A novel, rapid phenotypic assay for beta-lactam antibiotic susceptibility and an analysis of its theoretical limits

Thesis by
Eric Jer-Jiun Liaw

In Partial Fulfillment of the Requirements for the
Degree of
Biology

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended November 3, 2022

Eric Jer-Jiun Liaw
ORCID: 0000-0003-2244-8335

ACKNOWLEDGEMENTS

Thank you to Professor Rustem F. Ismagilov for welcoming me into his research group, securing financial support for this project, providing his scientific perspectives, successfully shouldering many responsibilities as our lab's principal investigator, and leading the lab (and campus) through the COVID-19 pandemic.

Thank you to Professors Richard M. Murray, Dianne K. Newman, and Long Cai for volunteering their time to serve on my thesis committee, providing mentorship, and supporting my degree progress. Thank you to Long Cai for advising me as a rotation student as well. Prof. Tom Chou at UCLA generously provided guidance on stochastic population modeling and Mathematica—you have my sincerest gratitude.

Many past and present members of the Ismagilov lab contributed technical advice, protocol validations, and equipment training that were invaluable to this thesis. Anna E. Romano was my primary sounding board during the long troubleshooting of filtration AST protocols. She also first suggested incubating lysis buffer to retrieve the lysate fraction, performed PCR experiments that I used, and kept the lab running smoothly as our lab manager. I am grateful for her kindness and dependability.

It was in productive conversations with Reid T. K. Akana during his lab rotation with me that we first derived the model for bulk filtration AST, augmented by his suggestion of using gamma distributed cell lifetimes. Reid and Alyssa M. Carter also kindly stepped in to perform PCRs when my pipetting workload was peaking. Sarah L. Simon confirmed the correlation of cell wall activity and filtration AST signal across antibiotic classes. I thank all of them for working with me.

Nathan G. Schoepp and Emily S. Savela had begun working on accessibility AST when I joined the lab and quickly brought me up to speed on lab procedures. Alex V. Winnett joined the project a year later, testing the method on clinical specimens in the field. I am grateful to all of them for the many hours we spent discussing each other's experiments as the pre-COVID AST subgroup. I am glad that our efforts successfully led to publications in PLoS Biology.

Natasha Shelby provided excellent feedback on every writing task and kept the lab together during tough times through her ebullient moral support, humor, and supreme organizational skills.

Eugenia M. Khorosheva, Emily Savela, and Mary V. Arrastia served as biosafety coordinators for the entire lab, while Justin Rolando, Sarah Simon, and Matt Cooper

took on the responsibility of chemical safety officers. Without their extra efforts to maintain the equipment and spaces that I and others relied upon, and to prepare biosafety SOPs for approval, this project would not be possible.

Justin C. Rolando, Said R. Bogatyrev, Roberta Poceviciute, and Octavio Mondragón-Palomino trained me on digital LAMP, microbiology, and microscopy protocols that helped advance my projects. Jacob T. Barlow, Jolena Jie Zhou, and Matt M. Cooper shared results about RapiDxFire activity in various conditions. Erik Jue gave advice on PCR troubleshooting, while Si Hyung Jin and Dmitriy V. Zhukov offered sanity checks on fabrication possibilities. Michael K. Porter was an outstanding tennis buddy and labmate.

I will cherish the interactions I've had with Roey Lazarovits, Ojas Pradhan, Joanne Lau, Tahmineh Khazaei, Natalie J. Wu-Woods, Minkyo Lee, Asher P. Steinberg, Matt Ratanapanichkich, Jenny Ji, Lealia Xiong, Tyler Nguyen, Heli Takko, Sohee Lee, and Mi Kyung Kim. Beyond the highlights I've listed, there are innumerable ways in which all lab members have contributed positively to my time at Caltech. I could not imagine more friendly and fun co-workers.

Dr. Omai B. Garner provided clinical isolates and urine samples from the UCLA Microbiology Lab and eagerly shared his extensive knowledge of clinical microbiology in practice. Enrica Bruno and her team worked diligently to file two patents on accessibility AST on behalf of me and the other inventors. I would like to thank Prof. Victoria J. Orphan for her support for my CEMI pilot grant application on bacteriophages. Without her sponsorship of my research guest status in the first half of 2023, finishing this thesis would have been difficult. I would also like to thank Lauriane Quenee for her expert input as the Caltech campus biosafety officer and Prof. Jared R. Leadbetter for pointers into the microbiology literature.

I would like to thank the Medical Scientist Training Program directors Drs. Olujimi Ajijola, David Dawson, Maureen Su, and Mitch Guttman for advocating on my behalf and building a sense of community within the MSTP. I would also like to thank Elizabeth Ayala, Susie A. Esquivel, Josephine Alviar, and the other MSTP office staff at UCLA and Caltech for their work.

I thank Victoria Sun, Katherine M. Sheu, Prashant Bhat, Shannon R. Esswein, Masami Hazu, Aditi Narayanan, Rory L. Williams, Rui Cheng, and many other classmates and friends at UCLA and Caltech for their support, feedback, and companionship.

Lastly, I wish to express my deepest gratitude to my parents, sister, and brother-in-law for their tremendous support, sacrifices, and encouragement during my academic journey and beyond.

This work was funded by:

- Defense Threat Reduction Agency (DTRA) award MCDC-18-01-01-007; an effort sponsored by the U.S. Government under Other Transaction number W15QKN-16-9-1002 between the MCDC, and the Government.
- Burroughs Wellcome Fund Innovation in Regulatory Science Award (BWF.1014981 awarded to R.F.I.)
- Caltech's Center for Environmental Microbial Interactions (CEMI) award in 2020
- Caltech's Jacobs Institute for Molecular Engineering for Medicine
- Caltech's Biology & Biological Engineering Division
- A fellowship from Joan and Jerry Doren
- National Institutes of Health award T32 GM008042; UCLA-Caltech Medical Scientist Training Program

ABSTRACT

Current management of bacterial infections is limited by the slow turnaround time of culture-based antibiotic susceptibility testing (AST). Culture-free phenotypic AST methods, though faster, are limited not only by analytical sensitivity but also by the low number, density, and purity of live pathogens present in clinical specimens before culturing. Separating and concentrating pathogens from clinical specimen matrices and improving the analytic sensitivity of phenotypic measurement technologies remain active areas of research. However, to date, the literature lacks consensus over what is a reasonable goal for the minimum number of pathogens in a clinical specimen needed to accurately perform phenotypic AST.

I describe "bulk filtration AST" and "digital filtration AST," two new filtration-based AST methods that improve an AST method previously published by others and myself. These methods use nucleic acid quantification to assess the activity of antibiotic classes (and only those classes) targeting peptidoglycan turnover, specifically the beta-lactams, which are the most frequently prescribed class of antibiotics. I use filtration AST to quantify the *in vitro* pharmacodynamics of beta-lactam antibiotics over time scales shorter than two hours, and I simultaneously validate the methods' accuracies on clinical isolates of Enterobacteriaceae. To analyze filtration AST results, either for fitting parameter values or for predicting susceptibility, I derive probabilistic models for the outcomes of each of the two filtration AST methods, then perform Bayesian parameter inference from my data.

I then propose a general mathematical framework for defining the concepts of the phenotypic assay and the ideal phenotypic assay. Within this framework, I calculate the ideal filtration AST performance as a function of the number of cells assayed, my fitted pharmacodynamic parameters, and other variables. Interestingly, the observed performance of my implementation of digital filtration AST is consistent with the implementation's approaching the ideal performance. I hope my demonstration of these new methods and my theoretical framework will help guide future research into rapid phenotypic AST.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Nathan G. Schoepp, Eric J. Liaw, Alexander Winnett, Emily S. Savela, Omai B. Garner, and Rustem F. Ismagilov. Differential DNA accessibility to polymerase enables 30-minute phenotypic β-lactam antibiotic susceptibility testing of carbapenem-resistant Enterobacteriaceae. *PLOS Biology*, 18(3):e3000652, March 2020. ISSN 1545-7885. doi: 10.1371/journal.pbio.3000652. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000652.

NGS, EJL, AW, ESS, and RFI contributed to conceiving the method, revising the manuscript, and interpretation of experimental results. NGS developed the sample handling workflow and performed all experiments for comparison of amplification methods, validation, and timed sample-to-answer experiments. NGS was the major contributor to manuscript preparation and prepared all figures. EJL performed filtration experiments, reviewed relevant medical literature, and contributed to manuscript writing. NGS and AW tested clinical samples using the modified workflow. ESS performed early experimental work to link beta-lactam exposure to differential nucleic acid readout, analyzed data from validation experiments, and developed TTPD metrics. OBG provided clinical guidance on the selection of clinical isolates and clinical samples and coordinated and provided oversight of clinical-sample collection at UCLA, including technical assistance to UCLA staff. RFI supervised and guided the project, and helped compose the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

BACKGROUND AND CONTRIBUTIONS

1.1 Open challenges in antibiotic susceptibility testing

1.1.1 The clinical need for rapid AST and phenotypic AST

Bacterial infections are common afflictions or complications in virtually every patient population around the world [1]. Antimicrobial susceptibility testing (AST), also often referred to as "culture and sensitivity," is an indispensable diagnostic procedure that guides modern management of bacterial infections. AST informs clinicians whether the causative bacterium is susceptible or resistant to a panel of antibiotics. Knowledge of susceptibility allows clinicians to choose an antibiotic with assured effectiveness, the narrowest range of target organisms, and the fewest adverse effects. However, clinicians today must rely on slow culture-based phenotypic AST assays for definitive answers. It takes a median of 11–12 hours [2, 3] for a typical venous blood draw to show positive bacterial growth by today's automated culture systems, another day (about 24 hours) for subculture on solid media to isolate the pure microorganisms, and then 18–24 more hours of culture to perform the AST itself [4]. The total sample-to-answer time for phenotypic AST in the USA was found to last a median of 2.57 to 3.1 days, depending on species (overall median 2.76 and interquartile range 2.49–3.12), in a 2018 study [5].

The long delay in susceptibility information necessitates empiric antibiotic regimens[1], in which clinicians choose antibiotics by prior experience or guidelines, but without precise knowledge of the cause of the suspected infection. Empiric therapy creates several problems. First, although clinicians choose broader-spectrum antibiotics to maximize the chance of covering the unknown pathogen, the empiric regimen will fail if a resistant organism is present. Second, the uncertainty of empiric therapy has prematurely eliminated some older antibiotics in outpatient settings where AST is too slow to be performed. When the community rate of resistance against a drug has become non-negligible (e.g., >5%), an unacceptable number of patients would be harmed by treatment failure if given that drug empirically, so no

[1] Empiric regimens are prescribed and administered based on prior experience instead of precise knowledge of the suspected infection's cause. The choice of antibiotics must contain sufficient antibiotic coverage of all reasonably likely organisms that match the patient's presentation and the levels of resistance recently observed in the community.

patients receive the drug even though most patients would still benefit. Third, several classes of antibiotics have predictable or frequent adverse affects that may harm the patient. However, if the chance is high enough that safer alternatives will not cover the reasonably expected diversity of resistant taxa, then the benefit of treating outweighs the risk. The riskier drug must still be given until proven extraneous. Lastly, excessive use of first-line antibiotics has been shown to be accelerating the emergence via natural selection of resistance in target pathogens, creating what public health agencies nearly unanimously have identified as an urgent global crisis [6].

Faster AST would reduce the duration of empiric therapy and thus mitigate all the above quandaries. Inpatient providers can switch to targeted monotherapy earlier, perhaps within the same hospital shift, simplifying treatment plans and slowing the spread of resistance. Outpatient providers may be enabled to wait until AST results return before sending prescriptions to patients seen earlier. In an ideal world, the turnaround time for AST would be shorter than the length of time it takes to physically prepare and administer intravenous antibiotics. Ultimately, a point-of-care device that returns results within 30 minutes would have a chance of eliminating empiric therapy altogether—but the possibility of such a device, at least for phenotypic AST, is questionable and one of this thesis's targets of inquiry.

AST methods can be classified as genotypic or phenotypic. Phenotypic tests measure bacterial response directly regardless of a strain's resistance mechanism. The current phenotypic gold standard methods of disk-diffusion and broth microdilution read population growth after 16–24 hours of incubation [4]. They also require pure isolates, which entails culturing in liquid media and subculturing on solid media for at least 2 days [5]. Phenotypic tests are cost effective and yield definitive answers. Their main drawback is their slow turnaround time.

Genotypic tests (e.g., nucleic acid amplification tests (NAAT)) detect the presence of known resistance genes, return results within hours without any culturing, and have become standard of care for a number of pathogens. Unfortunately, genotypic tests have failed to replace culture-based AST for most bacterial pathogens because they measure resistance indirectly and in most cases can only rule in, not rule out, resistance. Therefore, empiric antibiotics would still be continued even if a genotypic test for resistance returns negative. Only a rapid phenotypic test would be able to eliminate an antibiotic from a patient's regimen. Genotypic test use has mainly been limited to surveillance for resistance at large medical centers.

An exception is the genotypic test for the SCCmec cassette in *Staphlococcus aureus*. The SCCmec/mecA/mecC cassette is highly correlated with the methicillin-resistant (MRSA) phenotype and therefore can eliminate the use of empiric vancomycin in inpatients with suspected sepsis when the local MRSA prevalence is not too high. Unfortunately, no genotypic test exists for the most common Gram-negative pathogens. Yet, Gram-negative pathogens incur the greatest morbidity and mortality among the antibiotic-resistant pathogenic bacteria in bloodstream infections in developed regions [1, 6].

1.1.2 Obstacles to rapid phenotypic AST methods

The fundamental obstacle in speeding up phenotypic AST is the low signal-to-noise ratio (SNR) that can be elicited out of many types of clinical specimens. There are four specific obstacles that create both the low AST signal and high background of these clinical specimens. These four obstacles are also four tasks which any phenotypic AST must overcome to perform accurately on the breadth of clinical specimens on which it is deployed. First, a low density of bacteria may be present in the specimen. Since any phenotypic AST signal must arise proportionally from the bacteria assayed, at least in the AST methods envisioned to date, lower densities of bacteria imply lower signals. Secondly, the technology used to measure bacterial phenotypes will have a certain analytical sensitivity. Low analytical sensitivities will cause low signal. Thirdly, the non-bacterial components of the specimen, mostly human tissue, present a large mass of non-bacterial cells, many chemical species, a complex physical structure, and variability in their chemical composition due to the host's reactive physiology. This host tissue will contribute background signal. If the host tissue alters the bacterial phenotype, its unpredictable nature adds uncertainty to the signal. Lastly, as asserted by this thesis, if low numbers of bacteria are present, then the noise from biological stochasticity will obscure the signal from AST assays.

The traditional AST workflow overcomes each of these obstacles through the culturing of bacteria. In the first step of the workflow, a blood culture bottle containing growth media is cultured until bacteria growth is detected by continuous monitoring of gases produced by microbial metabolism. Each bottle is usually part of a set of bottles drawn at the same patient encounter. Allowing bacteria to replicate increases the number of bacteria and increases the amount of bacteria-specific signal from bacteria per unit volume relative to the non-replicating host tissue. At this stage, the background from hosts tissues is high enough that AST methods are not currently performed, though this is an area of current research [7]. The contents of

blood vary widely depending on the patient's physiology (e.g., antibiotics that were possibly administered, white blood cell counts, lipoprotein content, liver or kidney dysfunction) and contain thousands of different chemical species. Furthermore, about 17% of blood cultures in a clinical lab (across all conditions) contain more than one organism [5]. In the past, up to 25–31% [8, 9] of all polymicrobial cultures contained non-informative contaminants. For hospitalized inpatients in the USA, it is generally considered acceptable for a healthcare facility to have a contamination rate of 3% [4].

The next step of the workflow is typically subculturing to obtain pure isolates. It is during the act of streaking onto solid agar that host blood is diluted and physically separated from the bacteria. As the bacteria replicate into a solid mass from single colony-forming units (CFU), the separation is magnified to virtual completion, and the density of bacteria per unit volume is maximized. It is at this high signal-to-noise ratio that many analytical assays become usable, such as species identification by MALDI-TOF mass spectrometry and, of course, phenotypic AST. For broth microdilution specifically, the CLSI standard protocol requires a cell density of 0.5 McFarland, which is approximately 5×10^5 CFU/mL.

During the third step of the workflow, AST is performed. The gold standard measurement methods of agar dilution, disk/gradient diffusion, and broth macro/microdilution all rely on visual observation by certified technicians of the turbidity of liquid cultures or the opacity of bacterial lawns. The exponentially-increasing signal from the cultured bacteria, and the non-increasing signal of the growth media, is what eventually provides sufficient signal-to-noise ratios for the naked eye.

Although culturing very effectively overcomes the four necessary obstacles to phenotypic AST, it takes up the largest proportion of the turnaround time for traditional AST. This turnaround time is too long, as discussed in section 1.1.1. To speed up phenotypic AST, the four tasks performed by culturing need to be optimized, such as by automation, or to be performed by replacement techniques. (Automation will not be further discussed because improvement in that aspect generally involves mechanical and electrical engineering rather than knowledge of biology.)

The task for which humanity has made the most progress has been the substitution of the phenotypic readout technology. The current gold standard protocols require that the antibiotic exposures be incubated for 16–20 hours (for broth microdilution) or 16–18 hours (disk diffusion), and up to 24 hours for certain antibiotic and taxa pairings, to ensure that no positive growth is missed [4]. This is a comparatively

long turnaround time for assays in clinical chemistry. Newer automated systems are much faster than traditional assays, though generally at a higher cost. Besides cost, a major benefit of using disk diameter or turbidity as the detected system state variable is that a large body of data already exists for these traditional methods, simplifying adoption of new devices. Any other form of readout must demonstrate strong correlation to population growth or build up a substantial body of survey data in order to displace the current usage of population size and the MIC. This benefit is more economical and historical than scientific. End-point growth is not the most fundamental, uniform, or clinically relevant metric; rather, it was the most clinically-relevant and reproducible pharmacodynamic metric that could be measured with the easy and cheap gold standard assays at the time antibiotics were discovered. The guidelines for interpreting of disk diameter or minimal inhibitory concentration (MIC) themselves contain many exceptional cases to account for atypical behaviors in each taxa.

At least nine commercial automated systems have been developed so far that shorten the readout of growth to 4–8 hours, or as short as 3 hours in some cases. A review of these platforms can be found in Jacobs et al., 2021 [10]. The measurement technologies employed by these commercial offerings include conventional methods (microdilution, turbidometry, and disk diffusion) as well as newer modalities (laser nephelometry, live darkfield microscopy, volatile small molecule sensors). Additional readout methods have been proposed in the literature. These include improved photometric measurements (single-wavelength, non-spatial, and including colorimetric and fluorimetric assays requiring exogenous chemicals); Raman and other electromagnetic spectroscopy; other forms of live microscopy; electrochemical signals by electrodes; mass measurement by cantilever balances; atomic force microscopy; laser speckling patterns (for solid colonies); proteomic mass spectrometry; and nucleic acid quantification, often benefiting from microfluidic devices [11].

The task of host tissue separation has also been the subject of considerable research by engineers. Particular attention has been placed on separating bacteria from whole blood because of blood's clinical importance. Techniques such as centrifugation/ sedimentation, filtration, selective lysis of human cells by saponin, chemical binding (by antibodies or lectin), field-flow fractionation, capillary zone electrophoresis, dielectrophoresis, acoustophoresis, automatic cell sorting, and other microfluidic phenomena (e.g., hydrodynamic focusing, droplet sorting, sieves) have been at-

tempted in the literature, as reviewed in Pitt et al., 2016 [12]. However, many of the reported methods suffer from drawbacks in overcoming the high density of host cells and proteins in blood compared to other bodily fluids. Of the nine commercialized rapid phenotypic systems, only two—the Accelerate Pheno and the Q-linea ASTar—have been approved for use directly from positive blood cultures [10, 13]. The Accelerate Pheno device accomplishes host tissue removal through selective lysis of host cells followed by filtration by gel chromatography, but the exact methods of this device and the Q-linea ASTar device are trade secrets.

To address the obstacle of low bacterial spatial density, one may increase the density of bacteria by increasing the number of cells (e.g., culturing) or concentrating them into a smaller volume. The density is the important physical quantity because the background signal from most measurement sensors scales in proportion with the amount of space (or time) from which the sensor acquires signal. This statement is true for photodetectors, images (since the probability of imaging artifacts increases with the field of view), and Oftentimes, the signal measured by the measurement device is also proportional to the density of the analytes detected. This statement is true for colorimetric and electrochemical sensors. Of note, instead of physically altering the spatial density of cells, one can restrict the volume of space and time covered by one's measurement to the sparse cells in a dilute specimen, raising the effective density of the signal-generating bacteria [14, 15]. To do so, one must be able to analyze the discrete particles of the analyte, which is easier for cells than for molecules. In the literature, the task of increasing bacterial density without culturing has generally been attempted while attempting to separate host tissue. However, one of the reasons that some separation technologies have failed to progress to the clinic is that the separation dilutes bacteria to even lower densities, even though host cells density drops by many magnitudes more. Perhaps subsequent processing to increase (e.g., centrifugation) then causes the loss of too many cells.

The last obstacle of low cell numbers, regardless of density, has not received significant attention in the literature, though it has undoubtedly been encountered, and maybe recognized, in fields such as circulating tumor cell detection or single-cell AST.

Unlike the density of cells, which can be increased arbitrarily until one is making measurements from single cells in comparably-sized volumes, the absolute number of cells can only be increased by further collection or by culturing. During blood donation, it is widely accepted that no more than 100 mL of blood per 1 kg of

body weight, or 13% of an adult human's total circulating blood volume, can be removed within an 8 week window without risking symptoms of anemia [16]. For very young children, no more than 1% of the total circulating blood should be drawn [4]. The recommended volume per day of blood for blood cultures is 40 mL, or 4 sets of bottles, though there is evidence that larger volumes would improve yields [4]. Larger volumes would create more labor for laboratory staff, however. There is a practical limit to the amount of blood that can be collected for blood cultures, and therefore a limit to the numbers of cells that can be collected. To date, no one has demonstrated in humans the safe return of blood to the patient after the collection of bacteria, though the idea was once explored in rodents [17].

If large numbers of cells were collected by blood draws in standard practice, then the absolute number of cells would not be limiting to phenotypic AST assays. However, this is not the case. As the next section describes, the number of bacteria collected from the patient may be in the single digits.

1.1.3 The distribution of bacteria number in whole blood cultures

Whole blood is arguably the most important clinical specimen submitted for phenotypic AST today because it is used in the diagnosis of almost all infectious syndromes[2] that result in a systemic inflammatory response (which correlates to serious infections with risk of mortality), or syndromes in which the infection is not confined to a specific epithelium. Blood cultures are thus the most frequently ordered clinical specimen in inpatient settings. The utility of blood cultures arises from their ease of interpretation and the relative ease in obtaining them from peripheral draws. Blood is close to sterile in a healthy human being, and so any organism cultured from blood, besides a handful of common skin flora contaminants, is clinically significant. (Other serous bodily fluids such as cerebrospinal, pleural, pericardial, peritoneal, and synovial fluid are also expected to be sterile.)

Unfortunately, blood, and some serous fluids, typically have low densities of bacteria. One 1989 study found an average of 0.25 CFU/mL across positive blood cultures from 224 patients [18]. In several studies utilizing the pour plate method of quantitative culture for any bacteremic specimens, the percentage of blood spec-

[2] These syndromes include sepsis, catheter-associated bloodstream infections, and other bloodstream infections; endocarditis; pyelonephritis; meningitis; pericarditis; spontaneous bacterial peritonitis; pneumonia and empyema; gastrointestinal perforations/fistulas and septic pancreatitis; septic arthritis and prosthetic joint infections; some severe skin and soft tissue infections like osteomyelitis and gangrene; bacteria that invade the reticuloendothelial system like typhoid fever, legionellosis, and listeriosis; and fevers of unknown origin.

imens with <1 CFU/mL ranged between 35–54% , and the percentage of specimens with <10 CFU/mL ranged between 58–83%. The percentage of specimens with >100CFU/mL ranged between 6–10% [19]. The standard blood culture tube, into which venous blood is drawn directly into growth media, is designed to accept between 4 and 10 mL, which means that about half of all positive blood cultures begin with less than 100 bacteria contained inside. It is also common for replicate blood draws to yield positive cultures in only some of the replicates, explaining why it is standard to draw at least three sets of blood culture draws for suspected sepsis, each set containing one aerobic and one anaerobic culture [4]. Cases in which some blood draws return negative indicate that the density of bacteria is low enough that those draws returning positive have a high chance of having contained a single bacterium (the chance being calculable using the Poisson distribution). For example, a 1993 study, where six 5 mL replicates were drawn per case of suspected bacteremia, found that only 35% of *Escherichia coli* and 62% of *Staphylococcus aureus* cases had all 6 replicates return positive [20]. From the above studies, one can infer that the distribution of bacterial densities in blood has a strong positive skew, perhaps a power-law distribution. Many specimens contain few bacteria, often missing bacteria altogether while companion replicates turn positive. Decreasing fractions of specimens containing higher numbers of bacteria, ending with a thin tail of specimens with >100 CFU/mL.

The rough estimates above of circulating bacterial density make the assumption of a well-mixed, even distribution of bacteria in the blood. In reality, this assumption is not at all true, with bacteria possibly released from the site of infection in waves or aggregates. The precise distribution of densities in blood, unfortunately, has never been directly studied on a large scale with clinical samples due to labor costs and ethical limitations. Nonetheless, the numbers of cells in blood in the vast majority of blood specimens before culturing is simply too low for any current or proposed phenotypic AST method to achieve sufficient analytical sensitivity.

As a contrasting scenario, urine cultures are expected to have bacterial densities higher than 10^2 to 10^5 CFU/mL, depending on patient age and symptoms. Lower densities are in fact considered non-pathological[3].

[3]The majority of urine cultures are done for non-urgent suspected urinary tract infections. Such cultures are currently the most frequently requested type of culture [4], but AST is typically not performed unless an unusual and urgent pathogen is found. In the minority of urinary tract infections suspected of having escalated to pyelonephritis, the urine specimens would be accompanied by blood specimens too.

The specific type of infection and the specific taxa also affect the circulating density of bacteria, with bacteria residing inside white blood cells (e.g., typhoid fever) being sparser, while circulating densities in infectious endocarditis are 2–10 fold higher[4]than other bacteremias [19].

As the blood culture is incubated, classic models from *in vitro* cultures of pure isolates predict that the population will grow exponentially after a lag phase. No study has yet followed the growth of bacteria during blood culture bottle incubation however, due to the inability to non-destructively monitor growth until a large density is reached. Changes in turbidity are not distinguishable from background created by the opacity of whole blood (even if host cells are lysed). What is known is that a density of 7.5×10^6 to 5×10^7 CFU/mL is present at the time three commercial automated blood culture systems can confidently predict growth in the blood culture bottle. Growth is detected from measurements at 10 minute intervals of carbon dioxide and other metabolic gases in the headspace of the bottle [7].

By the time positive growth is detected in a blood culture, the number of cells has increased by 5 to 7 orders of magnitude. Since the initial blood culture is one of the longest steps in the traditional AST workflow, the question arises of whether blood culture can be shortened. The growth rate of wild bacteria is not a controllable variable, because it is believed that for the majority of human pathogens (those labeled "non-fastidious"), the maximum rate of growth allowed by the biophysics of macromolecule synthesis is already achieved by rich media, and because no amount of human engineering of the biosphere can guarantee that wild organisms have not deviated from expectations. Therefore, the duration of the initial blood culture is tied to the number of cells that have so far accumulated in the blood culture tube. Shortening or simply bypassing the duration of blood culture must address the question of whether there are enough bacterial cells in existence to perform subsequent analyses.

1.2 Thesis goals and intellectual contributions

In this thesis I will be examining the limits of low cells on the aspect of assay sensitivity, specifically how the number of cells in the assay system affects the outcome of the assay, the susceptibility calls.

[4]To clarify, fastidious agents of endocarditis like the HACEK organisms require prolonged incubations to turn positive and would be recorded as false negatives if the microbiology lab is not informed of clinical suspicion for endocarditis. Endocarditis from fast-growing pathogens like *Staphylococcus* or *Streptococcus* appear in intermittently higher densities, and also may be negative due to stochastic sample timing.

I assert that there is a minimum number of cells needed for sufficiently sensitive phenotypic measurements. Whatever processing is performed on a clinical specimen after collection must therefore produce at least that minimum number of cells sometime before the susceptibility calls are made to achieve a sufficient accuracy. Any attempt to remove host tissue and concentrate bacteria, such as the culturing of bacteria in liquid media found in current standard protocols, is inherently included in specimen processing and must also produce this minimum number. If the bacteria produced by the processing is too low, then the number of bacteria must be increased by further collection or by culturing somewhere before or after the host tissue separation, or—only if there are sufficient bacteria to begin with before the separation—the separation efficiency must be improved. Thus, the minimum number of cells determines in part the manner of specimen processing and especially the length of time required to process the specimen before (or possibly during) the assay.

One pertinent question remaining is how to calculate this minimum number of cells. One can easily theorize many factors controlling the minimum number of cells, and each one must be ruled out as negligible or be included in one's calculations. Of note, the minimum number of cells will be a function of inherent biological processes, such as the stochasticity in phenotypic responses, the pharmacodynamics of the phenotypic response, and the population dynamics that are part of the pharmacodynamic response. One would also expect these properties to vary between different bacterial taxa, antibiotic compounds, and possibly host physiology. The minimum number of cells also will depend on the measurement technology, by the limit of detection of the technology, whether the observations are destructive or non-destructive, and the type of physical quantities that can be measured. Lastly, included in the measurement process, but not limited to specific technologies, are factors such as the number of observations that can be made and the number of experimental conditions tested. A more rigorous discussion of the definition of the minimum number of cells and the definition of an ideal assay can be found in Chapter 5.

The goals of this thesis are to 1) derive a theoretical definition for the minimum number of cells for ideal and actual assays, 2) calculate the minimum number of cells for a clinically relevant bacterium and antibiotic drug using physical data, 3) describe a novel phenotypic AST assay for antibiotics that target peptidoglycan, and 4) assess the accuracy of the assay at low numbers of cells.

To achieve these goals, the following intellectual contributions were made:

1. Invention and reduction to practice of bulk filtration AST as a diagnostic method.
2. Preliminary validation of bulk filtration AST on clinical isolates of Enterobacteriaceae.
3. Application and successful fitting of a Bayesian mixed-effects compartment model to *in vitro* beta-lactam pharmacodynamics.
4. Invention of digital filtration AST as a diagnostic method.
5. Derivation of novel algorithms for the analysis of digital filtration AST results, including
    - the well population status probabilities of the simple Markov birth-death process and
    - the maximum likelihood estimators of the death probability and loading cell densities from well population status tallies in the simple Markov death process.
6. Preliminary validation of digital filtration AST on clinical isolates of Enterobacteriaceae.
7. Proposal for a definition and mathematical framework for phenotypic assays and ideal phenotypic assays.
8. Calculation of the minimum number of cells required by an ideal accessibility AST assay as a function of pharmacodynamic parameters, exposure duration, and a given assay performance.
9. Evidence showing that irreducible system stochasticity and dynamics are sufficient to explain the limits of digital filtration AST so far observed.

1.3 Phenotypic NAAT and Accessibility AST

The AST methods described in this thesis, and in the preceding work by members of the Ismagilov lab, use nucleic acid amplification (NAA) to measure antibiotic resistance phenotypes.

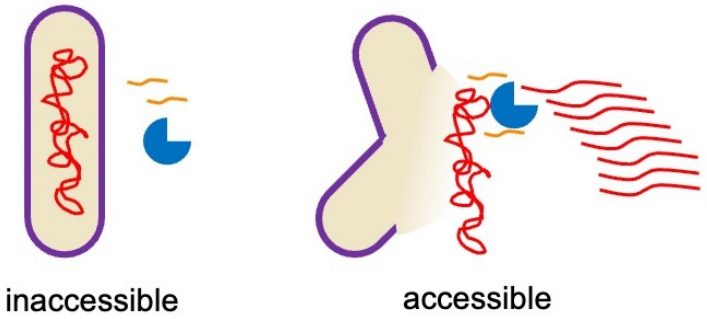


Figure 1.1: Accessibility of bacteria-derived molecules is a measurable phenotypic response to beta-lactam antibiotics.

The main rationale for using nucleic acid amplification over other readout technologies is that it is easier to separate and concentrate nucleic acids from host tissue than it is to separate and concentrate living bacteria from the same tissues. Many well-known protocols for extracting nucleic acids from human tissues have been developed over the decades. Their existence and relative robustness to host tissue complexity explains why non-phenotypic NAA tests have been successful in many aspects of laboratory medicine—except, of course, for antibiotic susceptibility testing.

Unfortunately, all of the currently used methods for nucleic acid extraction include chemicals such as alcohols, phenol, guanidinium chloride, and proteinases that kill living cells. Killing living cells destroys any future phenotypic signal. However, any phenotypic change in nucleic acid amounts or states elicited before extraction, and not obscured by the extraction, can be measured. For example, in Schoepp et al., 2017 [21], the difference in the total amount of uropathogenic *Escherichia coli* nucleic acids between a treated and a control antibiotic exposure was measured using fast (5–7 min) single-molecule digital LAMP. This method achieved phenotypic AST in 30 min from clinical urine samples without pre-culture. In Khazaei et al., 2018 [22], the changes in mRNA levels in *Neisseria gonorrhoeae* in response to ciprofloxacin was measured using ddPCR.

In Savela et al., 2019 [23] and Schoepp et al., 2019 [24] (latter is included as Chapter 2), a concept called "accessibility" was used as the phenotypic signal for beta-lactam antibiotics. The accessibility AST methods utilize the fact that beta-lactams cause physical disruptions in the peptidoglycan cell wall of bacteria [25]. It is believed that the covalent bonds of peptidoglycan exerts a tensile force that counteracts the osmotic pressure existing across all bacterial cell envelopes.

Beta-lactams bind to and irreversibly inhibit enzymes, collectively called penicillin-binding proteins (PBP), that remodel the peptidoglycan during normal growth [26]. When PBP activity is decreased, breakages in the peptidoglycan network appear. Osmotic pressure causes the cell envelope (which comprises the outer and inner lipid membranes as well as the peptidoglycan sacculus) to buckle and eventually burst, killing the cell. This mechanism of action has been demonstrated in many reports over the years [27, 28].

Any nucleic acid reagents introduced to the environment around the bacteria are unable to initiate amplification if they are topologically separated from their templates. However, if beta-lactams disrupt the bacterial cell envelope, then these reagents are able to initiate amplification, as depicted in Figure 1.1.

Accessibility AST differs from NAATs in other ASTs or other applications in that the extraction of nucleic acids must not damage the cell wall, unless the extracellular and intracellular nucleic acid fractions are first separated, as discussed in section 3.1.3.1. In prior existing methods or applications of nucleic acid extraction, disrupting the cell wall was a primary goal requiring enzyme activity, detergents, or vigorous physical shearing like bead beating. This requirement of no cell lysis means that nucleic acid extraction is not as simple as the measurement of total amounts. We believe that separating non-living nucleic acids from clinical specimens will prove easier than separating live cells from the same specimens nonetheless; this is an opinion still awaiting future investigation.

It should be noted that accessibility AST only measures the activity of antibiotics that cause lysis of the cell envelope. Thus, of the antibiotics currently in clinical use, the method should only work when testing the following classes: beta-lactams (comprising the penicillins, cephalosporins, monobactams, and carbapenems), fosfomycin, glycopeptides (e.g., vancomycin), and cyclic lipopeptides (e.g., daptomycin).

The limitation of accessibility AST to cell-wall agents does not preclude clinical adoption in the future because beta-lactams are the most widely prescribed class of antibiotics in clinical practice. Beta-lactams remain popular because they are well-tolerated with infrequent adverse reactions and have broad coverage. For example, in pneumonia in a patient without a history of recent hospitalization and severe enough to require hospitalization (i.e., inpatient community-acquired pnuemonia), the first-line antibiotics usually comprise one IV beta-lactam (ceftriaxone, cefotaxime, ceftaroline, ertapenem, or ampicillin-sulbactam) and either a macrolide or a tetracycline for coverage of atypicals (*Mycoplasma*, *Legionella*, and *Chlamydia*).

Alternatively, levofloxacin or moxifloxacin monotherapy, which would not include a beta-lactam, can be used; in practice both regimens are used with equal frequency by hospitals [29]. For bloodstream infections (which always warrant hospitalization), almost all empiric regimens include a beta-lactam. Even if the patient has a documented penicillin allergy (and the allergy is truly an allergy), it is still recommended to include another class of beta-lactam (e.g., aztreonam) due to the effectiveness of beta-lactams. For gram-negative bacteremias, a typical regimen includes a third or fourth generation cephalosporin (ceftriaxone, ceftazidime, or cefepime) or a beta-lactam/beta-lactamase inhibitor combination (e.g., piperacillin-tazobactam). If the patient is immunocompromised or shows septic shock, then pseudomonal coverage with another beta-lactam is added [30]. For Gram-positive bacteremias, a penicillin or first/second generation cephalosporin is used (nafcillin, oxacillin, or cefazolin). Vancomycin is added wherever MRSA is prevalent (which is frequent) [31]. If the Gram stain is not yet available, then the presumed bloodstream infection must be treated for both Gram-positive and Gram-negative possibilities, likely with a broad-spectrum beta-lactam covering both. Furthermore, though it is discouraged, some clinicians may opt to start with more powerful beta-lactams, including carbapenems, if there is a reason to suspect a patient has a resistant organism, such as a prior resistant infection or a prolonged hospital stay up to the point of infection.

Phenotypic AST for beta-lactams specifically would likely have a large clinical impact if successfully sped up, thanks to the ubiquity of beta-lactams. Demand for AST for beta-lactams is already ubiquitous in hospitals today, thanks to the widespread use of beta-lactams and the fact that resistance to earlier classes of beta-lactams is relatively common and increasing. As a specific example, if it were known that a Staphylococcus aureus isolate was penicillin-susceptible, penicillin G monotherapy would be the preferred approach instead of the broader first-line empiric regimen described above for Gram-positive bacteremia. Similarly, narrowing regimens or salvaging older antibiotics for Gram-negative bacteremia would also be encouraged if clinicians received earlier, definitive AST results that allow them to let go of more powerful options. For example, in practice, a de-escalation of meropenem, a broad-spectrum carbapenem used in intensive care units for empiric treatment, could be a switch to ertapenem, piperacillin-tazobactam, ceftriaxone and other third-generation cephalosporins, and finally penicillins like amoxicillin, if any of the latter choices were shown to be effective by phenotypic AST, with later choices being preferred over choices listed earlier [32].

The implementation and validation of the simplest form of accessibility AST, called polymerase AST or "pol-aAST," is described in the next chapter.

Chapter 2

DIFFERENTIAL DNA ACCESSIBILITY TO POLYMERASE ENABLES 30-MINUTE PHENOTYPIC β-LACTAM ANTIBIOTIC SUSCEPTIBILITY TESTING OF CARBAPENEM-RESISTANT ENTEROBACTERIACEAE

This chapter was published[1] as:

[1] Nathan G. Schoepp, Eric J. Liaw, Alexander Winnett, Emily S. Savela, Omai B. Garner, and Rustem F. Ismagilov. Differential DNA accessibility to polymerase enables 30-minute phenotypic β-lactam antibiotic susceptibility testing of carbapenem-resistant Enterobacteriaceae. *PLOS Biology*, 18(3):e3000652, March 2020. ISSN 1545-7885. doi: 10.1371/journal.pbio.3000652. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000652.

NGS, EJL, AW, ESS, and RFI contributed to conceiving the method, revising the manuscript, and interpretation of experimental results. NGS developed the sample handling workflow and performed all experiments for comparison of amplification methods, validation, and timed sample-to-answer experiments. NGS was the major contributor to manuscript preparation and prepared all figures. EJL performed filtration experiments, reviewed relevant medical literature, and contributed to manuscript writing. NGS and AW tested clinical samples using the modified workflow. ESS performed early experimental work to link beta-lactam exposure to differential nucleic acid readout, analyzed data from validation experiments, and developed TTPD metrics. OBG provided clinical guidance on the selection of clinical isolates and clinical samples and coordinated and provided oversight of clinical-sample collection at UCLA, including technical assistance to UCLA staff. RFI supervised and guided the project, and helped compose the manuscript.

2.1 Abstract

The rise in carbapenem-resistant Enterobacteriaceae (CRE) infections has created a global health emergency, underlining the critical need to develop faster diagnostics to treat swiftly and correctly. Although rapid pathogen-identification (ID) tests

[1] Contains sporadic differences in grammar and updated bibliographic references. The original "Notes and References" section has been combined with the references for the rest of the thesis.

Additionally, error bars have been added to Figure 2.3 to show ddPCR Poisson 95% confidence intervals after error propagation.

are being developed, gold standard antibiotic susceptibility testing (AST) remains unacceptably slow (1–2 days), and innovative approaches for rapid phenotypic ASTs for CREs are urgently needed. Motivated by this need, in this manuscript we tested the hypothesis that upon treatment with β-lactam antibiotics, susceptible Enterobacteriaceae isolates would become sufficiently permeabilized, making some of their DNA accessible to added polymerase and primers. Further, we hypothesized that this accessible DNA would be detectable directly by isothermal amplification methods that do not fully lyse bacterial cells. We build on these results to develop the polymerase-accessibility AST (pol-aAST), a new phenotypic approach for β-lactams, the major antibiotic class for gram-negative infections. We test isolates of the 3 causative pathogens of CRE infections using ceftriaxone (CRO), ertapenem (ETP), and meropenem (MEM) and demonstrate agreement with gold-standard AST. Importantly, pol-aAST correctly categorized resistant isolates that are undetectable by current genotypic methods (negative for β-lactamase genes or lacking predictive genotypes). We also test contrived and clinical urine samples. We show that the pol-aAST can be performed in 30 min sample-to-answer using contrived urine samples and has the potential to be performed directly on clinical urine specimens.

2.2 Introduction

The evolution and global spread of carbapenem-resistant Enterobacteriaceae (CRE) threatens to disrupt modern healthcare systems, which rely heavily on β-lactams (especially carbapenems, the last-resort treatments) to control bacterial infections [33–35]. Mortality rates for CRE infections are as high as 30%–49% [36–38], and thus the global emergence and spread of CRE infections represents a public health emergency [39–41]. The Centers for Disease Control and Prevention (CDC) places CRE in its highest ("urgent") category of antimicrobial-resistant pathogen threats [6, 40, 41], and the World Health Organization (WHO) labels CRE as a critical-priority pathogen [39]. *Escherichia coli*, *Klebsiella pneumoniae*, and *Enterobacter* spp. compose the majority of CRE infections and are the most commonly monitored Enterobacteriaceae [40, 42–44].

To halt the further spread of CRE, patients need to be treated swiftly and correctly at the point of care (POC); however, there is no fast and general method for determining antibiotic susceptibility [45–47]. The current clinical workflow for treatment of bacterial infections consists of an identification (ID) step followed by an antibiotic susceptibility test (AST). Although progress is being made to develop faster ID tests [48–50] and a rapid 20-min ID test is on the horizon [51–53], the gold-standard for

AST remains a culture-based workflow using broth or agar dilution that requires 1 to 2 days and is thus far too slow [54, 55]. Because AST results are so delayed, healthcare providers usually treat empirically, leading to inappropriate prescriptions and even life-threatening outcomes [56], as well as the further spread of resistance. To improve treatment and promote antibiotic stewardship, healthcare providers need a rapid phenotypic AST [57–59].

ASTs are either genotypic or phenotypic. Genotypic tests predict resistance by measuring the presence of genes known to be involved in resistance. Genotypic tests can be fast [60] but often have limited clinical utility because they target defined mechanisms of resistance. For example, rapid genotypic methods to detect gram-negative β-lactamase genes have been developed [61–64], but these tests only detect one of the many known β-lactamase classes and still require 30 to 40 min (estimated from described methods). Similarly, the commercial Cepheid Xpert Carba-R assay (Cepheid, Sunnyvale, CA), which detects 5 β-lactamase gene families, was shown to detect 50% of resistant isolates and took 88 min [65]. Moreover, although Carba-R is Food and Drug Administration (FDA) approved, its utility in treatment scenarios is limited (i.e., negative results are not actionable), so when prescribing antibiotics, it must be used in conjunction with a phenotypic AST [66, 67]. Rapid methods for measuring the activity of specific β-lactamases also exist [68–72]. However, these tests only detect one mechanism of resistance, and sample-to-answer times have not been reported.

Phenotypic ASTs are ideal because they determine susceptibility directly by exposing the sample to antibiotics and measuring the target organism's response. The gold-standard AST (broth microdilution [54, 55]) is a phenotypic test. Most phenotypic tests require the growth of viable organisms isolated from patient samples, a process that requires days and is thus too slow for the POC. Innovative, faster phenotypic tests for β-lactams were developed based on *in situ* nucleic-acid staining or fluorescence measurements [73–75], flow cytometry [76], microscopy [77–79], optical density [80, 81], and mass spectrometry [82]. However, the majority of the currently proposed methods still require 60- to 180-min antibiotic-exposure steps in addition to the time needed to perform the assay, and no method has emerged that achieves short (approximately 15 min) antibiotic exposure and short (approximately 15 min) assay time but does not require excessively complex or delicate instrumentation so the method can be deployed at the POC.

