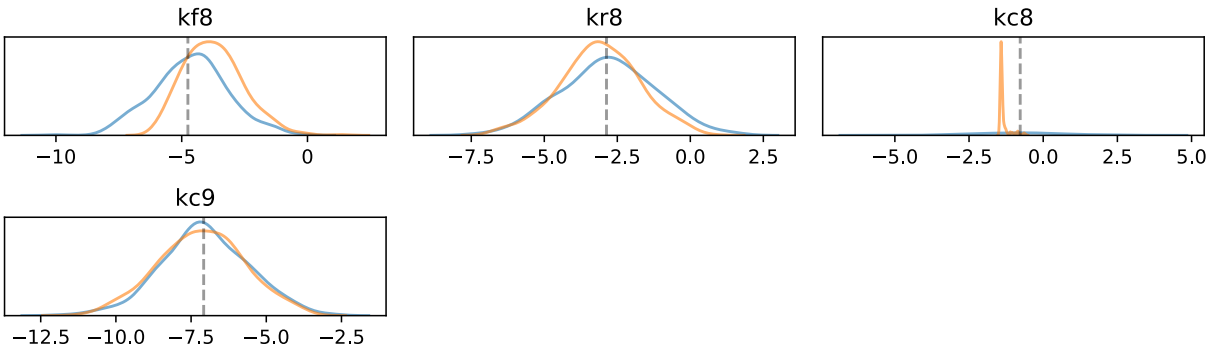


Figure B.17. (continued) Model parameters calibrated to Ordinal Dataset using Fixed *ad hoc* Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Fixed *ad hoc* values were applied to the measurement model parameters (see Table A.3. (Case 1))

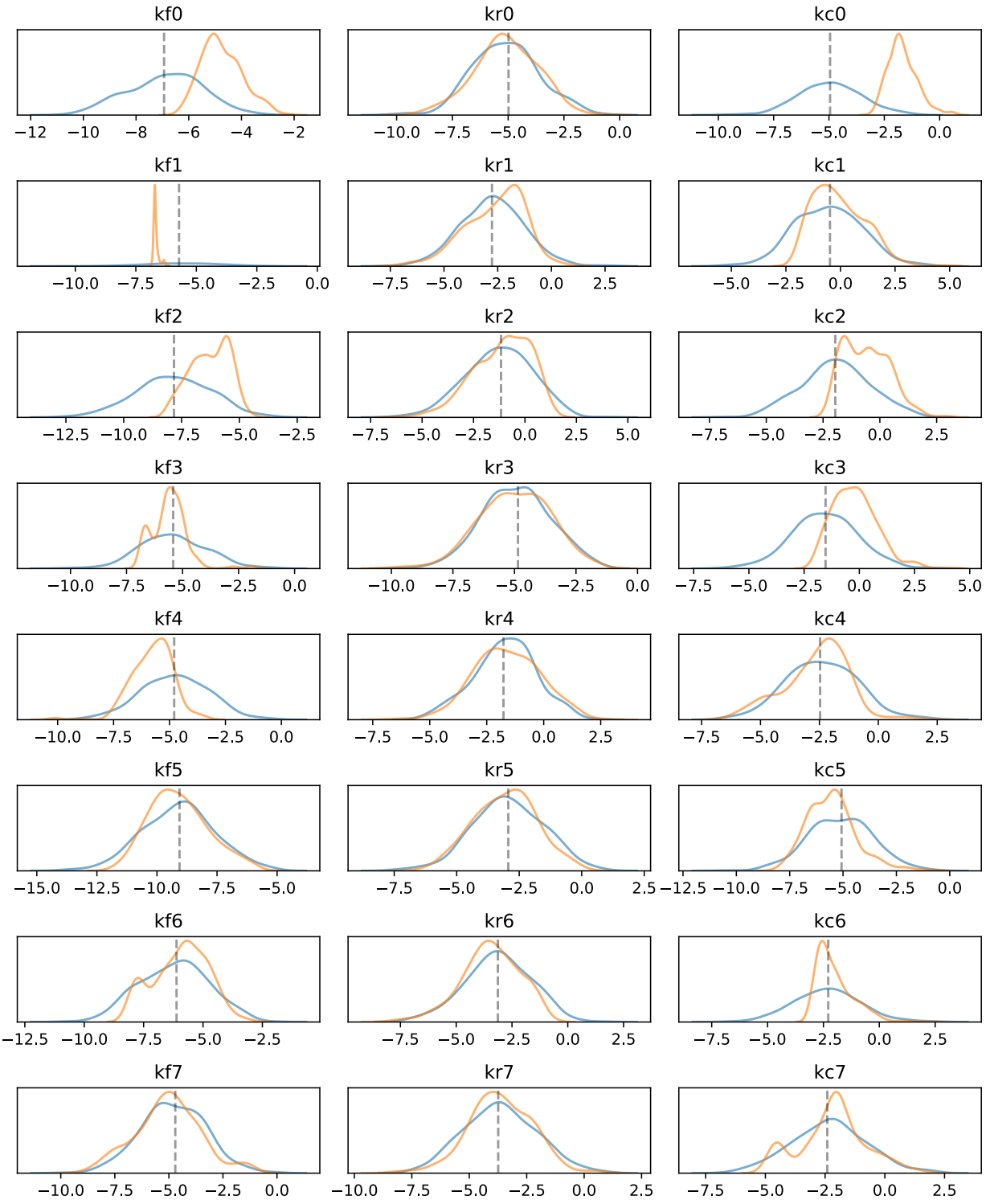


Figure B.18. Model parameters calibrated to Ordinal Dataset using Fixed *ad hoc* Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Fixed *ad hoc* values were applied to the measurement model parameters (see Table A.3. (Case 2))