Rapid phenotypic methods based on quantification of nucleic acids (NAs) have

shown great promise for a rapid POC AST due to the speed, specificity, and robustness of NA detection [83–88]. There is an additional advantage to using NA quantification as a readout of the bacterial response to antibiotic: because rapid pathogen ID from clinical samples is commonly performed via NA analysis, it would likely be easier to integrate an NA-based phenotypic AST into a combined ID/AST workflow performed from the same clinical sample. Additionally, the use of NA-based methods provides molecular specificity towards the target pathogen, which is important in clinical samples that can contain multiple organisms. For antibiotics that directly or indirectly impact NA replication on short timescales, we have demonstrated that the quantification of DNA [21, 89] or RNA [22] can be used to rapidly (30 min) and reliably determine susceptibility to nitrofurantoin and ciprofloxacin. Subsequent efforts have targeted the β-lactam class (the most widely prescribed class of antibiotic [33, 34]) using these methods [90]. However, because β-lactams do not directly impact NA replication on short timescales, this direct translation of the existing NA-based technique required a 2-h antibiotic exposure, which is not sufficiently rapid for the POC. For a POC AST to impact management of CRE infections, it must (i) determine susceptibility to β-lactams, including carbapenems; (ii) be rapid (<30-min sample-to-answer) [91, 92]; and (iii) be phenotypic [57, 58]. As discussed subsequently, rapid pathogen ID technologies are becoming available, and therefore pathogen ID is not the focus of this work.

Here, we hypothesized that a new NA-based approach could be used to develop a rapid phenotypic AST for multiple β-lactams. We hypothesized that upon treatment with β-lactam antibiotics, susceptible Enterobacteriaceae isolates would become sufficiently permeabilized so some of their DNA would become accessible to added polymerase and primers. Further, we hypothesized that this accessible DNA would be detectable directly by isothermal amplification methods that do not fully lyse bacterial cells. To differentiate between resistant and susceptible organisms, rather than measuring how total NA concentration is impacted by antibiotic exposure (as in previous NA-based ASTs), we hypothesized that we could measure the accessibility of NAs to polymerase following a short antibiotic exposure. Here, we test these hypotheses and use them to design a new AST method, termed polymerase-accessibility AST (pol-aAST). To validate the method, we performed 82 ASTs using clinical isolates of 3 major CRE pathogens exposed to each of 3 commonly prescribed β-lactams for gram-negative infections: ceftriaxone (CRO), ertapenem (ETP), and meropenem (MEM). To further demonstrate that this method has potential to be used clinically in POC relevant timescales, we (i) performed timed

sample-to-answer experiments using contrived urine samples to ensure that the whole assay can be performed in <30 min, and (ii) we performed a pilot study on clinical urine samples from patients with urinary tract infections (UTIs).

2.3 Results

The pol-aAST relies on differential accessibility of NAs to polymerases as a result of antibiotic exposure. In this manuscript, we define differential accessibility to polymerase as a difference in the measured rate of amplification between control and antibiotic-treated samples. In the first step of pol-aAST, a single sample is split into control and treated aliquots of equal volume, and the treated aliquot is exposed to a β-lactam. Antibiotic exposure is a critical step in any phenotypic AST because phenotypic tests measure the response of bacteria to antibiotics. If the bacteria in the sample are resistant, we hypothesized that no differences in NA amplification would be observed between control and treated aliquots. If the bacteria are susceptible, we hypothesized that antibiotic treatment would lead to a compromised peptidoglycan cell wall (Figure 2.1A) and partial release of NAs (Figure 2.1B). We hypothesized that both the compromised cell wall and partial release of NAs would increase the accessibility of NAs to polymerase in a treated antibiotic-susceptible aliquot. In the second step of pol-aAST, control and treated aliquots are exposed to polymerase in amplification conditions (Figure 2.1C), and the rate of amplification is measured.

To successfully differentiate susceptible and resistant samples, ideal amplification conditions must (i) not fully lyse cells, (ii) enhance alterations (damage) to the cell wall caused by exposure to β-lactams, and (iii) increase NA release only from antibiotic-damaged cells. The rate of amplification is dependent on the concentration of polymerase-accessible NA. In susceptible samples, more NAs are released in the treated aliquot, leading to faster amplification in susceptible treated aliquots (Figure 2.1D) relative to the controls. Resistant samples are not affected by the antibiotic, so control and treated aliquots have similar NA release and time-to-positive (TTP). In these samples, the low concentration of naturally occurring extracellular DNA is ultimately amplified, but at a slower rate. Amplification rate in an isothermal amplification reaction is quantified by measuring the TTP, the time it takes the reaction fluorescence to reach a predetermined threshold. We found that using pol-aAST, isolates susceptible to the β-lactam being tested show increased accessibility of NAs to polymerase, manifesting in an earlier TTP relative to the control. The TTPs of any two samples, such as the control and treated aliquots, can be compared to generate a TTP difference (TTPD) value, which can then be used to determine

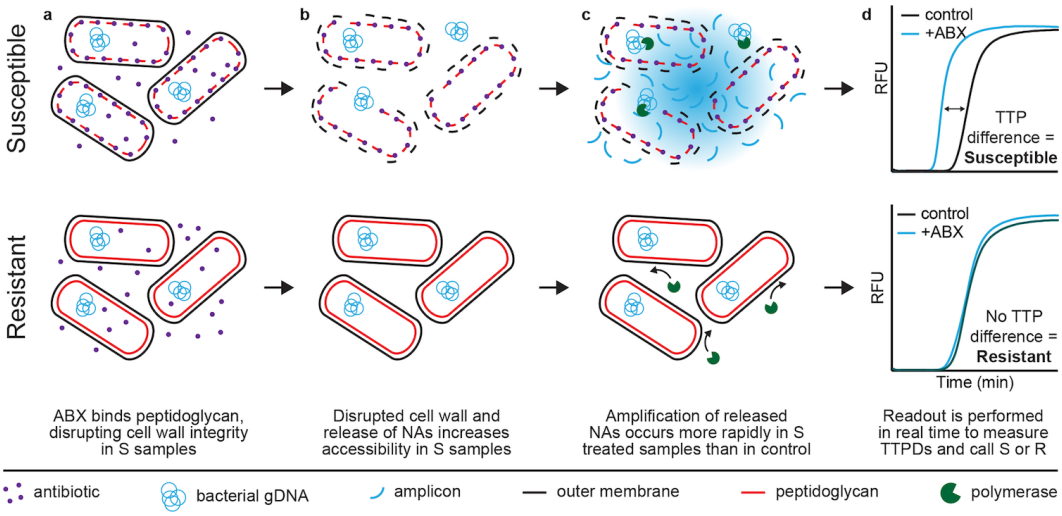


Figure 2.1: Overview of the pol-aAST shown for susceptible and resistant samples exposed to β-lactams. (a) Treated aliquots are exposed to a β-lactam. In susceptible samples, β-lactams compromise cell wall integrity. (b) NAs are released from compromised cells, increasing NA accessibility to polymerase.(c) Released NAs in the susceptible treated aliquot amplify faster than NAs from intact cells in the control aliquot, resulting in a difference in TTP. No difference in amplification between control and treated aliquots is observed in resistant samples. (d) TTPD between control and treated aliquots is used to assess susceptibility. ABX, antibiotic; AST, antibiotic susceptibility testing; gDNA, genomic DNA; NA, nucleic acid; pol-aAST, polymerase-accessibility AST; R, resistant; RFU, relative fluorescent units; S, susceptible; TTP, time-to-positive; TTPD, time-to-positive difference.

susceptibility by comparing to a susceptibility threshold. Here, we used the DNA polymerase Bst 3.0 (New England Biolabs [NEB], Ipswitch, MA) under loop-mediated isothermal amplification (LAMP) conditions.

We hypothesized that the chemical environment in which amplification occurs would significantly impact the result of pol-aAST and that—for pol-aAST to differentiate susceptible and resistant samples—amplification conditions should not be fully lysing. To test this, we performed pol-aAST using LAMP, as well as quantitative PCR (qPCR) (Figure 2.2). LAMP is performed at a single temperature (70°C), which we hypothesized would not be fully lysing, whereas qPCR is a thermocycled amplification technique reaching a maximum temperature of 95°C, which we hypothesized would be fully lysing. Indeed, we observed that pol-aAST was successful in differentiating susceptible and resistant isolates when performed using LAMP, but not when performed using qPCR (Figure 2.2). We tested qPCR with a total of 2 susceptible and 2 resistant isolates, none of which showed a statistically significant difference in quantitation cycle (Cq) between control and treated sam-

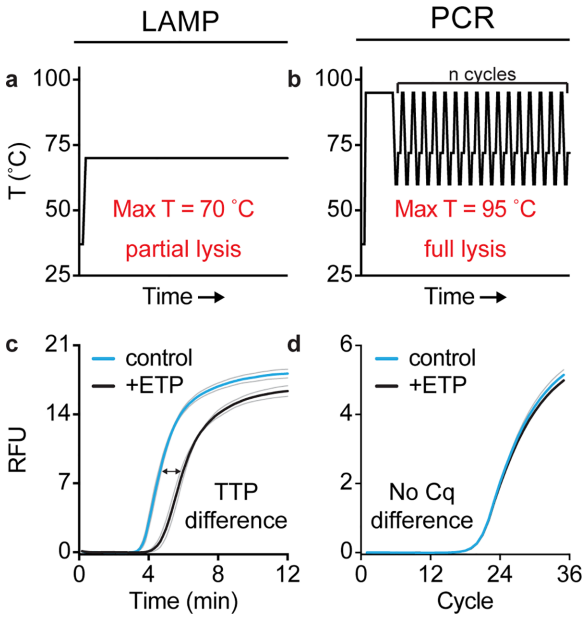


Figure 2.2: The pol-aAST requires non-lytic amplification conditions. (a–b) Thermal profiles of LAMP and PCR. (c–d) LAMP and PCR amplification curves for a susceptible *E. coli* isolate exposed to ETP for 15 min. Blue and black lines are the average of triplicate samples. Grey lines represent standard deviation of triplicates. A difference in TTP for control and treated aliquots is observed for susceptible isolates when quantifying NAs using LAMP, but not PCR. Raw data are provided in S5 Table of Schoepp et al. 2020 [24]. AST, antibiotic susceptibility test/testing; Cq, quantitation cycle; ETP, ertapenem; LAMP, loop-mediated isothermal amplification; NA, nucleic acid; PCR, polymerase chain reaction; pol-aAST, polymerase-accessibility AST; RFU, relative fluorescent units; TTP, time-to-positive.

ples. When using LAMP, detectable differences were observed between control and treated aliquots when using isolates susceptible to the target β-lactam (TTPD = 1.02 min). Additionally, the presence of cells not lysed during LAMP is evidenced by the shorter TTPs seen when an aliquot of the same sample is lysed using an extraction buffer prior to performing LAMP (explained in more detail subsequently). These differences confirm that choice of amplification chemistry is critical to the success of pol-aAST and are consistent with previous work evaluating thermal lysis [93].

To investigate the mechanism of pol-aAST, we performed experiments to separate free NAs from NAs contained within structurally intact cells or associated with cell debris. Susceptible and resistant clinical isolates were exposed to one or more β-lactams in parallel for 15 min, then filtered through 0.2 µm filters to remove cells from free NAs. NAs in the sample and eluate were then quantified using droplet digital PCR (ddPCR). We observed that following exposure to β-lactams, susceptible

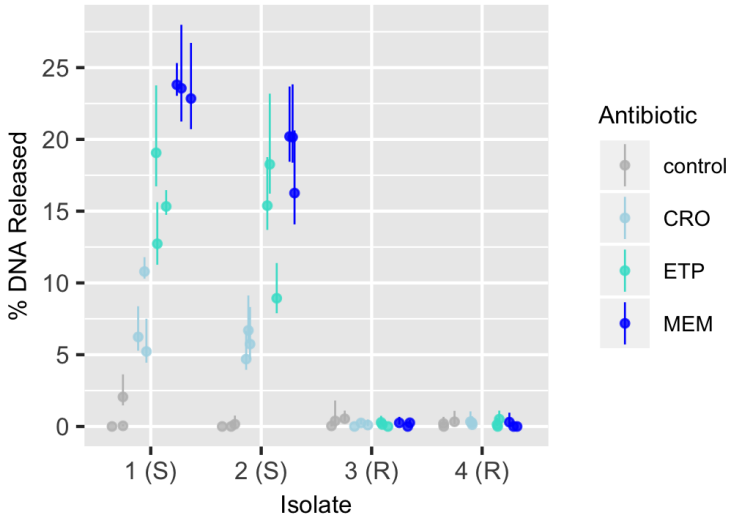


Figure 2.3: Two susceptible and two resistant *E. coli* isolates were exposed to no antibiotic (control), CRO, ETP, or MEM for 15 min before filtering to separate intact cells from extracellular DNA. Experiments were performed in triplicate for all isolate/antibiotic combinations. Each point represents a single experiment; lines represent the average and standard deviation of replicate experiments. Raw data are provided in S6 Table of Schoepp et al. 2020 [24]. ABX, antibiotic; CRO, ceftriaxone; ETP, ertapenem; MEM, meropenem; R, resistant; S, susceptible.

isolates treated with β-lactams released a significantly larger percentage of DNA than resistant samples (Figure 2.3). The amount of DNA released depended on the antibiotic being tested. Exposure to MEM resulted in an average of 21% of DNA being released from susceptible isolates, with a slightly smaller average percent (15%) released as a result of exposure to ETP. Interestingly, susceptible samples only released an average of 6% of DNA when exposed to CRO, demonstrating that NA release is dependent on choice of antibiotic and not, e.g., a universal stress response. These results also demonstrate that the magnitude of the effect of a β-lactam on cell wall integrity can be measured and is different depending on the antibiotic used, even on short exposure timescales.

To validate the pol-aAST method, we first performed 82 ASTs using 12 clinical isolates of *E. coli*, 8 clinical isolates of *K. pneumoniae*, 9 clinical isolates of 2 species of *Enterobacter* (*E. aerogenes* and *E. cloacae*, collectively "*Ebs*"), and the β-lactams CRO, ETP, and MEM. The set included isolates from each genus that were susceptible and isolates that were resistant to each of the three antibiotics. In addition to isolates obtained from the UCLA Clinical Microbiology Laboratory (CML; see Methods), those tested included *E. coli* and *K. pneumoniae* isolates from the CDC

Enterobacteriaceae Carbapenem Breakpoint panel [94], as well as all available *Enterobacter* spp. isolates from the same panel. All samples were amplified using quantitative LAMP, and categorical agreement was compared to gold-standard broth microdilution AST. Two approaches for determining susceptibility were investigated in all pol-aASTs performed.

The first approach we investigated was to compare the difference in TTP values of the control and treated aliquots in each pol-aAST. This difference was defined as TTPD control to treated ($TTPD_{CT}$) (Figure 2.4a). Using the $TTPD_{CT}$ method, we obtained 100% categorical agreement with gold-standard AST for all antibiotics tested with *E. coli* (Figure 2.4b), *K. pneumoniae* (Figure 2.4c), and *Ebs* (Figure 2.4d) isolates, even with resistant isolates for which the genotypic tests fail to correctly predict the resistance phenotype (red points in Figure 2.4). The values of $TTPD_{CT}$ were well-separated between susceptible and resistant isolates in all CRE-antibiotic combinations. Note that the threshold values separating $TTPD_{CT}$ of susceptible and resistant isolates depend on the antibiotic used (e.g., CRO gives a smaller response and therefore requires a lower threshold), as well as the pathogen tested (e.g., *K. pneumoniae* gives stronger response and requires a higher threshold). The area under the curve (AUC) of the receiver operating characteristic (ROC) curve was 1.00 for all isolates and antibiotics tested. There were no errors relative to gold-standard AST when determining susceptibility by $TTPD_{CT}$.

The second approach we investigated was to compare the difference in TTP values of a fully lysed aliquot and the antibiotic-treated aliquot in each pol-aAST. The fully lysed aliquot was created by extracting NA from the antibiotic-treated sample using a single-step, LAMP-compatible extraction buffer. This difference was defined as TTPD lysed-control to treated ($TTPD_{LT}$) (Figure 2.5a). It is important to note that $TTPD_{LT}$ only requires an antibiotic-treated sample during the exposure step (the method does not require the use of a no-antibiotic control during exposure), meaning that the original sample does not have to be split prior to exposure. Again, the thresholds were defined individually for each antibiotic and pathogen. Using the $TTPD_{LT}$ method, we obtained 100% categorical agreement with gold-standard AST for all antibiotics tested only with *E. coli* (Figure 2.5b) and *K. pneumoniae* (Figure 2.5c) isolates, and with resistant isolates for which the genotypic tests fail to correctly predict the resistance phenotype (red points in Figure 2.5). When testing *Ebs* (Figure 2.5d) isolates, we observed two errors in which an isolate classified as CRO resistant was called susceptible, resulting in an overall categorical agreement

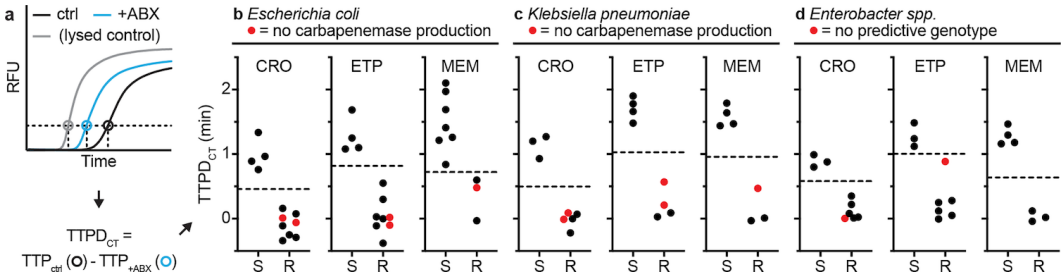


Figure 2.4: Validation of the pol-aAST method using control and antibiotic-treated aliquots. (a) Example calculation of TTPD between control and treated aliquots ($TTPD_{CT}$). The TTP (in minutes) of the control and treated aliquots are used to calculate $TTPD_{CT}$. (b–d) The pol-aAST results using *E. coli* (b), *K. pneumoniae* (c), and *Enterobacter* spp. (d) isolates exposed to CRO, ETP, and MEM for 15 min. Red points represent isolates with either no detectable carbapenemase genes (*Ec* and *Kp* isolates) according to a published genotypic assay [95] and commercial assay [96] or no predictive genotype (*Ebs* isolates) according to the whole genome sequencing by the CDC [94]. S/R thresholds (dashed lines) were set halfway between the lowest susceptible and the highest resistant $TTPD_{CT}$ values. Raw data are provided in S3 Table of Schoepp et al. 2020 [24]. +ABX, antibiotic-treated; AST, antibiotic susceptibility testing; CDC, Centers for Disease Control and Prevention; CRO, ceftriaxone; CT, control to treated; ctrl, control; *Ebs*, *E. aerogenes* and *E. cloacae* collectively; *Ec*, *E. coli*; ETP, ertapenem; *Kp*, *K. pneumoniae*; MEM, meropenem; pol-aAST, polymerase-accessibility AST; R, resistant; RFU, relative fluorescent units; S, susceptible; TTP, time-to-positive; TTPD, time-to-positive difference; $TTPD_{CT}$, TTPD control to treated.

of 88%. Because of these errors, the AUC for *Ebs* isolates tested with CRO was 0.94. Aside from these errors, susceptible and resistant isolates were well separated in all cases, with AUC = 1.000 for all antibiotics tested with *E. coli* and *K. pneumoniae*. Although we observed two errors, using the $TTPD_{LT}$ metric still gave excellent agreement with gold-standard AST and required no splitting of the sample prior to exposure.

To demonstrate one of the major differences between pol-aAST, a phenotypic method, and existing genotypic methods, we challenged the assay with 5 previously characterized isolates that had either (i) no detectable β-lactamase genes or (ii) lacked any genotypic signature predictive of β-lactam resistance. We tested 2 $_{\text{E. coli}}$ and 2 $_{\text{K. pneumoniae}}$ isolates with no detectable β-lactamase genes as measured by both a published genotypic assay designed to screen for 6 β-lactamase gene families [95], as well as the Cepheid Xpert Carba-R test (a commercial, FDA-approved genotypic assay designed to screen for 5 β-lactamase gene families) [96]. These 4 isolates did not test positive in either assay because they lack the genes these assays

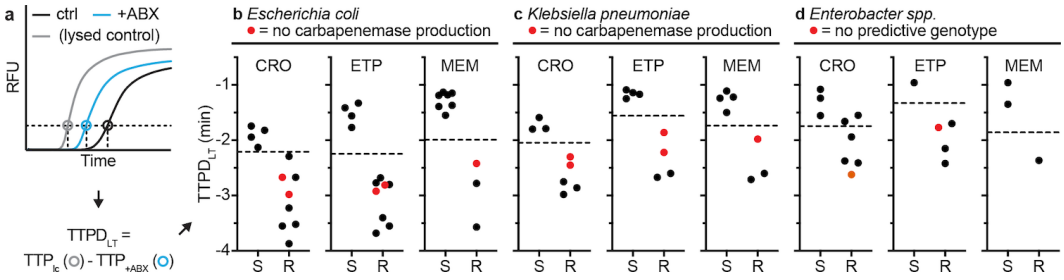


Figure 2.5: Validation of the pol-aAST method using lysed control and antibiotic-treated aliquots. (a) Example calculation of TTPD between the lysed control and antibiotic-treated aliquots ($TTPD_{LT}$). The TTP (in minutes) in the lysed control and antibiotic-treated aliquots are used to calculate $TTPD_{LT}$. (b–d) The pol-aAST results using *E. coli* (b), *K. pneumoniae* (c), and *Enterobacter* spp. (d) isolates exposed to CRO, ETP, and MEM for 15 min. Red points represent isolates with either no detectable carbapenemase genes (*Ec* and *Kp* isolates) according to a published genotypic assay [95] and commercial assay [96], or no predictive genotype (*Ebs* isolates) according to the CDC [94]. S/R thresholds (dashed lines) were set halfway between the lowest susceptible and the highest resistant $TTPD_{LT}$ values except in the case of *Enterobacter* spp. treated with CRO (see text). Raw data are provided in S3 Table. +ABX, antibiotic-treated; AST, antibiotic susceptibility testing; CDC, Centers for Disease Control and Prevention; ctrl, control; CRO, ceftriaxone; *Ebs*, *E. aerogenes* and *E. cloacae* collectively; *Ec*, *E. coli*; ETP, ertapenem; *Kp*, *K. pneumoniae*; lc, lysed control; MEM, meropenem; pol-aAST, polymerase-accessibility AST; R, resistant; RFU, relative fluorescent units; S, susceptible; TTP, time-to-positive; TTPD, time-to-positive difference; $TTPD_{LT}$, TTPD lysed-control to treated.

screen for, despite being resistant (as determined by gold-standard broth microdilution). These 4 tested isolates were resistant to CRO and ETP, and one isolate from each genus was also resistant to MEM. Additionally, we tested a single resistant $_{\text{Ebs}}$ isolate from the CDC Enterobacteriaceae Carbapenem Breakpoint Panel (AR-Bank #0007). Whole genome sequencing of this isolate (performed by the CDC) revealed no known resistance markers [94], meaning that the mechanism of resistance was uncharacterized. The pol-aAST performed excellently in all cases, and all 5 isolates were correctly categorized as resistant (Figs 2.4 and 2.5, red points).

To investigate the sample-to-answer time of the pol-aAST, we performed timed experiments using contrived urine samples (Figure 2.6). Sample-to-answer time is a critical metric for any assay designed to be used at the POC but is often not reported at all, even for methods claiming to be rapid. In timed experiments, we (i) reduced the exposure time from 15 to 13 min to ensure that all handling could be performed during the 15 min allocated for exposure and (ii) used an automated data-analysis spreadsheet to provide a susceptibility call as soon as the LAMP reactions reached

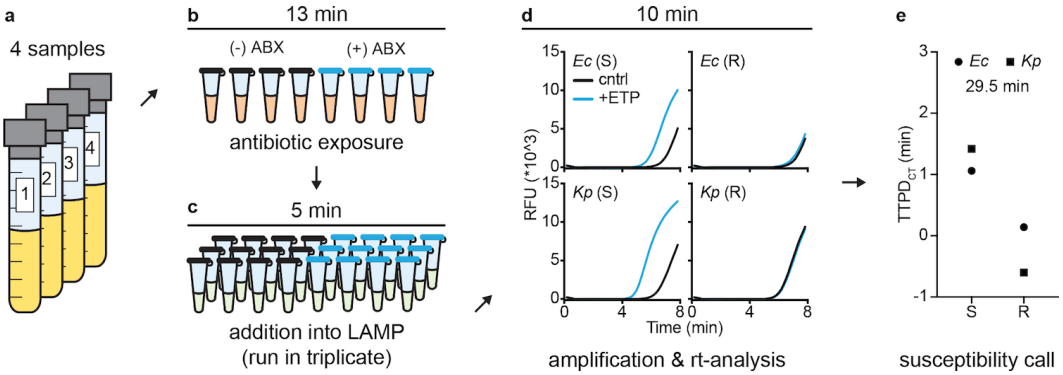


Figure 2.6: Timed sample-to-answer pol-aAST using contrived urine samples spiked with either *Ec* or *Kp*. (a) Because minimal sample handling is required for pol-aAST, all 4 contrived urine samples were run in parallel. (b) Urine samples were split into control and antibiotic-treated aliquots and incubated at 37°C for 13 min. A timer was started immediately after sample splitting. (c) All samples were added to pre-made LAMP mix and run in technical triplicate. (d) Samples were amplified using LAMP, and the fluorescence of reactions was monitored in real time. Once total fluorescence passed a predetermined threshold (indicating successful amplification), reactions were stopped and TTP values ported into an automated data-analysis spreadsheet. The timer was stopped as soon as the spreadsheet gave susceptibility calls. (e) Comparison of susceptibility calls with gold-standard AST categorization. Total assay time was 29.5 min. Raw data are provided in S3 Table of Schoepp et al. 2020 [24]. ABX, antibiotic; AST, antibiotic susceptibility test/testing; cntrl, control; Ec, *E. coli*; ETP, ertapenem; Kp, *K. pneumoniae*; LAMP, loop-mediated isothermal amplification; pol-aAST, polymerase-accessibility AST; R, resistant; RFU, relative fluorescent units; rt, real-time; S, susceptible; TTP, time-to-positive; TTPD, time-to-positive difference.

a predetermined threshold (indicating successful amplification). At the initiation of pol-aAST, a timer was started that ran for the duration of the experiment and was stopped once a susceptibility call had been made. The susceptibility of 4 isolates to ETP was tested simultaneously (Figure 2.6a). The pol-aAST consists of only 3 simple handling steps (Figure 2.6b–2.6d), which allowed us to perform pol-aAST in a total time of just 29.5 min, with results in agreement with gold-standard AST (Figure 2.6e).

We next ran the pol-aAST on clinical urine samples from patients diagnosed with UTI. These samples were confirmed to be Enterobacteriaceae-positive UTIs by the UCLA CML, and the pol-aASTs were run 3 to 5 days after collection. Initial experiments running the pol-aAST directly on clinical urine samples revealed an insufficient response to antibiotics in some samples. Because we analyzed urine samples that had been stored in a chemical preservative (see Methods) for 3 to 5

days after collection, some variation in the response to antibiotics was expected. However, we wished to test whether the delays in the response were indeed due to the phenotypic state of bacteria in these archived samples, and not due to the intrinsic biology of the bacterial strains in these samples. To test, we obtained 25 clinical urine specimens that exhibited an expected heterogeneity, as indicated by the wide range of urinalysis findings (see S2 Table of Schoepp et al. 2020 [24]): pH ranged from <5 to 8, specific gravities ranged from <1.005 to >1.060 (above and below the ranges detected in standard urinalysis), and protein, ketone, and bilirubin contents ranged from absent to the maximum measurable by urinalysis. Some samples contained red blood cells, leukocytes, and squamous epithelial cells. Two of the samples were polymicrobial. To ensure a response from bacteria in these specimens, we added a 30-min pre-incubation step of urine with media and increased the duration of antibiotic exposure to 45 min (see Methods). We did not optimize these conditions and did not attempt to identify the shortest possible incubation or exposure time. Eight samples were tested for ampicillin (AMP) susceptibility, and 17 samples were tested for ETP susceptibility. Prior to testing clinical samples using AMP, we tested 5 *E. coli* isolates using AMP (S1 Figure of Schoepp et al. 2020 [24]). Despite the heterogeneity in the urine matrix and the likely nutrient-deprived condition of the bacteria in the urine samples, pol-aAST experiments yielded clean separation between AMP-sensitive and -resistant *E. coli*. Additionally, we were able observe a response to ETP in 14 of 17 ETP-sensitive urine samples tested. Overall, we obtained 100% categorical agreement for determination of AMP susceptibility (4/4 susceptible and 4/4 resistant; Figure 2.7) and observed a response indicating susceptibility to ETP in 14 of 17 (82.4%) confirmed-susceptible samples (Figure 2.7), including the 2 polymicrobial samples. None of the samples received for testing by the pol-aAST method were ETP-resistant.

2.4 Discussion

The pol-aAST method enables rapid, organism-specific measurement of susceptibility to β-lactams—the most important class of antibiotic for gram-negative infections—thus providing the critically missing piece needed to develop a POC AST for this global health threat. The genera of isolates and the β-lactams used in this proof-of-concept study were intentionally chosen—*E. coli*, *K. pneumoniae*, and *Ebs*—and are responsible for the majority of CRE infections globally [40, 42–44] (in some areas of the US, *K. pneumoniae* is responsible for up to 90% of CRE infections [37]). It is for this reason that *E. coli*, *K. pneumoniae*, and *Ebs* together make up the

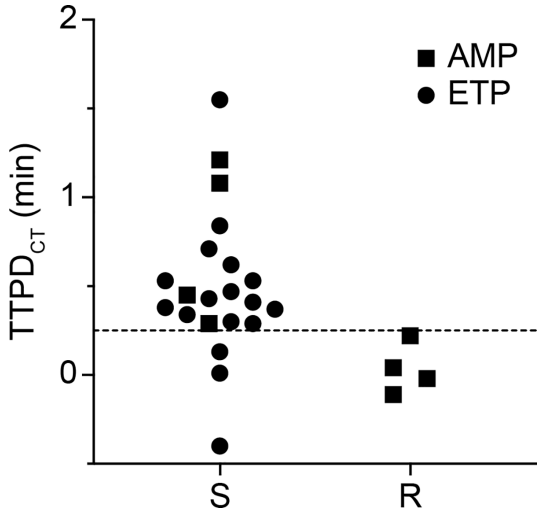


Figure 2.7: Pilot testing of pol-aAST with clinical UTI samples with a modified protocol (see Methods and Discussion). $TTPD_{CT}$ values for AMP and ETP susceptibility obtained by pol-aAST, with clinical UTI samples containing *E. coli*. Each point represents the $TTPD_{CT}$ value for one clinical sample tested once by pol-aAST (S2 and S4 Tables of Schoepp et al. 2020). LAMP was performed in technical triplicate, see S4 Table of Schoepp et al. 2020 [24] for values and statistical details. AMP, ampicillin; AST, antibiotic susceptibility testing; ETP, ertapenem; LAMP, loop-mediated isothermal amplification; pol-aAST, polymerase-accessibility AST; R, resistant; S, susceptible; TTPD, time-to-positive difference; $TTPD_{CT}$, TTPD control to treated; UTI, urinary tract infection.

majority of isolates in the CDC's Enterobacteriaceae Carbapenem Breakpoint panel, a collection of isolates designed specifically to challenge carbapenem-susceptibility tests in Enterobacteriaceae [94]. CRO, used broadly for a variety of infections because of its broad coverage and tolerability, was chosen as a representative third-generation cephalosporin. Similarly, ETP and MEM were chosen as clinically representative carbapenems [97]. When testing clinical samples, AMP was chosen because of its high resistance prevalence and thus availability of resistant samples (55.8% of clinical urine samples received by the UCLA CML are AMP resistant [98]). We chose ETP as a representative carbapenem.

The pol-aAST has two important requirements: (i) amplification conditions that are not fully lytic and (ii) release of NAs only from cells that are susceptible to the β-lactam to which they are exposed. If cells fully lyse, as they do in PCR, there is no difference in amplification between control and treated aliquots in susceptible isolates (Figure 2.2). It is only under partial-lysis conditions, as in LAMP, that cell integrity is preserved long enough to yield a substantial TTPD. Cell integrity, and rate and degree of lysis, will also depend on the identity of the organism, as well as

its growth rate. In partial-lysis conditions, most NAs are still protected inside cells in the control aliquot, whereas a significant portion of NAs are released and start amplifying immediately in the treated aliquot. We know from previous work [21] that the speed of an optimized bulk LAMP reaction makes it is difficult to linearly correlate TTP and NA concentration, unless very sensitive real-time measurements are made. Based on the magnitude of the differences in TTP observed here and the results measuring NA release (Figure 2.3), we suspect that both the state of NAs (inside intact cells versus inside or outside damaged cells) and the differences in concentration of free NAs contribute to the TTPDs observed. Cell-wall defects and damage are also likely to increase the penetration of amplification reagents into DNA trapped inside the remains of susceptible treated cells especially under the elevated temperature of the amplification reaction. We chose LAMP because we have shown previously that it is a rapid and specific isothermal amplification chemistry [21]. However, other non-lytic isothermal amplification chemistries could also be investigated. Additionally, DNA release (Figure 2.3) could be measured to determine susceptibility using PCR if combined with a filtration step; we have not evaluated the pros and cons of this approach in this paper. Lastly, alternative or modified accessibility-based AST approaches will likely need to be developed for different organisms, as we have done for *Neisseria gonorrhoeae* [23].

To demonstrate the flexibility of the pol-aAST method and the simplicity of the workflow, we investigated 2 approaches for determining susceptibility. The first, measuring $TTPD_{CT}$, gave 100% categorical agreement and uses a standard antibiotic-exposure step wherein one aliquot serves as the control and the other aliquot is exposed to an antibiotic. The second, measuring $TTPD_{LT}$, differs in that only a single aliquot of the original sample is used during the antibiotic-exposure step. After exposure, this aliquot is compared with a fully lysed control aliquot, which could be extracted at any point during the assay. Using only a single aliquot of the original sample during exposure reduces the challenges of fluid handling and metering, which will be valuable when developing fully integrated devices. When using a control and treated aliquot, both aliquots must have precisely metered volumes, and the heating required during exposure must be performed on both aliquots. Both methods showed excellent categorical agreement with gold-standard broth microdilution, and the choice of approach will be dictated by future device architecture.

To illustrate the value of phenotypic approaches, we evaluated pol-aAST using isolates that tested negative for β-lactamase genes and isolates that lack a predic-

tive genotype (e.g., no β-lactamase production, no modified porins, no modified penicillin-binding proteins), based on published and commercial genotypic assays [95], and CDC classification based on the ResFinder database [99], respectively. The antibiotic susceptibility of isolates lacking β-lactamases cannot be detected by current, FDA-approved genotypic methods, yet bacteria that do not produce β-lactamases can constitute 11% to 71% of CRE infections [36, 100, 101]. Using pol-aAST, all 5 of these isolates were correctly categorized as resistant.

Sample-to-answer time directly reflects the speed of diagnostics in practice and is a major factor in how likely a diagnostic is to be adopted. In general, the shorter the sample-to-answer time, the more valuable the test is and the more feasible for use at the POC. With urine as the contrived sample matrix, pol-aAST was able to be completed in <30 min. This timescale is on par both with suggested time-frames for rapid POC diagnostics [91, 92] and measured times of patient visits [102]. Additionally, because urine involves relatively simple sample-handling steps, we were able to perform four ASTs in parallel when testing contrived samples. The ability to run several samples in parallel demonstrates the potential to multiplex multiple antibiotics, which will be important for the next steps, including the design of integrated devices.

We have demonstrated direct testing of 25 clinical UTI samples using the pol-aAST with changes to the workflow (see Methods). However, even with the heterogeneity of clinical urine specimens (see urinalysis in S2 Table of Schoepp et al. 2020 [24]), including 2 polymicrobial samples that were correctly classified as ETP-S, the pol-aAST demonstrated good agreement with gold-standard broth dilution. The ability to handle polymicrobial samples was predictable based on the molecular specificity of NA-based methods. We expect this work to set the foundation for future improvements when using clinical samples.

The pol-aAST method demonstrates a rapid NA-based phenotypic AST for β-lactams and CREs. As with any academic report of an innovative diagnostic technology development, this work has limitations in the breadth of its scope and level of technological maturity. The following work would further extend the clinical applicability of this study and will be necessary for translation into a system suitable for regulatory approval and clinical use. First, the pol-aAST needs to be further developed and evaluated with fresh clinical urine samples from patients; here, we have used chemically preserved samples that were 3 to 5 days old, which likely decreased the response time of bacteria to antibiotics. We expect fresh clinical samples to show

more rapid and consistent responses; this hypothesis remains to be tested. We note that many state-of-the-art phenotypic AST methods are initially published without validation of performance directly on clinical samples, e.g., a recent breakthrough demonstrating phenotypic AST on isolates and on blood cultures [88]. Urine is a relevant matrix for a CRE diagnostic because UTIs are the most common source of CRE isolates [103], and because of the large number of hospital-acquired infections that involve catheters or other long-term indwelling medical devices [42], where CRE infections cause major problems. Second, to expand the scope of this approach, other sample types such as blood and blood cultures should be tested (in combination with appropriate pathogen-isolation and pathogen-enrichment technologies). Third, only categorical (S/R) agreement with the gold-standard method was tested here. While in the majority of cases a rapid categorical AST is clinically actionable, testing samples with a range of minimum inhibitory concentrations (MICs), including those with intermediate resistance, would further broaden the scope of the method. Fourth, we have not tested pol-aAST against heteroresistant samples. However, these are more common in gram-positive organisms [104] and are not common in gram-negative organisms. Fifth, the pol-aAST chemistry should be integrated with microfluidic devices so the AST can be performed directly on clinical samples with minimal user intervention. Sixth, the performance of these integrated devices will need to be evaluated in preclinical and clinical studies.

We emphasize that the specific pol-aAST described in this paper, just like other innovative rapid ASTs [21, 105–108], is not intended to be the sole test to guide treatment. Even though pol-aAST is based on detection of pathogen-specific NAs and can therefore provide pathogen ID, we anticipate that in a clinical workflow pol-aAST would be performed after a separate rapid pathogen ID step [48, 49, 51]. This ID step would then allow an unambiguous choice of the appropriate rapid AST. Furthermore, pol-aAST would likely be combined with rapid AST for other antibiotics, such as fluoroquinolones that can be used to treat CRE infections. AST methods that rely on similar underlying chemistries are more likely to be successfully integrated together. Isothermal amplification of pathogen-specific NAs appears to be a promising approach for AST, and we have already shown how a rapid fluoroquinolone AST can be performed in 30 min using digital LAMP [21]. Integration of pol-aAST with these complementary methods and translation to a distributable diagnostic will enable (i) improved antibiotic stewardship by reducing empiric use of carbapenems for Enterobacteriaceae, (ii) improved patient outcomes by detecting CRE infections for which carbapenems would be ineffective, and (iii)

more cost-effective surveillance of CRE outbreaks.

We envision that exploratory and mechanistic research inspired by pol-aAST will lead to a new generation of AST diagnostics. Additional mechanistic studies, such as those involving visualizing bacterial response to antibiotics [109, 110], would clarify the effects of different antibiotics on the responses measured in pol-aAST for different pathogens. To evaluate whether pol-aAST can be broadened beyond CREs and β-lactams, these studies would include organisms with cell envelopes that differ from Enterobacteriaceae (e.g., gram-positives) and other antimicrobials that affect the cell envelope, such as antimicrobial peptides [111] or vancomycin. It would also be desirable to evaluate pol-aAST with more amplification chemistries, including modified LAMP assays [112, 113] and other isothermal chemistries [114–116], such as recombinase polymerase amplification (RPA), that are actively being developed and can be performed at lower temperatures. Ultimately, this new generation of AST diagnostics will be integrated with the rapid ID methods being developed [48, 49, 51] and with future rapid NA-based AST methods for additional antibiotics and pathogens. For example, we have developed the nuclease-accessibility AST (nuc-aAST) [23], which measures accessibility of DNA to nucleases and was used to perform a rapid test of antibiotic susceptibility on the fastidious organism *N. gonorrhoeae*. In contrast to the pol-aAST, the nuc-aAST enhances antibiotic-induced damage using surfactants after the antibiotic-exposure step and performs full cell lysis. Ultimately, to address the broad diversity of antibiotic-resistant pathogens, it is clear that integrated, multiplexed POC devices that incorporate multiple rapid phenotypic AST methods are needed. Innovative methods based on antibiotic-induced accessibility of NAs to enzymes are promising for generating such ASTs for multiple antibiotics and pathogens in an approach that is intrinsically compatible with other rapid AST methods [21] and with rapid pathogen ID [48, 49, 52, 53].

2.5 Materials and methods

Ethics statement

Remnant urine samples from patients with confirmed UTI were received by UCLA CML and released to the Caltech researchers under UCLA IRB #19–001098. The UCLA IRB waived the requirement for informed consent and/or assent and/or parent permission under 45 CFR 46.116(d) for the entire study. No identifying information was obtained by the Caltech team, and the research was determined to be exempt by Caltech IRB (applications #18–0858 and #19–0909).

Study design

The objective of this study was to develop a rapid phenotypic AST for β-lactams based on DNA accessibility to polymerase for use with Enterobacteriaceae. To calculate the sample size necessary to validate the method (Figs 2.4 and 2.5), the Methods and Equation 5 from Banoo and colleagues [117] were used as described previously [21], namely, we suspected that the specificity and sensitivity of the nuc-aAST method would be 95% with a desired margin of error of ±10%. Under these conditions, 18.2 (or 19) samples must be tested with the nuc-aAST method and compared to the gold standard. We performed 36 ASTs with isolates susceptible to the antibiotic being tested and 46 ASTs with isolates resistant to the antibiotic being tested.

Isolates, growth conditions, and antibiotic exposure conditions

We obtained 25 de-identified clinical isolates from the UCLA CML and the CDC's Enterobacteriaceae Carbapenem Breakpoint panel [94]. In the case of isolates obtained from the UCLA CML, MICs were determined as described previously [89]. Genotypic testing of the 2 *E. coli* and 2 *K. pneumoniae* isolates selected for their lack of known β-lactamase genes was performed by UCLA CML using a previously published assay [95] and separately at the Keck School of Medicine of USC using the FDA-approved Cepheid Xpert Carba-R test. Whole genome sequencing of the single *Ebs* isolate selected for its lack of known resistance genes was performed by the CDC [94]. All isolates were stored as glycerol stocks at -80°C. Glycerol stocks were streaked onto Trypticase Soy Agar with 5% sheep's blood (Becton Dickinson, Franklin Lakes, NJ) and grown overnight at 37°C or resuspended directly in liquid media. Prior to experiments, a small clump of cells was resuspended from plates or glycerol stocks in 2 mL Brain Heart Infusion Broth (BHI; Becton Dickinson) at 37°C + 5% CO_2 with 500 rpm shaking for 2 to 4 h until visibly turbid. OD_{600} of the cultures was then measured, and working cultures were prepared at an OD_{600} of 0.01–0.07 and grown for 50–145 min at 37°C + 5% CO_2 with 500 rpm. Working cultures were then diluted 10X into control and treated aliquots for antibiotic exposure. For validation experiments, antibiotic exposure was performed in 100 µL volumes consisting of 80 µL Mueller Hinton II Broth (MHB; Becton Dickinson), 5 µL nuclease-free H_2O (NF-H_2O), 5 µL 20X antibiotic stock solution, and 10 µL of working culture. In control aliquots, antibiotic stock solution was replaced with NF-H_2O. For filtration experiments, antibiotic exposure

was performed in 100 µL volumes consisting of 65 µL MHB (Becton Dickinson), 21 µL NF-H_2O, 4 µL 25X antibiotic stock solution, and 10 µL of working culture. In control aliquots, antibiotic stock solution was replaced with NF-H_2O.

Antibiotic stocks

CRO disodium salt hemi(heptahydrate) (Sigma, St. Louis, MO), ETP sodium salt (Research Products International, Prospect, IL), and MEM trihydrate (TCI, Portland, OR) were used to create 1.0 mg/mL antibiotic stock solutions in NF-H_2O based on manufacturer-reported purity, aliquoted, and stored at -80°C. AMP sodium salt (Sigma, St. Louis, MO) was used to create 10.0 mg/mL antibiotic stock solutions in NF-H_2O based on manufacturer-reported purity, aliquoted, and stored at -80°C. Aliquots were only thawed and used once on the days of experiments.

Comparison of amplification methods

In order to compare amplification using LAMP and PCR, *E. coli* isolates were exposed to 0.5 µg/mL ETP for 15 min. Samples were then transferred directly into either PCR or LAMP mix on ice. Amplification was started immediately. qPCR was performed on a Roche LightCycler 96 using SsoFast EvaGreen Supermix (BioRad, Hercules, CA); 10 µL reactions were used. 10% of the final reaction volume was template. Published primers targeting the 23S rRNA genes of Enterobacteriaceae were used [118] at a final concentration of 500 nM. Cycling conditions consisted of 3.0 min at 95°C, followed by 35 cycles of 95°C for 10 s, 60°C for 10 s, and 72°C for 15 s. Fluorescence was measured using the SYBR Green channel after each 72°C extension step. LAMP was performed on a BioRad CFX96 using the following conditions: 10 µL reaction volume containing 1X Isothermal Reaction Buffer II (NEB), 5 mM $MgSO_4$ (NEB), 1.4 mM dNTPs (NEB), 320 U/mL Bst 3.0 (NEB), and 2 µM Syto-9 (Thermo Fisher); 10% of the reaction volume was template. Primer sequences (designed to target the 23S rRNA genes of Enterobacteriaceae) and concentrations have been described previously [21]. Cycling conditions consisted of 2.0 min at 12°C (while lid was heating), followed by 120 cycles of 70°C for 10 s. Fluorescence was measured using the SYBR Green channel every 10 s (after each cycle). We also ran an analogous LAMP reaction in the absence of Tween-20 (which is normally present in Isothermal Reaction Buffer II; NEB), to test for a potential difference in lysis efficiency; however, the resulting reaction rates were substantially lower than when Tween-20 was included.

Filtration experiments

Filtration experiments were performed using *E. coli* isolates exposed to 0.5 µg/mL ETP for 15 min. Immediately following exposure, cultures were passed through 0.22 µm, 1.5 mL cellulose acetate centrifuge tube filters (Corning Costar Spin-X, Corning, NY). DNA retention by the filters was <7% when measured by quantifying purified Lambda phage DNA (NEB) before and after filtration. Quantification was performed using ddPCR (QX200, BioRad). In filtration experiments, 50 µL of sample was added to the filter and centrifuged for 4 min at 1,000 rcf. DNA was extracted from both the feed and filtrate using QuickExtract DNA Extraction Solution (Lucigen, Middleton, WI). Samples were diluted 10X into extraction buffer and extracted according to manufacturer instructions. The concentration of the single copy *E. coli uidA* gene was then quantified in the feed and filtrate extractions. The percentage of *E. coli* DNA in the filtrate was calculated as the filtrate concentration divided by the feed concentration. ddPCR was performed using QX200 ddPCR Supermix for EvaGreen (BioRad); 10% of the final reaction volume was template. Published primers targeting the *uidA* gene in *E. coli* were used [119] at a final concentration of 500 nM. Cycling conditions consisted of 5.0 min at 95°C, followed by 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s, with final dye stabilization steps of 4°C for 5.0 min followed by 90°C for 5.0 min.

pol-aAST validation with clinical isolates

For pol-aAST validation experiments, *E. coli* and *Enterobacter* spp. isolates were exposed to either 2.0 µg/mL CRO, 0.5 µg/mL ETP, or 1.0 µg/mL MEM. *K. pneumoniae* isolates were exposed to either 2.0 µg/mL CRO, 1.0 µg/mL ETP, or 1.0 µg/mL MEM. Some isolates were run multiple times on different days. If this was the case, the average $TTPD_{CT}$ and $TTPD_{LT}$ are reported for that isolate. All isolates were exposed to antibiotics for 15 min in 100 µL reaction volumes in 200 µL PCR tube strips. After 15 min of antibiotic exposure, 10 µL of samples were transferred as template to LAMP reaction mix (as described earlier) on ice in technical triplicate. Amplification was immediately started.

Timed sample-to-answer using contrived urine samples

Timed sample-to-answer experiments were performed in the same fashion as pol-aAST validation experiments, except with the following modifications. Following initial growth and measurement of OD, isolates were resuspended in fresh, never-

frozen, pooled human urine from healthy donors (Lee BioSciences). Additionally, a timer was started as soon as samples were added to the antibiotic exposure conditions. *E. coli* and *K. pneumoniae* isolates were exposed to 0.5 and 1.0 μg/mL ETP (respectively) for 13 min. The duration of 13 min was chosen to ensure that all handling steps could be completed within the first 15 min of the assay. Amplification was performed until all reactions reached a fluorescence value of 1,000 relative fluorescent units (RFU) or greater. Amplification was then stopped, and TTP values were copied into a spreadsheet pre-populated with formulas to automatically output susceptibility calls. The timer was stopped once a susceptibility call had been determined.

Testing of pol-aAST with clinical samples

UCLA CML performed urinalysis, confirmation of UTI, pathogen isolation and ID, and subsequent gold-standard AST using broth microdilution. Gold-standard AST results were sent to Caltech researchers on the same day samples were received. Enterobacteriaceae-positive samples were shipped at ambient temperature to Caltech in BD Vacutainer Plus C&S preservative tubes (Becton Dickinson, Catalog Number 364951) containing a boric acid preservative. The pol-aAST experiments were performed directly on these samples within 3–5 days of their collection at UCLA. Urine samples were first warmed up to 37°C without shaking for 30 min, to approximate temperature of freshly collected urine. Then, 30 μL of urine was diluted into 70 μL of Cation-adjusted MHB (BD) containing 0.1% Tween-20 (Teknova, Hollister, CA) and placed at 37°C with shaking at 750 rpm for 3 min. Samples were then centrifuged at 5,000 rcf for 2 min. The supernatant was removed, and the sample was resuspended in 100 μL of MHB. Samples were then incubated for 30 min at 37°C with 750 rpm shaking. Antibiotic exposure was performed in a final volume of 100 μL, after transfer of 20 μL of incubated sample to 80 μL of the exposure condition: 75 μL of MHB and 5 μL of 20X antibiotic stock solution in NF-H_2O for treated aliquots, or 75 μL of MHB with 5 μL of NF-H_2O alone for control aliquots. For measurement of ETP susceptibility, the exposure condition contained a final concentration of 1 μg/mL of ETP. Aliquots were incubated at 37°C with shaking for 20 min. For measurement of AMP susceptibility, the antibiotic-exposure condition contained a final concentration of 16 μg/mL of AMP, and aliquots were incubated at 37°C with shaking for 45 min. The control and treated aliquots were subjected to a set of dilutions to account for variable bacterial load of the samples and resolution within the working range of the LAMP reaction. Following dilution, 1 μL

of the control and treated aliquots was added to each LAMP reaction well. There were 3 technical replicates (3 LAMP reaction wells) for each condition (control and treated). We measured the TTP for the reactions at each dilution, and then selected the dilution that yielded a control TTP value later than 4.7 min. The TTP results from this dilution were used to calculate $TTPD_{CT}$ (and determine susceptibility). Samples with a $TTPD_{CT} > 0.25$ min were considered susceptible, while samples with $TTPD_{CT} \leq$ to 0.25 min were considered resistant. The susceptibility determination of the pol-aAST method was then compared to the gold-standard culture results obtained by the UCLA CML to measure assay performance.

Statistical analysis

Significance referenced in the text for Figure 2.2 were calculated using GraphPad Prism 8.0 software from an unpaired, two-tailed t test comparing the averages of 3 replicate Cq values of each control sample to each treated sample. A significance value of 0.02 was used for statistical significance. All percent release values (Figure 2.3) and TTPD values (Figs 2.4–2.6) were calculated using Microsoft Excel. Data were plotted using GraphPad Prism 8.0 software. Thresholds for determining susceptibility in $TTPD_{CT}$ and $TTPD_{LT}$ plots were set halfway between the lowest S and highest R values for each organism/antibiotic combination. For preliminary tests with clinical samples, we defined a $TTPD_{CT}$ of above 0.25 min for a susceptible determination; this value would be further defined in a subsequent larger-scale clinical trial.

2.6 Conflicts of interest

The technology described in this publication is the subject of a patent application filed by Caltech. RFI has a financial interest in Talis Biomedical Corp.

2.7 Acknowledgements

We thank Sukantha Chandrasekaran, Shelley Miller, Romney Humphries, Marisol Trejo, Catherine Le, and Lyna Chheang at the UCLA Clinical Microbiology Laboratory for providing isolates and clinical urine samples and for discussion of gold-standard practices. We thank Jennifer Dien Bard at the Keck School of Medicine of USC for performing Cepheid Xpert Carba-R tests. We also thank Natasha Shelby for help with writing and editing this manuscript.

Chapter 3

BULK FILTRATION AST

In this chapter, I describe a new form of accessibility AST that improves upon the method described in Chapter 2, wherein the accessible and inaccessible nucleic acids are physically separated. I also examine *in vitro* beta-lactam pharmacodynamics using this new method of bulk filtration AST.

3.1 The bulk filtration AST protocol

3.1.1 Overview of protocol of bulk filtration AST

As practiced in this thesis, bulk filtration AST comprises the steps illustrated in 3.1.

The first step of bulk filtration AST is exposure of live bacterial cells to a chosen dose of antibiotics. During the antibiotic exposure, cells which are susceptible to the antibiotic at the dose provided lyse and die. Cells resistant to the dosage undergo less or no lysis. A more detailed quantitative description of how often and how quickly cells lyse was worked out from experiments described in the below sections 3.2.

After a chosen duration of antibiotic exposure, the filtration process converts the original sample, or "feed" fraction, into at least two new fractions: the "filtrate"

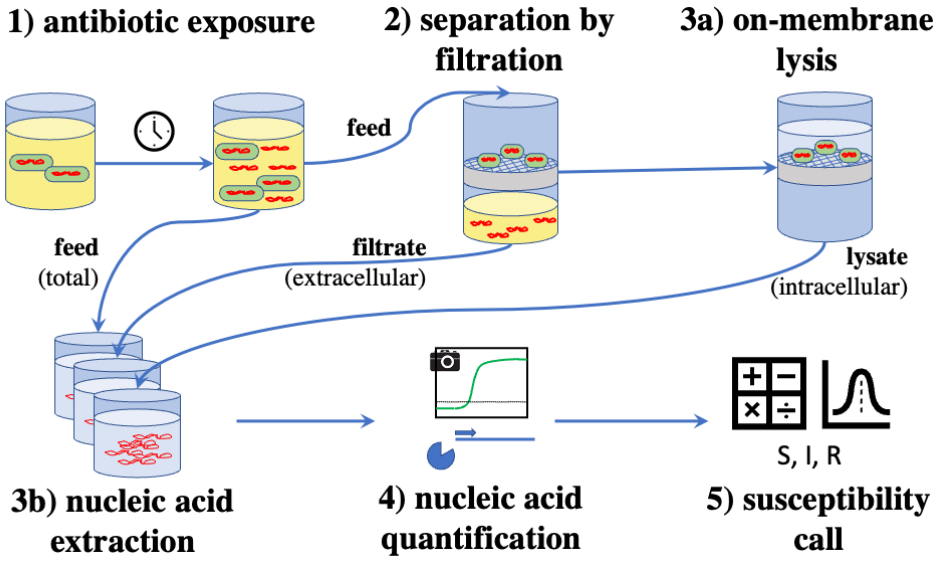


Figure 3.1: Schematic of the bulk filtration AST protocol.

containing extracellular nucleic acids and the "lysate" containing nucleic acids that were intracellular at the time of filtration. I will use the term "feed" to refer to any volume of the antibiotic-exposed specimen that is not filtered and which has its nucleic acids quantified, because a nucleic acid concentration measured in this way is generally representative of the concentration in the specimen prior to filtration (as discussed later in section 3.1.2). The liquid that passes through the filter is a fraction known as the "filtrate." The filtrate will contain freely-dissolved nucleic acids released into the extracellular fluid space by the antibiotic-induced lysis of bacterial cells that previously contained them. The second new fraction, the "lysate," is created by treating the filter cake of intact cells retained by the filter with a lysis buffer. Since in our experiments, the cake was not visible macroscopically and was only detected after the deliberate application of a lysis buffer to the filter membrane, I will call this fraction the "lysate" instead of the "cake" as is done in traditional chemical engineering terminology. Note that the filtrate and the lysate both contain the contents of cells that have been lysed—by antibiotics or by lysis buffer, respectively—and so would both fit the typical definition of a lysate as "a preparation of the products of lysed cells." However, in this thesis, I use the term lysate to refer to the material retained on the filter membrane only.

Each fraction, or at least two of the three fractions, undergoes a nucleic acid extraction by QuickExtract lysis buffer, followed by quantification by a nucleic acid amplification chemistry such as droplet digital polymerase chain reaction (ddPCR).

Lastly, the numeric data from the nucleic acid quantification are used by a susceptibility classification algorithm to generate either a final categorical susceptibility call, with the possibility of first calculating a real-valued susceptibility metric en route to the final categorical susceptibility call.

The rationale and specifications for each step of the protocol are now described in detail below.

3.1.2 Detailed protocol for bulk filtration AST and design rationales

3.1.2.1 Contrived clinical samples by bacterial culture

The contrived clinical samples used in all bulk filtration AST experiments comprised clinical isolates cultured in Brain Heart Infusion media (BD Diagnostics).

Up to a week prior to the experiment, a glycerol stock of the chosen strain was streaked onto solid agar media, either Luria Broth (BD Diagnostics, for *E. coli*

K12 only) or Tryptic Soy Agar + 5% sheep's blood (BD Diagnostics) and incubated overnight. The colonies were inoculated into 2 mL of Brain Heart Infusion (BD Diagnostics) in an amount that was undetectable by a portable cell density turbidometer (Biochrom Ultrospec 10) at 600 nm. This batch culture was incubated at 37°C with shaking at 300 rpm. The optical density was monitored approximately every 30 minutes until an OD_{600} between 0.2 and 0.5 was obtained. If the desired turbidity was reached, then the antibiotic exposure proceeded immediately, with a stopwatch already having been started. If proceeding to the antibiotic exposure protocol was not possible immediately after reading turbidity in a biological safety cabinet at a room temperature of approximately 22°C, then the batch culture was incubated uninterrupted at 37°C for at least 5 more minutes before a second attempt.

The isolates in Table 3.1 below were included in order to have both susceptible and resistant strains. As the mechanisms of resistance for the isolates from UCLA were not known, resistance mechanisms were not used to select the isolates.

Strain Name	ETP MIC	CRO MIC	Notes
E. coli K12 sub-strain MG1655	0.012, S (Etest)	≤1, S	Obtained commercially. MIC values reported in [120, 121]
E. coli #2 (UCLA Study 15-04A-02)	≤0.5, S	≤1, S	From urine on 12/29/2014.
E. coli #11 (UCLA Study 15-31-001)	≤0.5, S	16, R	Isolated 2015.
E. coli #38 (UCLA Study 15-31A-001)	>4, R	>32, R	From bile in 2017.
E. coli CDC AR Bank 0001 (#44)	8, R	>32, R	Enterobacterales Carbapenem Breakpoint panel. β-lactamases: KPC-3, OXA-1. Other resistance loci: aac(6')-Ib-cr, aadA5, ACRF, catB4, dfrA17, MDF(A), mph(A), sul1, tet(A), tet(R)

Table 3.1: Clinical isolates used in bulk filtration AST. *E. coli*=*Escherichia coli*, ETP=ertapenem, CRO=ceftriaxone, MIC=minimum inhibitory concentration. MICs obtained by broth microdilution unless noted. UCLA = University of California Los Angeles Clinical Microbiology Lab.

3.1.2.2 Antibiotic exposure

To begin the antibiotic exposure step of the bulk filtration AST protocol, a stopwatch was started when the batch culture of bacteria was removed from the incubator for

the last time. 500 µL of the culture was diluted 1:2 in BHI or cation-adjusted Mueller-Hinton Broth (MHB) growth media in a 1 cm cuvette, the batch culture was returned to 37°C, and then the batch culture OD_{600} was measured in the cuvette again at least three times. The density of bacterial CFUs/mL was calculated using a ratio of 1.5×10^8 CFU/1.0 unit absorbance. A serial dilution using the remainder of the sterile batch culture was promptly performed in pre-warmed MHB to yield a target density of 5×10^5 CFU/mL in the antibiotic exposure. Each antibiotic exposure comprised 14 µL of MHB, 1 µL of antibiotic dissolved in nuclease-free water, and 10 µL of the serial dilution of cells in MHB. In most of the exposures, the volume was mixed by repeated pipetting when the serial dilution containing bacteria was added. The time that the cells were first added to the exposure was rounded to the nearest second and recorded as the start of the exposure. The exposures were then vortexed and placed at 37°C, 700 rpm in a shaking block heater (Benchmark Scientific, MultiTherm Shaker or Eppendorf ThermoMixer C) for the remaining duration of the antibiotic exposure.

Antibiotic stocks were prepared by measuring out the ertapenem sodium (Research Products International, E32100) or ceftriaxone disodium hemiheptahydrate salt to at least two significant digits on a mass balance using spatula washed with ethanol and deionized water (Millipore), then dissolving in nuclease-free water to create a 1 mg/mL or 8 mg/mL stock. Only the weight of the antibiotic anion was considered, and the weight was scaled by the percent purity listed on the lot's certificate of analysis. This stock was filter-sterilized, separated into 100 µL aliquots, and frozen at -80°C. Aliquots were used within 24 hours of thawing.

In my experiments, I chose to expose cells to antibiotics by incubating cells in rich liquid media in which the antibiotics are dissolved. This arrangement is the most obvious choice for the following several reasons.

First, of the culture methods characterized for the majority of human pathogens, rich liquid media achieves the highest known rate of growth. In most environments supporting life, the growth of bacteria is limited by diffusion of aqueous nutrients to the cell surface, due to the small size of bacteria. Most bacteria are cylindrical rods <4 µm in length and about 1 fL in volume [122], and thus their environmental context exhibits a low Reynolds number. The turbulence of shaking during liquid media culture helps thin the narrow shell of nutrient depletion around planktonic bacteria and thus speeds their growth.

Antibiotic activity is generally faster when cells are actively growing and dividing.

This is because antibiotics generally work by disrupting a specific cell process, especially those involved in the Central Dogma processes of DNA, RNA, or protein replication. Quiescent cells have fewer active processes to disrupt beyond their maintenance of a membrane proton gradient, and so conventional antibiotics do not affect them.

Beta-lactam antibiotics target peptidoglycan cell wall turnover. A bacteria's rate of peptidoglycan turnover and synthesis is far greater when it is actively elongating or synthesizing a new septum during cytokinesis than when it is not changing its shape as a quiescent cell. Thus, beta-lactam antibiotic activity is also proportional to the rate at which a bacteria gains biomass.

Second, exposure to antibiotics in liquid media has already been adopted as a standard method for AST in the form of broth microdilution assays. Thus, in this thesis, the antibiotic exposures were designed to most closely mimic the CLSI standard conditions for broth microdilution: an inoculum no denser than 5×10^5 CFU/mL in cation-adjusted Mueller-Hinton Broth, shaking for at least 300 rpm at 37°C, with oxygenation (since the taxa tested are not strict anaerobes).

3.1.2.3 Separation and nucleic acid extraction

At the chosen time points, 10 µL of the antibiotic exposure was transferred, as the feed fraction, to an equal volume of QuickExtract™ DNA Extraction Solution (LGC/Lucigen QE09050, abbreviated as "DEB") and vortexed. The remainder of the antibiotic exposure was transferred to a 0.22 µm pore size filter unit and centrifuged to create the filtrate fraction. Earlier experiments utilized a cellulose acetate (Corning) membrane and were centrifuged at 2000 rpm for 2 minutes (Eppendorf A-2-MTP rotor), while later experiments used a hydrophilic polyvinylidene fluoride (Millipore-Sigma MultiScreenHTS MSGVS2210) membrane in a 96-well plate spun at at 2200 rcf (Eppendorf A-2-MTP) for 5 minutes. 10 µL of the collected filtrate was then vortexed with an equal volume of DEB. In certain exposures, the actual volume transferred differed from the target 10 µL, but always by a known volume.

Due to time constraints, the vortexed feed and DEB buffer was left at room temperature for 10–30 minutes while other steps were performed. At the earliest possible time, the feed and DEB was heated to 65°C for 6 minutes, then 98°C for 4 minutes, in a thermocycler with a heated lid (BioRad C1000). The sample was held at 4°C in the thermocycler, then transferred to a refrigerator or ice bucket as soon as possible

(both of which provide a temperature close to 4°C). Likewise, the vortexed filtrate and DEB was also incubated as soon as possible for 6 minutes at 65°C, 4 minutes at 98°C, and then held at 4°C. At the end of the day's experiment, the DEB extractions were frozen at -80°C to ensure stability of both DNA and RNA in future nucleic acid quantifications.

The timing of every action performed on the cells or nucleic acids during the above manipulations was recorded by the stopwatch. Qualified actions included pipette transfers; mixing or vortexing; heating, cooling, or other changes in environment temperature; or centrifugation (including brief un-timed spins to collect liquid at the bottom of tubes). For the feed fraction, the end of the exposure was set to be the time of vortexing in DEB, because it was presumed, without evidence, that cells would sense the harsher environment and down-regulate growth or die some short time later. The start of the centrifugation was recorded as the end of the exposure for that filtrate fraction measurement. It was presumed that no further increase or decrease in the extracellular nucleic acids would occur in the absence of living cells (and nucleases), and that the majority of the filtrate would be collected within the first few seconds of centrifugation. In some centrifugations of unwashed polyvinylidene fluoride filters interrupted at 1–2 minutes after reaching top speed, the filtrates were either mostly filtered or completely unfiltered, suggesting that the passage of fluid does occur suddenly, but that the commencement of fluid movement is variable and random.

Lucigen's proprietary QuickExtract buffer was employed for the nucleic acid extraction step because it does not employ any further physical separations. Other methods adsorb nucleic acids to a solid silica surface or partition them by liquid phases in phenol-chloroform extractions. QuickExtract buffer contains a surfactant and a proteinase and which stabilizes both DNA and RNA. Guanidinium thiocyanate is another reagent employed in nucleic acid extraction, but the compound interferes with downstream enzymatic reactions if it is not removed by another extraction method.

3.1.2.4 Nucleic acid quantification

In the bulk filtration AST experiments reported herein, nucleic acids were quantified by droplet digital PCR using the BioRad QX200 system. In 20 out of 30 amplification batches, the DEB extractions from the AST protocol were used directly as the templates in the ddPCR reaction. In the other 10 amplification batches, the

DEB extractions were used as templates in a reverse transcription reaction, and the reverse transcription reaction was the template of the ddPCR reaction. Each reverse transcription reaction comprised 1.45 µL of nuclease-free deionized water, 0.5 µL of 10X RapiDxFire buffer (LGC/Lucigen), 0.25 µL of 10 mM deoxyribonucleotides (New England Biolabs), 0.2 µL of primer (reverse strand only, final concentration 0.4 nM), 0.1 µL of 3 U/µL RapiDxFire™ Thermostable Reverse Transcriptase (LGC/Lucigen 30250-1), and 2.5 µL of the DEB extraction template. A master mix of all components except the templates was created and distributed to separate tubes. Individual templates were then added. The reactions were vortexed, then incubated in the following thermocycler protocol: extension at 60°C for 5 minutes, heat inactivation at 95°C for 5 minutes, and a hold indefinitely at 4°C.

Each ddPCR reaction comprised 10 µL of BioRad EvaGreen 2X ddPCR supermix, 0.8 µL of forward and reverse primer (final concentration 0.4 µm), 8.2 µL of nuclease-free deionized water, and 1.0 µL of template. A master mix of all components except the templates was created and distributed to separate tubes. Individual templates were then added. The reactions were vortexed, then formed into an emulsion in QX200 droplet generation oil using the QX200 droplet generator, eight wells at a time. The QX200 droplet generator creates 0.85 nL droplets. The droplet emulsion was then incubated in the following thermocycler protocol: enzyme activation and droplet stabilization at 95°C for 5 minutes; 40 cycles of melting at 95°C for 30 seconds, then annealing and extension at 60°C for 1 minute; a post-cycling cooling to 4°C for 5 minutes, 90°C for 5 minutes, then an indefinite hold at 4°C. The droplets were read within 4 hours of the end of cycling in the QX200 droplet reader using both FAM and HEX wavelength channels.

The droplet calls were made manually using the manufacturer's QX200 QuantaSoft software, which allows the counts of up to 4 labels to be exported at one time. The labels were chosen to be 1) "positive" droplets that belonged to the roughly Gaussian cluster of positive droplets 2) "negative" droplets that belonged to the roughly Gaussian cluster of negative droplets, 3) "rain" droplets located along a trajectory between the negative and positive clusters, and 4) "artifact" droplets with extreme high or low values in either channel. The threshold between "positive" and "rain" droplets was drawn using the FAM channel only. The threshold between "negative" and "rain" droplets was a sloped line using both channel values. The gating to assign labels was performed on groupings of the wells that showed similar positions of their positive and negative droplet clusters. The thresholds between

"positive" and "rain" droplets and between "negative" and "rain" droplets were therefore shared between wells of the same gating group. For example, it was common, especially when 48 or more samples were being amplified by ddPCR, for the two columns of 8 wells whose droplets were generated earliest to have a higher negative cluster fluorescence in the FAM channel than the rest of the wells. Evaporation of the droplet generation oil was observed to be most noticeable in these two columns of wells, due to up to 30 minutes passing between emulsion generation and thermocycling. Therefore, the first two columns were placed in their own groups, while the remaining columns were placed in a third group. Any wells with aberrant fluorescence averages were separated into their own groups, if positive and negative droplets could be discerned, or ignored.

The "*uidA*" primers (Forward: 5'-CAACGAACTGAACTGGCAGA-3', Reverse: 5'-CATTACGCTGCGATGGAT-3') used for ddPCR in the absence of reverse transcription were taken from Chern et al., 2011 [119]. This primer pair targets the beta-glucuronidase gene, which is located at position 1,695,547- 1,695,667 in the 4,641,652 bp *E. coli* K12 reference genome (NCBI Genbank U00096.3) [123].

The "23S" primers (Forward: 5'-GGTAGAGCACTGTTTTGGCA-3', Reverse: 5'-TGTCTCCCGTGATAACTTTCTC-3'), used for reverse transcription and ddPCR in amplification batches, targeted a 88 bp amplicon in all seven 23S ribosomal RNA genes in the *E. coli* genome and were adapted from Chern et al., 2011 [119]. Only the reverse primer was used for the RT reaction.

3.1.2.5 Data processing

The concentration of nucleic acids in the ddPCR reaction C_{PCR} was calculated from the number of "negative" droplets N_{neg}, the total number of "negative," "rain," and "positive" droplets N_{tot}, and the droplet volume $V_{\mathrm{drop}} = 8.5\times10^{-4}$ µL using equation 3.1. In this thesis, the "rain" droplets were considered positive (counted in N_{tot} but not N_{neg}. The "artifact" droplets were not counted in either N_{tot} or N_{neg}. The inclusion or exclusion of "rain"-labeled droplets in the negative droplets was found to have a noticeable influence on the final concentrations only in a few samples with very low "positive" or "rain" droplet counts.

$$C_{\mathrm{PCR}} = -\frac{1}{V_{\mathrm{drop}}} \ln\left(\frac{N_{\mathrm{neg}}}{N_{\mathrm{tot}}}\right) \tag{3.1}$$

The concentration of nucleic acids in the antibiotic exposure C_{expos} was calculated as $C_{\mathrm{expos}} = \frac{C_{\mathrm{PCR}}}{d}$, where d is the fraction of the exposure present in the ddPCR reaction.

For example, in one of the experiments shown in Figure 3.6b, the untreated sample at 120 minutes exposure underwent the following dilutions: 25 µL of exposure + 25 µL of DEB, a 2:20 dilution in water, a 2:800 dilution in water, 2.5 µL diluted DEB into a 5 µL RT reaction, a 1:50 dilution of cDNA in water, and a 1:20 dilution of template into ddPCR. The total dilution was 0.0000000625, or 1:16,000,000. Therefore, the C_{PCR} =212.02333 copies/µL inferred by ddPCR corresponds to $C_{\text{expos}} = 3.392 \times 10^9$ copies/µL. Subsequent analysis in this chapter in section 3.3 used this value of C_{expos}.

Unfortunately, the information of the number of each type of dilution is lost during the above data processing. Inclusion of this information may improve statistical modeling of nucleic acid amplification artifacts in the future, as mentioned in section 3.2.3.

3.1.3 Validation of filtration AST

In the included bulk filtration AST experiments, only the feed and the filtrate were measured. This choice was done because I believed that measuring the feed and filtrate only was sufficient, and because lysate measurements would be more affected by variation from manual operation during nucleic acid extraction than would the other two fractions. (The lysate is measured in experiments involving the digital filtration AST protocol, described in Chapter 4.)

To infer the lysate fraction outputs, it was assumed that the concentration in the feed fractions T should be the sum of the concentrations E and I in the filtrate and lysate fractions, respectively.

$$T = I + E \tag{3.2}$$

This equation only holds if the concentration of nucleic acids is high enough that stochasticity in molecule loading is negligible (see section 4.1.3). Furthermore, I cannot be solved for if other destinations for nucleic acids exist. In the actual bulk filtration AST experiments herein, the density of cells was higher than 1×10^7 CFU/mL, so stochastic loading of cellular nucleic acids is highly unlikely to be measurable.

Retention of nucleic acids by the filter was assessed in Figure 3.2 by measuring the difference before and after filtration of spike-in nucleic acid controls. Decreases and (smaller) increases were noted. Whether or not the observed decrease is due to filtration needs to be further verified. Other effects such as the retained volume of wash liquid (since the filters were washed with water to remove the manufacturer's surfactant coating), PCR inhibition by MHB media, improved PCR efficiency from

DEB, the effects of DEB surface tension on pipette volume, and other pipette volume errors also could appear as retention of nucleic acids on filter materials. In a separate experiment where only the lambda phage DNA spike-in in water was filtered, a difference of -6.7% and 2.4% (n=1, relative to 160 copies/µL unfiltered template) were observed by ddPCR for cellulose acetate and PVDF filters, respectively. Given these results as well as the results of the bulk filtration AST experiments in section 3.2 (i.e., the highest doses of ETP and longest exposure durations), the loss of bacterial nucleic acids in the filtrate from retention by the filter is likely to be $\leq 6\%$ for cellulose acetate filters, and close to 0 for PVDF filters, during the bulk filtration ASTs performed here.

Another deviation from equation 3.2 would be the destruction of nucleic acid molecules, which would appear in neither the filtrate nor the lysate fractions. Exogenous nucleases exist in the laboratory environment or may not be properly removed during fabrication of laboratory plasticware. Bacteria also can produce non-specific nucleases that are secreted or stored in the periplasm. In Figure 3.2, *E. coli* K12 cells and DNA and RNA spike-ins (or water) were incubated in 1 µg/mL ertapenem for 15 minutes, then filtered. The feed was extracted 1:4 into DEB, while the filtrates were not placed in DEB. One-step reverse transcription-ddPCR was performed using NEB WarmStart RTx placed into QX200 supermix. Exposures without a spike-in during the exposure were spiked with the same control during the preparation of the RT-ddPCR reaction. To facilitate comparison, the observed concentrations were corrected by the 0.025 dilution from the exposure (for feed fraction bacterial 23S and spike-ins during the exposure) or by the 0.1 dilution into the RT-ddPCR reaction (for all other conditions). There was no significant trend in the spike-in concentrations that correlated with cell number, suggesting that the bacteria do not non-specifically degrade extracellular (or intracellular) DNA or RNA. The small variation in spike-ins also argues against contamination by exogenous nucleases.

3.1.3.1 Differences between previous accessibility AST methods and filtration AST

As mentioned in section 1.3, the accessibility AST methods previously published relied on the topological, steric accessibility (or inaccessibility) of nucleic acids to dissolved reagents such as polymerases or nucleases. The separation of the intracellular and extracellular fractions occurred on a microscopic level of the individual cells distributed randomly within the volume of the assay system.

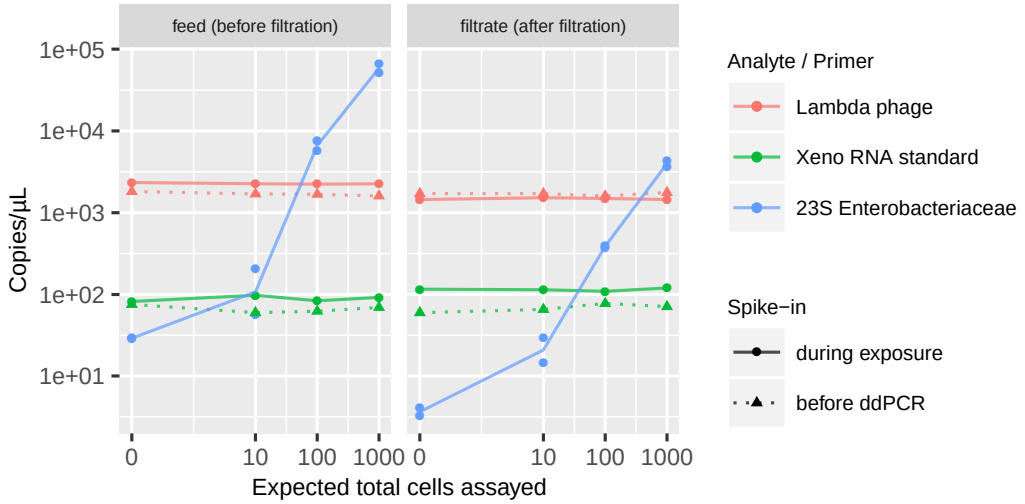


Figure 3.2: Spike-in control nucleic acids are not affected by the presence of bacteria.

Bulk filtration AST uses a physical separation process to spatially separate the intracellular and extracellular fractions into volumes that can be located and independently manipulated. Not only is the existence of the cell wall as a topological barrier taken advantage of, but the physical coupling of intracellular nucleic acids (and physical de-coupling of extracellular nucleic acids) is exploited when a separation process is performed. After an efficient separation, the cell wall of the bacteria is no longer necessary for maintaining the signal of susceptibility; unless one plans to perform additional separations later in a time series. Thus, performing a physical separation allows flexibility in how nucleic acids and bacteria are manipulated during the AST protocol.

Another fundamental difference from the published methods is that the separation process does not destroy or discard the nucleic acids of any of the specimen fractions, or at least the majority of each fraction. Being able to infer the total nucleic acid content of the specimen or replicate partitions of the specimen offers advantages that are discussed later in section 4.1.3.

3.1.3.2 Filtration improves the signal-to-noise ratio of "polymerase accessibility AST"

In Schoepp et al., 2019, no separation process is performed after the antibiotic exposure. This circumvents the time and complexity of performing an efficient separation process. However, the subsequent steps of nucleic amplification results in some background cell lysis. This lysis becomes indistinguishable from the true

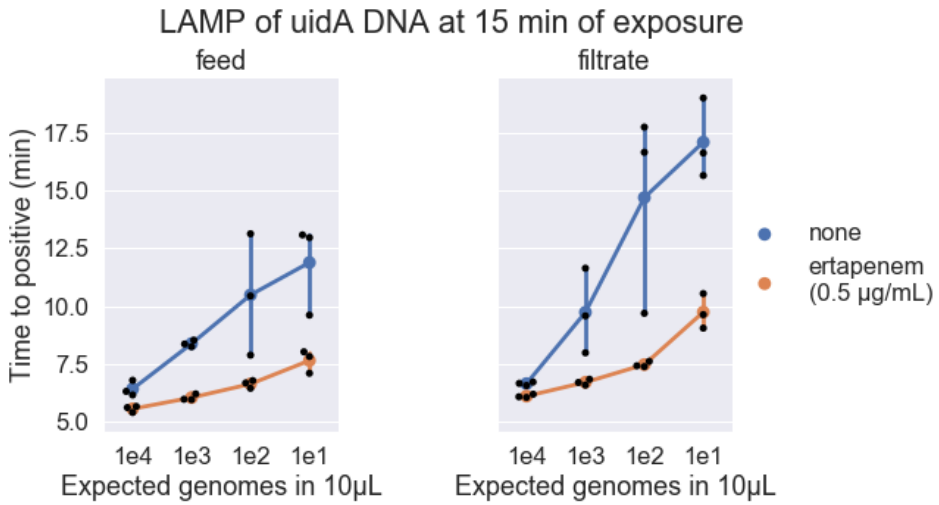


Figure 3.3: Bulk AST by LAMP ± filtration. Filtration reduces the background during LAMP pol-AST and proves that extracellular nucleic acids are sufficient for pol-AST signal.

beta-lactam signal and lowers the signal-to-noise ratio.

If Figure 3.3, pol-AST was performed using the antibiotic exposure without filtration or the filtrate of the same antibiotic exposures. With filtration, any intact cells are removed from the system prior to addition to the LAMP reaction and so cannot contribute to a background signal. A lower TTP indicates more template DNA is present and accessible. As the total number of cells in the exposure decreases, the background also decreases, and the signal-to-noise ratio increases. The low TTPD seen at the highest inoculum can be explained by the non-linearity of LAMP; the TTP of 5 minutes is already nearing the maximum speed at which LAMP can operate.

To help illustrate how sensitive LAMP is to the presence of template molecules, one can compare the LAMP TTPs to a similar experiment measured by ddPCR in Figure 3.4. The amplification by ddPCR of two different genomic DNA target loci shows a background of 0 copies/µL in the filtrate when 3.7510^{-6} CFU/mL of *Escherichia coli* K12 cells were exposed to 1 µg/mL of ertapenem antibiotic for 15 minutes. Meanwhile a signal is seen in the filtrate that represents about 15% of the feed fraction's DNA.

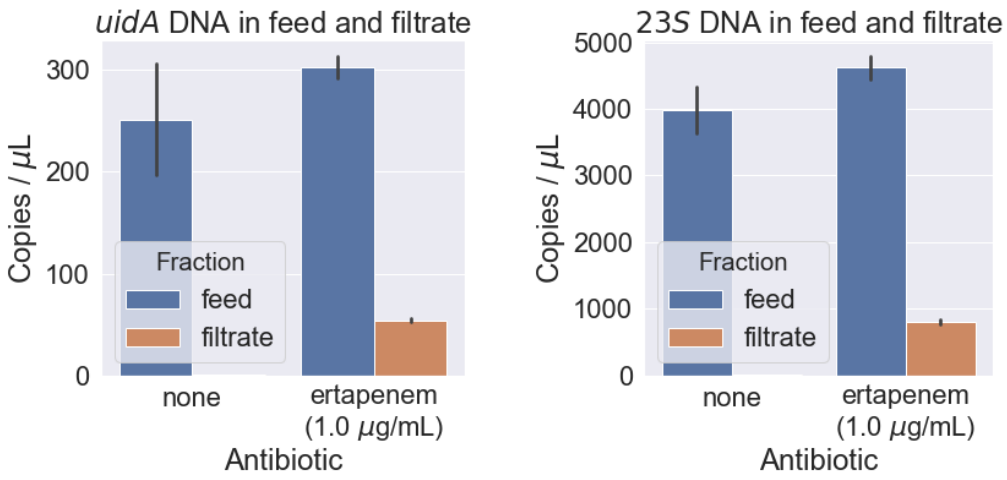


Figure 3.4: Bulk filtration AST does not depend on genome fragmentation and accessible bacterial DNA has an effective diameter of less than 0.2 µm.

3.1.3.3 Complete lysis of cells is sufficient to explain bulk filtration AST signal

The idea of polymerase accessibility was discovered when a beta-lactam treated and an untreated culture of *Escherichia coli* K12 were placed into qPCR and ddPCR reactions after heating to 70°C in Tris-EDTA buffer instead of a customary nucleic acid extraction kit. It was expected that the nucleic acid readout of the treated condition would be lower due to the cessation of growth [21]. However, in some of the experiments, the treated condition gave a higher signal (Nathan Schoepp, unpublished data). The same antibiotic exposures heated in DNA Extraction Buffer did not show a large difference between treated and control conditions in one experiment (n=3). In one experiment (n=3) increase was higher when ddPCR was performed instead of qPCR.

At this time, there were several possibilities considered. One hypothesis posited that the genome of the bacteria were fragmenting due to antibiotic stress. Since the *E. coli* genome possesses multiple copies of the 23S gene, a fragmented genome would yield more copies than an unfragmented genome in digital PCR. Another hypothesis proposed that the ddPCR polymerase entered the cell through increased permeability of the cell envelope. In this scenario, the openings in the cell envelope would be large enough to allow enzymes to pass, but small enough that the intracellular nucleic acids remained physically coupled to the rest of the cell's mass. In a third hypothesis, catastrophic lysis of the cell would create a hole large enough for the

intracellular nucleic acids to become physically uncoupled from the rest of the cell and accessible to polymerases.

The results of Figure 3.3 show that extracellular are sufficient for a robust accessibility signal in pol-AST. In the treated condition, a large TTP difference is present both with and without filtration. This difference already indicates that the majority of the signal remains after filtration, and so those template molecules are extracellular. Since LAMP is only semiquantitative, quantifying the amount of signal lost from filtration in this experiment should be interpreted carefully. Unfortunately, using a more precise non-isothermal protocol like qPCR or ddPCR would raise questions about lysis of cells from the multiple heating cycles instead of from antibiotics.

The TTP signal is slightly lower with filtration, especially at lower cell densities. The slight prolonging of TTP indicate that some of the nucleic acids that contribute signal during LAMP were physically coupled to the intact cells removed by filtration. But this does not prove the existence of accessible but intracellular nucleic acids. Instead, these nucleic acids could be inside intact cells prior to addition to the LAMP reaction in an inaccessible and intracellular state, and then convert into an accessible and extracellular state due to background lysis of cells at 70°C in the LAMP reagents. Indeed, the existence of background lysis not due to antibiotics is seen by the decrease in TTP caused by filtration in the untreated condition, a decrease similar in magnitude to that of the treated condition.

From this experiment it is not possible to prove that accessible but intracellular nucleic acids play no role in the effect measured by polAST. Since the accessible but intracellular effect supposedly can only exist during LAMP, one would need an experiment comparing the LAMP protocol without filtering to a condition in which intact cells are collected and the extracellular nucleic acids removed. The absence of signal would then show that being extracellular is necessary for nucleic acids to be accessible. Later experiments with the bulk filtration AST protocol in Figure 3.6l show that in the highest antibiotic doses and longest durations, the concentration of nucleic acids in the filtrate essentially equals that in the feed. By that point in time at least, the vast majority of all nucleic acids are both accessible and extracellular.

Figure 3.4 shows an attempt to measure the extent of genome fragmentation during bulk filtration AST. The same DEB extraction of a 15-minute bulk filtration AST were quantified using two different primer pairs without reverse transcription. In ddPCR, if multiple templates are physically coupled, they will only be loaded into one droplet. If the same templates were physically uncoupled, then more droplets

would turn positive after amplification.

The *uidA* primer pair amplifies one genomic loci located opposite of the origin of replication by sequence distance in the circular *E. coli* genome. Meanwhile, the 23S primers amplify an amplicon shared by the seven 23S rRNA loci found in the *E. coli* genome (mostly near the origin of replication by sequence distance). In a rapidly-dividing *E. coli*, genome replication requires more time than cytokinesis, and so a single cell dividing every 40 to 20 minutes may contain up to 4–8 origins of replication. Therefore, in each bacterium, there are up to ~7-fold more 23S templates in existence than *uidA* templates. If genome fragmentation occurs then the 23S signal would be expected to be higher than the *uidA* signal.

The genome of *E. coli* can fragment in two conceivable ways. The two strands of DNA can separate and fail to re-hybridize, creating two molecules from one. In addition, there could be breaks in the DNA strands that would separate different loci in the same genome. Additionally, in this protocol, there were two conceivable causes of genome fragmentation: antibiotic action, and the DEB extraction protocol. It is unlikely for the DEB buffer to cause genome fragmentation by strand breaks, since such non-specific DNA cleavage would eventually destroy all nucleic acids in the buffer despite its marketing as a nucleic acid preservative. Despite varying incubation times at 65°C during extractions of bacteria (data not shown), I have not observed any decrease in yield over time, ruling this mechanism out. However, it is plausible that the 98°C heating step does cause separation of complementary DNA strands. The genome fragmentation has been documented during exposure to fluoroquinolones, an antibiotic class that targets DNA gyrase. Beta-lactams do not target gyrase, but if there were any effect, it would be seen in Figure 3.4.

In actuality, the 23S primers resulted in a remarkably consistent increase in apparent concentration regardless of antibiotic exposure or extracellular location. This increase is consistent with the multiple copies of the 23S gene being physically uncoupled. However, another explanation is that some primer sequence characteristic causes this increase in efficiency. Even though digital PCR should be robust to slight primer inefficiencies, if the templates were stochastically prevented from participating in amplification, having multiple copies in the same droplet may increase the chance of successful amplification.

There is an increase (n=1) between the untreated to treated feed fractions. Interestingly, the same increase in the total DNA measured in the treated feed fraction is observed repeatedly in ertapenem bulk filtration ASTs in Figures 3.6a-3.6c, 3.6g,

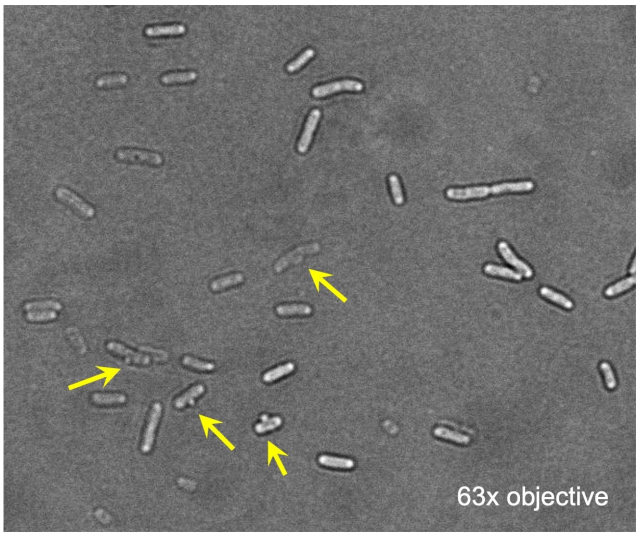


Figure 3.5: Light microscopy of unfixed *E. coli* exposed to 1.0 µg/mL ertapenem for 15 minutes shows large, midline cell envelope defects (yellow arrows).