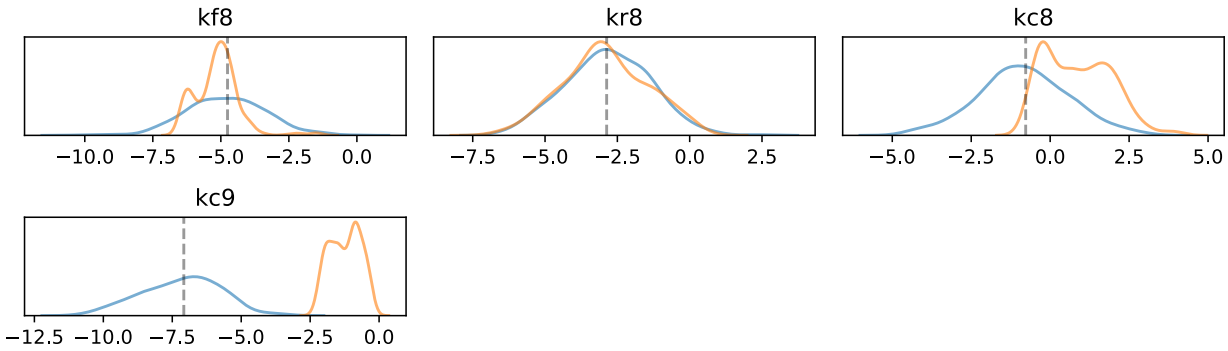


Figure B.18. (continued) Model parameters calibrated to Ordinal Dataset using Fixed *ad hoc* Measurement Model Parameters. Parameters for aEARM were calibrated to ordinal values of tBID and cPARP abundance at every 60s interval. Prior (blue) and posterior (orange) distributions $\log_{10}$ of the value of parameter are shown. Fixed *ad hoc* values were applied to the measurement model parameters (see Supplemental Table A.3. (Case 2))