3.6h, and 3.6j at time points earlier than than 30 minutes. The magnitude of this early increase is within the range of measurement error, but the bias away from 50% increases indicates a real phenomenon. This trend could be explained by the failure of DEB to lyse all cells and by antibiotic-induced lysis boosting the extraction efficiency. In Figure 3.4, the magnitudes of the increase and the amount of extracellular DNA coincides, consistent with this hypothesis. Alternatively, the presence of DNA strand breaks also cannot be definitively ruled out. In this case, only the fragmentation must be partial, since the increase is not close to the expected 1:7 ratio. In any case, the trend eventually reverses when the faster, non-negative cell growth in the untreated condition outpaces the increase from beta-lactam treatment.

The literature reports that catastrophic lysis is readily observed during live cell microscopy for several beta-lactam antibiotics [28], and the biophysics of this lysis have been analyzed in detail [124]. Figure 3.5 also shows the results of microscopy of a 15 minute ertapenem exposure performed in the same facility as all other experiments in this thesis. The midline blebs reported in the literature were observed, suggesting that the same conditions studied by live microscopy were also present during bulk filtration AST.

None of the three proposed mechanisms for the bulk filtration AST signal have been definitively ruled out with the experiments included above. However, the mechanism of catastrophic lysis must be ruled in. It is sufficient to explain all of the above data, and it is in line with the majority of the literature of beta-lactam mechanism of

action. The other two mechanisms require more assumptions than the mechanism of catastrophic lysis, and they are at odds with the existing literature. Therefore, I believe that lysis of cells, with openings in the cell wall larger than any internal molecule or particle, is the most likely, or at least the largest, mechanism by which accessibility AST operates.

3.2 Bulk filtration AST reveals pharmacodynamics of beta-lactam antibiotics at short time scales

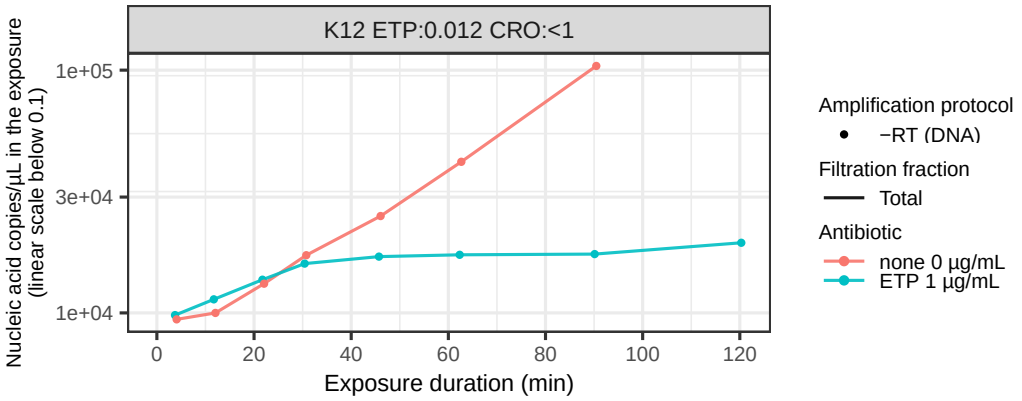
A single bulk filtration AST condition yields one set of observations of the amount of live and dead biomass at one time. One can use bulk filtration AST to understand the state of the *in vitro* culture system, aside from its use as a diagnostic assay. To better understand and optimize tradeoffs in assay design, 1186 ddPCR measurements of 507 independent bulk filtration AST exposures, grouped into 18 AST runs and 30 ddPCR experiment batches, were performed on 5 strains of *Escherichia coli* and 2 antibiotics. For each measurement, the choice of antibiotic compound, the antibiotic dose, the duration of the antibiotic exposure, the starting number of bacteria (the inoculum), and inclusion of reverse transcription were controlled by the experimenter. The mapping between ddPCR The results are discussed below.

3.2.1 Pharmacodynamics at short time scales is a balance between cell growth and cell death

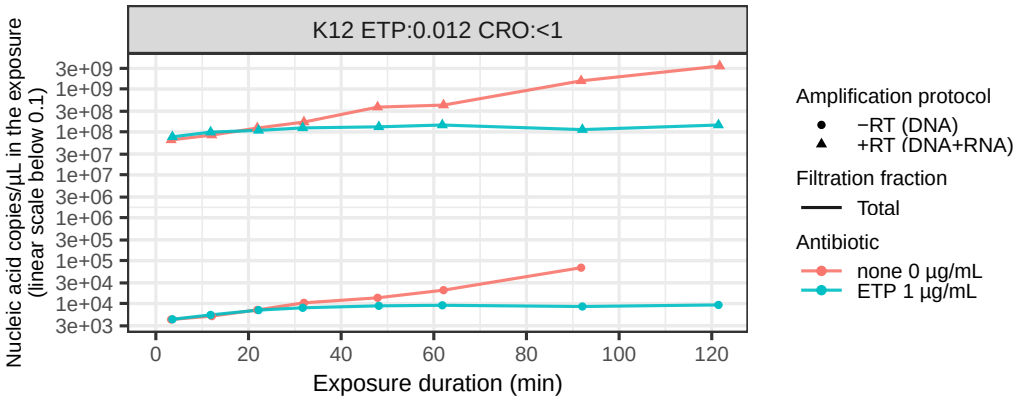
Phenotypic ASTs, by their definition of being phenotypic, involve the culturing of live bacteria in antibiotics, usually in controlled conditions that promote maximum growth rates and thus fastest antibiotic action. Thus, during phenotypic ASTs, there is necessarily continued population growth while antibiotic killing commences. Studies in the literature not focusing on pharmacodynamics often ignore this balance of growth and death [125].

In antibiotic treated exposures where the bacteria were susceptible, an initial increase in the total nucleic acids and a lag in the extracellular nucleic acids was observed. Eventually, total population growth slowed and both fractions plateaued at nearly the same concentrations, indicating the cessation of population growth.

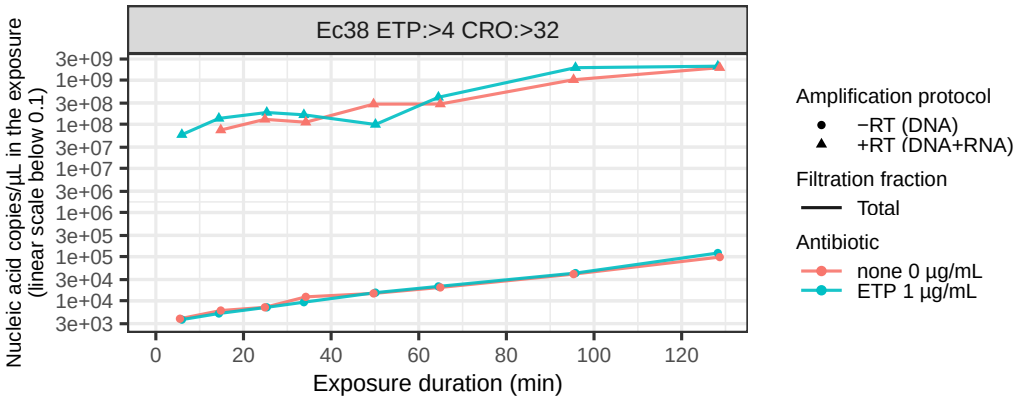
In exposures lacking antibiotics, the population of untreated cells increased exponentially. In 8 out of 11 relevant AST runs, no lag phase was seen despite a transition to a different rich growth media and a room temperature interruption in the 37°C incubation. A lack of an increase in the total nucleic acids in the initial 1–2 time points was seen in 3 AST runs and could be explained as a lag phase or as



(a) AST Run 1. No filtrate fraction was collected.

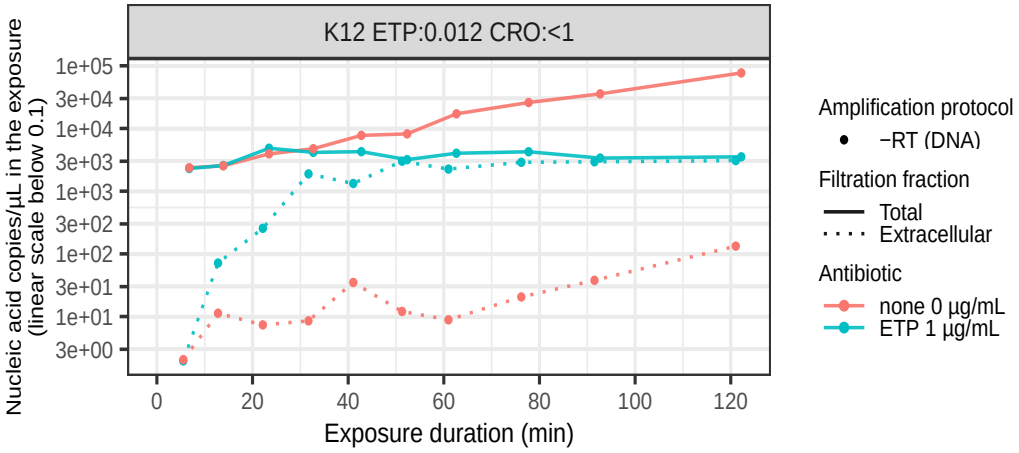


(b) AST Run 2. No filtrate fraction was collected. Both RT and no RT shown.

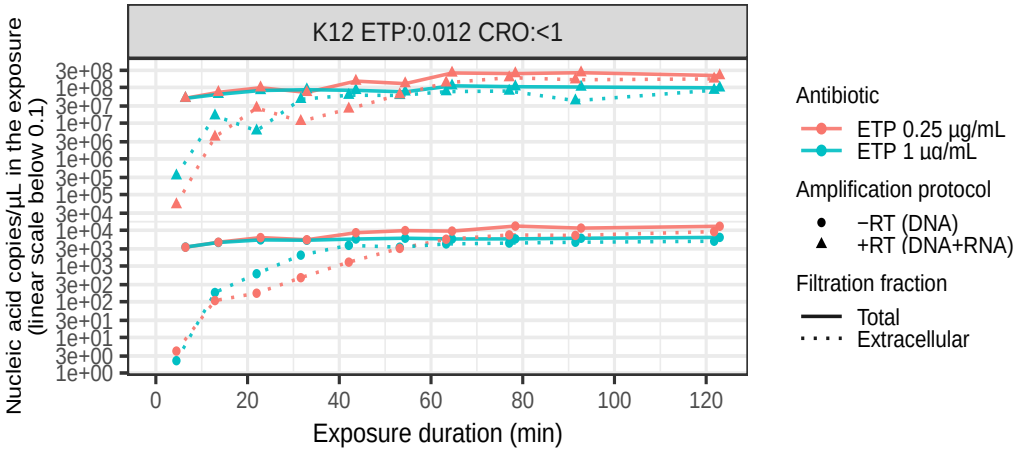


(c) AST Run 3. No filtrate fraction was collected. Both RT and no RT shown.

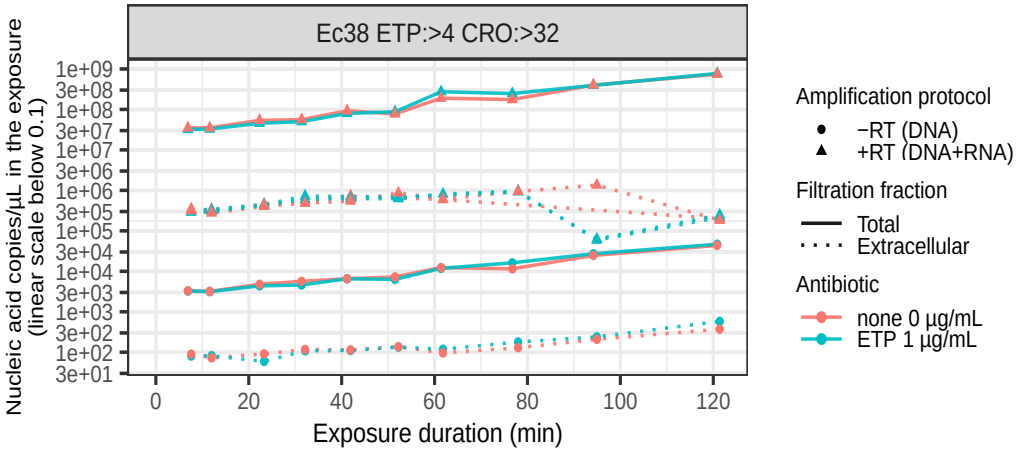
Figure 3.6: Bulk filtration AST measures total and extracellular gDNA or rRNA over time during *in vitro* antibiotic exposure.



(d) AST Run 4.

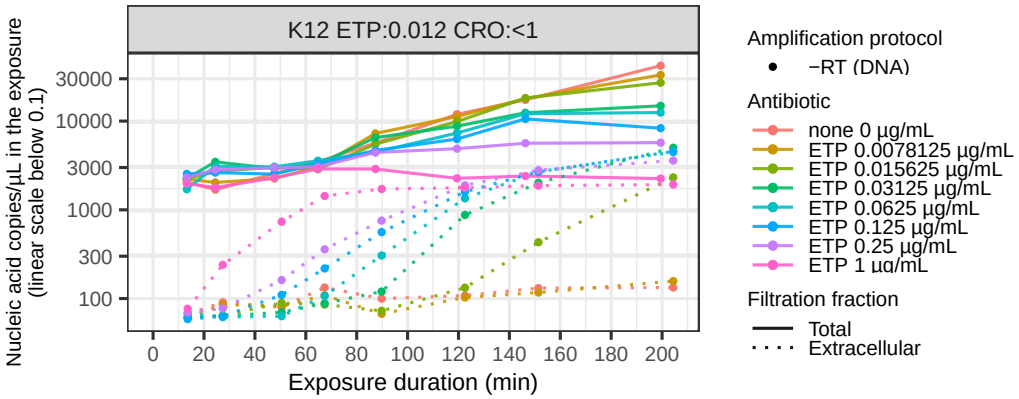


(e) AST Run 5. No untreated condition was included. Both RT and no RT shown.

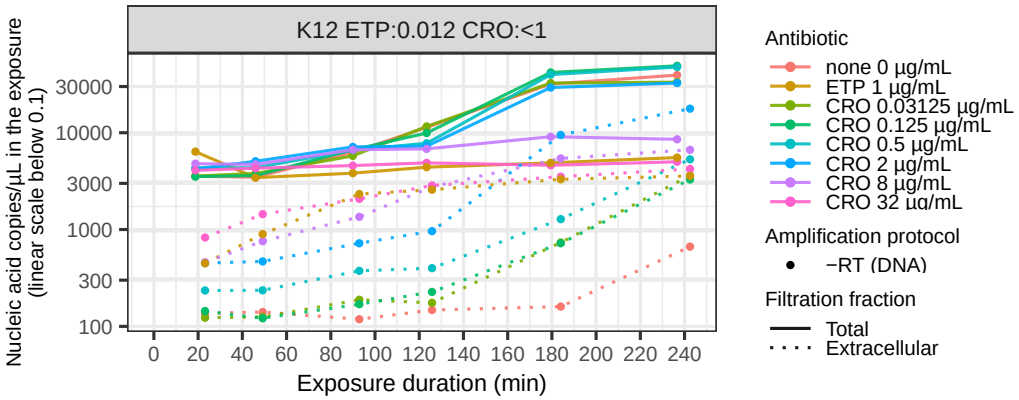


(f) AST Run 6. Both RT and no RT shown.

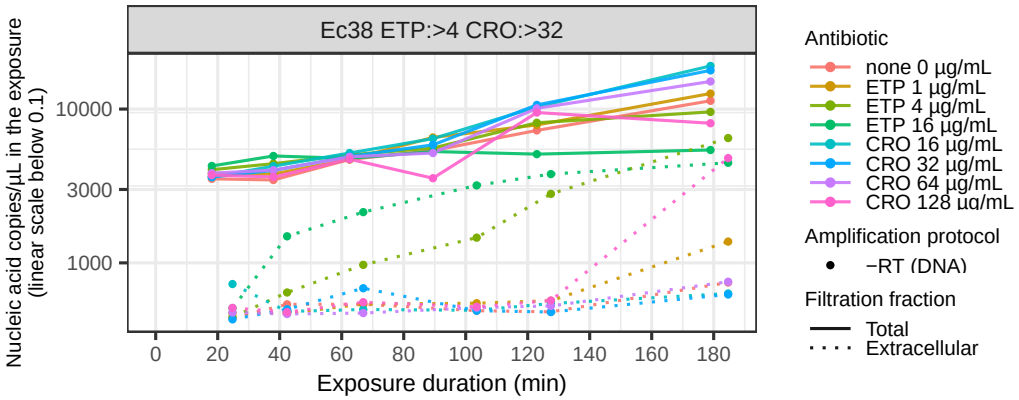
Figure 3.6: Bulk filtration AST measures total and extracellular gDNA or rRNA over time during *in vitro* antibiotic exposure.



(g) AST Run 7.

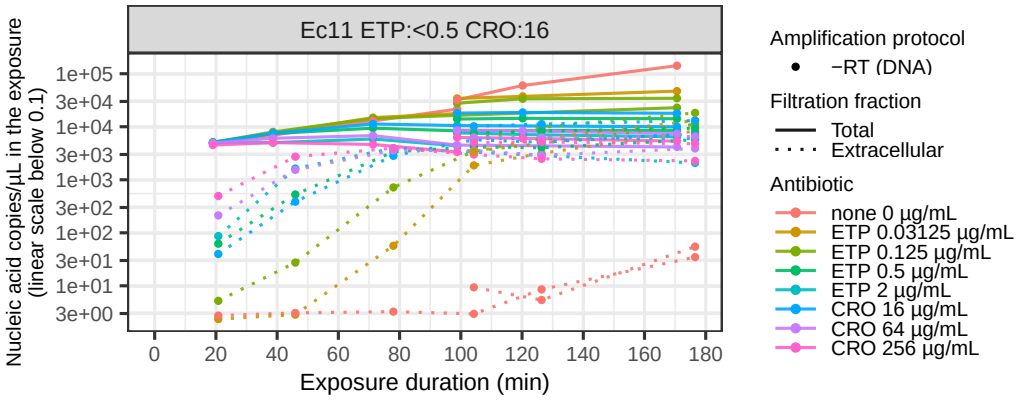


(h) AST Run 8.

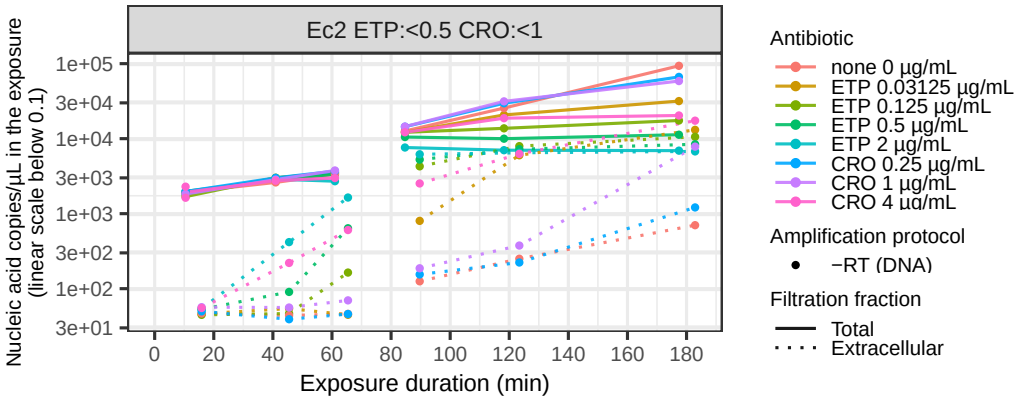


(i) AST Run 9.

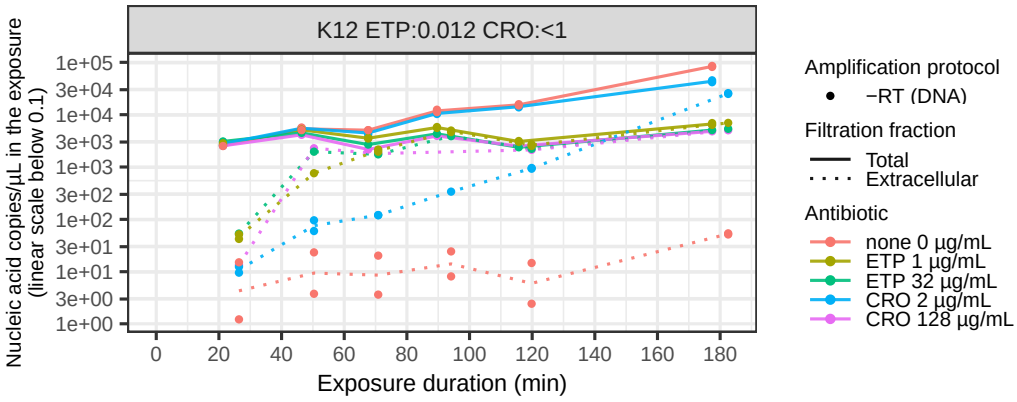
Figure 3.6: Bulk filtration AST measures total and extracellular gDNA or rRNA over time during *in vitro* antibiotic exposure. Lines connect points that were observed during the same PCR run; each point is an independent antibiotic exposure.



(j) AST Run 10.



(k) AST Run 11.



(l) AST Run 12.

Figure 3.6: Bulk filtration AST measures total and extracellular gDNA or rRNA over time during *in vitro* antibiotic exposure. Lines connect points that were observed during the same PCR run; each point is an independent antibiotic exposure.

measurement error. Slowing of population growth at longer durations, which would be expected during entry into stationary phase or diauxic growth, was not observed within the examined durations of exposure.

In Figure 3.6, the total nucleic acids increase before plateauing in all treated conditions with susceptible strains. The increase is initially comparable to the rate of untreated conditions. Thus, to model bulk filtration AST it will be necessary to assume continued growth by surviving bacteria.

3.2.2 Lysis efficiency and background lysis

When comparing untreated and susceptible treated conditions, a slight increase in the total amount of nucleic acids was observed in the latter in early time points. One explanation for this effect is that lysis from beta-lactams contributes to a higher nucleic acid extraction efficiency than from the lysis buffer alone. If true, it follows that the lysis buffer's efficiency is less than 100%. The uncertainty in the lysis buffer efficiency will create a bias in the estimation of pharmacodynamic parameters, but since the magnitude of the lysis efficiency should be small, the effect was not included in the model. An alternative explanation is that the presence of antibiotic salts changed the surface tension of the antibiotic-exposed specimen and increased the volume drawn into the tips of micropipettors. In either case, the effect was not modeled and allowed to be part of the stochastic noise term.

In untreated exposures, the amount of extracellular nucleic acids was always $\leq 1\%$ of the total nucleic acids, with the amount increasing over time, paralleling the population growth. Since no antibiotics were present, the amount of extracellular nucleic acids must be interpreted as a background lysis rate. Certain strains such as *E. coli* #38 consistently had a higher background lysis rate than other strains such as *E. coli* K12.

The nature of this background lysis rate was not determined in detail. It may represent the small rate of death and then subsequent cell envelope degradation from stochastic reasons during cultivation. The background lysis rate may also represent a physiological extrusion of nucleic acids by living cells. Gram-negative bacteria, such as the well-studied *Pseudomonas aeruginosa*, are known to secrete DNA extracellularly (eDNA) as part of biofilm formation [126]. In *Escherichia coli* and *Klebsiella pneumoniae*, only certain strains produce biofilms, and protein and carbohydrate polymers form the bulk of the extracellular polymeric substance [127]. At least to the naked eye in the short time scales examined, only planktonic growth

was seen in the strains examined. It is claimed in the literature that many bacteria secrete extracellular vesicles that may contain ribosomal RNA; these extracellular vesicles would also contribute to an apparent background lysis rate. No matter the mechanism, the parameter for extracellular nucleic acid production in the absence of antibiotics is captured in our models as a background lysis rate.

3.2.3 Cell density is directly proportional to gDNA and rRNA amplification, without inoculum effects

When the starting inoculum of the antibiotic exposure was varied by a serial dilution, the observed concentration of the targeted nucleic acid species measured after the antibiotic exposure was found to vary according to a linear function. Furthermore, the proportion of extracellular to total nucleic acids remains almost constant as initial inoculum varies and the antibiotic dosage and exposure durations were held constant.

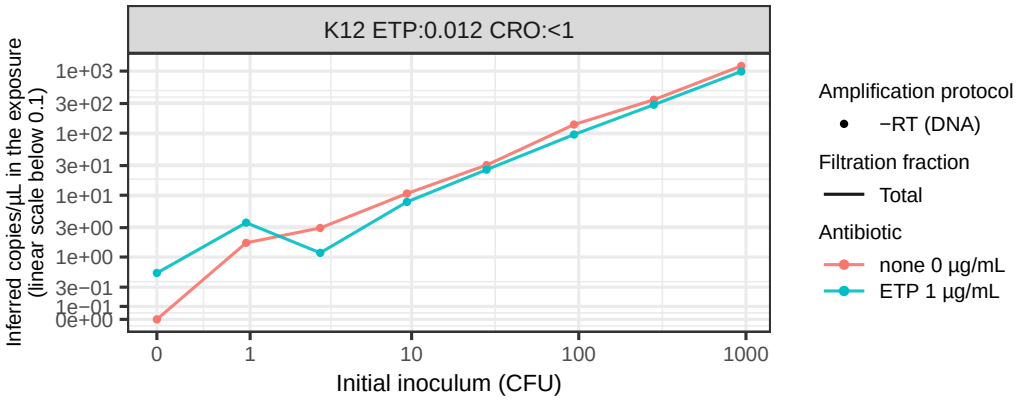
These two relationship are expected if the copy number of the targeted nucleic acid does not change in response to antibiotic exposure and if the magnitude of the response to antibiotic is not affected by the density of cells at the start of the antibiotic exposure. The latter condition is discussed in the next section (3.2.4).

The first condition is a result of the linearity of expectations and the law of large numbers. The amount of variability in the copy number per cell of a nucleic acid species does not affect our prediction that the concentration of nucleic acids we measure, y_{obs}, is linearly proportional to the number of cells N. So long as the copy number per cell x_i is finite (which it is) and that the offspring of bacteria remain similar enough to be in the same species (which they do) so that the x_i have the same distribution, then there will exist an average copy number per cell $\overline{x}$. Our bulk measurement, normalized by N, will then always measure that average copy number per cell in the limit of large inoculums, thanks to the law of large numbers. (The copy numbers of both of these nucleic acid species are generally proportional to the size of the cell anyways, and their variance is finite too.) In other words,

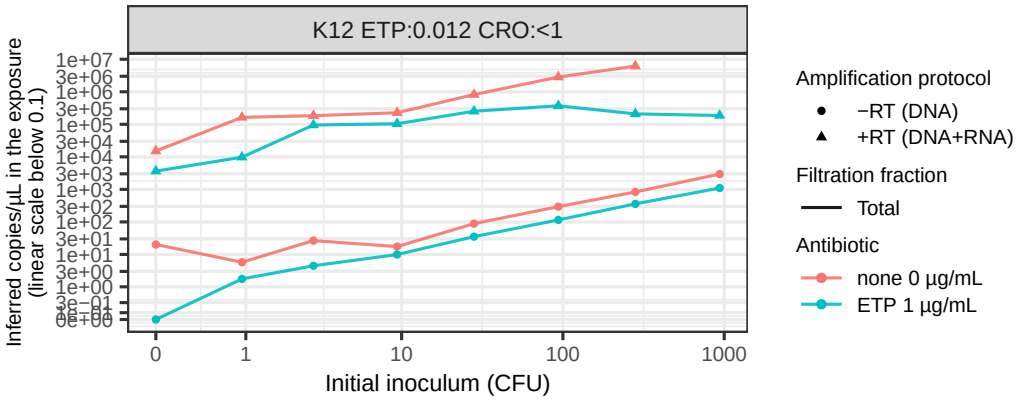
$$E\left[y_{\text{obs}}\right] = E\left[\sum_{i=1}^{N} x_i\right] = \sum_{i=1}^{N} E\left[x_i\right] = \overline{x}N \tag{3.3}$$

$$\lim_{N\to\infty} \frac{y_{\text{obs}}}{N} = \lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_i}{N} = \overline{x}. \tag{3.4}$$

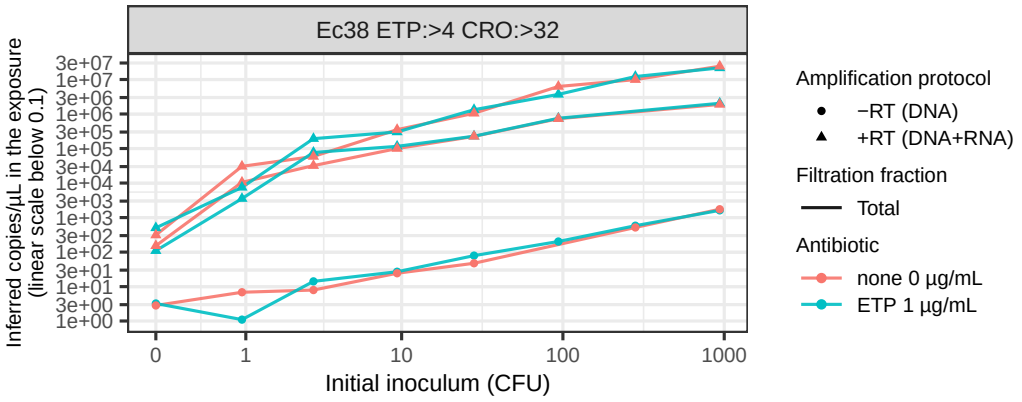
However, if the average copy number per cell changes as a result of antibiotic exposure, then $\overline{x}$ becomes a variable, and y_{obs} becomes a function of both $\overline{x}$ and



(a) AST Run 13. Filtrate fraction was not collected.

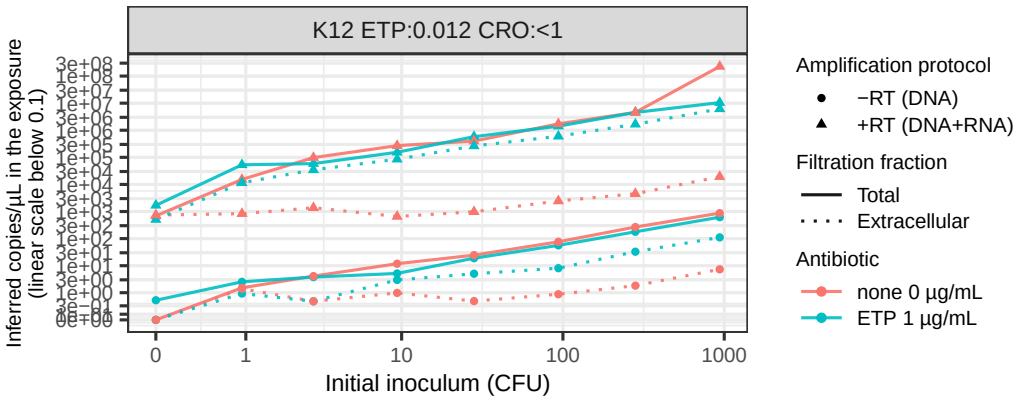


(b) AST Run 14. Filtrate fraction was not collected. Both RT and no RT shown.

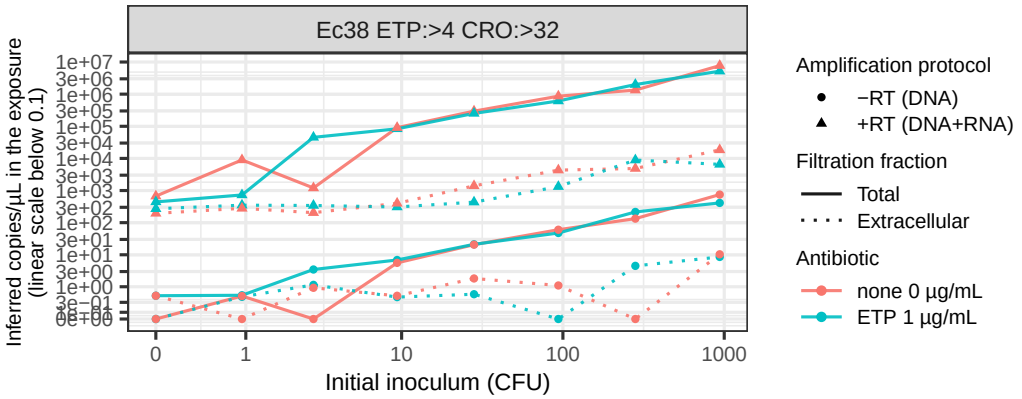


(c) AST Run 15. Filtrate fraction was not collected. Both RT and no RT shown.

Figure 3.7: Total and extracellular gDNA or rRNA in bulk filtration AST as a function of starting inoculum.



(d) AST Run 16. Both RT and no RT shown.



(e) AST Run 17. Both RT and no RT shown.

Figure 3.7: Total and extracellular gDNA or rRNA in bulk filtration AST as a function of starting inoculum (continued).

N. The two nucleic acid species targeted in Figure 3.7 were the *uidA* locus in the genomic DNA and the 23S ribosomal RNA. An increase or decrease in the number of genomes per cell as a response to beta-lactam antibiotics has not been described in the literature. Likewise, there is no report of the average number of ribosomes in a bacterial cell changing with beta-lactam exposure. I found it reasonable to assume that neither gDNA nor rRNA copy number per cell were affected by antibiotic dosage. Even if this were the case, the amount of increase would have to coincide exactly with the magnitude of any inoculum effect present in order to achieve the observed results. The principle of parsimony (Occam's Razor) would lead us to prefer the former hypothesis.

If the observed nucleic acid concentration was perfectly proportional to the bacterial biomass, then the linear function between the two would have no intercept. In actuality, at low inoculums, the measured concentrations for certain experiments

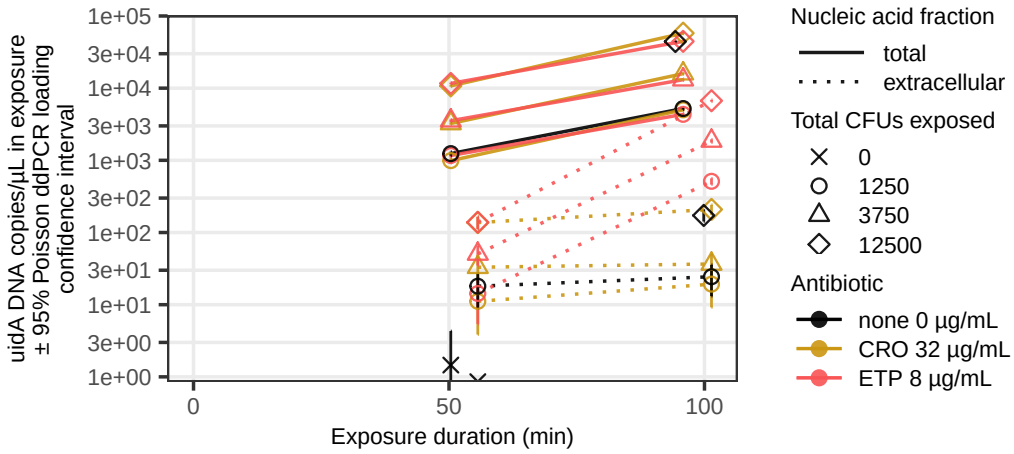
plateaued at a constant low concentration, including when no cells were added. This minimum concentration is readily apparent when reverse transcription was performed. Rather than reflecting a biological phenomena, I interpret this as a result of contamination and presumptive ddPCR artifacts sometimes observed.

For low concentrations, the possible set of ddPCR results is discrete. This means that zero positive droplets is calculated as a concentration of 0, while 1 positive drop due to ddPCR artifacts in approximately 15,000 droplets will jump the concentration to about 0.078 copies/µL in the ddPCR reaction. When this concentration is corrected for a typical dilution of 0.125 (no RT) or 0.0001 (RT), the resulting artifact is 0.63 and 784 copies/µL, respectively.

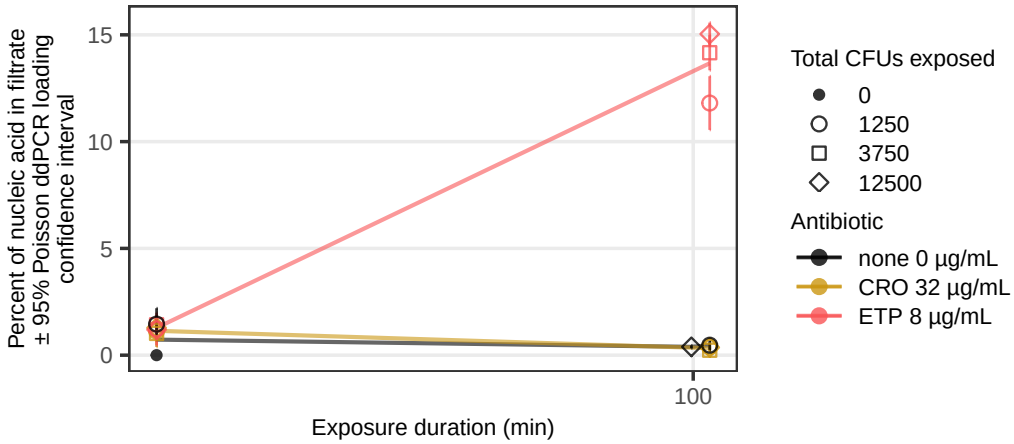
In the experiments with reverse transcription, it was noticed in template-less reactions (separate from the data shown) that amplification using *E. coli* rRNA primers detects an extrapolated 2500 copies/µL of contaminating rRNAs in the commercial enzyme stock itself, due to the enzyme being produced in *E. coli* cultures. In the bulk filtration AST experiments herein, the enzyme stock was diluted 1:100 in to the RT reaction, which was further diluted 1:20 when creating the ddPCR mastermix. We thus expect 1.25 copies/µL of contamination in the ddPCR. The observed contamination concentration to which the model was fit is corrected by the dilutions of the sample from the exposure to the ddPCR reaction. For example, the 0 CFU samples in AST Run 16 were diluted 1:2 in DEB, 1:10 in water, 1:2 in the RT reaction, and 1:20 in the ddPCR reaction, so we would expect $1.25(2)(10)(2)(20) = 1000$ copies/µL to be shown in the above figures. The observed background signal is indeed at 1000 copies/µL.

3.2.4 Inoculum effects do not affect bulk filtration AST

An inoculum effect, specifically an increase in apparent susceptibility when the inoculum is lowered, has been well documented in the literature for bacteria exposed to beta-lactam antibiotics in liquid culture. Thus it is puzzling why no inoculum effect was seen during bulk filtration AST in Figures 3.7a–3.7e. Of note, the strain *Escherichia coli* CDC #1 is a known producer of 4 different beta-lactamase genes, yet the proportion of lysed cells at 100 minutes of ertapenem exposure was constant across 2 orders of magnitude of starting inoculum in Figure 3.8a. Interestingly, replicate exposures cultured overnight did show an inoculum effect. The lowest inoculum of 1250 CFU (1×10^4 CFU/mL) in both the ertapenem and ceftriaxone doses failed to produce a cell pellet, in contrast with the two higher inocula.



(a) AST Run 18. Inoculum, dose, and exposure duration were all varied.



(b) A similar proportion of cells lysed regardless of cell density

Figure 3.8: The inoculum effect of beta-lactamases does not affect short-term antibiotic kill rates.

The currently accepted explanation of the inoculum effect is that certain strains of bacteria produce beta-lactamases. The large the inoculum, the faster the beta-lactam antibiotics in the exposure volume are degraded. When the beta-lactam antibiotic concentration decreases past a certain point (possibly the concentration at which the rate of killing is lower than the growth rate to a degree that the probability of eventual extinction is small), the surviving cells rebound if not yet extinct, and that particular antibiotic concentration is considered to not be inhibitory. Thus, at the MIC, defined by the CLSI for an inoculum of about 5×10^5 CFU/mL, larger inoculums than the standard inoculum will survive, while smaller inoculums do not. A detail not often stated explicitly in the inoculum effect literature is the stochastic nature of whether the population of bacteria goes extinct before sufficient antibiotic

is degraded. A smaller inoculum may reach the same proportion of cells surviving as a larger inoculum, for example 10^{-5} fold fewer cells, but since cells are discrete entities, it is more likely for a starting inoculum of 10^6 cells to go extinct than it is for a 10^7 inoculum [128].

If this model is correct, then one explanation for the constant proportion of lysis I observed lies in the time scales examined and the experimental protocols followed. In the bulk filtration AST experiments herein, the longest exposure times are between 2 and 3 hours, and the resolution of the time series is about 5 minutes. In contrast, the typical time-kill curve in the literature examines up to 24 hours of incubation with a resolution of 1 hour. Nucleic acid amplification also does not distinguish between very low amounts of intact cells the way plating a serial dilution on solid media can distinguish between 10^{-4} and 10^{-5} CFU/mL.

The MIC inoculum effect manifests only for endpoint measurements (i.e., the MIC) of antibiotic exposures of long duration. The proportion of lysis, in contrast, reflects a kinetic phenomena unaffected by the slight decrease in otherwise saturating antibiotic concentration within the time frame examined.

3.2.5 The antibiotic death rate at short time scales is correlated with antibiotic dosage

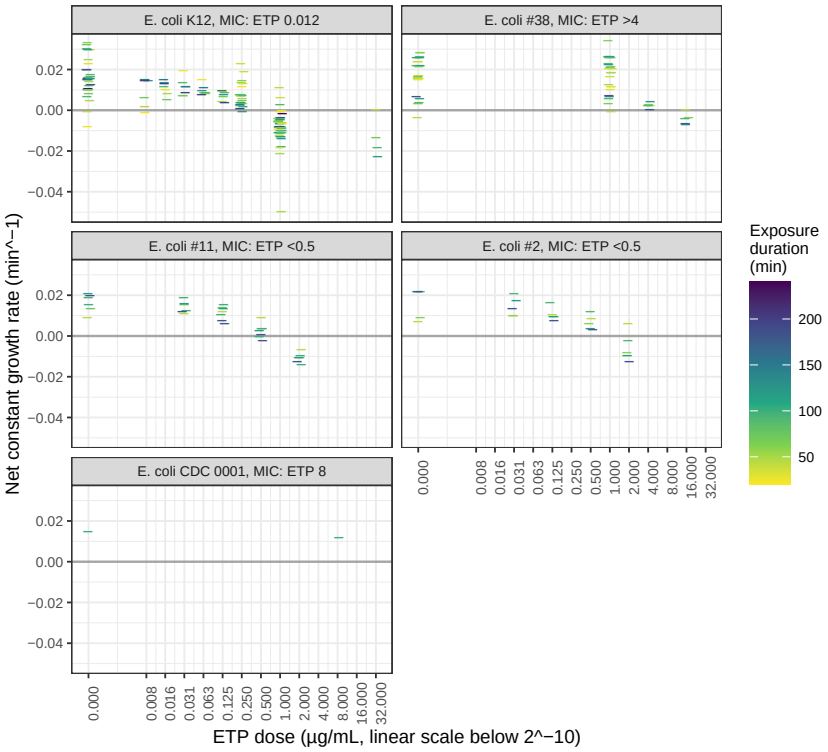
Another important function to understand about *in vitro* antibiotic exposure is the dose-response curve, the function of some metric of antibiotic response to the dosage of antibiotics in the exposure.

In Figures 3.9a and 3.9b, a statistic μ^* that I will call the "constant net growth rate" (known as the net growth rate in Regoes et al., 2004 [129]) is plotted as a function of antibiotic dosage for all bulk filtration data. The constant net growth rate is the quantity

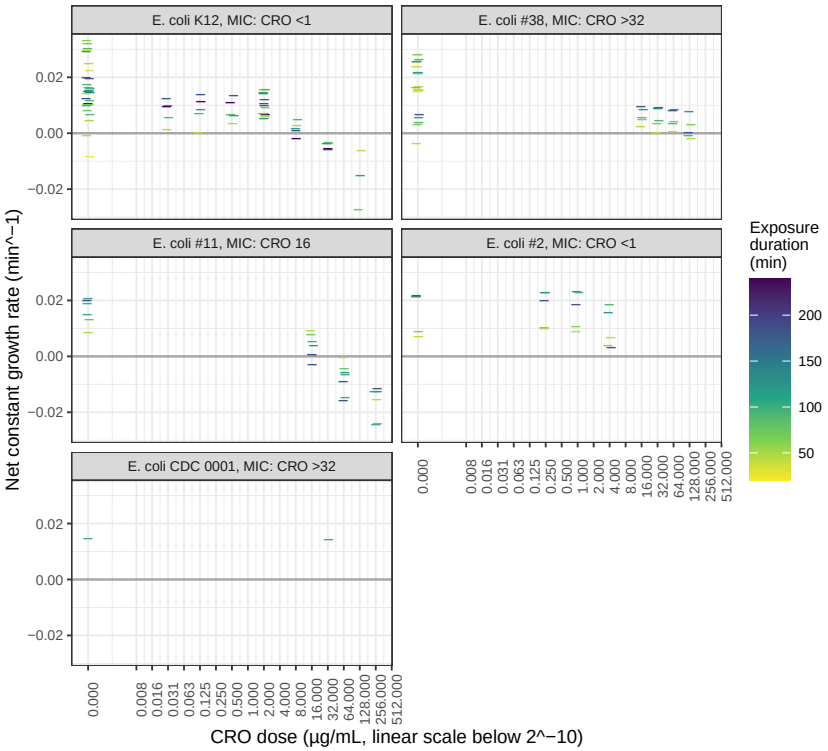
$$\mu^* = \frac{1}{t_i - t_1} \log \frac{T[t_i] - D[t_i]}{T[t_1] - D[t_1]}, \tag{3.5}$$

where t_i is the i-th time point, t_1 is the earliest time point of the AST run, $T[t]$ is the total/feed nucleic acid amount at time t, and $D[t]$ is the extracellular/filtrate nucleic acid amount at time t. The constant net growth rate statistic makes the rough assumption that the amount of live cells follows exponential growth in the following manner: $L[t_i] = L[0]e^{\mu^* t}$, where μ^* is assumed to be constant. One can further interpret $\mu^* = \mu - \beta$, the difference between a true growth rate μ and a death rate β.

Later, I show that the constant net growth rate model is a special case of a more



(a) Ertapenem



(b) Ceftriaxone

Figure 3.9: Constant net growth rate as a function of antibiotic dose.

general dynamic model which we use to fit the entire bulk filtration AST data set. The constant net growth rate is nonetheless displayed here to compare with my dose-response model fit because the former can be calculated arithmetically for all but the first time point in each AST run and makes fewer assumptions. In contrast, my later model only yields a single set of parameters for each pair of strain and antibiotic compound included.

3.3 A dynamical compartment model quantifies the limiting parameters for phenotypic assays of beta-lactams in general

As discussed above, there are many variables at play during the *in vitro* antibiotic exposure. Qualitative trends of each variable that can be deduced from the data have already been described above. In this section, I further propose a mathematical model to quantitatively relate these variables and thus describe the state of the *in vitro* antibiotic exposure, and I also model the observable, noisy output of the assay. There are several purposes of modeling. First, being able to describe and predict the state of the assay system is necessary to rationally optimize trade-offs between duration, dose, and inoculum during in assay design. Second, predicting the outcome of the assay, subjected to measurement noise, is necessary to rationally specify the limits of assay performance. Third, the model can be used as the susceptibility classification algorithm that is necessarily part of any diagnostic assay, and I hypothesize that an algorithm that can remove the confounding variables of duration, dose, inoculum, and other pharmacodynamic parameters from the variable of susceptibility will show improved accuracy over algorithms that do not take the former other variables into consideration.

Philosophically, any mathematical model of nature can only offer an approximate level of accuracy. The possibility always remains that a model does not take into consideration all existing and relevant variables, variables whose inclusion would make the model more accurate but more complex. Furthermore, simplicity of interpretation often decides which functional forms are chosen in the model to relate the included variables in a model, even if a more complex equation would also be plausible. As one makes more accurate and more frequent measurements, it becomes increasingly more useful to model additional variables. With the experiments in this thesis, it would already be plausible to assume variables or phenomena such as medium richness (for growth rate), medium osmolarity (increase kill rates), carrying capacity, aggregation or biofilm formation, gene regulation, or population age structure when modeling *in vitro* antibiotic exposures. However, these variables

were not necessary to make important conclusions and so their exploration was postponed as future work.

3.3.1 Solving ordinary differential equations for *in vitro* antibiotic incubation

3.3.1.1 Compartment model assumptions

The main concern of our AST assay is the assessment of susceptibility. Susceptibility is currently defined by changes cell growth or number caused by antibiotics, and antibiotics must have some time to act. Thus, the number of bacterial cells and the changes in that number over time are important to include in a model. By definition, such a model is a pharmacodynamic model.

Ordinary differential equations, and in particular compartment models, are a popular model for dynamical systems and pharmacodynamics in particular. The following system of ordinary differential equations captures the interaction between cell growth and cell death seen in my experiments.

$$\begin{aligned} \frac{dL[t]}{dt} &= (\mu - h[t])L[t] \\ \frac{dD[t]}{dt} &= h[t]L[t] \\ D[0] &= 0 \\ L[0] &= L_0 \end{aligned} \tag{3.6}$$

To the equations for Malthusian exponential growth, we add a term to represent the death from antibiotics. The quantity $h[t]$, in units of time^{-1}, represents the proportion of cells alive at time t that die in the next instant of time due to antibiotics. $L[t]$ and $D[t]$ are the amount of live and dead bacteria, respectively, as functions of time, μ is the intrinsic rate of growth in units of time^{-1}, and L_0 is the initial inoculum of live cells.

Several assumptions were made in writing this system of equations. First, it was assumed that the number of cells is a continuous, real-valued quantity. This assumption is strongly violated when the number of cells is low, but it is reasonable when the number of cells is large and the probability of stochastic extinction is low. We discuss in Chapter 4 alternative population models where the number of cells is an integer.

It is reasonable assumption that the absolute amount of live cells generated over a period of time is proportional to the amount of live cells at the start of the period of time, an assumption known as the law of mass action. Similarly, it is a reasonable

assumption that the The exponential growth of bacterial cultures is indeed modeled well by the Malthusian growth model (or simple exponential growth model), one of the earliest models in the Western literature for population growth.

Both cell growth and cell death occur simultaneously during the *in vitro* antibiotic exposure. Live cells only arise from other live cells, according to the long-standing cell theory of biology. Dead cells are generated from live cells and do not produce more dead cells. We assume that DNA and RNA are not degraded or lost in the system, as no nuclease activity was observed in our batch cultures. We also assume that nucleic acids are not secreted by the cells. Therefore, no further transitions between the live and dead compartments are included in the model.

We assumed that cells during our assay did not enter stationary phase or diauxic growth within the durations tested. Otherwise, the constant intrinsic growth rate could be replaced with any of the density-dependent growth rate models known in the literature. We also assumed no lag phase in growth.

Only two cell states, live and dead, were assumed in the compartment model. Entry into a quiescent stage is described for bacteria exposed to bacteriostatic antibiotics. Beta-lactam antibiotics are considered bacteriocidal antibiotics, and so quiescence was not included in the model. Very small populations of quiescent cells called persisters have been claimed in the literature to exist during beta-lactam exposure. However, they were not modeled in this work due to the small proportion persister cells represent and because of the short time durations examined.

Age of individuals in a population is another important variable often considered in population dynamics. The evolution of age-structured populations can be described by the McKendrick-von Foerster partial differential equations, but for this thesis, no age structure was assumed because of the large, unsynchronized inoculums examined.

3.3.1.2 General solution for antibiotic exposures

Without making assumptions about the form of h[t], one can derive a closed form general solution to the system of equations 3.6. To do so, one first assumes that $h[t]$, the proportion of cells that die within the period of time $t + dt$, is equal to the probability of a given cell dying within $t + dt$. $h[t]$ now meets the definition of the hazard rate of death from antibiotics, as defined in the survival analysis literature. As discussed later, the assumption above is only valid when the number of live and

dead cells is large and can be approximated as a continuous quantity.

In survival analysis, the hazard rate is related to other quantities by the following identities.

$$h[t] = \frac{f[t]}{S[t]} = \frac{f[t]}{1 - F[t]} = -\frac{1}{S[t]}\frac{dS[t]}{dt} \tag{3.7}$$

$$S[t] = e^{-H[t]} \tag{3.8}$$

$S[t]$ is the probability of a cell surviving to time t from start of drug exposure (t=0) and called the survival function.. $F[t]$ is cumulative probability function of cell lifetimes from the start of exposure ($t = 0$), with $S[t] = 1 - F[t]$. $f[t]$ is the probability density function of cell lifetimes from t=0, with $f[t] = \frac{dF[t]}{dt}$. $H[t]$ is the cumulative hazard function, with $H[t] = \int_0^t h[\tau]d\tau$.

Then, since the ODE system is linear and separable,

$$\begin{aligned} L[t] &= L_0 e^{\int_0^t (\mu - h[\tau])d\tau} \\ &= L_0 e^{(\mu t - 0) - (H[t] - 0)} \\ &= L_0 e^{\mu t} S[t] \end{aligned} \tag{3.9}$$

$$\begin{aligned} D[t] &= \int_0^t h[\tau] L[\tau] d\tau \\ &= \int_0^t h[\tau](L_0 e^{\mu\tau} S[\tau])d\tau \\ &= L_0 \int_0^t e^{\mu\tau}(h[\tau]S[\tau])d\tau \\ &= L_0 \int_0^t e^{\mu\tau} f[\tau] d\tau. \end{aligned} \tag{3.10}$$

Because a small increase in the extracellular DNA was observed in untreated conditions, one can add a term k to represent a background rate of extracellular DNA creation independent of antibiotics and constant with respect to time. Equations 3.6 become

$$\begin{aligned} \frac{dL[t]}{dt} &= (\mu - k - h[t])L[t] \\ \frac{dD[t]}{dt} &= (k + h[t])L[t] \\ D[0] &= 0 \\ L[0] &= L_0. \end{aligned} \tag{3.11}$$

First solving for $L[t]$, then solving $D[t]$ using integration by parts, the solutions are

$$\begin{aligned}
\int_{L_0}^{L} \frac{dL[t]}{L[\tau]} &= \int_0^t (\mu - k - h[\tau])\, d\tau \\
\ln|L[t]|\Big|_{L_0}^{L} &= (\mu - k)t - H[t]\Big|_0^t \\
L[t] &= L_0 e^{(\mu-k)t} e^{-H[t]} \\
&= L_0 e^{(\mu-k)t} S[t],
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
\int_{D_0=0}^{D} dD[t] &= \int_0^t (k + h[\tau])\, L[\tau] d\tau \\
D[t] &= \int_0^t (k + h[\tau])\, L_0 e^{(\mu-k)\tau} S[\tau] d\tau \\
&= kL_0 \int_0^t \left[S[\tau] e^{(\mu-k)\tau} d\tau \right] + L_0 \int_0^t \left[h[\tau] e^{(\mu-k)\tau} S[\tau] d\tau \right] \\
&= kL_0 \left[\frac{e^{(\mu-k)t}}{\mu-k} S[t]\Big|_0^t - \int_0^t \frac{e^{(\mu-k)\tau}}{\mu-k}(-f[\tau]) d\tau \right] \\
&\quad + L_0 \int_0^t e^{(\mu-k)\tau} f[\tau] d\tau \\
&= \frac{kL_0}{(\mu-k)} \left[e^{(\mu-k)t} S[t] - 1 \right] + \frac{kL_0}{(\mu-k)} \int_0^t e^{(\mu-k)\tau} f[\tau] d\tau \\
&\quad + \frac{(\mu-k)L_0}{(\mu-k)} \int_0^t e^{(\mu-k)\tau} f[\tau] d\tau \\
&= L_0 \left[\left(\frac{k}{\mu-k} \right) \left(e^{(\mu-k)t} S[t] - 1 \right) + \left(\frac{\mu}{\mu-k} \right) \int_0^t e^{(\mu-k)\tau} f[\tau] d\tau \right],
\end{aligned} \tag{3.13}$$

$$D[t] = \left(\frac{k}{\mu-k} \right) (L[t] - 1) + L_0 \left(\frac{\mu}{\mu-k} \right) \int_0^t e^{(\mu-k)\tau} f[\tau] d\tau. \tag{3.14}$$

Further generalization is possible. Notice that the constant background rate of death is mathematically equivalent to if a second antibiotic was present whose dosage never changes and whose hazard function is a constant. Generalizing, one can solve the following system of equations.

$$\begin{aligned}
\frac{dL[t]}{dt} &= L[t] \left(\mu - \sum_i h_i[t] \right) \\
\frac{dD[t]}{dt} &= L[t] \sum_i h_i[t] \\
D[0] &= 0 \\
L[0] &= L_0
\end{aligned} \tag{3.15}$$

Because in this work, I did not study antibiotic interactions, the analysis of this model was left for future work. Synergistic and antagonistic effects of antibiotics, however, are well characterized and often used (or avoided) in clinical practice.

The assumption of constant growth rate in the above equations will break down for longer antibiotic exposures or large initial inoculum since the nutrients in the growth media will deplete, and the bacteria will enter stationary phase. The literature contains many models for density-dependent population growth that generalize the equation for $L[t]$ into a more complex function, such as the logistic function or the Gompertz equation. Since this work focused on short exposure durations and low numbers of cells, the analysis of models with carrying capacities was left for future work.

3.3.2 The multi-hit hazard rate and dose response curve

3.3.2.1 Single-hit hazard rate

Different choices of $h[t]$, which uniquely determine $f[t]$ and $S[t]$, yield different solutions to equations 3.6 and 3.11.

For example, when $h[t] = \beta$ is constant over time, the cell lifetimes are exponentially distributed.

$$h[t] = \beta \tag{3.16}$$

$$S[t] = e^{-\beta t} \tag{3.17}$$

$$f[t] = \beta e^{-\beta t} \tag{3.18}$$

Equations 3.9 and 3.10 then take the following forms:

$$L[t] = L_0 e^{(\mu-\beta)t} \tag{3.19}$$

$$\begin{aligned} D[t] &= L_0 \left(\frac{\beta}{\mu-\beta}\right)\left(e^{(\mu-\beta)t} - 1\right) \\ &= L_0 \left(\frac{\beta}{\beta-\mu}\right)\left(1 - e^{-(\beta-\mu)t}\right). \end{aligned} \tag{3.20}$$

Similarly, equations 3.12 and 3.13 simplify to:

$$L[t] = L_0 e^{(\mu-\beta-k)t} \tag{3.21}$$

$$\begin{aligned} D[t] &= L_0 \left(\frac{\beta+k}{\mu-\beta-k}\right)\left(e^{(\mu-\beta-k)t} - 1\right) \\ &= L_0 \left(\frac{\beta+k}{\beta+k-\mu}\right)\left(1 - e^{-(\beta+k-\mu)t}\right). \end{aligned} \tag{3.22}$$

The hazard rate $h[t]$ from antibiotics is the only term where antibiotic action takes place. It follows that $h[t]$ should be a function, the dose-response curve, of the antibiotic dose and the strain's resistance. This function should not increase to infinity as antibiotic dose increases, since an infinite kill rate is physically implausible for all currently known antibiotic drugs. The function should also reach a value of zero when no antibiotics are present. Of the possible functional forms for $h[t]$, the Hill function is a simple, popular choice for dose-response curves that satisfies the two limiting behaviors. The Hill function is

$$h[t] = \beta = \frac{\beta_{max}[Abx]^{\gamma}}{[Abx]^{\gamma} + \mathrm{EC}_{50}^{\gamma}} \tag{3.23}$$

where β_{max} is the maximum rate of antibiotic killing in units of time^{-1}, $[Abx]$ is the dosage or concentration of antibiotic in units of µg/mL or equivalent, and γ is the Hill coefficient, a dimensionless parameter which controls the steepness of the Hill function's sigmoidal shape. The EC_{50}, or "effective concentration 50," is the antibiotic concentration at which the antibiotic killing is half of the maximum β_{max}. It is a metric of susceptibility of a given strain. In Mouton et al., 2005 [130], the EC_{50} is related to the MIC by the equation 3.24, assuming that a broth microdilution well is inoculated with 5×10^5 CFU/mL, becomes visibly turbid only at 10^8 CFU/mL, and is read at 18 hours of incubation.

$$\begin{aligned} MIC &= \left(\frac{\mu - \frac{1}{t} \ln \frac{L_{\text{visible}}}{L_0}}{\beta_{\max} - (\mu - \frac{1}{t} \ln \frac{L_{\text{visible}}}{L_0})} \right)^{\frac{1}{\gamma}} \mathrm{EC}_{50} \\ &= \left(\frac{\mu - 0.29\mathrm{hr}^{-1}}{\beta_{\max} - (\mu - 0.29\mathrm{hr}^{-1})} \right)^{\frac{1}{\gamma}} \mathrm{EC}_{50}. \end{aligned} \tag{3.24}$$

With a constant antibiotic hazard rate, the bacterial lifetimes are exponentially distributed. A constant hazard rate would arise if discrete events of antibiotic damage occurred to the with a constant probability per time β, and if a single event of antibiotic damage causing the death of the bacterium. I do not have evidence that beta-lactam antibiotics cause discrete events of antibiotic damage, even though mechanical gaps in the peptidoglycan cell wall caused by beta-lactam action would be a plausible manifestation of discrete events of antibiotic damage. Nonetheless, it is correct to say that a choice of a constant hazard rate implies an equivalence to a single-hit model of antibiotic action.

3.3.2.2 Multi-hit hazard rate

In my data, I observed a significant delay in antibiotic killing despite no evidence of a significant lag phase in growth in the untreated conditions. To model this delay in antibiotic killing, we generalize the above single-hit model of antibiotic action, where the hazard rate is constant, to a multi-hit model. If a bacterium dies upon incurring α events of antibiotic damage, and the events occur with a constant probability per time β, then the lifetimes of the bacteria would be gamma distributed with a shape parameter of α and a rate parameter of β. Setting $h[t]$ to be the gamma distribution hazard rate introduces time-dependence to $h[t]$:

$$\begin{aligned} h[t] &= \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma[\alpha] Q[\alpha, \beta t]}, \\ \beta &= \frac{\beta_{max}[Abx]^\gamma}{[Abx]^\gamma + \mathrm{EC}_{50}^\gamma}. \end{aligned} \tag{3.25}$$

$$S[t] = Q[\alpha, \beta t]. \tag{3.26}$$

$$f[t] = \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma[\alpha]}. \tag{3.27}$$

where $\Gamma[\alpha] = \int_0^\infty u^{\alpha-1} e^{-u} du$ is the gamma function, and $Q[\alpha, x] = \frac{1}{\Gamma[\alpha]} \int_x^\infty u^{\alpha-1} e^{-u} du$ is the regularized upper incomplete gamma function.

Notice that when $\alpha = 1$, we recover the single-hit model where $h[t] = \beta$.

It would be elegant to interpret α as truly representing the effective number of damaging events that a cell needs to incur from antibiotic exposure to lyse and die, since it is reasonable to assume that multiple defects in the cell wall must accumulate to lead to catastrophic failure during beta-lactam activity. However, alternative mechanisms would appear in the same way. A possible alternative would be a pause in overall metabolic activity at the start of the exposure due to changes in the growth media composition or temperature. Another possible alternative would be the influence of cell age on the hazard rate, if the hazard rate only increases when cells reach an older age. Thus, it is safer to conclude that the α parameter only captures the empirical or apparent magnitude of the delay in antibiotic killing, without knowledge of the underlying mechanism.

It should be noted that there exists substantial literature on the dynamics of age-structured populations. The generalization of the compartment model of section 3.3.1 is the McKendrick-von Foerster equation. This approach was not further explored since the effects of cell age seemed adequately accounted for by the α parameter of the multi-hit hazard rate.

Combining equations 3.12, 3.14, 3.26, and 3.27, we arrive at the following solutions for equation 3.11.

$$L[t] = L_0 e^{(\mu-k)t} Q[\alpha, \beta t] \tag{3.28}$$

$$D[t] = \left(\frac{k}{\mu-k}\right)(L[t]-1) + L_0\left(\frac{\mu}{\mu-k}\right)\int_0^t e^{(\mu-k)\tau} f[\tau]d\tau \tag{3.29}$$

$$= L_0\left[\left(\frac{k}{\mu-k}\right)\left(e^{(\mu-k)t}Q[\alpha,\beta t]-1\right) + \left(\frac{\mu}{\mu-k}\right)\int_0^t e^{(\mu-k)\tau} f[\tau]d\tau\right] \tag{3.30}$$

$$\begin{aligned} = L_0\Bigg[&\left(\frac{k}{\mu-k}\right)\left(e^{(\mu-k)t}Q[\alpha,\beta t]-1\right) \\ &+ \left(\frac{\mu}{\mu-k}\right)\left(\frac{\beta}{\beta+k-\mu}\right)^{\alpha} P[\alpha,(\beta+k-\mu)t]\Bigg] \end{aligned} \tag{3.31}$$

where $P[\alpha, x] = \frac{1}{\Gamma[\alpha]}\int_0^x u^{\alpha-1}e^{-u}du = 1 - Q[\alpha, x]$ is the regularized lower incomplete gamma function. Note that when $\mu > \beta + k$, equation 3.31 generates two complex numbers. For numerical calculations on a computer, it may be easier to implement the integral of equation 3.30 with existing libraries than to work with complex numbers.

When $\beta + k > \mu$, then the population eventually goes extinct, and the total amount of nucleic acids plateaus as follows:

$$\lim_{t\to\infty} L[t] = 0 \tag{3.32}$$

$$\lim_{t\to\infty} D[t] = L_0\left[\left(\frac{\mu}{\mu-k}\right)\left(\frac{\beta}{\beta+k-\mu}\right)^{\alpha} - \left(\frac{k}{\mu-k}\right)\right]. \tag{3.33}$$

When $\beta + k < \mu$ and $\mu, \beta, k > 0$, the population grows indefinitely because no stationary phase is modeled.

Example trajectories from equations 3.28 and 3.31 are shown in Figure 3.10.

3.3.3 Hierarchical error model for batch effects

The above model now describes the assay system, but not the technical observations made by the experimenter. To fit to the nucleic acid quantification data, one also needs to specify the function relating the system state, namely the amount of live and dead bacteria, to the measured outcome of nucleic acid in copies/µL. The following hierarchical mixed-effects model in equation 3.36 was proposed.

The amount of both 23S genomic DNA loci and of 23S ribosomal RNA, Y, was set to be a linear function of the amount of bacterial biomass X with slope or

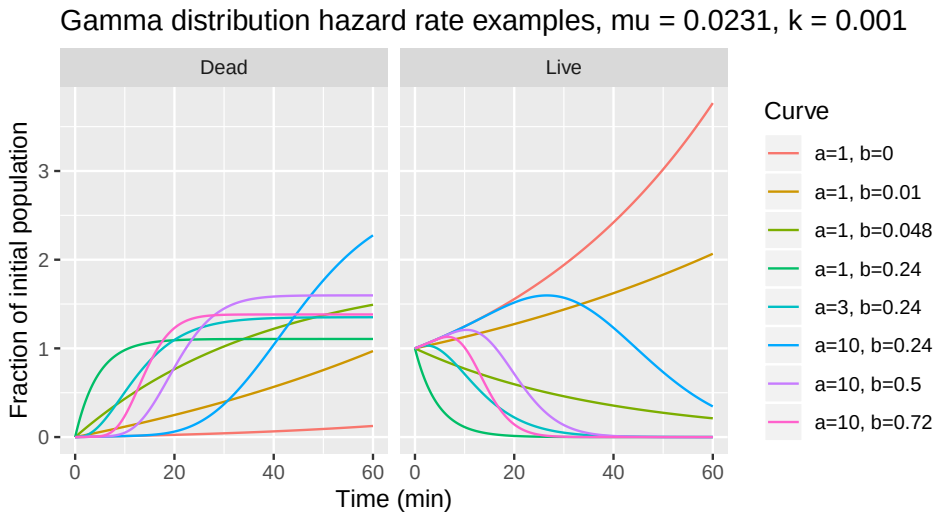


Figure 3.10: Examples of population trajectories with a multi-hit hazard rate

proportionality constant m and intercept b. The biomass X represents either the feed or filtrate fraction biomass. In the former case, $X = D + L$, while in the latter, $X = D$. D and L were calculated specifically with equations 3.28 and 3.30. The slope m represents the amplification efficiency. The intercept b is proportional to the limit of detection (LOD) of the amplification and to the background level of amplification. A non-zero limit of detection would promote a lower, more negative intercept value, while a background signal, like contamination, would promote a higher, more positive value. Whether the intercept overall is negative or positive depends on which of these two effects is larger. However, if the LOD was reached for a given amplification protocol (i.e., primer and reverse transcription combination), and the intercept was negative, negative concentrations would not be seen. Instead, either artifacts of ddPCR would be seen, causing a positive concentration to be estimated, or a concentration of 0 copies/µL would be seen. Therefore, my model assumes that the observed concentration Y is the maximum of a constant artifact concentration c or the concentration expected by the linear function of the bacteria population, $mX + b$. Thus, the model so far, without stochasticity, is written as

$$Y_i = \max(mX_i + b, c). \tag{3.34}$$

where Y_i and X_i are the observed concentration and unobserved bacterial population values of the i-th AST observation.

It is important to note that any deviations in the starting inoculum from the intended target inoculum will proportionally change the value of m. The starting inoculum L_0 factors out of X as a constant, so $Y = mX \propto mL_0$. The deviations from the target

inoculum could plausibly arise because of an incorrect conversion ratio between the batch culture cell density and the measured OD_{600}; this average deviation would be shared by all experiments. Variations between batches could plausibly arise form the variable time between the measuring of the OD_{600} and the start of the serial dilution, which could approach up to half of a 30-minute doubling time, as well as pipetting errors.

Each bulk filtration AST ddPCR batch, by definition, comprised the PCR measurements of one AST run performed on the same day using a shared PCR protocol. There were two such PCR protocols: reverse transcription + 23S primers, or no reverse transcription + *uidA* primers. The ddPCR batches therefore fall into two exchangeable groups indicating shared PCR protocols. Accordingly, to fit the available data, the slope and intercept of Equation 3.34 was split into two different amplification efficiencies m_p and two intercepts b_p, where $p = \{1, 2\} = \{\text{"-RT"}, \text{"+RT"}\}$ indicates which protocol was used.

The model so far now has one fixed effect of unknown magnitude representing protocol effects, but it does not have a term for the stochastic noise one expects in any real life measurement. Therefore, a random variable ϵ is added to represent stochastic noise. For our data, two separate noise terms ϵ_p were introduced to capture any difference in the relative noise across the two PCR protocols. The error ϵ was assumed to be multiplicative and log-normally distributed with a multiplicative standard deviation of σ_ϵ:

$$Y_i \sim \text{Log-normal}\left(\max(m_{p[i]}X_i + b_{p[i]}, c), \sigma_\epsilon\right). \tag{3.35}$$

A multiplicative error model was chosen because errors in the serial dilution of the nucleic acid extractions caused by retention of a constant proportion of the volume in the pipette tip would accumulate multiplicative errors, rather than additive errors. Furthermore, the concentration of nucleic acids is always non-negative, while the traditional assumption of normal noise predicts the possibility of negative values. The log-normal distribution admits no negative values, and thus is a more natural model for non-negative data. The log-normal distribution, however, does not admit zero values. The inclusion of the artifact parameter c was thus necessary for the error to be modeled as log-normal.

The true distribution of ddPCR results should be a binomial distribution in the number of positive and negative droplets, after the dilution from exposure to ddPCR and other sources of droplet positivity are inputted. For computational simplicity,

the droplet counts were not directly fitted, only the most likely estimate of the concentrations.

The observed nucleic acid concentrations from the compartments of a single exposure could be modeled as a multivariate random variable, such as a multivariate Gaussian. This model was examined, and the estimated covariance between the two compartments was almost zero. Thus, for computational simplicity, the noise of each compartment's nucleic acid concentrations were assumed to be independent.

The exact amplification efficiency and intercept appeared to vary between different PCR batches run on different days, even among batches using the same PCR protocol. While the variation was within an order of magnitude, the fit of the lines exhibited multimodality in the posteriors estimated for the amplification efficiency m_p. Therefore, each PCR batch was fitted to its own amplification efficiency m_e and intercept b_e, where $e = \{1, \ldots, 30\}$ indicates the PCR "experiment" or batch. These experiment-specific parameters themselves were distributed around position parameters of m_p^* and b_p, with scale parameters of $\sigma_{m_p}^*$ and σ_{b_p}, thus creating the hierarchical mixed-effects model below.

$$Y_i \sim \text{Log-normal}\left(\max(c, m_{e[i]} X_i + b_{e[i]}), \sigma_{\epsilon_{p[i]}}\right) \tag{3.36}$$

$$m_{e[i]} \sim \text{Log-normal}\left(m_{p[i]}^*, \sigma_{m_{p[i]}}^*\right) \tag{3.37}$$

$$b_{e[i]} \sim \text{exGaussian}\left(b_{p[i]}, \sigma_{b_{p[i]}}, \lambda_{b_{p[i]}}\right) \tag{3.38}$$

Note that m_p^* is the mean of the logarithm of the m_e values, not the mean value of the m_e parameters, which we denote m_p. Instead, $m_p = \exp\left(m_p^* + \frac{(\sigma_{m_p}^*)^2}{2}\right)$.

In eight pairs of experiments/PCR batches, the same nucleic acid extraction sample was quantified by PCR using two different primers. Of the thirty PCR batches, three pairs comprised replicate PCR batches from different days but using the same PCR protocol and nucleic acid extraction samples. The sharing of the DEB nucleic acid extraction step by these pairs of PCR batches would represent a known source of correlation between data. However, in the above model, these correlations are not encoded. All PCR experiments are assumed to be independent, whether they arose from different or the same nucleic acid extraction event. An additional hierarchical level indexed by the nucleic acid extractions could have been introduced to reflect these presumed correlations. However, for the 15 out of 18 batches of extraction samples without PCR replicates using the same protocol and extraction sample batch, there would be no way to separate the random effects from before the extraction

from the random effects occurring after the extraction, except from the information borrowed (via the hierarchical error structure) from the 3 pairs where the two kinds of effects can be separated. Thus, adding correlated hierarchical parameters with uninformative priors would create non-identifiabilities and geometries of high curvature in the parameter space, and these phenomena disrupt the HMC algorithm. Because the gain in accuracy about $m_{e[i]}$ and $b_{e[i]}$ would be outweighed by the increased difficulty in interpreting the fitting results, this model was not pursued to completion. A principled comparison of the possible hierarchical structures was left as future work.

Plausible sources of PCR batch variation would include differences in the timing of manual operations and thus the amounts of evaporation of the reactions, different pipettors used, and different batches of reagents (especially the age of the droplet generation oil). These noise sources affect the bulk filtration AST protocol both before and after the nucleic acid extraction step. A source of noise occurring before the nucleic acid extraction step would be errors in setting the starting inoculum of the batch of antibiotic exposures during the serial dilution of a single batch culture. The net effect of all these noise sources are empirically represented by the assumed hierarchical error model above.

The literature does not offer reasons for choosing a particular way in which the amplification efficiencies should vary around the mean amplification efficiency, other than the need for a central tendency and a support restricted to the positive real numbers. Therefore, I chose the log-normal distribution for the m_e parameters because it admitted a non-centered parameterization that Stan's HMC algorithm fits more efficiently.

For the intercepts, I elected to use the exponentially-modified Gaussian distribution (exGaussian). The intercepts can be negative, should have a mode near 0, but on occasion reach high values, as seen by visual inspection of figures 3.6. The exGaussian distribution possesses these qualities[1] and is available or can be derived from most statistical software libraries.

The other pharmacodynamic parameters were not hierarchically modeled due to the

[1] The normal distribution and the Student's t-distribution (with a small, possibly fixed, degrees-of-freedom parameter) could be considered. However, the normal distribution would be sensitive to the outliers in the positive tail. The Student's t-distribution would be robust to outliers, but its location would still be biased by the presence of only positive outliers. Furthermore, Stan's HMC algorithm encountered issues with numerical stability and inefficient posterior exploration if a fat-tailed distribution were used.

small number (2) of antibiotics tested so far. After inclusion of more antibiotic compounds in the future, especially if the data for some of the antibiotic compounds and/or strains is limited, the modeling may benefit from hierarchical modeling of the $B_{\max}$, γ, α, and EC_{50} parameters. Modeling the growth rate μ and death rate k hierarchically by strain or by experiment may also improve the model fitting by loosening constraints on the other pharmacodynamic parameters, which may reduce bias from underfitting in exchange for wider posteriors and more influential priors.

With two amplification protocols (two combinations of primer and reverse transcription), 5 strains, 2 antibiotics, and 30 PCR batches, there are 93 parameters in total.

3.3.4 Prior distributions for parameters of the model

In the previous sections of 3.3, the likelihood function of the nucleic acid concentration data given model parameters was defined. To make the model a complete Bayesian model, one must define the priors for the 31 of 91 parameters that do not have ancestors in the graph formed by the conditional dependencies. The priors are listed in Table 3.2. The probability distributions are parameterized using the standard forms in Appendix A of Gelman et al., 2014 [131], and justifications are given below.

Parameter ~ Prior	Parameter interpretation
$\mu \sim \text{Gamma}\left(\alpha = \frac{0.023^2}{0.005^2}, \beta = \frac{0.023}{0.005^2}\right)$	intrinsic growth rate (min^{-1})
$k \sim \text{Gamma}\left(\alpha = \frac{0.000023^2}{0.000023^2}, \beta = \frac{0.000023}{0.000023^2}\right)$	intrinsic death rate (min^{-1})
$\alpha[\text{ETP}] - 1 = \alpha^*[\text{ETP}] \sim \text{Exponential}\,(\lambda = 1)$ $\alpha[\text{CRO}] - 1 = \alpha^*[\text{CRO}] \sim \text{Exponential}\,(\lambda = 1)$	Antibiotic activity delay. Effective number (unitless) of antibiotic events, but re-parameterized as α^*.
$\beta_{\text{max}}[\text{ETP}] \sim \text{InvGamma}\left(\alpha = 2 + \frac{0.1^2}{0.08^2}, \beta = 0.1 + \frac{0.1^3}{0.08^2}\right)$ $\beta_{\text{max}}[\text{CRO}] \sim \text{InvGamma}\left(\alpha = 2 + \frac{0.1^2}{0.013^2}, \beta = 0.1 + \frac{0.1^3}{0.013^2}\right)$	maximum kill rate (min^{-1}).
$\gamma[\text{ETP}] - 1 = \gamma[\text{ETP}]^*[A] \sim \text{Exponential}\,(\lambda = 0.7)$ $\gamma[\text{CRO}] - 1 = \gamma[\text{CRO}]^*[A] \sim \text{Exponential}\,(\lambda = 0.7)$	Hill coefficient of the dose-response curve.
$\text{EC}_{50}[\text{ETP, K12}] \sim \text{Gamma}\left(\alpha = \frac{0.0131^2}{0.00654^2}, \beta = \frac{0.0131}{0.00654^2}\right)$	EC_{50} (µg/mL) for ertapenem + *E. coli* K12 (S, MIC 0.012)
$\text{EC}_{50}[\text{CRO, K12}] \sim \text{Gamma}\left(\alpha = \frac{0.102^2}{0.0523^2}, \beta = \frac{0.102}{0.0523^2}\right)$	EC_{50} (µg/mL) for ceftriaxone + *E. coli* K12 (S, MIC $\leq$ 1)
$\text{EC}_{50}[\text{ETP, Ec38}] \sim \text{Gamma}\left(\alpha = \frac{13.1^2}{13.1^2}, \beta = \frac{13.1}{13.1^2}\right)$	EC_{50} (µg/mL) for ertapenem + *E. coli* #38 (R, MIC >4)
$\text{EC}_{50}[\text{CRO, Ec38}] \sim \text{Gamma}\left(\alpha = \frac{209^2}{105^2}, \beta = \frac{209}{105^2}\right)$	EC_{50} (µg/mL) for ceftriaxone + *E. coli* #38 (R, MIC >32)
$\text{EC}_{50}[\text{ETP, Ec11}] \sim \text{Gamma}\left(\alpha = \frac{0.0131^2}{0.00654^2}, \beta = \frac{0.0131}{0.00654^2}\right)$	EC_{50} (µg/mL) for ertapenem + *E. coli* #11 (S, MIC $\leq$ 0.5)
$\text{EC}_{50}[\text{CRO, Ec11}] \sim \text{Gamma}\left(\alpha = \frac{26.2^2}{13.1^2}, \beta = \frac{26.2}{13.1^2}\right)$	EC_{50} (µg/mL) for ceftriaxone + *E. coli* #11 (R, MIC 16)
$\text{EC}_{50}[\text{ETP, Ec2}] \sim \text{Gamma}\left(\alpha = \frac{0.0131^2}{0.00654^2}, \beta = \frac{0.0131}{0.00654^2}\right)$	EC_{50} (µg/mL) for ertapenem + *E. coli* #2 (S, MIC $\leq$ 0.5)
$\text{EC}_{50}[\text{CRO, Ec2}] \sim \text{Gamma}\left(\alpha = \frac{0.1022^2}{0.0523^2}, \beta = \frac{0.1022}{0.0523^2}\right)$	EC_{50} (µg/mL) for ceftriaxone + *E. coli* #2 (S, MIC $\leq$ 1)
$\text{EC}_{50}[\text{ETP, Ec44}] \sim \text{Gamma}\left(\alpha = \frac{8^2}{4^2}, \beta = \frac{8}{4^2}\right)$	EC_{50} (µg/mL) for ertapenem + *E. coli* #44/CDC #1 (R, MIC 8)

Parameter ~ Prior	Parameter interpretation
EC_{50}[CRO, Ec44] ~ Gamma $\left(\alpha = \frac{64^2}{32^2}, \beta = \frac{64}{32^2}\right)$	EC_{50} (µg/mL) for ceftriaxone + *E. coli* #44/CDC #1 (R, MIC >32)
$b_{p=1} \sim$ Normal $(\mu = 0, \sigma = 3)$	Mean intercept (copies/µL) of nucleic acid amplification. Either LOD or contamination (multiplied by volume dilution from the exposure), without RT and with *uidA* primers.
$b_{p=2} \sim$ Normal $(\mu = 1000, \sigma = 100)$	Mean intercept (copies/µL) of nucleic acid amplification. Either LOD or contamination (multiplied by volume dilution from the exposure), with RT and with 23S rRNA primers.
$\sigma_{b_{p=1}} \sim$ HalfNormal $(\mu = 0, \sigma = 3)$	Standard deviation of experiment-level intercepts of nucleic acid amplification / LOD and contamination, without RT and with *uidA* primers.
$\sigma_{b_{p=2}} \sim$ HalfNormal $(\mu = 0, \sigma = 300)$	Standard deviation of experiment-level intercepts of nucleic acid amplification / LOD and contamination, with RT and with *uidA* primers.
$\lambda_{b_{p=1}} \sim$ InvGamma $\left(\alpha = 2 + \frac{100^2}{50^2}, \beta = 100 + \frac{100^3}{50^2}\right)$	Scale/mean of the exponential component of experiment-level intercepts of nucleic acid amplification, with RT and with *uidA* primers.
$\lambda_{b_{p=2}} \sim$ InvGamma $\left(\alpha = 2 + \frac{3000^2}{1000^2}, \beta = 3000 + \frac{3000^3}{1000^2}\right)$	Scale/mean of the exponential component of experiment-level intercepts of nucleic acid amplification, without RT and with *uidA* primers.
$m^*_{p=1} \sim$ Normal $\left(\mu = \ln\left(\frac{0.04^2}{\sqrt{0.04^2+0.04^2}}\right), \sigma = \sqrt{\ln\left(1 + \frac{0.04^2}{0.04^2}\right)}\right)$	Mean of the log nucleic acid amplification efficiencies without RT and with *uidA* primers.

Parameter ~ Prior	Parameter interpretation
$m^*_{p=2} \sim \text{Normal}\left(\mu = \ln\left(\frac{1000^2}{\sqrt{1000^2+1000^2}}\right), \sigma = \sqrt{\ln\left(1+\frac{1000^2}{1000^2}\right)}\right)$	Mean of the log nucleic acid amplification efficiencies with RT and with 23S rRNA primers.
$\sigma^*_{m_{p=1}} \sim \text{Gamma}\left(\alpha = 2, \beta = \frac{4}{\ln 2}\right)$	Log fold-change (unitless) of the experiment-level nucleic acid amplification efficiencies, without RT and with *uidA* primers.
$\sigma^*_{m_{p=2}} \sim \text{Gamma}\left(\alpha = 2, \beta = \frac{4}{\ln 2}\right)$	Log fold-change (unitless) of the experiment-level nucleic acid amplification efficiencies, with RT and with 23S rRNA primers.
$\sigma_{\epsilon_{p=1}} \sim \text{HalfNormal}\left(\mu = 0, \sigma = \ln 2\right)$	Fold-change in overall observed nucleic acid concentration due to random log-normal error, without RT and with *uidA* primers.
$\sigma_{\epsilon_{p=2}} \sim \text{HalfNormal}\left(\mu = 0, \sigma = \ln 2\right)$	Fold-change in overall observed nucleic acid concentration due to random log-normal error, with RT and with 23S rRNA primers.
$c \sim \text{Exponential}\left(\lambda = \frac{-\ln\frac{15000-0.5}{15000}}{0.125(0.00085)}\right)$	Typical lower bound on the observed nucleic acid concentration (copies/µL) caused by droplet digital PCR artifacts.

Table 3.2: Priors for the bulk filtration AST pharmacodynamic model.

For the growth rate μ, OD_{600} measurements using an automatic plate reader and during the batch culture incubation suggest a doubling time of about 30 minutes. These OD_{600} measurements are independent of the fitted data, except for the last measurement, which is used to calculate the target starting inoculum. Therefore, a moderately informative prior was chosen so that one (non-Gaussian) standard deviation spanning the doubling times of 20 to 40 minutes.

The background death rate k must be several orders of magnitude less than the growth rate for a gut bacterium to survive natural selection. Therefore, for the mean background lysis rate k, a weak prior was chosen with a mean at 0.1% of the growth rate and a standard deviation at least 2 times the mean. The mean was chosen based on ratio of filtrate to feed in untreated conditions incubated for >1.5 hours.

The difference between population trajectories decreases to 0 as α goes to infinity (or to 0). Such high and low values are physically absurd and cause numerical issues during HMC sampling, as well as poor efficiency since the amount of noise precludes α being identifiable in those regions. Therefore the prior for α was set to be moderately informative to avoid having a large cumulative probability at those extremes. Most of the prior was kept close to 1, with the chosen mean of 2 being the next highest integer. Furthermore, values below 1 were excluded by the reparameterization to $\alpha - 1$ because their interpretation was contradictory, and because doing so improved HMC convergence. When $\alpha < 1$, the rate of antibiotic activity would initially be faster than an exponential decay, then drop asymptotically to zero. While a small fraction of surviving bacteria evokes the notion of persister cells, persisters have been reported in percentages ($< 10^{-5}$ [132]) far smaller than would be visible in bulk filtration AST experiments. Heteroresistance would not be expected in pure isolates. Induced resistance is a possibility in some strains, but the rebound in growth should be apparent in the population trajectories, and the value of α would need to be close to 1 anyways to generate trajectories on the time scale of gene transcription and translation.

For $\beta_{\max}$, an inverse gamma distributed prior was chosen to restrict $\beta_{\max} > 0$. The literature reports values, all based on colony counts in time-kill experiments, of $0.0532 \pm 0.0232\text{min}^{-1}$ for piperacillin and *E. coli* [133]; $0.042 \pm 0.011\text{min}^{-1}$, $0.023 \pm 0.004\text{min}^{-1}$, and $0.027 \pm 0.004\text{min}^{-1}$ for benzylpenicillin, cefixime, and ceftriaxone and *Neisseria gonorrhoeae* [134]; 0.0598min^{-1} and 0.0428min^{-1} for meropenem and ceftazidime in *Pseudomonas aeruginosa* [130]; $0.079 \pm 0.014\text{min}^{-1}$ for ampicillin and *E. coli* [129]; and 0.14 to 0.540min^{-1} for 5 penicillins and *E. coli*

[135]. A value up to 10 times the growth rate seemed reasonable. The mean and standard deviation of the prior was chosen to cover all these values, and the shape of the gamma distribution was set to be greater than 1 to avoid extreme values while remaining a weakly informative prior.

The Hill coefficient for ceftazidime and meropenem in *Pseudomonas aeruginosa* was reported by Mouton et al., 2005 [130] to be 3.32 and 3.5 h^{-1}, respectively. In Foerster et al., 2019 [134], the Hill coefficient for ceftriaxone, cefixime, and benzylpenicillin and *Neisseria gonorrhoeae* were reported as 1.1, 1.6, and 1.7, respectively. Nolting et al., 1996 [133] fixes the value at one. Similar to α, extreme values near 0 and ∞ are indistinguishable by HMC and physically unlikely. Values less than 1 are problematic to interpret and may have led to HMC numerical instability, while a value of 1 has the interpretation of a lack of cooperativity in any underlying binding kinetics of the antibiotic to target molecules in the bacterium. Therefore, the exponential prior (a special case of the Gamma distribution) was chosen to have the highest probabilities near a value of 1 and a mean of 1.7 to sit in the middle of the literature values. Values below 1 were implicitly given a prior probability of 0.

The priors for the 8 MIC values covered by the bulk filtration AST experiments were determined by first choosing a mean MIC value, then converting the MIC to an EC_{50} value using the formula in Mouton et al., [130]. The gamma distribution was employed since EC_{50} values must be positive. To avoid numerical instability from extreme EC_{50} values during HMC sampling, the shapes of the gamma distributions were all set to a value of 2. Moderately informative priors were appropriate given the availability of external data on MICs. The external data comprised gold standard MICs from the UCLA clinical laboratory (or the literature for strain K12) and non-standard overnight incubations of replicate conditions during the bulk filtration AST experiments themselves.

The prior for the intercept for no RT, $b_{p=1}$, was centered at 0 due to the absence of evidence for systematic contamination by *E. coli* genomic DNA. The intercept can also reflect the limit of detection (LOD), which is a function of both the enzyme and of stochastic loading of molecules. In my case, the standard deviation of the mean of the intercept was set at 30 copies/µL to cover up to an LOD of 90 CFU/25 µL exposure within the 99-th percentile (assuming an amplification efficiency of 1.0 copy/CFU).

One can also justify $b_{p=1}$ being on the order of ±30 cells by the approximately 1/27 chance of a molecule being included in the final ddPCR template volume.