C. Necrotic Cell Death in a Nanophysiometer

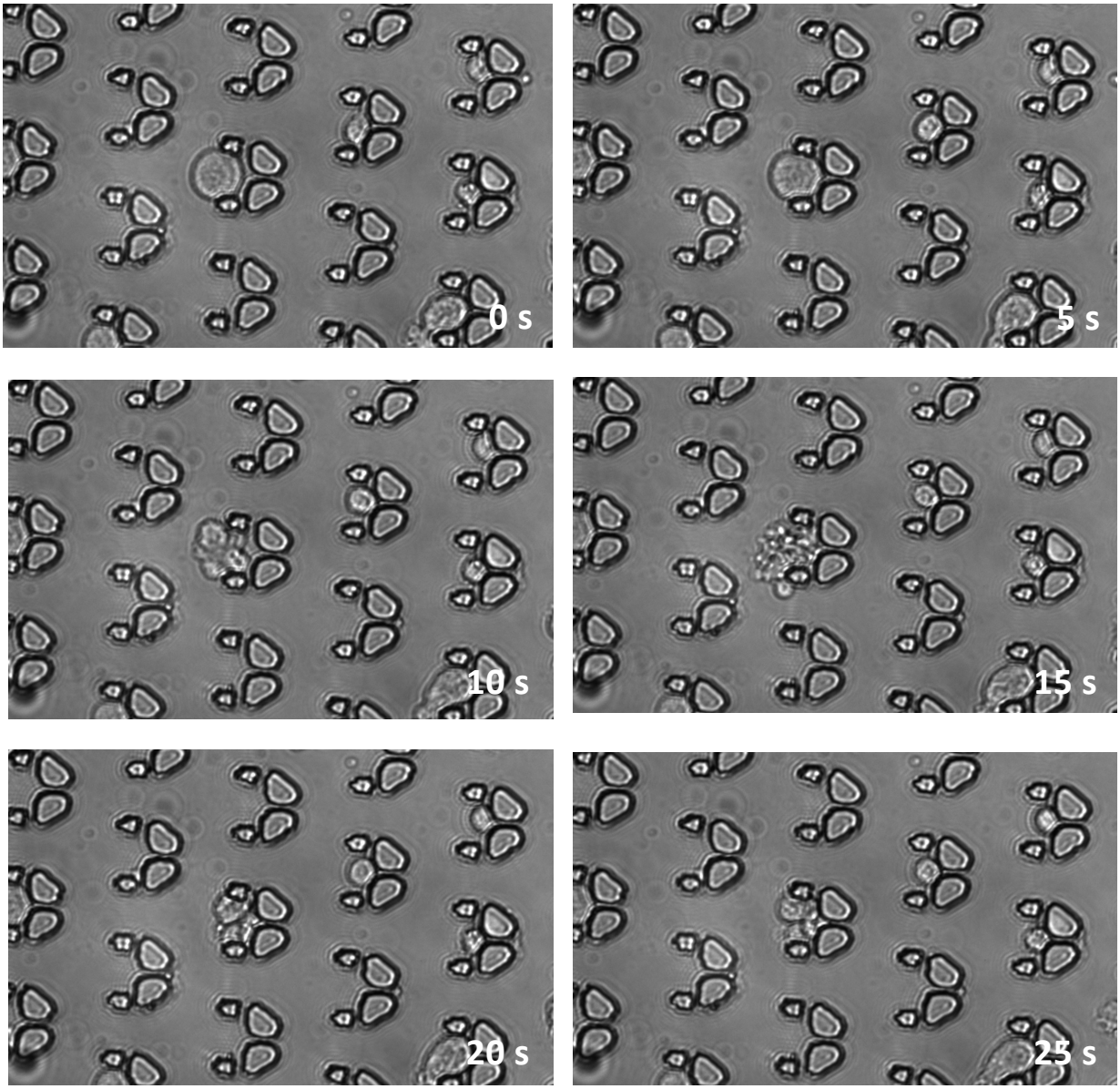


Figure C.1. Necrotic Cell Death Observation in a Nanophysiometer Necrotic cell death in L929 cells that were cultured in a nanophysiometer with 50ng TNF. The nanophysiometer is a microfluidic device[00] that can subject individual cells to unique and knowable experimental stimuli. By monitoring these cells we can generate the data necessary to support a measurement model of necrotic cell death outcomes.

D. List of Equations

$$y_{obs} \in \{0, 1, \ldots J\} \tag{1}$$

$$y_{obs,i} \neq y_{obs,j} \longrightarrow y_i \neq y_j$$

$$y_{obs} \in \{0, 1, \ldots J\} \tag{2}$$

$$y_{obs,i} \preccurlyeq y_{obs,j} \longrightarrow y_i \leq y_j$$

$$y_{obs} = f_M(y, \boldsymbol{c}) \tag{3}$$

$$y_{obs} = y \tag{4}$$

$$x(\boldsymbol{t}) = \boldsymbol{f}(\boldsymbol{t}, \boldsymbol{\theta}) \tag{5}$$

$$y(t_i, \boldsymbol{\theta}) = f_M\big(\boldsymbol{x}(t)\big) \tag{6}$$

$$d = \sum_{i=1}^{n} w_i\big(\hat{y}(t_i) - y(t_i, \boldsymbol{\theta})\big)^2 \tag{7}$$

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}) = c + \sum_{i=1}^{n} \frac{-1}{2\sigma_i{}^2}\big(\hat{y}(t_i) - y(t_i, \boldsymbol{\theta})\big)^2 \tag{8}$$

$$\{P(y_i = c_1|x_i(t)), P(y_i = c_2|x_i(t)), \ldots P(y_i = c_K|x_i(t))\} = f_M\big(\boldsymbol{\theta}_M, x_i(t)\big) \tag{9}$$

$$P(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\theta}_M) = \prod_{i=1}^{N} \sum_{j=1}^{K} P(\hat{y}_i|y_i = c_j) P\big(y_i = c_j\big|x_i(t)\big) \tag{10}$$

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}) = \sum_{1}^{N} \sum_{t}^{T} -1/2\sigma_i(t)^2 \times \big(\hat{y}_i(t) - y_i(t, \boldsymbol{\theta})\big)^2 \tag{11}$$

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\boldsymbol{i}}^{\boldsymbol{N}} \sum_{\boldsymbol{t}}^{\boldsymbol{T}} \log \sum_{j}^{K} P\big(\hat{y}_i(t)\big|y_i(t) = c_j\big) P\big(y_i(t) = c_j\big|x_i(t, \boldsymbol{\theta}), \alpha_i, \beta_{i,j}\big) \tag{12}$$

$$\log \mathcal{L}(\hat{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\boldsymbol{m}}^{\boldsymbol{M}} P(\hat{y}_m|y_m = c_1) \log \varphi \left(\alpha\big(\beta + \sum_{l}^{L} \beta_l x_{l,m}\big)\right)$$

$$+ \sum_{\boldsymbol{m}}^{\boldsymbol{M}} \big(1 - P(\hat{y}_m|y_m = c_1)\big) \log \left[1 - \varphi \left(\alpha\big(\beta + \sum_{l}^{L} \beta_l x_{l,m}\big)\right)\right] \tag{13}$$

$$K \leq 0.7 \times 2^{EBR}, \qquad EBR = -(SNR + 1.76)/6.02 \tag{14}$$

$$SNR = 20 \log_{10} q_{0.95} \mathrm{rms}(d)/(\max d - \min d) \tag{15}$$

$$\mathcal{L}(D|\, \theta, \varphi, \tau) = k \exp\left(-N\, KL\big(\hat{\pi}\big||\pi\big)\right) \tag{16}$$