About 10 µL of the 25 µL exposure is transferred to each of the feed and filtrate fractions. The average (between feed and filtrate) DEB volume after extraction was about 18 µL, of which 1 µL was used as template. Thus, $\frac{10}{25} \times \frac{1}{18} = \frac{1}{45}$ of the exposure is actually measured. The chance that none of N cell genomes is inside that 1/45 volume fraction is $(44/45)^N$. For a ≥ 0.5 chance of no detection, $N = \ln(0.5)/\ln(44/45) \leq 30.8$ cells. This calculation ignores the fact that in some exposures, genome replication will occur. The calculation also ignores the PCR reagent characteristics, though this is justified since ddPCR is purported to be able to detect single molecules (the effective concentration of template molecules relative to the background amplification has a floor at 1 molecule per 0.85 nL droplet).

For $b_{p=2}$, the intercept of PCRs with a preceding RT step, contamination was assumed to be a major factor as discussed in section 3.2.3. The expected contamination was 1.25 copies/µL, and the most common dilution in experiments with RT was $\frac{1}{800}$, so the prior for $b_{p=2}$ was given a mean of 1000 copies/µL. The standard deviation was assigned to be about one order of magnitude less than the mean.

In the future, the known dilution volumes after the antibiotic exposure should be included in the model for more accurate model fitting, especially if the ddPCR droplet counts are fitted instead of the extrapolated nucleic acid concentrations as done herein.

The average nucleic acid amplification efficiency without RT, $m_{p=1}$, was given a mean of 1/25 copies/CFU/µL because there are between 1 and 2 genomes in one undivided cell and one *uidA* locus per genome, and because each CFU's DNA is dispersed into the 25 µL volume of the exposure. The standard deviation was set to be the same as the mean. For $m_{p=2}$, the mean and standard deviation were set to 1000 copies/CFU/µL, since this was the average copies/µL seen during one digital filtration AST experiment where ddPCR was performed. This concentration implies 25,000 copies/CFU are created, and indeed, approximately 26,000 ribosomes are reported to exist in *E. coli* growing at a doubling time of 40 minutes [122]. The variance of the number of ribosomes per cell is also well studied in the literature, but to factor in additional amplification errors, the standard deviation was set to the higher value of 1000. The parameters present in the model, however, are the mean and standard deviation of the log-transformed nucleic acid amplification efficiencies m_p^*. By setting the mean and standard deviation of m_p^* to be as shown in Table 3.2, the mean and standard deviation of the un-transformed, experiment-specific m_e parameters will be the un-transformed values.

The two parameters σ_{ϵ_p} for the stochastic observation noise represent the typical fold-change in in the observed nucleic acid concentrations, due to the assumption of log-normal noise. A 2-fold change was deemed a reasonably broad representation of the bulk filtration AST data, which varies over several orders of magnitude.

The mean of the final parameter, the ddPCR artifact c, was chosen as the concentration expected if 1 droplet out of a typical yield of 15,000 otherwise negative droplets was erroneously called positive, but only in half of the ddPCRs run. The droplet size is 0.85 nL, and the most common dilution from the exposure across all bulk filtration AST measurements was 0.125.

3.4 Bayesian Hamiltonian Monte Carlo provides fitted parameter values

The above Bayesian model was fitted to the data using Stan version 2.29, CmdStanR version 0.4.0, and R version 4.1.1 running on Ubuntu 16.04.7 LTS (GNU/Linux 4.4.0-210-generic x86_64). Stan is an open source Bayesian statistics library that performs Hamiltonian Monte Carlo (HMC). The R package bayesplot 1.8.1 was used to visualize some Stan results. The use of Bayesian statistics instead of a frequentist maximum likelihood estimator or other heuristic optimizer was due to the ability of Stan to fit models of arbitrary complexity without the need to re-derive formulas for frequentist estimators, the ability to easily switch variables from being parameters to being fixed constants in code, and the more intuitive interpretation of Bayesian posteriors compared to frequentist frameworks.

The workflow for fitting the bulk filtration data comprised three steps: prior predictive checks, HMC sampling, and posterior predictive checks.

3.4.1 Prior predictive checks

Prior predictive checks were performed to assess the appropriateness of the priors chosen in the previous section 3.3.4.

Figure 3.11 shows the expected observed nucleic acid concentrations given the prior for all combinations of hyperparameters and the lowest and highest values of the three input variables. The density was calculated by Hamiltonian Monte Carlo sampling for 1000 iterations. While any nucleic acid concentration value in the interval $(0, \infty)$ is within the support of the prior distribution, the majority of the values fall within the orders of magnitude expected from a liquid bacterial culture like bulk filtration AST. The extremes of the resulting outcomes under the prior distributions did not stray into unrealistic values that are physically impossible.

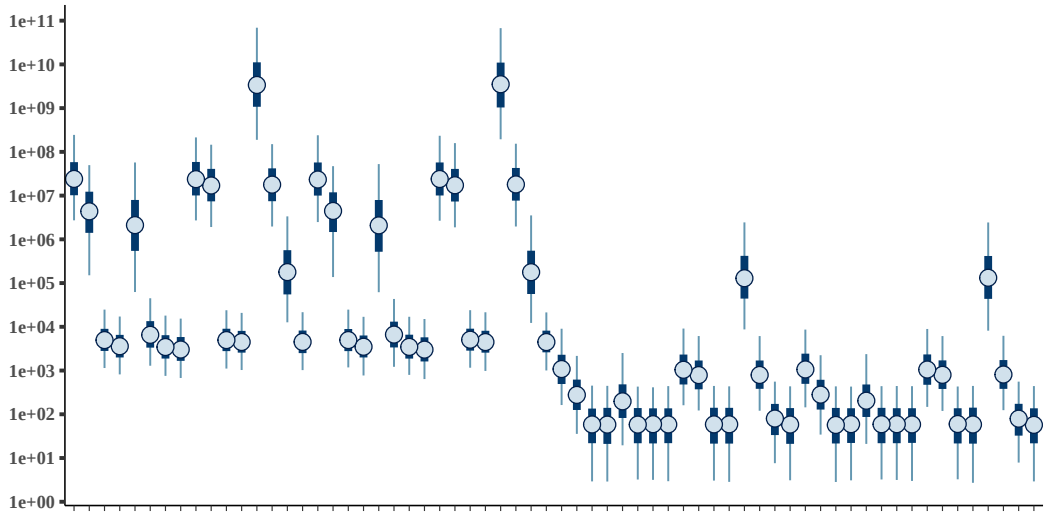


Figure 3.11: Prior predictive checks for modeling bulk filtration AST results. All combinations (x-axis, labels omitted for space) of a duration of 10 or 240 minutes, a dose of 0 or 256 µg/mL, and an inoculum of 1 or 20000 were applied to all combinations of antibiotic (ETP or CRO), amplification protocol (± RT), and five strains' EC_{50} priors (only 1 strain shown for space; effect was indistinguishable at this scale). Y-axis shows the range of nucleic acid concentrations reached under the prior assumptions. The minimum and maximum Y-axis ranges reached lie within reasonable physical limits, assuring that the chosen priors do not preclude accurate inference.

The trace (Figure 3.12) and autocorrelation (Figure 3.13) plots of the prior sampling show no issues as well.

3.4.2 Hamiltonian Monte Carlo No-U-Turn Sampler calculates the parameter posteriors

Five Monte Carlo chains were run for 5000 iterations each after a 3000 iteration adaptation period. The convergence of the chains was usable but not perfect, with some $\hat{R}$ statistics being larger than 1.02. It appears in the trace plots of Figure 3.14b that the posterior of the intercepts for some experiments are multimodal, and that running more iterations may be necessary. Poor sampling is seen in the trace for chain 1 of parameter $\sigma_{b_{p=\text{-RT}}}$.

The majority of parameters showed low or slight autocorrelation, indicating efficient sampling (Figure 3.15). Unfortunately, the parameters $m_{p=\text{-RT}}$, $m_{p=\text{+RT}}$, $\sigma_{m_{p=\text{-RT}}}$, $\sigma_{m_{p=\text{+RT}}}$, $b_{p=\text{-RT}}$, and $\sigma_{b_{p=\text{-RT}}}$ exhibit moderately high autocorrelation in all chains. The slope and intercepts for batch 6, PCR 2 (Figures 3.6f, 3.17j) batch 6, PCR 3 (Figures 3.6f, 3.17k) batch 10, PCR 1 (Figures 3.6j, 3.17q), batch 10, PCR 2 (Figures

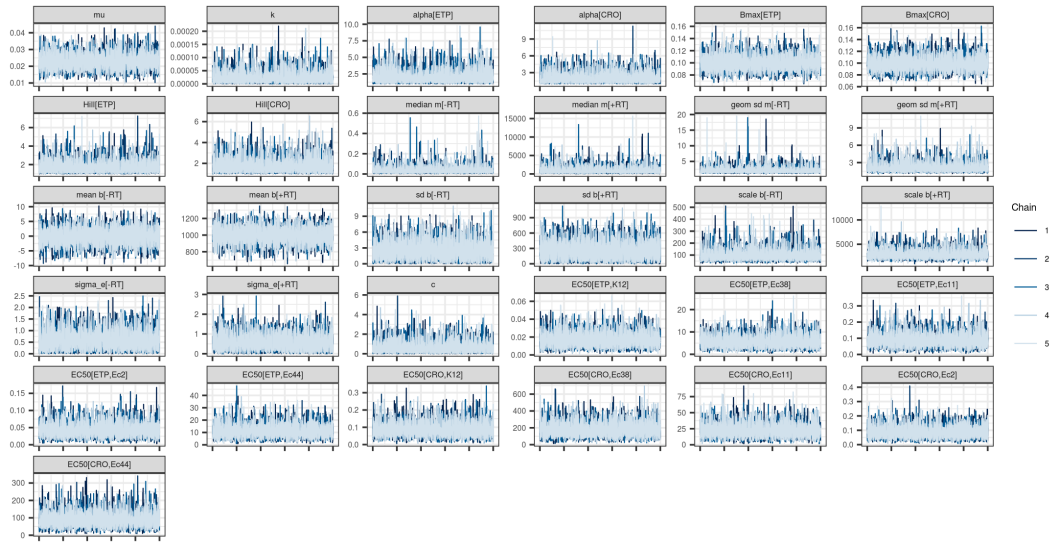


Figure 3.12: Monte Carlo chain traces of the prior distribution for the bulk filtration AST dynamic model.

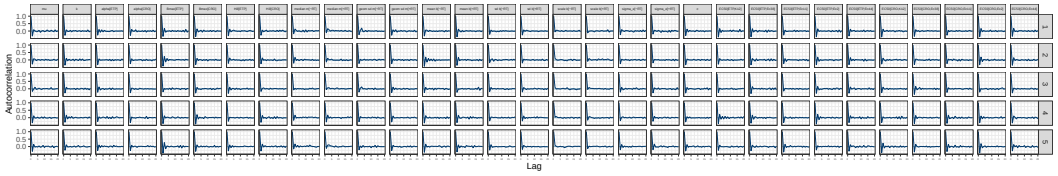


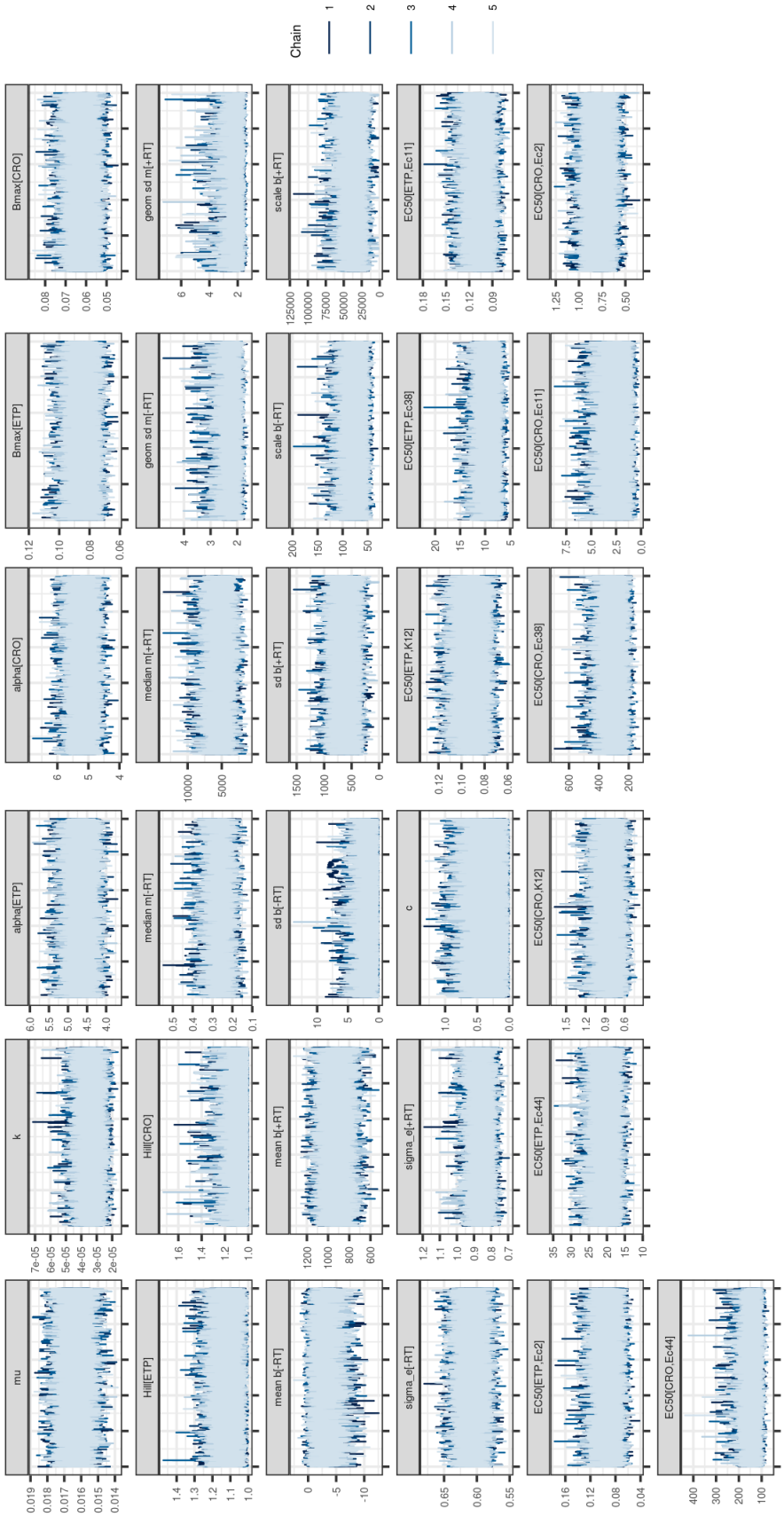
Figure 3.13: Autocorrelation of Monte Carlo samples of the prior distribution for the bulk filtration AST dynamic model.

3.6j, 3.17r), batch 13, PCR 1 (Figures 3.7a, 3.17u), and batch 15, PCR 2 (Figures 3.7c, 3.17y) exhibited moderate autocorrelation as well. Chain 1 exhibited high autocorrelation in additional parameters. High autocorrelation does not indicate a bias in the posterior, simply a lower computation efficiency.

3.4.3 Parameter posteriors and posterior predictive checks

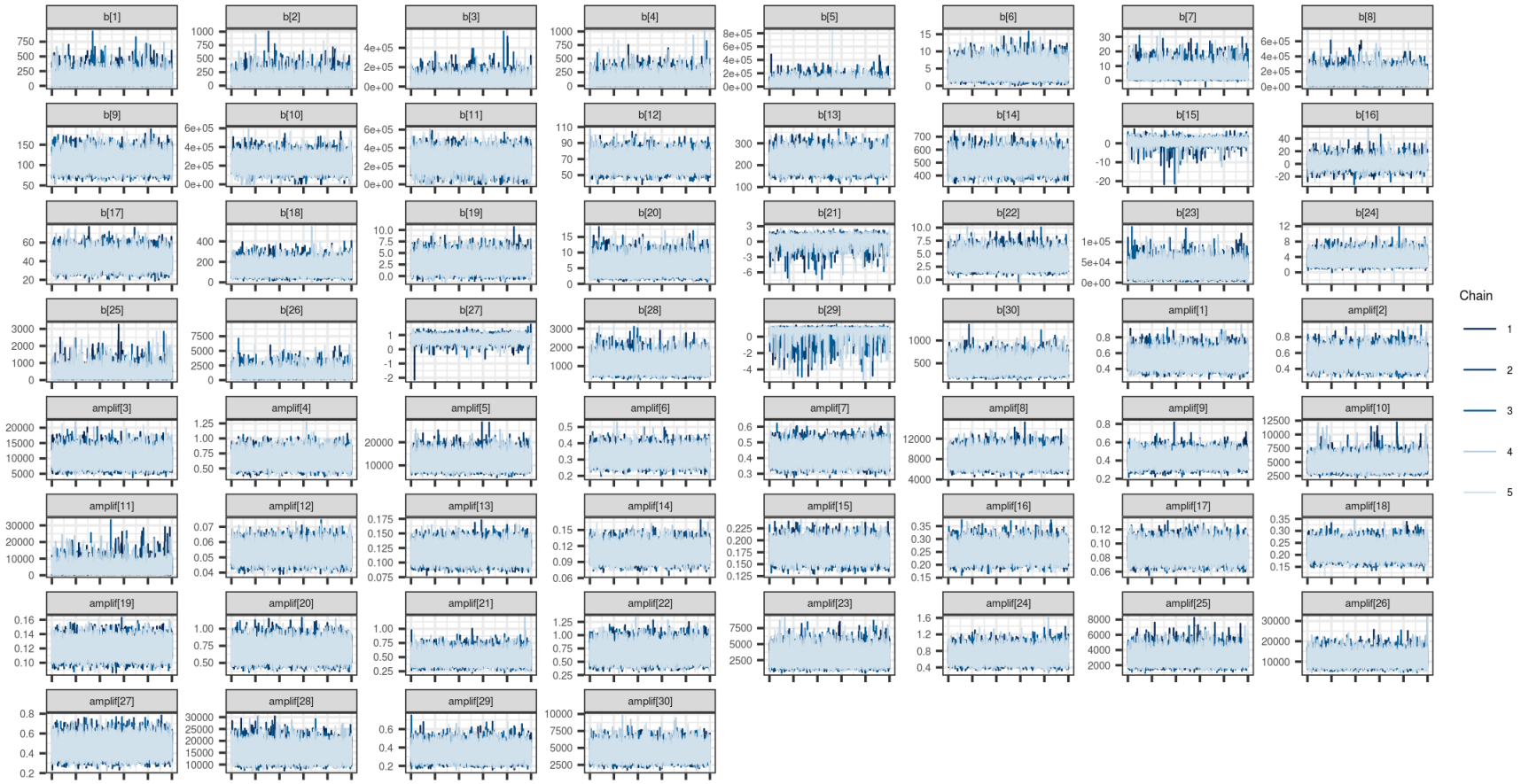
The Monte Carlo samples of the posterior obtained by Stan represents the multidimensional joint probability distribution of all the model parameters.

To summarize this joint distribution, the marginal means of the Monte Carlo sampling are reported in Table 3.3.



(a) Shared hyperparameters

Figure 3.14: Monte Carlo chain traces for parameters in the bulk filtration AST dynamic model.



(b) Experiment-specific parameters

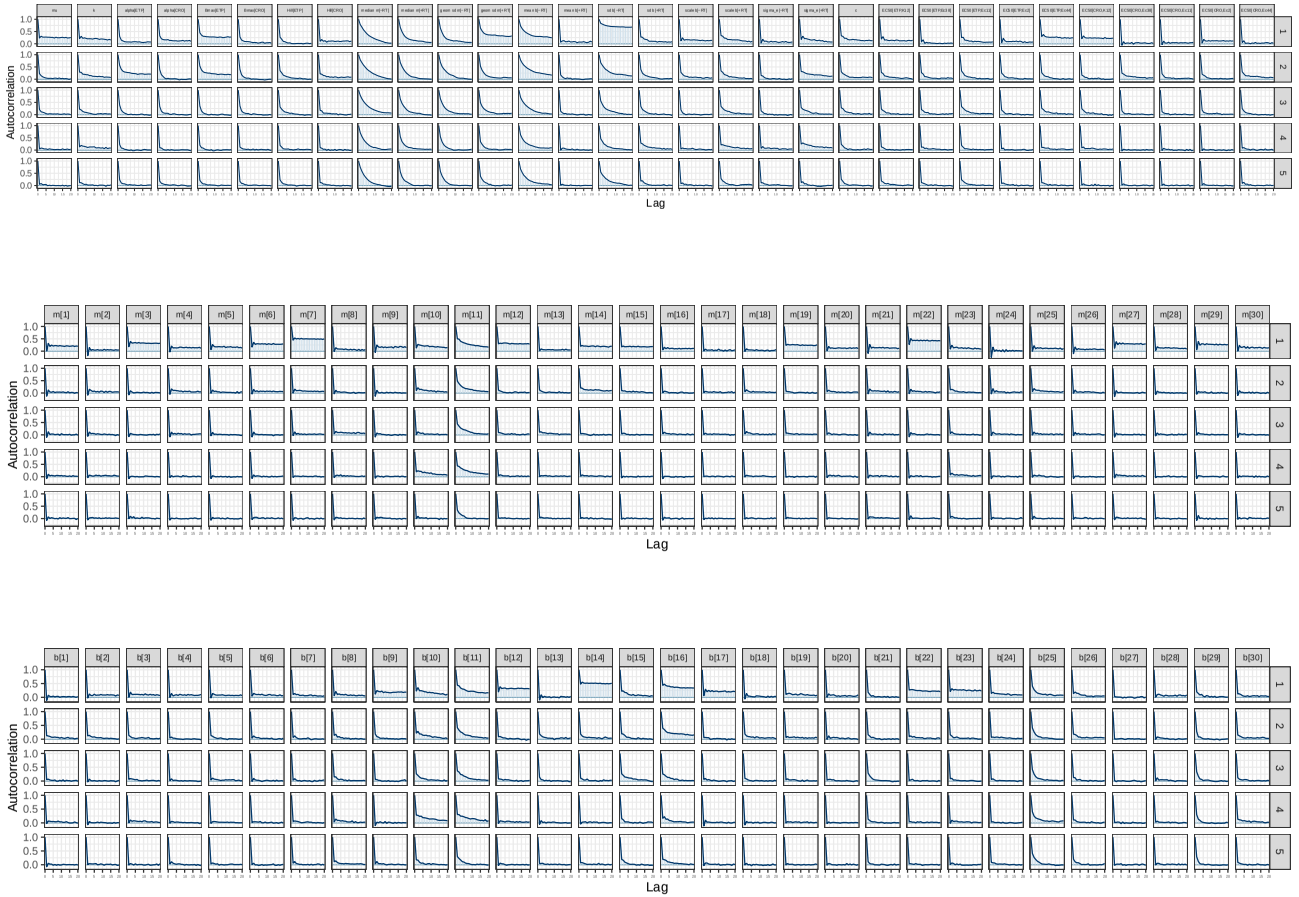


Figure 3.15: Autocorrelation of Monte Carlo samples for parameters of the bulk filtration AST dynamic model (a) Shared hyperparameters (b) Amplification efficiency (c) Intercept parameters

Parameter	Marginal 5% quantile	Marginal median	Marginal 95% quantile	Marginal mean	Geometric median	Units
μ	0.0152	0.0163	0.0174	0.0163	0.0186	min^{-1}
k	0.0000251	0.0000336	0.0000444	0.0000341	0.0000320	min^{-1}
α[ETP]	4.22	4.63	5.15	4.65	4.54	
α[CRO]	4.68	5.17	5.69	5.17	5.13	
B_{max}[ETP]	0.0730	0.0840	0.0972	0.0843	0.0849	min^{-1}
B_{max}[CRO]	0.0536	0.0616	0.0706	0.0617	0.0668	min^{-1}
γ[ETP]	1.02	1.10	1.21	1.11	1.10	
γ[CRO]	1.01	1.06	1.20	1.08	1.16	
EC_{50}[ETP, K12]	0.0758	0.0903	0.108	0.0908	0.0935	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[ETP, Ec38]	6.97	9.29	12.3	9.39	8.74	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[ETP, Ec11]	0.0877	0.106	0.130	0.107	0.135	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[ETP, Ec2]	0.0695	0.0904	0.118	0.0917	0.113	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[ETP, Ec44]	16.1	20.1	25.2	20.3	21.2	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[CRO, K12]	0.603	0.811	1.09	0.822	1.05	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[CRO, Ec38]	210.	291	405	298	329	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[CRO, Ec11]	0.947	2.27	4.39	2.43	6.69	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[CRO, Ec2]	0.606	0.776	0.962	0.778	0.855	$\mu\text{g} \cdot \text{mL}^{-1}$
EC_{50}[CRO, Ec44]	102	138	198	142	170	$\mu\text{g} \cdot \text{mL}^{-1}$
$m_{p=\text{-RT}}$	0.190	0.256	0.348	0.261	0.180	copies/µL/CFU

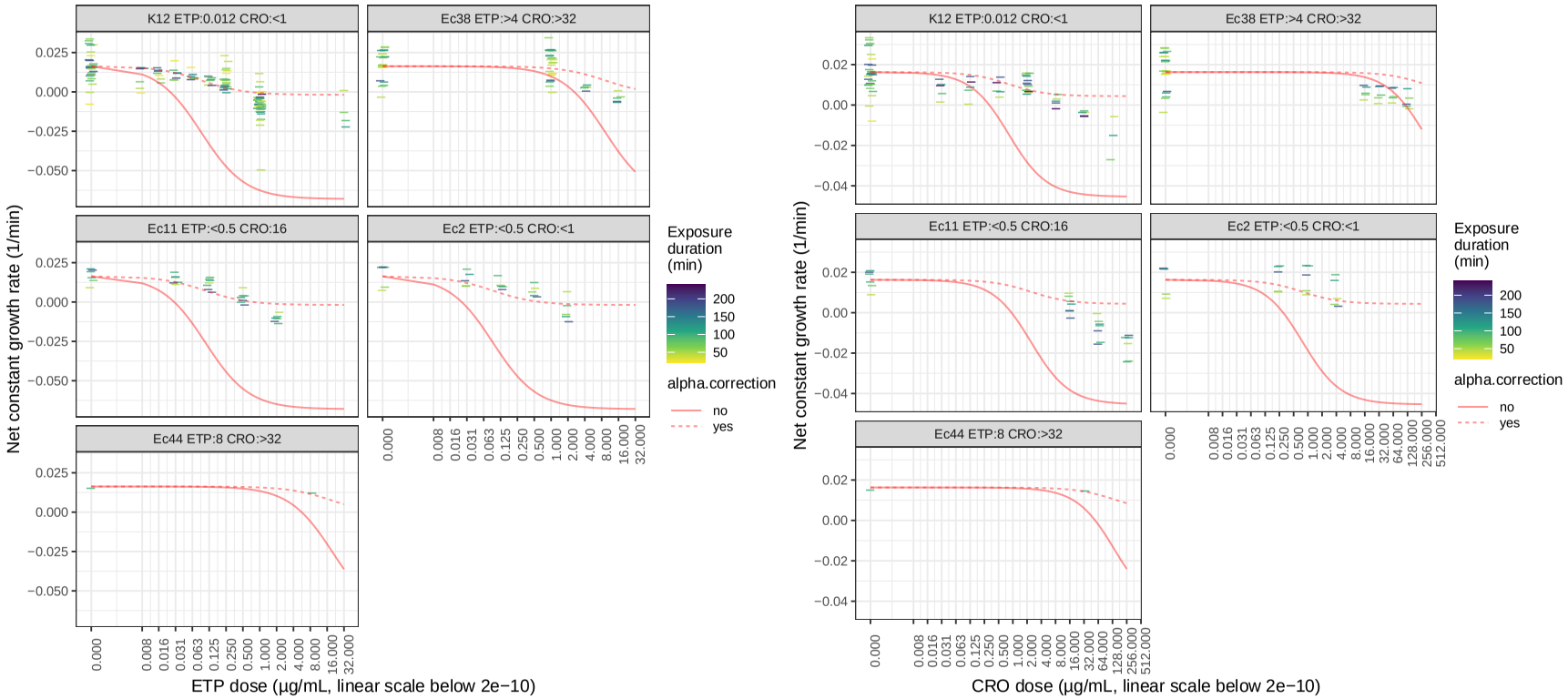
Parameter	Marginal 5% quantile	Marginal median	Marginal 95% quantile	Marginal mean	Geometric median	Units
$m_{p=\text{+RT}}$	3083	5266	7711	5308	4661	copies/µL/CFU
$\sigma_{m_{p=\text{-RT}}}$	1.85	2.18	2.80	2.24	2.02	copies/µL/CFU
$\sigma_{m_{p=\text{+RT}}}$	1.59	2.07	3.14	2.19	2.41	copies/µL/CFU
$b_{p=\text{-RT}}$	-6.86	-3.06	-3.45	-3.29	-1.45	copies $\cdot \mu L^{-1}$
$b_{p=\text{+RT}}$	755	913	1082	917	1031	copies $\cdot \mu L^{-1}$
$\sigma_{b_{p=\text{-RT}}}$	0.130	1.48	5.53	1.93	0.622	copies $\cdot \mu L^{-1}$
$\sigma_{b_{p=\text{+RT}}}$	339	584	908	598	743	copies $\cdot \mu L^{-1}$
$\lambda_{b_{p=\text{-RT}}}$	51.2	72.5	105	74.6	64.9	copies $\cdot \mu L^{-1}$
$\lambda_{b_{p=\text{+RT}}}$	16300	32300	56000	33700	45900	copies $\cdot \mu L^{-1}$
$\sigma_{\epsilon_{p=\text{-RT}}}$	0.592	0.608	0.633	0.608	0.631	$\ln(\text{copies} \cdot \mu L^{-1})$
$\sigma_{\epsilon_{p=\text{+RT}}}$	0.786	0.859	0.948	0.862	0.834	$\ln(\text{copies} \cdot \mu L^{-1})$
c	0.0251	0.279	0.772	0.321	0.255	copies $\cdot \mu L^{-1}$

Table 3.3: Fitted marginal posterior mean values of the bulk filtration AST pharmacodynamic model.

Figures 3.17a–3.17ad compare all data points used to fit the model to the posterior predictive distribution. The posterior predictive distribution are the values predicted by the model given the corresponding input variables of time, dose, inoculum, strain identity, antibiotic identity, and PCR protocol identity as the original data point.

Underfitting of the model manifests as systematic mismatches between model predictions and observations within the same experiment or PCR batch of an experiment. There are differences in the slopes of lines in some of the untreated conditions, indicating that the growth rate differed between experiments. A difference in lysis efficiency between feed and filtrate fractions, as seen in Figure 3.4, may explain why the growth rate was constrained to be the average of the slopes of the feed and filter fractions. Another explanation is that the apparent contamination in experiments with RT will be greater for higher nucleic acid concentrations due to the larger dilution factor performed on those samples prior to ddPCR, yet all observations in the experiment are modeled with an intercept that is more informed by the earlier timepoints, when the background lysis cannot explain any filtrate signal seen in the untreated conditions. A third explanation would be the inability of the model to accurately capture the dose-response parameters due to lack of data at high antibiotic doses. In Figure 3.16, uncertainty in the values of $\beta_{\max}$ may have arisen due to a lack of data at the highest doses. This uncertainty manifested as a correlation between the fitted values of α and $\beta_{\max}$, but was mitigated by the moderately informative priors set for $\beta_{\max}$.

While some model mis-specification or underfitting is apparent, the overall agreement between model and data suggests that the mean parameter values in Table 3.3 can be used with reasonable confidence.



(a) ertapenem

(b) ceftriaxone

Figure 3.16: Fitted dose-response curves for *in vitro* antibiotic exposure. The data shown are the constant net growth rates calculated for each data point relative to the earliest data point of the experiment batch. The solid red line shows the fitted values of $\mu - \frac{\beta_{\max} A^{\gamma}}{\mathrm{EC}_{50}^{\gamma} + A^{\gamma}}$ for the dose A, which is the expected instantaneous growth rate at longer exposure durations. The dotted red line shows the same function using $\beta^{*}_{\max} = \frac{\beta_{\max}}{\alpha}$. This function is the average instantaneous net growth rate during shorter exposure durations and will be similar in magnitude to the empirical net growth rate.

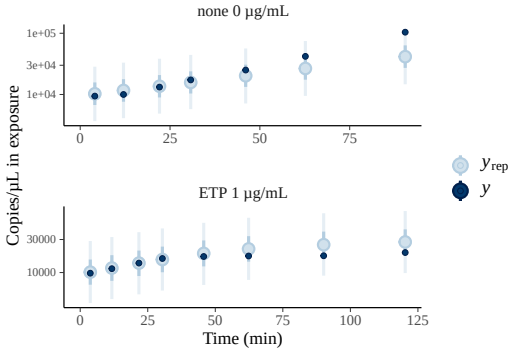
3.5 Future work

As mentioned above, the inclusion of information about dilution volumes to model the dilution of reagent contamination and the dilution of inhibiting MHB media may negate the need for the exponentially-modified Gaussian distribution of amplification intercepts.

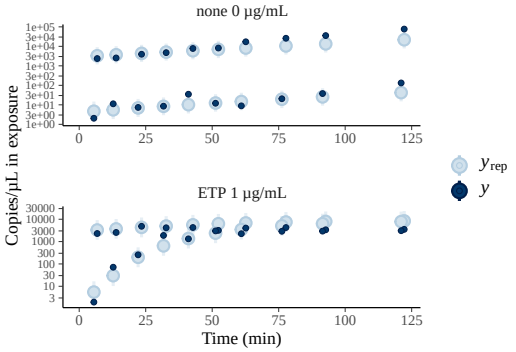
Model fits also improve with improved data. Increasing the number of antibiotic doses examined would better resolve the values of $\beta_{\max}$ and α. The evaluation of standard quality control strains, whose MICs are known with greater precision, would help to improve the estimates of the pharmacodynamic parameters.

Exploration of additional taxa, including pseudomonads and Gram-positive Firmicutes, would answer questions about whether these other clinically important taxa are amenable to accessibility AST in general. Specific concerns include the release of extracellular nucleic acids or secretion of extracellular nucleases.

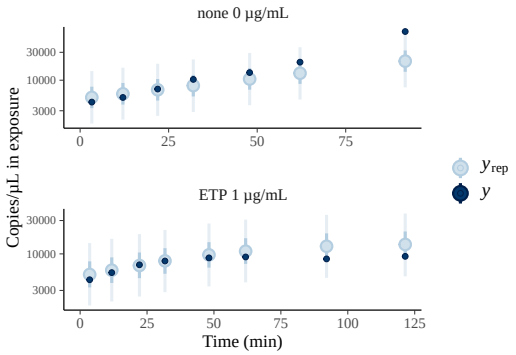
Exploration of additional antibiotics that target the cell wall would also answer questions about the distribution of the α parameter among the beta-lactams. A database of parameter values would be needed should accessibility AST ever be adopted into clinical use.



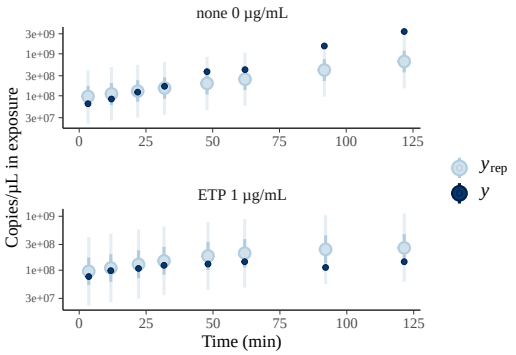
(a) Batch 1, PCR 1, -RT. Feed only, 18750 CFU.



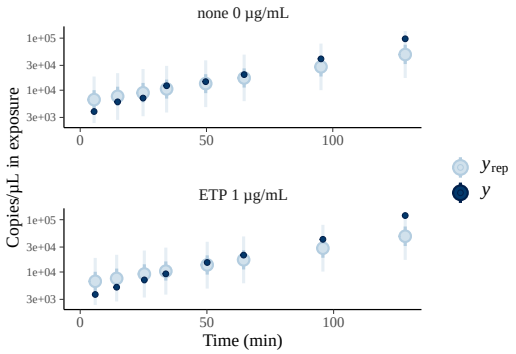
(b) Batch 4, PCR 1, -RT. Feed and filtrate shown, 9375 CFU.



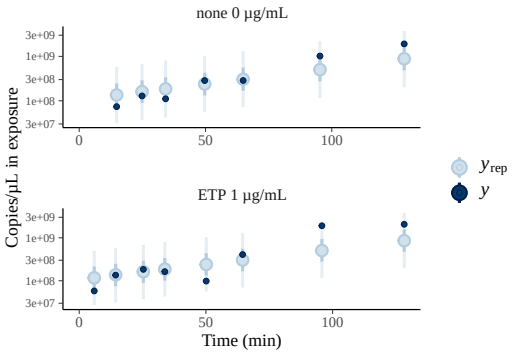
(c) Batch 2, PCR 1, -RT. Feed only, 9375 CFU.



(d) Batch 2, PCR 2, +RT. Feed only, 9375 CFU.

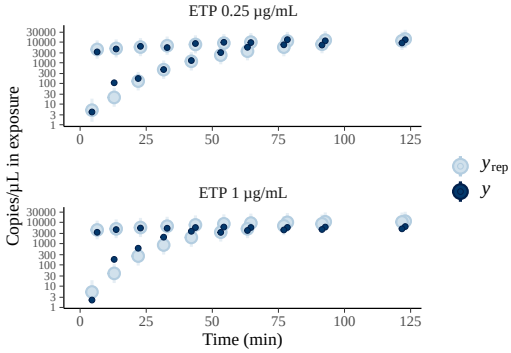


(e) Batch 3, PCR 1, -RT. Feed only, 9375 CFU.

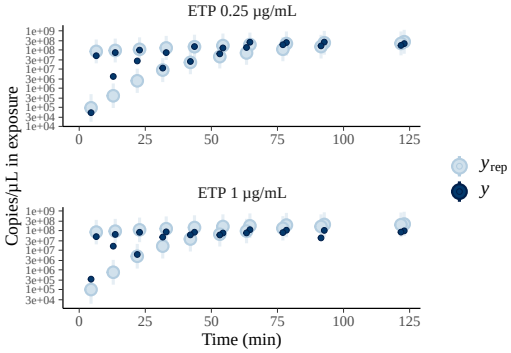


(f) Batch 3, PCR 2, +RT. Feed only, 9375 CFU.

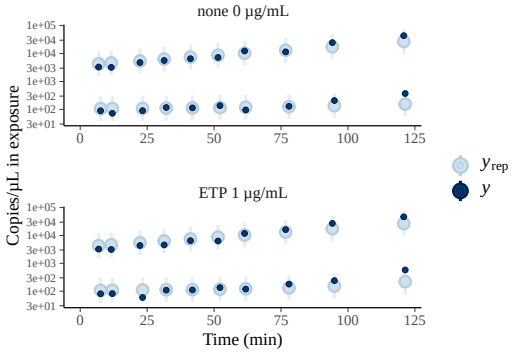
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model.



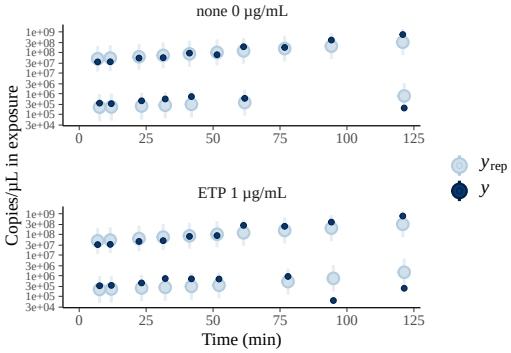
(g) Batch 5, PCR 1, -RT. 9375 CFU.



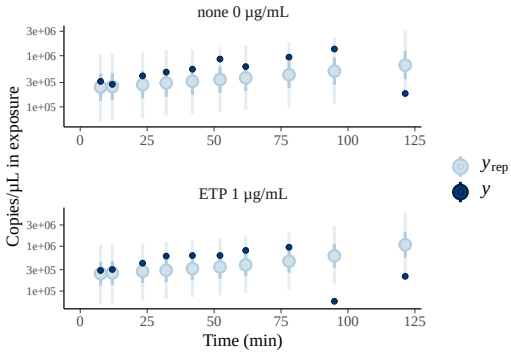
(h) Batch 5, PCR 2, +RT. 9375 CFU.



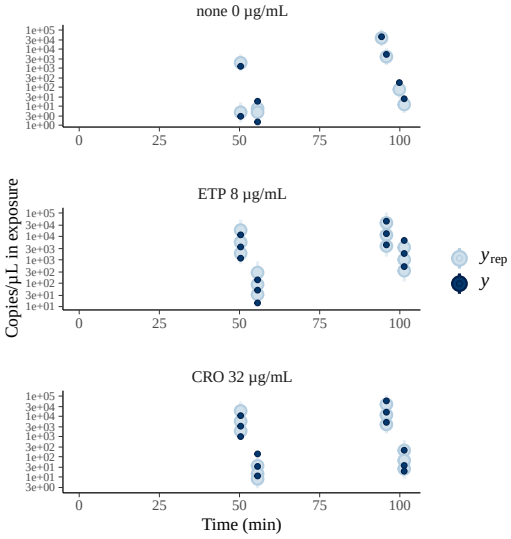
(i) Batch 6, PCR 1, -RT. 9375 CFU.



(j) Batch 6, PCR 2, +RT. 9375 CFU.

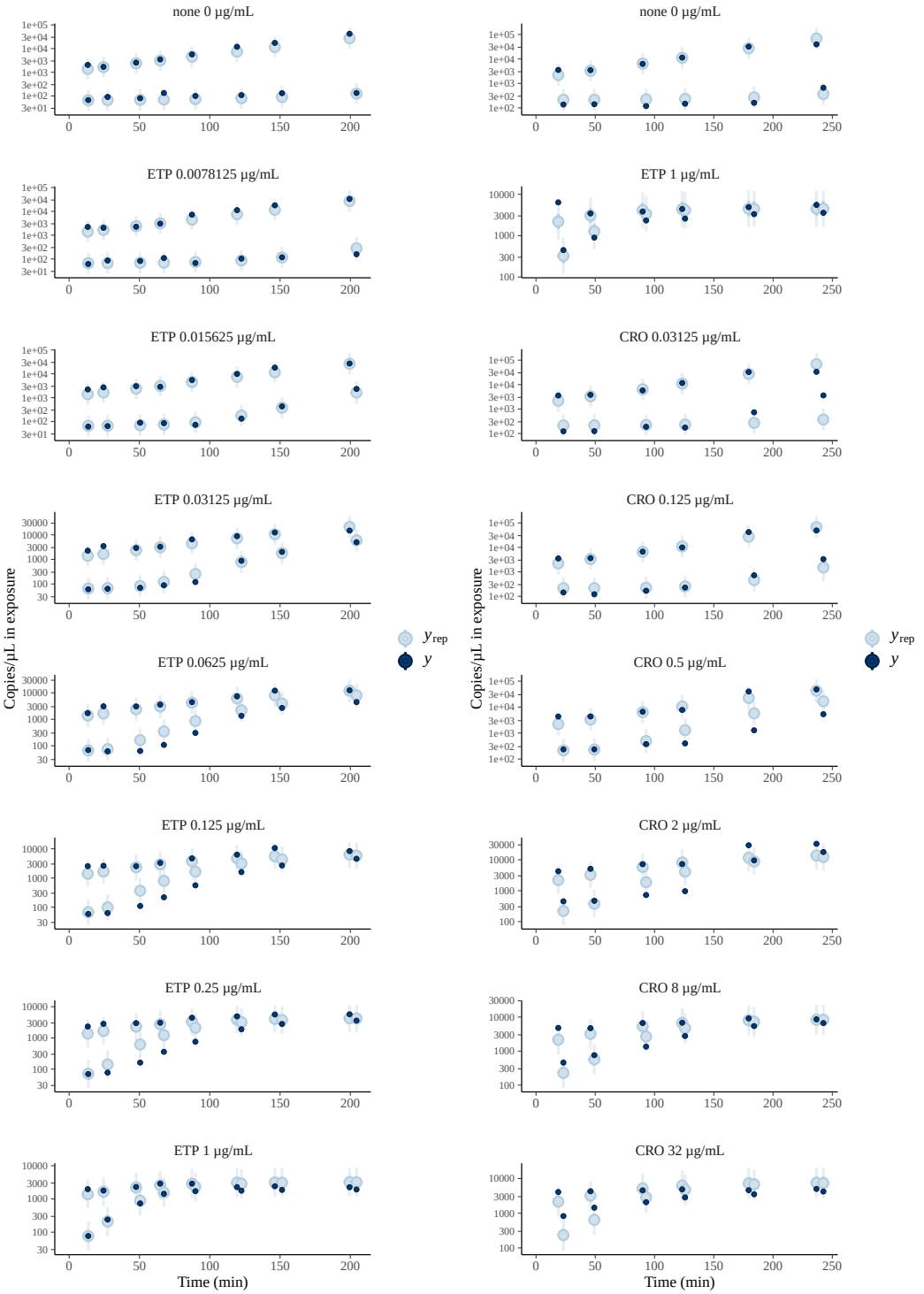


(k) Batch 6, PCR 3, +RT. Filtrate only, 9375 CFU.



(l) Batch 18, PCR 1, -RT. 12500, 3750, or 1250 CFU.

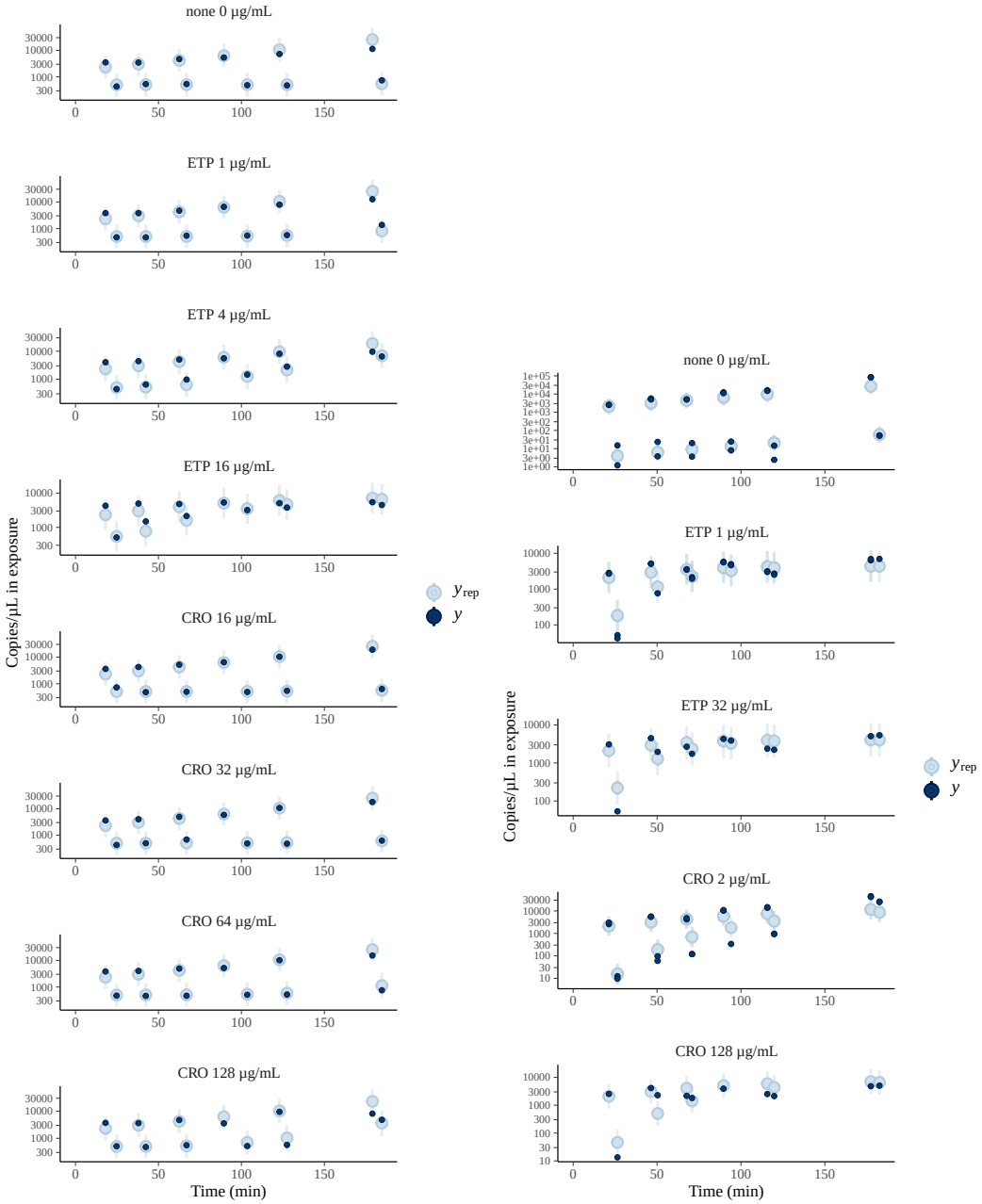
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



(m) Batch 7, PCR 1, -RT. 20000 CFU.

(n) Batch 8, PCR 1, -RT. 12500 CFU.

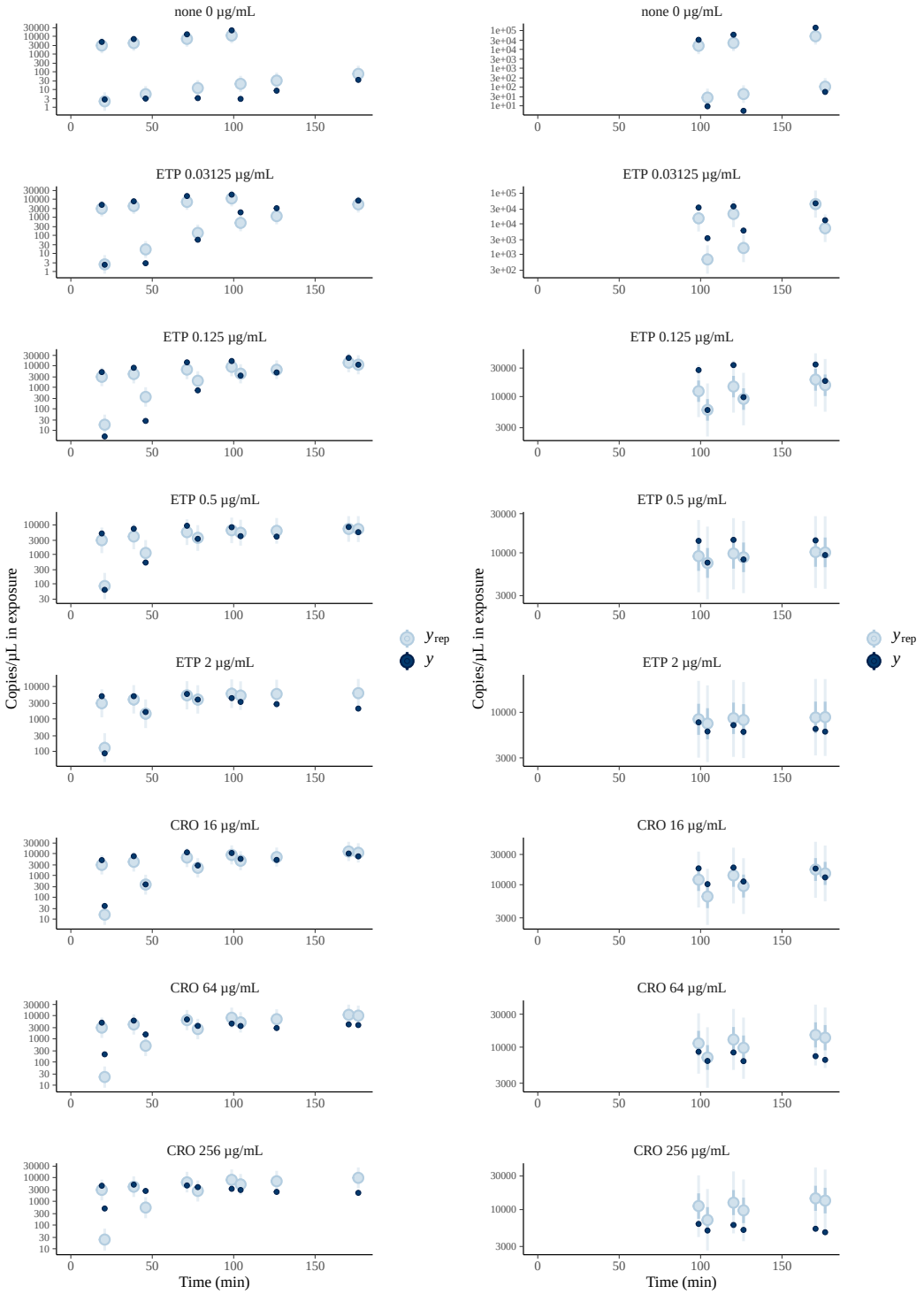
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



(o) Batch 9, PCR 1, -RT. 12500 CFU.

(p) Batch 12, PCR 1, -RT. 12500 CFU.

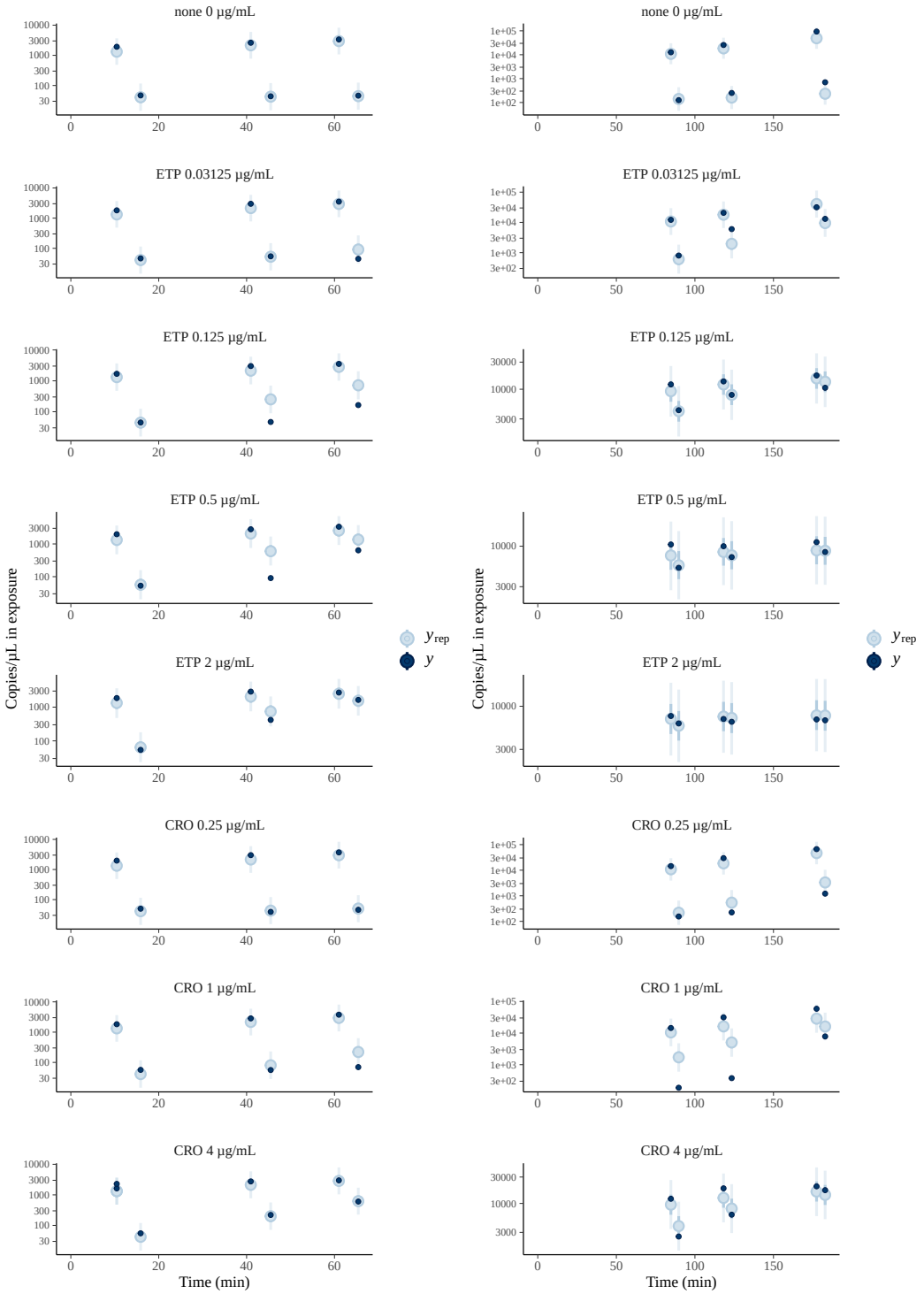
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



(q) Batch 10, PCR 1, -RT. 12500 CFU.

(r) Batch 10, PCR 2, -RT. 12500 CFU.

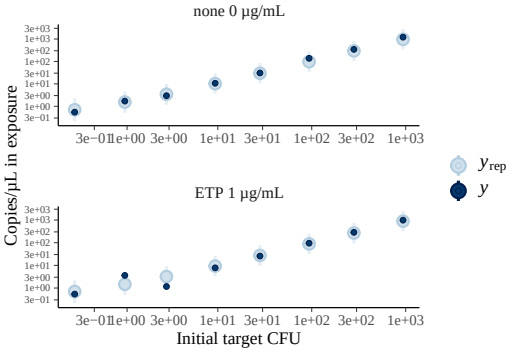
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



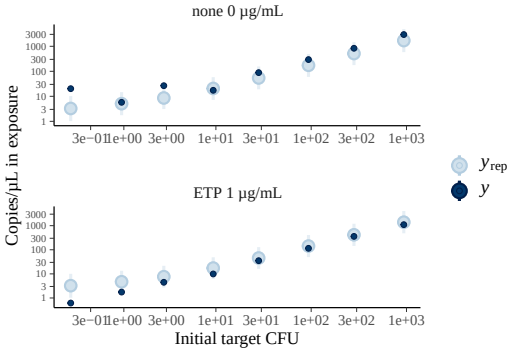
(s) Batch 11, PCR 1, -RT. 12500 CFU.

(t) Batch 11, PCR 2, -RT. 12500 CFU.

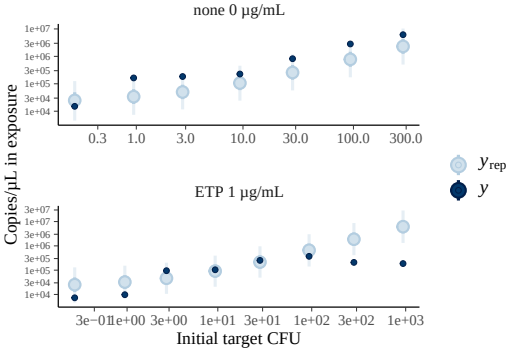
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



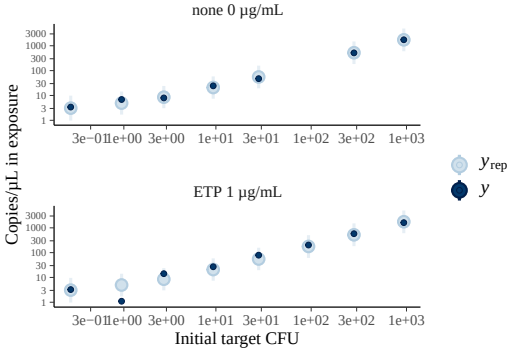
(u) Batch 13, PCR 1, -RT. Feed only, 45 min.



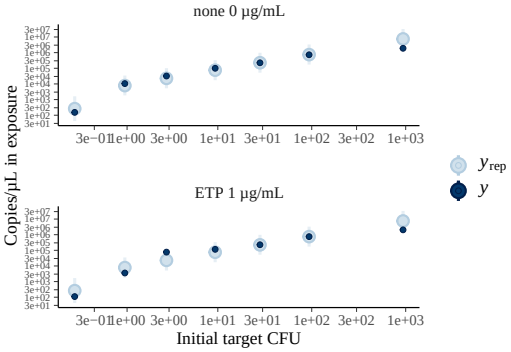
(v) Batch 14, PCR 1, -RT. Feed only, 60 min.



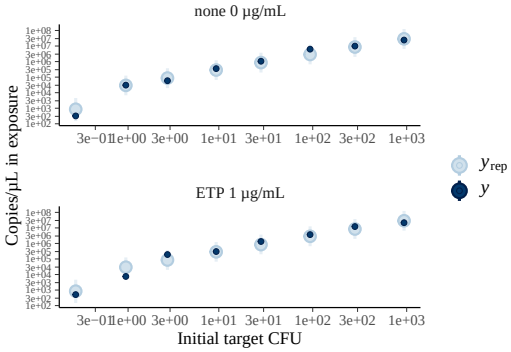
(w) Batch 14, PCR 2, +RT. Feed only, 60 min.



(x) Batch 15, PCR 1, -RT. Feed only, 60 min.

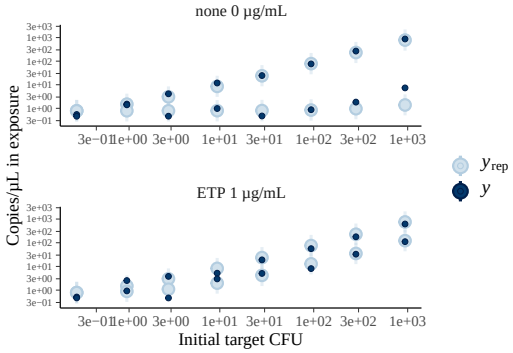


(y) Batch 15, PCR 2, +RT. Feed only, 60 min.

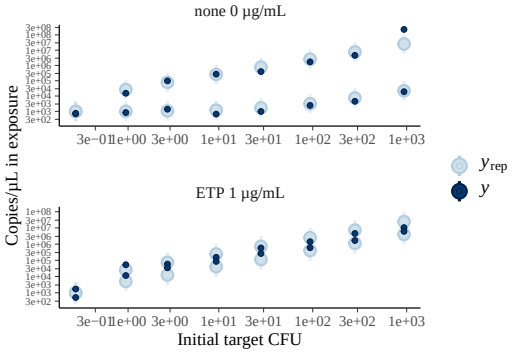


(z) Batch 15, PCR 3, +RT. Feed only, 60 min.

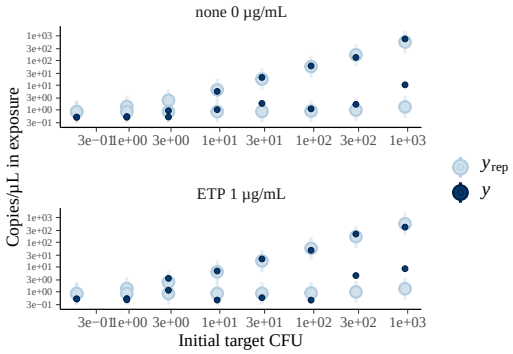
Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).



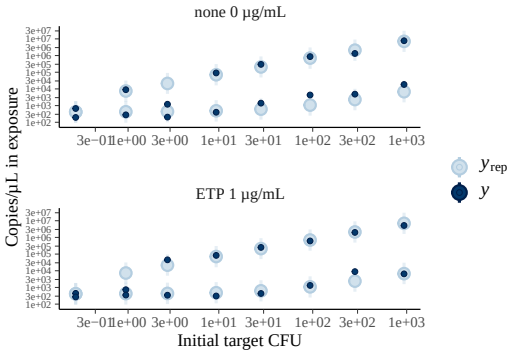
(aa) Batch 16, PCR 1, -RT. 30 min.



(ab) Batch 16, PCR 2, +RT. 30 min.



(ac) Batch 17, PCR 1, -RT. 30 min.



(ad) Batch 17, PCR 2, +RT. 30 min.

Figure 3.17: Posterior predictive distributions for the fitted bulk filtration AST dynamic model (cont'd).

Chapter 4

DIGITAL FILTRATION AST

In this chapter, I describe a variation of filtration accessibility AST wherein batches of parallel, digitally-loaded antibiotic exposures are employed and analyzed together, increasing the amount of information gathered per experimental condition. The analytic sensitivity of this new method of digital filtration AST approaches single cell resolution.

4.1 Bulk assays are limited by low inoculums

The experimental protocol and analysis used in Chapter 3 assumed that the amount of bacteria is a continuous, infinitely divisible quantity. This assumption was reasonable when the total number of cells in the bulk filtration AST was large. At low numbers of cells, however, the bulk filtration AST results diverge from the expectations. The range of AST metrics increases, incorrect AST calls are more frequent, and sometimes the impossible situation of more extracellular nucleic acids than total nucleic acids arises, as seen in panels B and F of Figure 3.7.

It is important to understand the performance of phenotypic ASTs at low cell numbers due to the low cells present in clinical samples like blood. There are three phenomena in particular that I believe are likely to contribute the breakdown of bulk AST results at low inocula. First, the ratio of nucleic acid amplification measurement errors to signal is lower when the measured nucleic acids are more dilute. Second, the inherent discreteness and stochasticity of bacterial cells is not captured by the mean expected value present in the compartment model. Lastly, the physical splitting of a liquid specimen, performed to create parallel experimental conditions, introduces uncertainty in the total number of cells in the resulting partitions, and the uncertainty is relatively larger at low cell counts.

It will always be possible, and likely, that additional unidentified phenomena affect the the output of filtration AST than those listed here. For now, they are grouped conceptually with the empirically defined "nucleic acid amplification measurement errors." Future efforts to improve the implementation of filtration AST will investigate these processes.

4.1.1 Signal-to-noise ratios

The first phenomena, or group of phenomena, that may cause bulk filtration AST errors at low cell numbers is the presence of random or otherwise uncontrolled noise in nucleic acid concentration quantification, called nucleic acid amplification measurement errors here. Source of noise in nucleic acid amplification, specifically the ddPCR used in Chapter 3, include non-specific amplification from primer dimers, contamination of samples or reagents by templates not derived from the sample, artifacts that produce signal in the measured fluorescence channel but not from the EvaGreen dye, inhibitors of the PCR reaction, and thermal noise. Processes outside of the ddPCR reaction occur too, such as errors in pipette volume due to differing surface tensions or improper pipette usage, evaporation of water from the sample during the AST protocol, contamination of the sample by nucleases, omission of mixing steps, and errors in the recording of other variables such as time or antibiotic dosage. All of these nucleic acid amplification measurement errors are not functions of target cell number, so their magnitude will be constant as cell number decreases. (The exception are inhibitor compounds, which may be produced by the bacteria, and whose dilution is also correlated with cell number. Analysis of these inhibitor compounds may be performed as future work.)

The signal in bulk filtration AST comes only from the bacteria of interest, and its magnitude therefore decreases as the cell number decreases. It follows then that the probability of an incorrect assay call increases as the number of cells decreases.

The discrete of cells and the stochastic loading of discrete cells are two sources of noise that can be separated logically from the rest of the empirically observed measurement noise listed above. They do depend on cell number, and they cannot be overcome by improving nucleic acid amplification noise or precision in preceding steps of the bulk filtration AST protocol. While the statistical analysis of limits-of-detection and signal-to-noise ratios from nucleic acid amplification measurement errors are frequent in the NAAT literature, most papers do not separate out effect of discreteness when measuring LOD with serial dilutions. Hence, the two phenomena are discussed below.

4.1.2 Discrete, stochastic models versus continuous, deterministic models

The compartment model of section 3.3 makes an key assumption that the amount of bacteria is a continuous, infinitely divisible quantity. However, bacterial cells are actually discrete entities, varying in size within a range of 75–225 femptograms

[136]. In liquid culture, bacteria are located randomly within the liquid specimen volume, and when the specimen is split into smaller volumes, each bacteria must wholly reside in one or the other volumes.

The inherent discreteness of cells makes it impossible for the system to assume all the ratios predicted by the continuous model. In other words, if the model predicts that 0.012 of total biomass will be extracellular, but there are only 4 cells, then the closest possible system states will have 0.0 or 0.25 of the cells dead. Even if one assumes that cells can have different masses, the situation with the least difference between lysed and unlysed cells would require the dead cell to have a size of 0.012 units while the other 3 surviving cells each have a size of $(1 - 0.012)\frac{1}{3} = 0.329$ units. But it is not biologically implausible for one cell to be $0.012/0.329 = 0.036$ the size of another. In an ideal population, the youngest bacteria are $1/2$ the size of the oldest, and in experiments, the coefficient of variation of cell sizes from the ideal age distribution is only about 0.20 in an exponential phase culture [137]. Thus, the discrete nature of cells imposes a minimum granularity on the possible system states. This minimum granularity cannot be overcome by reducing measurement noise.

4.1.3 Stochasticity in sample partitioning limits bulk assays with low inocula

In a separate phenomena, the discrete nature of cells violates the assumptions made when splitting liquid samples to create experimental conditions. Although the concept of stochastic loading is intuitive once pointed out and not novel, I attempt to describe it in detail here because a comprehensive treatment is lacking in the recent literature.

In the particular bulk filtration AST protocol used in Chapter 3, there were two situations in the protocol where bacteria cells were split amongst multiple physical partitions. Splitting was performed to create different experimental conditions to test in parallel. First, a partitioning of the diluted batch culture was performed each time a pipette tip was used to draw up some of the liquid culture, the pipette tip's contents then being used to inoculate and start one antibiotic exposure with a target inoculum. Every batch of bulk filtration AST experiments included at least two experimental conditions inoculated in the above fashion. Secondly, at the end of each antibiotic exposure, a small partition of each antibiotic exposed specimen was set aside to quantify the feed fraction's nucleic acids, while the rest of the specimen

proceeded as the feed of the filtration process.

In the first situation, it is common (but not necessary) to compare, across different experimental conditions, the outputs (nucleic acid concentrations) from one of the fractions. For example, this kind of simple analysis was done in Chapter 2 and in other published rapid AST methods [21–24, 88, 125, 138]. It is assumed that the comparison will reflect strain susceptibility if the output differs from the control condition. In the second situation, the outputs of different fractions were compared across two partitions of one antibiotic-exposed specimen (within one experimental condition). It is assumed that the output in each partition is the same as it would have been before partitioning.

In both situations, the comparisons were informative only because a certain assumption was being made: that the amount of bacteria (and thus nucleic acids) was a continuous quantities that were infinitely divisible, so that splitting a well-mixed sample physically results in partitions with exactly the same density of cells. In other words,

$$C_{\text{before}} = C_1 = \ldots = C_{\text{i}} = \ldots = C_{\text{k}}, \tag{4.1}$$

where C_i is the concentration in partition i, which was always a pipette tip. Only when the tip's contents were expelled into a new vessel does the concentration in the new volume after mixing obey the dilution equation:

$$C_{\text{tip}} V_{\text{tip}} = C_{\text{final}} V_{\text{final}}. \tag{4.2}$$

However, when the discrete nature of cells is considered, the above equations must be modified, and the quantity C must be redefined for the model to be consistent and to be in terms of an integer numbers of cells. First, because the cells are discrete, when the liquid specimen is split into smaller volumes, each bacteria must wholly reside in one or the other volumes. Furthermore, in liquid culture, bacteria are located randomly within the liquid specimen volume. The numbers of bacteria in the partitions become random variables instead of fixed constants as in Chapter 3.

As a random variable, the number of cells in each partition is no longer known with certainty (unless the splitting is done a process that tracks and controls the partitioning of cells, like automated cell sorters). The unknown total cell amount in each partition becomes a confounding factor in the susceptibility calling algorithm. By assuming different values of total cell amount, any outcome can be attained. Such non-identifiabilities will lead to incorrect susceptibility calls.

For example, in the first situation where it is assumed that experimental conditions are inoculated with the same inoculum, if the strain is resistant in a particular experimental condition, but the control condition began with more cell biomass, then the treated condition's output may appear to be lower than the control's, indicating susceptibility. Likewise, if the strain was susceptible, but the control condition began with less cell biomass, then the treated condition's total cell biomass will appear similar to the control's, indicating resistance. In the second situation, fluctuations in the total number of cells in the two partitions would increase or decrease the ratio between the feed and filtrate fraction outputs, potentially causing incorrect susceptibility calls or even nonsensical results if the filtrate output exceeds the feed output. Total cell amount in different experimental conditions is thus a limiting factor in previous accessibility ASTs cited above.

The magnitude of the uncertainty caused by sample partitioning in any situation can be calculated from theory. In the simplest model defining an integer number of cells $N \in \mathbb{Z}^+$, cell size is assumed to be a constant, so that the observed nucleic acid amount $Y \in \mathbb{R}$ is a function of N alone, as follows: $Y = kN$, where $k \in \mathbb{R}$ is a constant. Each cell has an independent chance of being in each of the partitions, with the probability of being in a partition equal to the proportion of that partition's volume $V_{\text{part}} \geq 0$ to the total volume $V_{\text{total}} \geq V_{\text{part}} \geq 0$ currently occupied by the bacteria. In this model, the number of bacteria in the partitions will be multinomially distributed random variable with the probability parameters as described. If there are only 2 partitions, as is the case for the bulk filtration AST experiments of Chapter 3, and one examines the behavior of one of the partitions at a time, then the multinomial distribution is equivalent to the binomial distribution. We can calculate the coefficient of variation to measure the degree of expected dispersion. (For simplicity, the moments of N are discussed, since the moments of $Y = kN$

follow directly.)

$$N_{\text{part}} \sim \text{Binomial}\left(n = N_{\text{total}}, p = \frac{V_{\text{part}}}{V_{\text{total}}}\right) \tag{4.3}$$

$$p\left(N_{\text{part}} = x\right) = \binom{N_{\text{total}}}{x}\left(\frac{V_{\text{part}}}{V_{\text{total}}}\right)^{x}\left(1 - \frac{V_{\text{part}}}{V_{\text{total}}}\right)^{N_{\text{part}}-x} \tag{4.4}$$

$$E\left[N_{\text{part}}\right] = N_{\text{total}}\frac{V_{\text{part}}}{V_{\text{total}}} \tag{4.5}$$

$$\text{Var}\left[N_{\text{part}}\right] = N_{\text{total}}\frac{V_{\text{part}}}{V_{\text{total}}}\left(1 - \frac{V_{\text{part}}}{V_{\text{total}}}\right) \tag{4.6}$$

$$\text{CV}\left[N_{\text{part}}\right] = \frac{\sqrt{\text{Var}\left[N_{\text{part}}\right]}}{E\left[N_{\text{part}}\right]} = \sqrt{\left(\frac{V_{\text{total}} - V_{\text{part}}}{V_{\text{part}}}\right)\frac{1}{N_{\text{total}}}} \tag{4.7}$$

If one redefines the density of cells to be $C = \frac{N}{V}$, then instead of equations 4.5 to 4.7, we have

$$E\left[C_{\text{part}}\right] = \frac{1}{V_{\text{part}}}E\left[N_{\text{part}}\right] = \frac{N_{\text{total}}}{V_{\text{total}}} = C_{\text{total}} \tag{4.8}$$

$$\text{Var}\left[C_{\text{part}}\right] = \frac{1}{V_{\text{part}}^2}\text{Var}\left[N_{\text{part}}\right] = \frac{N_{\text{total}}}{V_{\text{total}}}\left(\frac{1}{V_{\text{part}}} - \frac{1}{V_{\text{total}}}\right) = C_{\text{total}}\left(\frac{1}{V_{\text{part}}} - \frac{1}{V_{\text{total}}}\right) \tag{4.9}$$

$$\text{CV}\left[C_{\text{part}}\right] = \sqrt{\left(\frac{V_{\text{total}} - V_{\text{part}}}{V_{\text{part}}}\right)\frac{1}{N_{\text{total}}}}. \tag{4.10}$$

The coefficient of variation for both N_{part} and C_{part} (and thus Y) goes to 0 as the total number of cells increases. (As a side note, the volume of the partition also affects the CV, with a CV of 0 if the partition is the entire specimen or a CV approaching ∞ if the partition has no volume.) Thus, equation 4.8 becomes 4.1 as $N \to \infty$, explaining why the assumption in equation 4.1 holds for large inoculums. At low inoculums, however, the coefficient of variation for a given partition increases as N decreases.

A more accurate model would account for cell size when relating the observed nucleic acid amount Y to the number of cells: $Y = KN$, where, instead of a constant, K is another random variable representing the distribution of cell sizes. In this case, the same conclusions about $\text{CV}[Y]$ as a function of N apply because N and K are independent, and because $E[K]$, $\text{Var}[K]$, and $\text{CV}[K]$ are all constant with respect

to N.

$$E[Y] = E[KN] = E[K]E[N] \tag{4.11}$$

$$\mathrm{Var}[Y] = \mathrm{Var}[KN] = E[N]^2\mathrm{Var}[K] + E[K]^2\mathrm{Var}[N] + \mathrm{Var}[K]\mathrm{Var}[N] \tag{4.12}$$

$$\mathrm{CV}[Y] = \frac{\sqrt{\mathrm{Var}[Y]}}{E[Y]} = \sqrt{\mathrm{CV}[K]^2 + \mathrm{CV}[N]^2 + \mathrm{CV}[K]^2\mathrm{CV}[N]^2} \tag{4.13}$$

$$= \sqrt{\mathrm{CV}[K]^2 + \left(1 + \mathrm{CV}[K]^2\right)\left(\frac{V_{\mathrm{total}} - V_{\mathrm{part}}}{V_{\mathrm{part}}}\right)\frac{1}{N_{\mathrm{total}}}} \tag{4.14}$$

To understand how the result of our comparison is affected by stochastic splitting of the specimen's N_{total} cells, we must now consider the joint distribution of the partitions' counts. Our comparison will be a function $f[Y_A, Y_B] : \mathbb{R}^2 \to \mathbb{R}$ of the outputs of partitions A and B. The simplest functions are the difference function $f_{\mathrm{diff}}[Y_A, Y_B] = Y_A - Y_B$ and the ratio function $f_{\mathrm{ratio}}[Y_A, Y_B] = \frac{Y_A}{Y_B}$, though for analytic reasons, I will discuss the relative difference $f[Y_A, Y_B] = \frac{Y_A - Y_B}{Y_A + Y_B}$. We now wish to know the spread of a new random variable $Z = f[Y_A, Y_B]$ as a function of N_{total}. For space, we again consider the case where $Y_A = kN_A$ so that we redefine Z as $Z = f[N_A, N_B]$. We also will only consider the simple case where partitions A and B together take up the entire specimen; that is, $V_A + V_B = V_{\mathrm{total}}$ and $N_A + N_B = N_{\mathrm{total}}$.

Because the ratio function can take on a value of ∞, the moments of its random variable Z_{ratio} are not defined. The difference function remains finite, but it can take on negative values that make the coefficient of variation unsuitable as a metric of how its Z is dispersed. For the difference function, one could consider the variance of Z_{diff} itself (instead of relative variance via the coefficient of variation), but $Var[Z_{\mathrm{diff}}] = 4N_{\mathrm{total}}\left(\frac{V_A}{V_{\mathrm{total}}}\right)\left(1 - \frac{V_A}{V_{\mathrm{total}}}\right)$ scales with N_{total}, so it is hard to understand how much more $Var[Z_{\mathrm{diff}}]$ increases as a function of N_{total}, as was done in equations 4.3 using the coefficient of variation. Therefore, I will discuss instead the relative difference function. The relative difference is the difference function, but normalized by N_{total}. When a volume is split into equal halves, the expected value of the difference and of the relative difference between the counts of the two halves is 0. But, unlike the difference, which is bounded by and grows with N_{total}, the relative difference always takes on a value between -1 and 1. Thus, the (non-relative) variance of the relative difference of $Y_A \propto N_A$ and $Y_B \propto N_B$ is a suitable measure of how much relative deviations from the mean value will occur during splitting as a

function of N_{total}.

$$E[Z] = \sum_{N_A}\sum_{N_B} f[N_A, N_B]\text{Prob}\,(N_A, N_B) \tag{4.15}$$

$$= \sum_{N_A=0}^{N_{\text{total}}} \frac{(N_A - (N_{\text{total}} - N_A))}{N_{\text{total}}} \binom{N_{\text{total}}}{N_A} \left(\frac{V_A}{V_{\text{total}}}\right)^{N_A} \left(1 - \frac{V_A}{V_{\text{total}}}\right)^{N_{\text{total}}-N_A} \tag{4.16}$$

$$= 2\left(\frac{V_A}{V_{\text{total}}}\right) - 1 \tag{4.17}$$

$$\text{Var}[Z] = E[Z^2] - E[Z]^2 \tag{4.18}$$

$$= -E[Z]^2 + \sum_{N_A}\sum_{N_B} (f[N_A, N_B])^2\,\text{Prob}\,(N_A, N_B) \tag{4.19}$$

$$\begin{aligned} = -E[Z]^2 + \sum_{N_A=0}^{N_{\text{total}}} & \frac{(N_A - (N_{\text{total}} - N_A))^2}{N_{\text{total}}} \\ & \times \binom{N_{\text{total}}}{N_A} \left(\frac{V_A}{V_{\text{total}}}\right)^{N_A} \left(1 - \frac{V_A}{V_{\text{total}}}\right)^{N_{\text{total}}-N_A} \end{aligned} \tag{4.20}$$

$$= \frac{4}{N_{\text{total}}}\left(\frac{V_A}{V_{\text{total}}}\right)\left(1 - \frac{V_A}{V_{\text{total}}}\right) + \left(2\left(\frac{V_A}{V_{\text{total}}}\right) - 1\right)^2 - E[Z]^2 \tag{4.21}$$

$$= \frac{4}{N_{\text{total}}}\left(\frac{V_A}{V_{\text{total}}}\right)\left(1 - \frac{V_A}{V_{\text{total}}}\right) \tag{4.22}$$

The variance of $Z = \frac{N_A - N_B}{N_A + N_B}$ thus decreases to 0 as N_{total} increases, but increases as N_{total} decreases, the exact value being a function of the relative evenness of the partition volumes (i.e., $\frac{V_A}{V_{\text{total}}}$) and N_{total}.

In the worst case, only one cell is present in a specimen: $N_{\text{total}} = 1$. There will never be an equal number of cells in both conditions of the specimen is split into a treated and an untreated condition, and the results of the AST will be nonsensical. This is apparent because if the specimen is split evenly ($V_A = \frac{1}{2}V_{\text{total}}$), then the $\text{SD}[Z] = \sqrt{\text{Var}[Z]} = 1$, which spans the entire range of -1 to 1. (For uneven splitting, the variance will be less, simply as a reflection of the Shannon information gained.)

We have now discussed in detail the magnitude of the contribution of stochastic liquid specimen partitioning on the observed measurement variation in any phenotypic assay where the measurement technology assesses the population size of bacteria in some way.

It is important to remember that if the total amount of cells in each partition is somehow measured reliably, then in fact, the measurement error from stochastic

liquid specimen partitioning can be subtracted, provided there is a model for how the output measurement varies as a function of the total amount of cells.

For example, if the lysate fraction had been measured in the bulk filtration ASTs of Chapter 3, then any discrepancies in cells between the specimen partition used to measure the feed and the specimen partition undergoing filtration could be inferred. Likewise, if the total inoculum of two different experimental conditions is measured before the start of the antibiotic exposure during bulk filtration AST, then the output measurements can be scaled by the starting inoculum. The model in equation 3.12-3.13 in particular states that the final output is simply linearly proportional to the starting inoculum. Alternatively, by comparing an intrinsic metric such as the percent extracellular (in which the total population amount is normalized out) or the EC_{50} parameter in equation 3.25 instead of an extensive metric like the filtrate concentration (which does scale with the total population amount), any contribution from unequal inoculations is normalized out.

Of course, the magnitude of other error sources such as inefficiencies during the nucleic acid extraction steps or nucleic acid amplification errors may be large enough that the contribution solely from stochastic liquid specimen splitting cannot be measured directly. If the number of observations is too low, or the design of control measurements and blocking is inadequate, the ability to deconvolve different error sources is also impaired. In such cases, one would need to model the total error empirically.

4.2 Digital filtration AST is a more informative generalization of bulk filtration AST

4.2.1 Overview of protocol for digital filtration AST

We have discussed three ways in which low numbers of inoculating bacteria limit any phenotypic AST assay where the measurement technology assesses the population size of bacteria, such as in bulk filtration AST, where the population size is measured by nucleic acid amplification. There is a technique, however, that addresses all three of the obstacles posed by low inocula. This technique is called digitization, and it has been used in several contexts such as in digital PCR or single-cell RNA-seq.

A digitized assay uses partitioning of a bulk liquid sample to make multiple individual measurements of (nearly) individual particles instead of one measurement for the population of particles. Digitizing assays takes advantage of the discrete nature of analyzed particles, turning discreteness from a confounding factor into

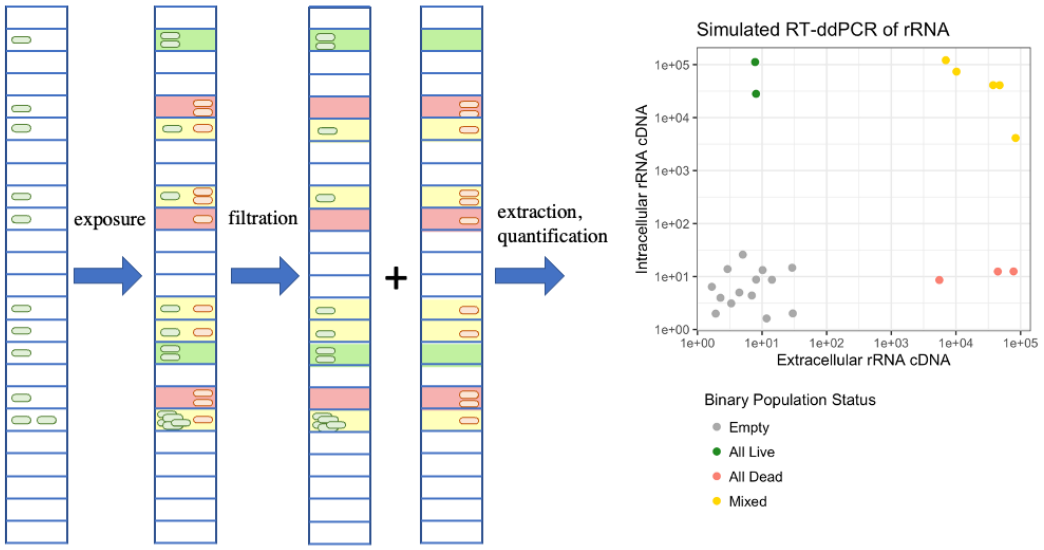


Figure 4.1: Schematic of the digital filtration AST protocol.

a source of additional information. Specifically, the discrete nature of bacterial cells prevents cells from being split into fractional cells during the splitting of the specimen containing them. Secondly, the minimum size of bacterial cells enables one to distinguish the presence or absence of at least one bacterial cell in a partition of the specimen.

In order for an assay to be digital, its partitions must be loaded "digitally" or in the "digital range." Generally speaking, a set of partitions is digitally loaded when the volume of the partitions is small enough relative to the density of bacteria that a substantial number of partitions do not receive any bacterial cells and the remainder receive a low number of cells that is Poisson distributed. A detailed definition and discussion of the digital range is discussed in the following section 4.2.2.

I developed a digitally-loaded variation of filtration AST, which I will call digital filtration AST. Digital filtration AST differs from bulk filtration AST in that the bacteria are digitally loaded when the bacteria-containing liquid specimen is split into multiple partitions.

In the digital filtration AST protocol performed herein, the bacteria-containing liquid specimen was partitioned into a 96-well plate, where the antibiotic exposure occurs. Each of the 96 partitions then undergoes a protocol similar to bulk filtration AST, using a 96-well filter plate. Each partition undergoes a physical separation process, namely filtration, and the resulting fractions from each partition are collected and

tracked. Each fraction's nucleic acids are extracted and then quantified by nucleic acid amplification. Lastly, the set of nucleic acid concentration measurements are used to estimate both the density (and total number) of bacteria and the susceptibility of the strain. To achieve both aims, algorithms different than those used in bulk filtration AST must be used to analyze the outcome of a digital filtration AST. The inference of cell density uses the tally of empty wells, while the inference of susceptibility includes information from the live-dead status of each population descended from mostly single cells.

A detailed protocol of digital filtration AST is expounded in section 4.2.3.

Furthermore, by detecting single cells, or at least small numbers of cells per partition, digitally loaded same-sample AST may make it easier to detect the detection of low frequency heterogenous resistance phenomena (i.e., heteroresistance and persister cells). Demonstrating this advantage remains as future work.

4.2.2 Defining the digital range

The basic statistical analysis of digital assays has been explained many times in the literature on digital PCR. However, I will describe it in detail here for the reader's reference, and because some aspects I describe, such as the continuum between digital and bulk assays, have not been explicitly addressed. Furthermore, most derivations of equation 4.24 in the literature are superficial, and even some more detailed treatments of digital PCR theory [139] do not fully explain how the Poisson approximation of the binomial applies to equation 4.24.

A sample of discrete particles is "digitally" loaded into a given set of physical partitions when there is a reliable chance that one or more of the partitions will not receive any of the discrete particles in the sample, due to random chance and the indivisibility of the particles. A sample is said to be in the digital range when it has been diluted sufficiently to be digitally loaded into a given device of interest. When a sample is loaded at a density higher than the digital range, the resulting loading is a bulk loading condition.

How does one determine the probability by which partitions receive cells? Consider a sample of N discrete particles randomly distributed in a volume V, then split completely into a given set of M physical partitions, each with a volume of V_m, so that $\sum_m V_m = V$. (Without loss of generality, one can consider one of the partitions, say $m = 1$, to represent the remaining sample volume when only some of the sample is split into the other partitions of interest, $m = 1$ being differentiated by a change

in intention only.) The probability of each particle ending up in a given partition p_m is the ratio of that partition's volume to the total volume of all the partitions, $p_m = \frac{V_m}{V}$. It is reasonable to assume that the destination of one particle does not affect the destination of the next particle loaded, so the number of particles in each volume N_m is multinomially distributed with size N and probabilities p_m. This is a generalization of equation 4.3.

When N is given as a fixed number, the N_m are dependent on each other, as can be easily imagined if $N = 1$. However, if the total number of particles N is assumed to be Poisson distributed, as opposed to being fixed, then it can be proven that each N_m becomes an independent Poisson-distributed random variable [140]. In many situations of interest to us, N is indeed not controlled or fixed, and instead the assumption that N is a Poisson random variable is justified for the following reasons. We consider that some of the sample will remain behind in the partition $m = 1$. Now, instead of N, we become interested in the total number of particles $N_{m\neq 1} = \sum_{m=2}^{N} N_m$ loaded into the other partitions of interest ($m = 2, \ldots, M$). $N_{m\neq 1}$ is binomially distributed as $N_{m\neq 1} \sim \text{Binomial}\left(n = N, p = \frac{\sum_{m=2}^{N} V_m}{V}\right)$. When N is large and $\sum_{m\neq 1} V_m << V$, then the binomial distribution is approximated by the Poisson distribution as $N_{m\neq 1} \sim \text{Poisson}\left(\lambda = N\frac{\sum_{m\neq 1} V_m}{V}\right)$. Therefore, each N_m where $m > 1$ is Poisson distributed, and the number of cells in each partition of interest m is related to the density of cells $C = N/V$ as follows:

$$N_m \sim \text{Poisson}(\lambda = V_m C) \Leftrightarrow \text{Prob}(N_m = k) = \frac{(V_m C)^k e^{-V_m C}}{k!}. \tag{4.23}$$

The quantity $V_m \times C$ represents the density of particles per partition, as opposed to density per unit volume. Finally, the probability of an empty well is

$$\text{Prob}(N_m = 0) = e^{-V_m C}. \tag{4.24}$$

Notice that in equation 4.23, there is no distinct mathematical difference between the digital and bulk ranges; instead, the two loading schemes are two ends of a continuous spectrum separated by a subjective threshold of what is a "reliable chance." One can view a digital loading as a set of bulk loadings of low density, complicated by the discrete nature of cells. One can also view a bulk loading as an attempted digital loading that failed due to limitations on the number and size of partitions available. Higher densities have a lower chance of a well being empty, and lower densities have a higher chance of empty wells.

A practical threshold for when bulk loading transitions to digital ranges is when there are less than 3.5 cells per partition on average. At this density, there is a $e^{-3.5} = 0.0302$ chance of a given well being empty.

Note that the density in cells per partition is also function of the partition volume V_m. Controlling either the density of the cells or the volume of the partitions brings the loading in or out of the digital range. Lowering the density or decreasing the volume of the partitions moves one towards digital loading and vice versa.

During digital loading, it is preferable to have a high density of cells when loading but a low density of cells per partition. Assuming that the volume of the specimen has an upper limit, a high density of cells at loading will always mean that more cells are available for analysis. Meanwhile, a lower density of cells per partition makes it more likely for partitions containing some cells to be containing only one cell, which makes interpretation of the partition's contents easier. To achieve a preferred digital loading, a practitioner usually cannot increase the density of cells in the original specimen. One can decrease the density easily by dilution, include more partitions, or reduce the size of the partitions. Devices intended to be digitally loaded therefore should maximize the number of partitions and minimize the size of the partitions, up to the limit imposed by manufacturing and hardware. Should the digital range still not be reached by such a device design, then dilution of the sample can be performed.

In contrast, during bulk loading, high density of cells and a high density of cells per partition at loading is desired. At high densities of cells per partition, the stochasticity of loading cells into partitions is overcome by the central limit theorem, which states that the variance in well loading decreases relative to the mean well loading as the mean increases. Thus, it is preferable to use a smaller number of partitions in bulk loading, compared to a maximum number of partitions in digital-loading.

Loadings that are only slightly above the digital range fall into a gray zone where the stochasticity of loading is poorly overcome by inference using either Poisson statistics or by the central limit theorem. Such loadings are not preferred, but nonetheless can be analyzed as a bulk loading if they occur. The practitioner should decide beforehand whether to use a bulk loading or a digital loading, then design one's device towards one end of the spectrum or another.

4.2.3 Detailed protocol of filtration AST

The most typical digital filtration AST protocol is outlined below. Different runs may have incorporated variations in the volumes used; these are noted in the attached data set.

4.2.3.1 Contrived clinical samples by bacterial culture

Serial dilutions of 2 mL liquid batch cultures of pure isolates were used as contrived clinical samples in the digital filtration AST experiments herein. The batch cultures grown to an OD_600 of 0.3 in the same way as bulk filtration AST experiments. Typically, three dilutions in MHB media were performed to achieve a target CFU density of 20–80 CFU/mL representing a clinical sample with a low density of cells. To initiate an antibiotic exposure, 10 µL of this final dilution yields 0.2–0.8 CFU per 25 µL antibiotic exposure partition.

Bacterial strains used in the digital filtration AST experiments are listed in Table 4.1.

Strain Name	ETP MIC	CRO MIC	Notes
E. coli K12 sub-strain MG1655	0.012, S (Etest)	≤1, S	Obtained commercially. MIC values reported in [120, 121]
E. coli #1 (UCLA Study 15-04A-01)	≤0.5, S	≤1, S	From urine on 12/29/2014.
E. coli #38 (UCLA Study 15-31A-001)	>4, R	>32, R	From bile in 2017.
K. pneumoniae #3 (UCLA Study 15-31A-155)	≤0.5, S	≤1, S	Isolated 5/27/2017.
K. pneumoniae CDC AR Bank 0005	>8, R	>32, R	Enterobacterales Carbapenem Breakpoint panel. Resistance loci: aac(6')-Ib, aph(3')-Ia, catA1, dfrA12, EMRD, KDEA, KPC-2, mph(A), Omp35, OmpK35, oqxA, oqxB, OXA-9, SHV-11, sul1, TEM-1A

Table 4.1: Clinical isolates used in digital filtration AST. E. coli=Escherichia coli, K. pneumoniae=Klebsiella pneumoniae, ETP=ertapenem, CRO=ceftriaxone, MIC=minimum inhibitory concentration. MICs obtained by broth microdilution unless noted. UCLA = University of California Los Angeles Clinical Microbiology Lab.

4.2.3.2 Antibiotic exposure

To initiate the antibiotic exposures, the final dilution of bacteria was transferred using a multichannel pipette from a divided sterile reservoir to a 96-well plate whose wells contained 15 µL of media and antibiotics. The 96 or fewer antibiotic exposures created above were grouped into 2 or more experimental conditions containing different antibiotic concentrations (or sometimes target inocula). In most cases, the exposure volume was mixed by repeated pipetting upon addition of cells, and the timing of the second pipette tip expulsion was recorded as the start of the exposure. The median start time across all exposures in one experimental condition was used during statistical analysis. Once all exposures had been started, the plate was sealed with an adhesive seal (BioRad DNA/RNAse-free, not gas-permeable), then incubated in a shaking thermoblock at 37°C, 700 rpm. The lid of the thermoblock was not heated. In most AST experiments, a sterile 96-well plate was loaded with bacteria seconds to minutes before, after, or both before and after the addition of bacteria to the antibiotic exposures. The same volumes of the contrived clinical sample was used during these transfers as during the antibiotic exposure creation, and care was made to not introduce trace amounts of antibiotics in residual tip volumes. The plate was later sealed and incubated overnight at 37°C until pellets of cells were visually observed in some wells. The number of turbid and clear wells were recorded. This "CFU count" plate serves as an orthogonal measurement of the true density of bacteria in the contrived clinical sample, which often differed from the target density by up to a factor of 2.

The antibiotic stock solutions were prepared in the same manner as for bulk filtration AST experiments.

4.2.3.3 Separation and nucleic acid extraction

Prior to the start of the antibiotic exposure, a 96-well filter plate (Millipore-Sigma MSGVS2210) was blocked with 30 µL of 50 µg/mL salmon-sperm DNA dissolved in nuclease-free water. The blocking solution was centrifuged through the filters, and then the filters were dried at 65°C for several hours or at room temperature overnight. In the last 2 experiments, no blocking was performed.

Shortly before the target end time, the exposures were removed from 37°C incubation. The sealing membrane, with condensation that cannot be removed by brief centrifugation, was carefully removed. The entire volume of each exposure was then transferred to the 96-well filter plate and sealed with a new sealing membrane.

The filter plate was sealed to a clean 96-well plate by tape, then centrifuged. The start of the 5 minute, 2200 rcf centrifugation (Eppendorf A-2-MTP rotor) of the filter plate was recorded as the end of the antibiotic exposure. After centrifugation, the collected filtrates were set aside and covered to slow evaporation while the next steps were performed. Any cells that were not lysed by the time the majority of the filtrate passed through the filter would be measured in a subsequent filtration fraction, and so they were considered intact at the recorded time point.

20–30 µL of sterile MHB media was placed into each filter plate well as wash fluid, and a new 96-well plate was attached to the filter plate. The filters centrifuged for 5 minutes, 2200 rcf, and the collected wash fluid discarded, or treated in the same manner as the filtrate for quantification as an additional fraction of the sample.

The lysates were collected using two different protocols. First, 20 µL of DEB was added to each washed filter membrane. Then, in some experiments, the filter plate was placed upon a nuclease-free foil (BioRad #1814040) to protect the filter plate nozzles from contamination, then placed on a flat-bed heating block. The filter plate was then heated to 65°C for 6–20 minutes without shaking and with a non-airtight plastic lid without heating. The filter plate then was attached to a clean 96-well plate and centrifuged. In other experiments, the filter plate with DEB was attached to the clean 96-well plate first, then sealed with a sealing membrane. The filter plate and collecting plate were then heating to 65°C for at least 40 minutes at 1200 rpm with a heated lid (Eppendorf C). The lysates were then harvested by centrifugation. In both protocols, the lysates were collected by centrifugation for 10 minutes total at 2200 rcf.

A known, recorded volume, typically 10 µL, of the filtrates were transferred as early as possible (typically >30 minutes) to a new 96-well plate containing a known amount of DEB. The volume of filtrate transferred was adjusted based on the amount of collected filtrate, since <10 µL may have been available. The filtrate in DEB were vortexed, then heated in a thermocycler to 65°C for 6 min and then 98°C for 4 minutes. The lysates were heated at 98°C for 4 minutes, the 65°C step having been performed by heating the filter membranes. Both filtrate and lysate extractions were then frozen at -80°C.

4.2.3.4 Nucleic acid quantification

In the digital filtration AST experiments reported herein, nucleic acids were quantified by reverse transcription and real time qPCR. Quantification with ddPCR yielded

similar results for the same templates.

In the first 10 experiments, the reverse transcription reaction comprised 1.708 µL of nuclease-free deionized water, 0.40 µL of 10X RapiDxFire buffer (LGC/Lucigen), 0.20 µL of 10 mM deoxyribonucleotides (New England Biolabs), 0.16 µL of primer (reverse strand only, final concentration 0.40 nM), 0. 032 µL of 3 U/µL RapiDxFire reverse transcriptase (LGC/Lucigen), and 1.50 µL of the DEB extraction template. In the other 4 experiments, the reverse transcription reaction comprised 3.00 µL of nuclease-free deionized water, 0.50 µL of 10X RapiDxFire buffer (LGC/Lucigen), 0.25 µL of 10 mM deoxyribonucleotides (New England Biolabs), 0.20 µL of primer (reverse strand only, final concentration 0.40 nM), 0.05 µL of 3 U/µL RapiDxFire reverse transcriptase (LGC/Lucigen), and 1.00 µL of the DEB extraction template. A master mix of all components except the templates was created and distributed to separate tubes. Individual templates were then added. The reactions were vortexed, then incubated in the following thermocycler protocol: extension at 60°C for 5 minutes, heat inactivation at 95°C for 5 minutes, and a hold indefinitely at 4°C. In the last 5 experiments, a 15 second interval at 75°C preceded the 60°C extension step to encourage rRNA denaturation.

4.2.3.5 Susceptibility classification after pre-processing to well population status calling

Multiple algorithms for calculating a susceptibility metric or call were employed in parallel in this work to gauge their accuracy. The definitions of these algorithms are discussed in section 4.5.1. A taxonomy of these and other algorithms, and the pharmacodynamic models each one implies, are further discussed in sections 4.5.1 and 4.4, respectively.

One major dichotomy of the possible algorithms is discussed in this section: the use of counts of an intermediate, categorical "well population status" call instead of the real-valued nucleic acid concentration measurements. So far, all of the algorithms used in this thesis did not utilize the results of the RT-qPCR directly, but rather the tallies of well population statuses. The calling of well population status in this thesis therefore serves as a data-processing stage upstream of the algorithms discussed in 4.5.1. This section details the procedures by which the well population statuses were called (i.e., estimated) from the data. Section 4.4 discusses the theoretical models that were assumed in order to motivate and explain the results of the well population status calling algorithm discussed here. Section 4.5.1 explains how the tallies of

well population status were interpreted by subsequent algorithms.

The intermediate step of calling well population statuses was done because the spread in the qPCR and ddPCR values from non-empty wells differs by up to 30X fold and is not normally distributed. It is difficult (though perhaps not impossible) to distinguish the integer number of cells present in the population from such noisy data from a manually-performed assay. It is easier, however, to distinguish whether any nucleic acids (and thus cells) were present in each digital partition's filtration fractions, or if no nucleic acids and cells were present during the exposure.

Specifically, data pre-processing into well population statuses reduced the information in the raw outputs of the many nucleic acid quantifications into four integers, in exchange for noise reduction and the ability to easily compare different digital filtration AST experiment batches. The raw output of the digital filtration AST, as performed herein, are C matrices with dimensions $N_C \times F_C$ containing nucleic acid concentrations (absolute copies/µL or threshold cycles), and a pair of integers from the inoculum control incubation. Here, C is the number of experimental conditions, N_C is the number of partitions in condition C, and F_C is the number of fractions observed in condition C. Usually $F_C = 2$ but occasionally wash fractions were observed. The input variables of duration, antibiotic dose, and starting inoculum (measured, respectively, by a stop-watch; a mass balance and micropipettors; and a spectrophotometer) can also be considered outputs if a measurement error model is used. The melt curves from qPCR measurements were not used in this analysis but could be used in future analyses. Together, the melt curves form an $N_C \times F_C \times T$ matrix, where T is the number of temperatures observed.

Therefore, in this thesis, the digital partition's bacterial populations were classified as being empty, containing only intact cells, containing only antibiotic-lysed cells, and containing both intact and antibiotic-lysed cells. If a well had no nucleic acids in either its filtrate or lysate fractions, then its well population status was interpreted to be empty, the well having received no cells during digital loading. If a well had nucleic acids only in the filtrate fraction, then its well population status was interpreted as having only antibiotic-lysed cells. If a well had nucleic acids only in the lysate fraction, then its well population status was interpreted as having only intact cells. Lastly, if a well had nucleic acids in both fractions, then its well population status was interpreted as having only antibiotic-lysed cells

The number of wells of each well population status was tallied. For the rest of this thesis, the integer number of wells with with no nucleic acids (empty) is denoted as

O. The number of wells with no deaths (all intracellular and intact) is I, the number of wells with both intra- and extracellular nucleic acids present is B, and the number of wells with extinct populations (all extracellular) is E.

The well population statuses were called using either manual-drawn thresholds near or at 35 qPCR cycles, or, if indicated, by K-means clustering.

The use of manual thresholds differing from 35 cycles was justified because of batch effects in the chemical environment of the samples that led to different backgrounds. Evaporation of water from samples and DEB extractions could be identified as a source of such batch effects in at least two experiments because of the correlation of Cq with the proximity of the well to the edge of the 96-well plate.

For the few experiments where a wash step was not included, there was a correlation between the measured extracellular nucleic acid concentrations and the measured intracellular nucleic acid concentrations. The intracellular nucleic acid concentrations were interpreted as representing extracellular nucleic acids carried over from the retained volume of exposure that could not be removed from the filter membrane. In these experiments, wells that would have been called by manual thresholds as having nucleic acids present in both fractions were instead called as only having extracellular nucleic acids if the intracellular nucleic acid concentration was less than half the extracellular nucleic acid concentration.

If K-means clustering was used, then the threshold depicted is the midpoint of the two closest points along the axis in question. The choice of algorithm was decided after visualizing the data, and the chosen thresholds are visible in Figures 4.4a through 4.4d.

An automated composite algorithm could be written to attain the same end performance, but this was not pursued for simplicity.

4.3 Validation and optimization of digital filtration AST

To enable digital filtration AST, one needs to be able to robustly detect single cells by nucleic acid amplification.

As discussed in Chapter 3, the 23S rRNA loci in genomic DNA are present in approximately 14 copies per cell. Obtaining nucleic acid amplification from single cells. Amplification and detection of single copies of DNA have been reported. However, these reactions require very low amplification backgrounds. In my protocol, I did not test if I could amplify low copies of DNA, since it appeared that in bulk

filtration AST, the signal-to-noise ratio was too low at low inocula. Instead, I chose to amplify a nucleic acid species with higher copy number per cell. Of all the nucleic acids in a typical bacteria, the transfer RNAs have the highest numerical copy number. However, amplification of tRNAs is difficult due to their modified bases and strong secondary structure. The second most abundant nucleic acid species are the ribosomal RNAs, which are the most abundant nucleic acid by mass as well. It is estimated, from biochemical fractionation and from cryo-electron microscopy, that an Escherichia coli cell with a fast doubling time of 24 minutes contains 72,000 ribosomes on average (and 26,000 ribosomes for a doubling time of 40 minutes). Thus, the ability to amplify the ribosomal RNA itself would offer a pre-existing amplification in signal of about 10,000-fold.

4.3.1 The efficiency of reverse transcription of rRNAs is a function of amplicon size

Several commercial RT kits based on retroviral reverse transcriptases were unable to create cDNA detectable by ddPCR when the digitally-loaded cells were at an effective concentration of 40 CFU/mL. However, the proprietary RapiDxFire reverse transcriptase was able to reverse transcribe rRNA at an estimated efficiency of 25000/CFU. I speculate that reverse transcriptases descended from the Avian Myeloblastosis Virus or Moloney Murine Leukemia Virus reverse transcriptases are not efficient for reverse transcribing rRNAs, as these are the shared characteristics of the enzymes tested that did not show detectable activity. RapiDxFire is claimed to be isolated from a bacteriophage [141].

To investigate the limitations of rRNA-targeted reverse transcription, we assessed the length distribution of cDNAs using RT-ddPCR using three different antisense (complementary to the rRNA) primers, 2 targeting the 23S rRNA and 1 targeting the 16S rRNA. For each antisense primer, a set of 5 sense primers with similar melting temperatures was chosen so as to evenly span a distance of about 400 base pairs. The same cDNA sample, generated by the antisense primer, was then used as a template in a ddPCR reaction. The difference in apparent nucleic acid concentration would therefore be attributable to the ability of the sense primer to hybridize with the cDNA. Hybridization of the sense primer in turn is a function of whether the cDNA was extended to the position of the sense primer, as well as any intrinsic properties of the sense primer that occur despite attempts to keep primer lengths, melting temperatures, and the absence of predicted secondary structures consistent.

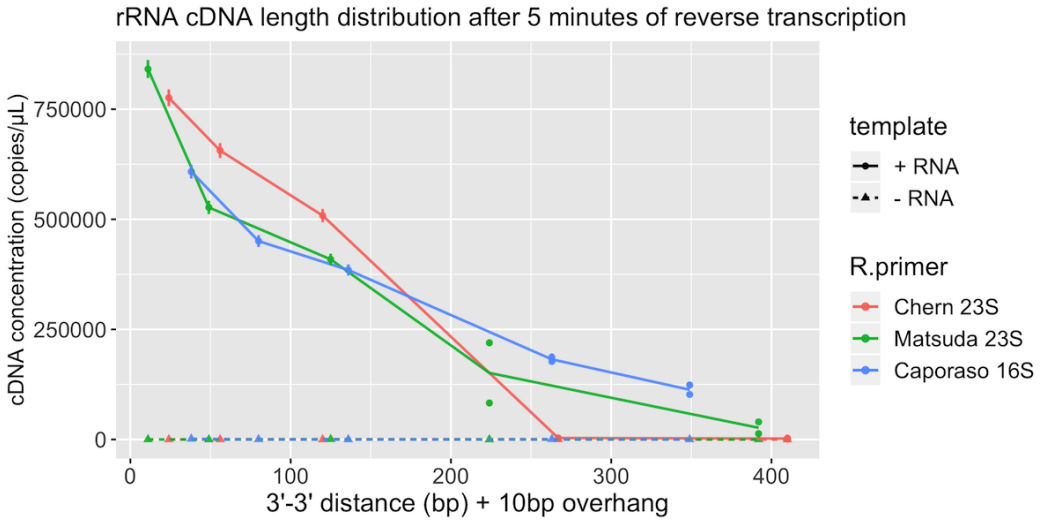


Figure 4.2: Ribosomal RNA cDNA creation is a function of amplicon length, not sequence position.

All three antisense primers generated exponential distributions with similar mean constants. This is consistent with a model where the secondary structure found throughout the rRNA molecules imposes a constant probability per base pair of enzyme dissociation, rather than three unique stepwise-decreasing distributions expected if enzyme processivity were inhibited by distinct locations along the rRNA molecule.

Further studies should include additional sense primers to better average out any intrinsic properties of the sense primers, additional sets of antisense and sense primers, and additonal replicates.

The effect of temperature during the extension phase and denaturants could also be investigated, with the prediction that decreasing the strength of rRNA secondary structure will increase RT efficiency, until balanced by a decrease in enzyme activity or primer hybridization kinetics.

4.3.2 Single cells can be detected by rRNA reverse transcription

In Figure 4.3, the amplification of digitally-loaded Escherichia coli K12 was performed without filtration. Instead, MHB media containing bacteria was directly placed into to QuickExtract buffer at a 1:1 ratio, vortexed, and heated. Within 3 minutes, the same volumes of the same dilution of bacteria were transferred to a sterile 96-well plate, sealed, and incubated overnight. The nucleic acid extractions underwent reverse transcription with RapiDxFire enzyme and quantification

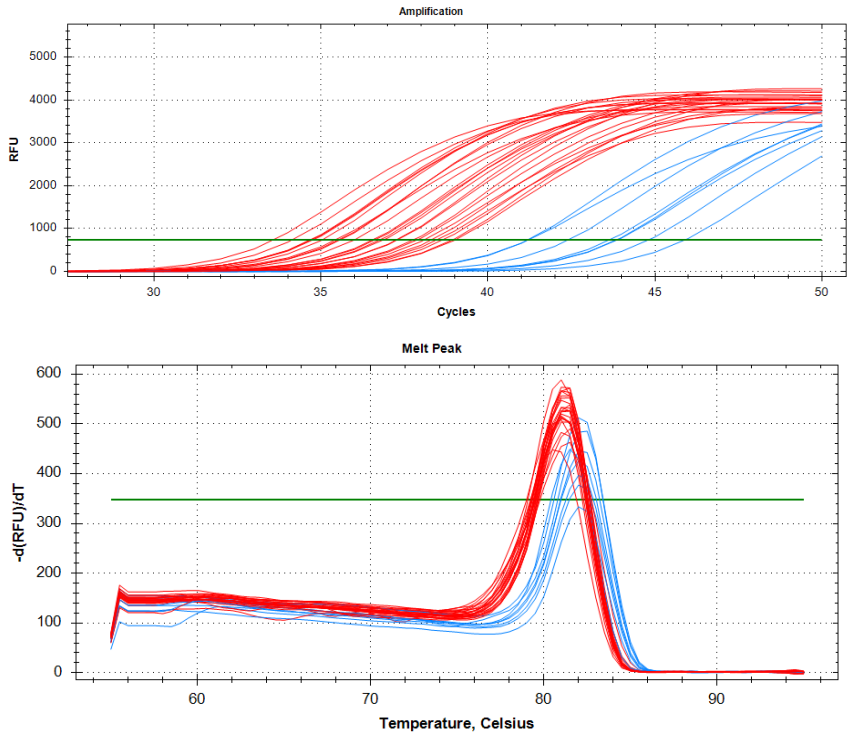


Figure 4.3: Empty and loaded partitions can be distinguished by threshold cycle and by melt curves in qPCR.

by qPCR.

A clear separation by qPCR threshold cycles (or by melt curve peaks) enabled manual calling of 8 empty wells and 24 loaded wells out of 32 total wells. (The high Cq values in Figure 4.3 were rationalized by the accidental evaporation of DNA extraction buffer, which inhibited qPCR polymerase activity.) Visual inspection of the overnight incubation showed 20 clear wells and 76 turbid wells with cell pellets out of 96 total wells. The hypothesis that these two loadings arose from the same density of cells can be tested using Barnard's unconditional test, which yields a p-value of 0.727823 (R package 'Barnard' vers 1.8). Thus, it is highly consistent that the qPCR experiment detected the same cell density as the overnight incubation, neither failing to amplify some cells or falsely amplifying empty wells.

4.3.3 Digital filtration AST results are consistent with gold-standard MIC and with bulk filtration AST results

2464 qPCR or ddPCR measurements of 1247 digitally-loaded wells or AST exposures, grouped into 52 conditions and 14 batches (i.e., sharing the same contrived clinical specimen), were performed on 5 strains of Escherichia coli or Klebsiella

pneumoniae and 2 antibiotics. 11 of the AST runs were accompanied by a CFU count control. 15 batches containing 2848 qPCR or ddPCR measurements of 1247 digitally-loaded antibiotic exposures, grouped into 45 conditions, were performed on 6 strains and 2 antibiotics. For each measurement, the choice of antibiotic compound, the antibiotic dose, the duration of the antibiotic exposure, and the starting number of bacteria (the inoculum) were controlled by the experimenter.

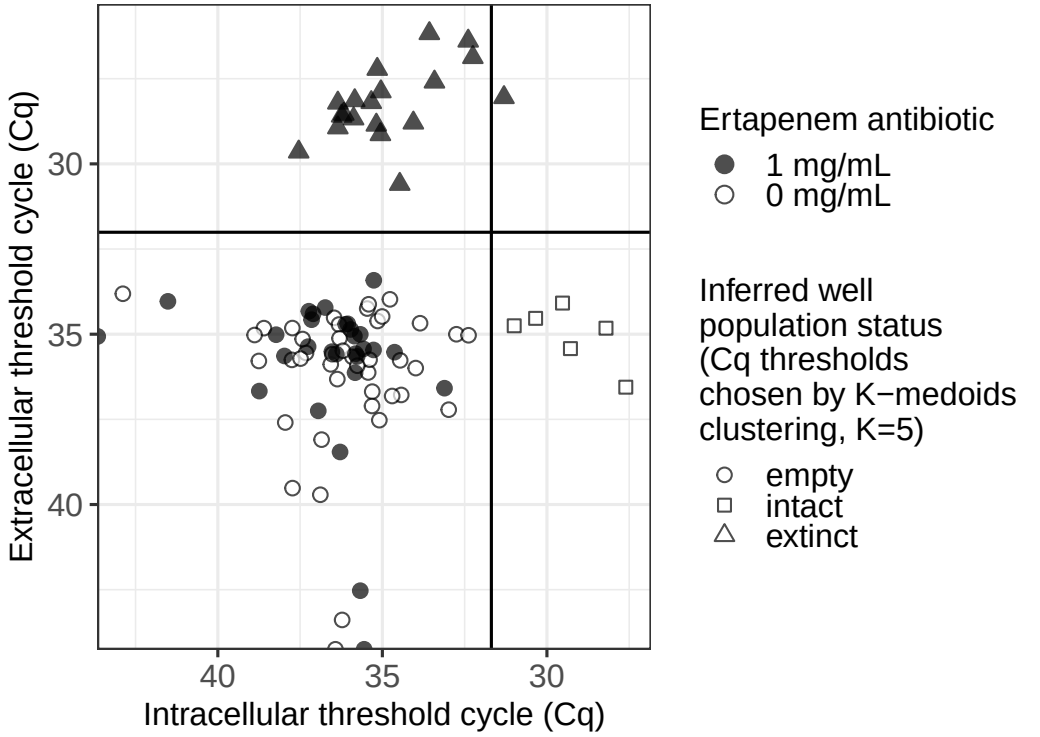
First, the validity of digital filtration AST was examined using two antibiotics and a pair of strains, one susceptible and one resistant to the two antibiotics. Breakpoint concentrations were used, and the antibiotic exposure lasted 60–70 minutes. Of the eight conditions tested, only two involve a treated, susceptible strain. The other six conditions are not expected to show any antibiotic killing.

The qPCR or ddPCR results in all experiments were easily separable into the four well statuses. As expected, antibiotic-lysed cells were only seen in the conditions where E. coli K12 was treated with antibiotics. At 70 minutes, in the susceptible and treated conditions, all detected cells were killed by ertapenem, but only one well showed killed cells when ceftriaxone was used. The most likely explanation is that ceftriaxone's activity is delayed until approximately 60 minutes after the start of the exposure.

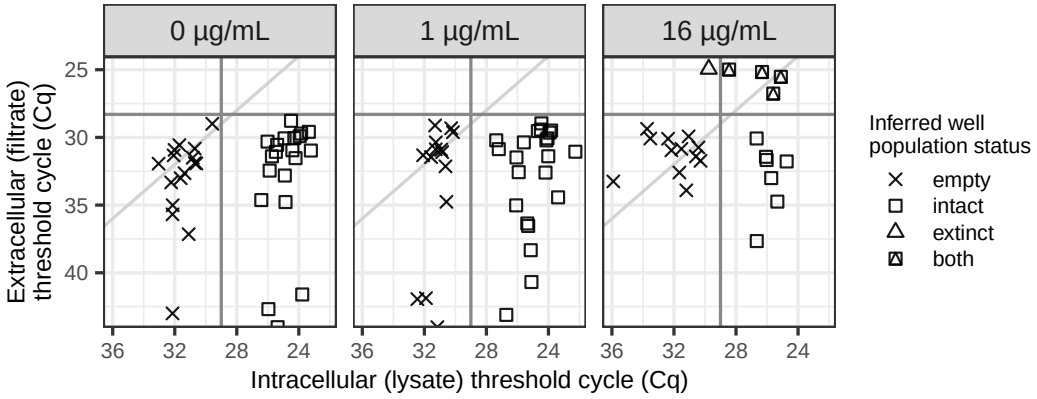
The effect of the starting inoculum on assay loading was also assessed in Figure 4.5. A serial dilution of one batch culture was prepared prior to the digital filtration AST to create 4 target inocula: 0, 0.5, 1, and 2 CFU/well. The antibiotic exposure totaled 80 minutes. As expected, the number of wells with detected cells increased as a function of the starting inoculum, ruling out the possibility that the observed nucleic acid concentrations were experimental artifacts. All ertapenem-treated non-empty wells contained only killed cells, while non-treated non-empty wells contained only intact cells, consistent with the true susceptibility.

The effect of exposure duration also was apparent in the digital filtration AST results. Figure 4.6 shows one experiment where multiple durations of antibiotic exposure (10, 35, 60, and 100 min) were performed using the same batch culture.

As expected, the untreated conditions contained only intact cells. The treated (and susceptible) conditions contained only killed cells, except for 1/5 wells in the 35 minute exposure where only intact cells were seen. These findings are consistent with the pharmacodynamics seen in the bulk, as the fraction of killed cells should rise from 0 to 1 over time. The section 4.4 more precisely discusses how consistent

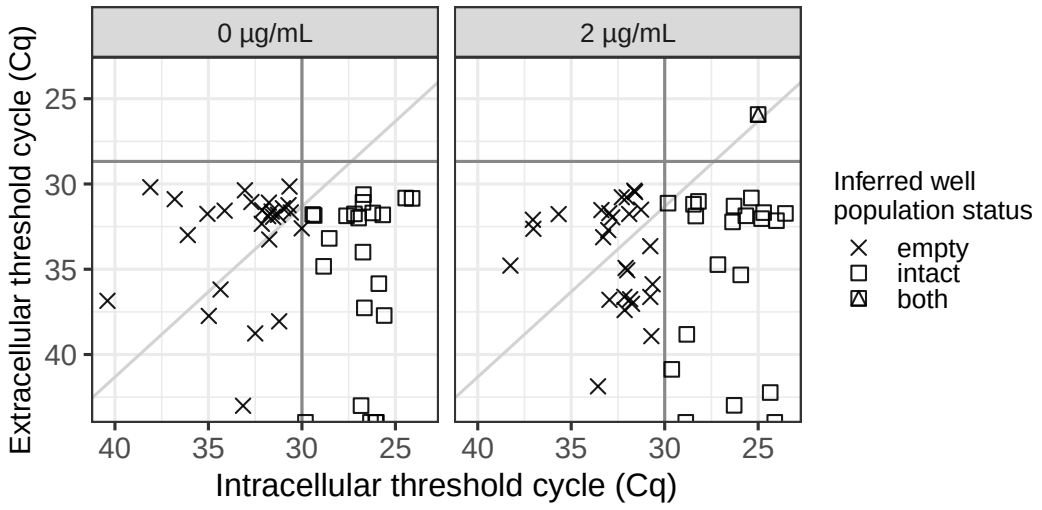


(a) E. coli K12 (MIC: 0.012, S), 70 minutes, ertapenem

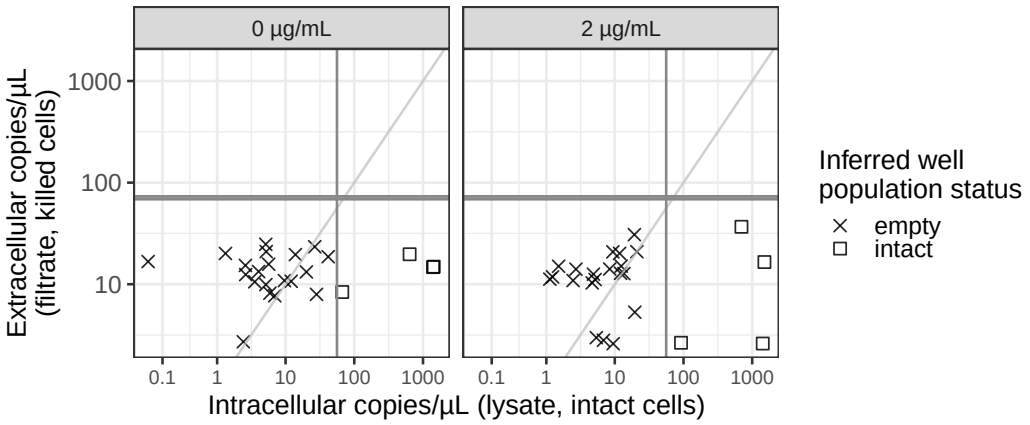


(b) E. coli #38 (MIC: >4, R), 70 minutes, ertapenem

Figure 4.4: Digital filtration AST at 60–70 minutes.



(c) E. coli K12 (MIC: <1, S), 60 minutes, ceftriaxone



(d) E. coli #38 (MIC: >32, R), 70 minutes, ceftriaxone. Quantified by ddPCR.

Figure 4.4: Digital filtration AST at 60–70 minutes (continued).

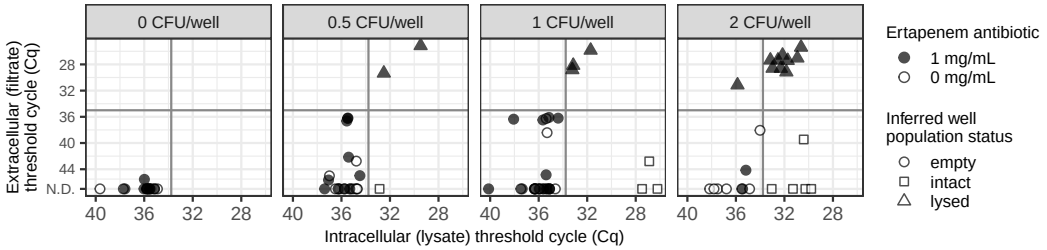


Figure 4.5: Effect of inoculum on digital filtration AST. E. coli K12 (MIC: 0.012, S), 80 minutes, ertapenem. The untreated (white) and treated (black) loaded wells separate by well population status.

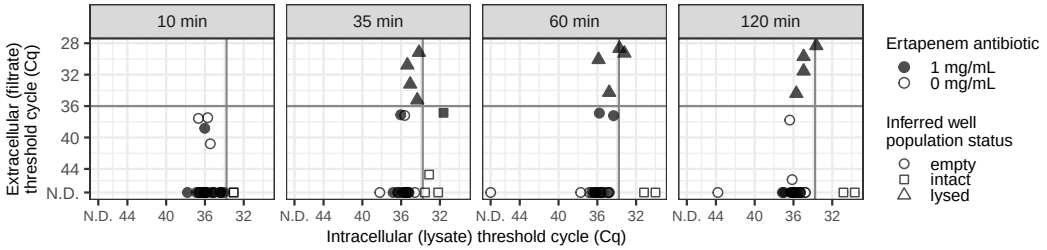


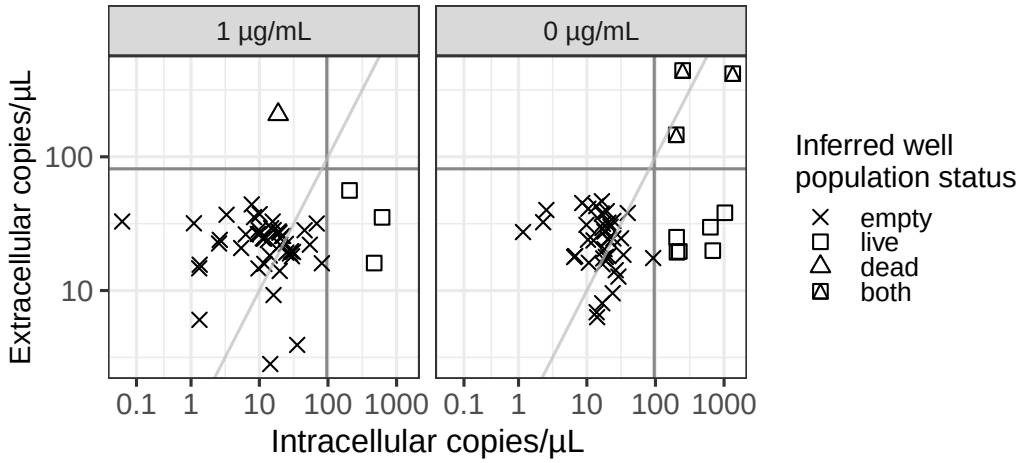
Figure 4.6: Effect of exposure duration on digital filtration AST. E. coli K12 (MIC: 0.012, S), 80 minutes, ertapenem. One of five treated wells at 35 minutes remained in a state with no deaths.

the results are using a quantitative modeling.

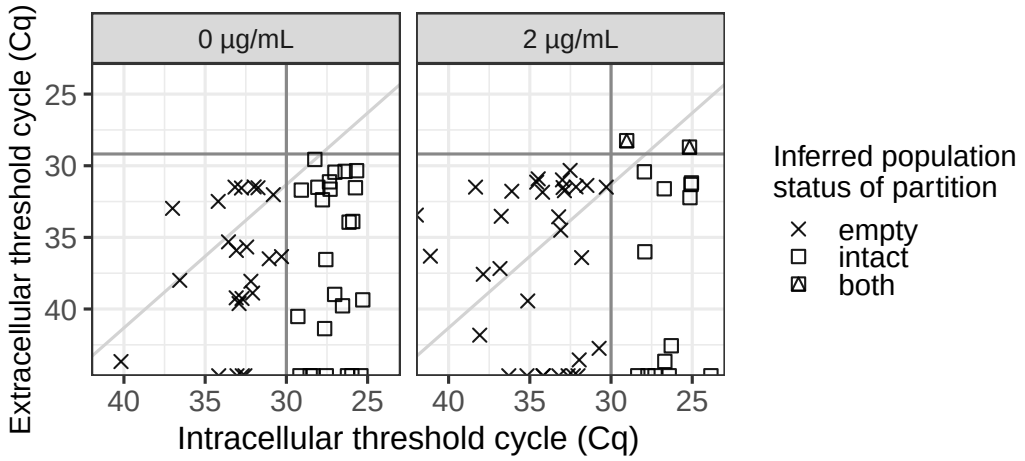
The digital filtration AST results at earlier times sometimes had equivocal results. Two experiments with 70 minute exposures were repeated with 40 minute exposures. In one experiment (Figure 4.7a), extracellular rRNA was detected in the ertapenem treated and untreated conditions of a resistant strain. The lysis in the untreated condition cannot be due to antibiotics, and the amount of lysis in the treated condition was not significantly more than the background lysis that must have occurred in this batch (or strain).

In another experiment (Figure 4.7b), E. coli K12 exposed to ceftriaxone showed 2/15 wells with some killed cells, more than the 0/25 wells in the untreated condition. However, at 60 minutes exposure in Figure 4.4c, 1/21 wells contained some killed cells, which is less than at 40 minutes. The difference though is easily attributable to random stochasticity, as the difference is not statistically significant (p=0.421, two-sided Barnard's test).

The effect of antibiotic dose was also examined in Figures 4.8 and 4.9. A susceptible



(a) E. coli #38 (MIC: >32, R), 40 minutes, ertapenem. Quantified by ddPCR. Extracellular rRNA was seen in both treated and untreated conditions in this batch.



(b) E. coli K12 (MIC: <1, S), 30 minutes, ceftriaxone

Figure 4.7: Digital filtration AST at 30–40 minutes.

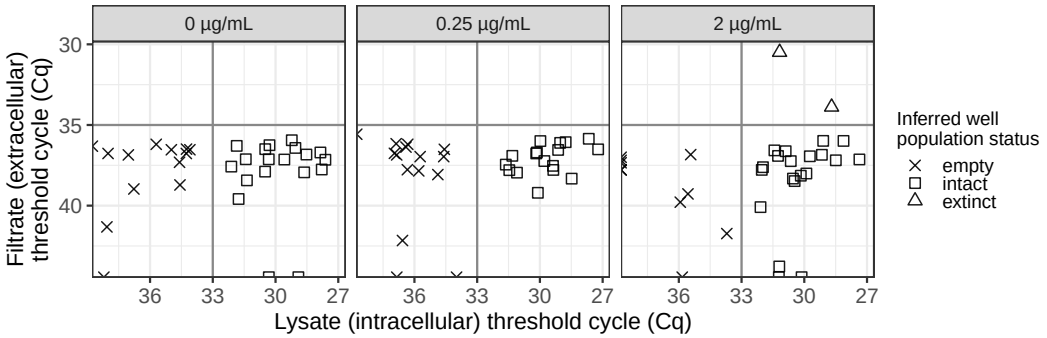


Figure 4.8: Digital filtration AST with different ertapenem doses at 20 minutes, E. coli K12.

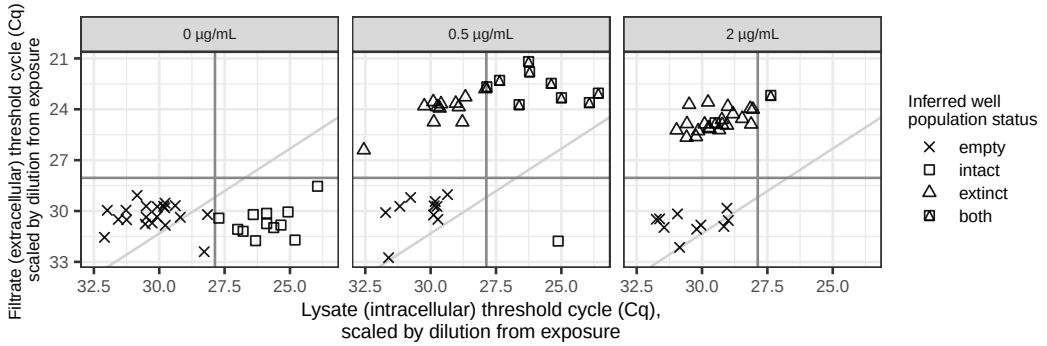


Figure 4.9: Digital filtration AST with different ceftriaxone doses at 40 minutes, E. coli K12.

strain was exposed to two doses of ertapenem (in addition to a third dose of 0 µg/mL) for 20 and 60 minutes. As expected, the higher dose caused a faster rate of antibiotic killing overall. At 20 minutes, the higher dose caused extracellular rRNA to appear in 2 wells, while the lower dose did not show any extracellular rRNA. At 60 minutes, most of the wells in the higher dose had only killed cells, while most wells in the lower dose still had both intact and killed cells.

Additional strains other than K12 and UCLA #38 were tested by digital filtration AST in Figures 4.10 through 4.11. The results were all consistent with the known susceptibilities of the strains.

The qPCR, ddPCR, and well population status tallies of all digital filtration ASTs used in this thesis can be found in the accompanying data files.

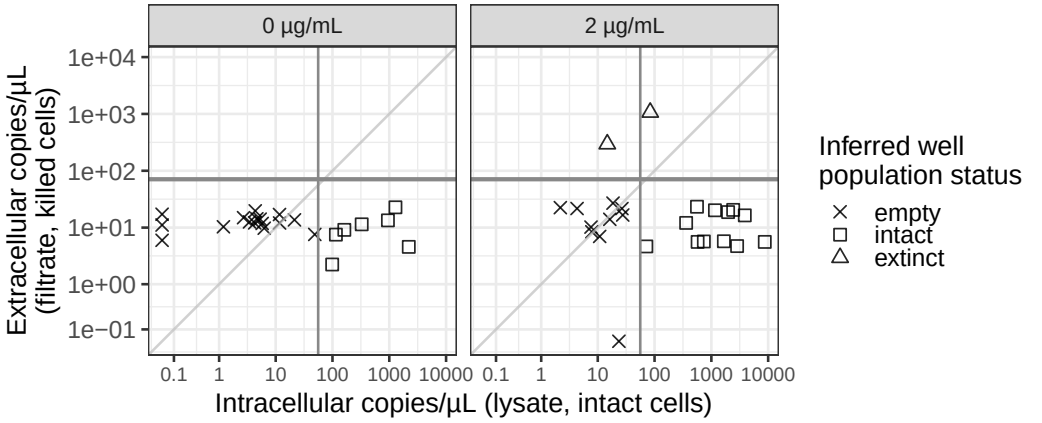


Figure 4.10: Digital filtration AST for *Escherichia coli* UCLA #1 (MIC: <1, S), ceftriaxone, 70 minutes.

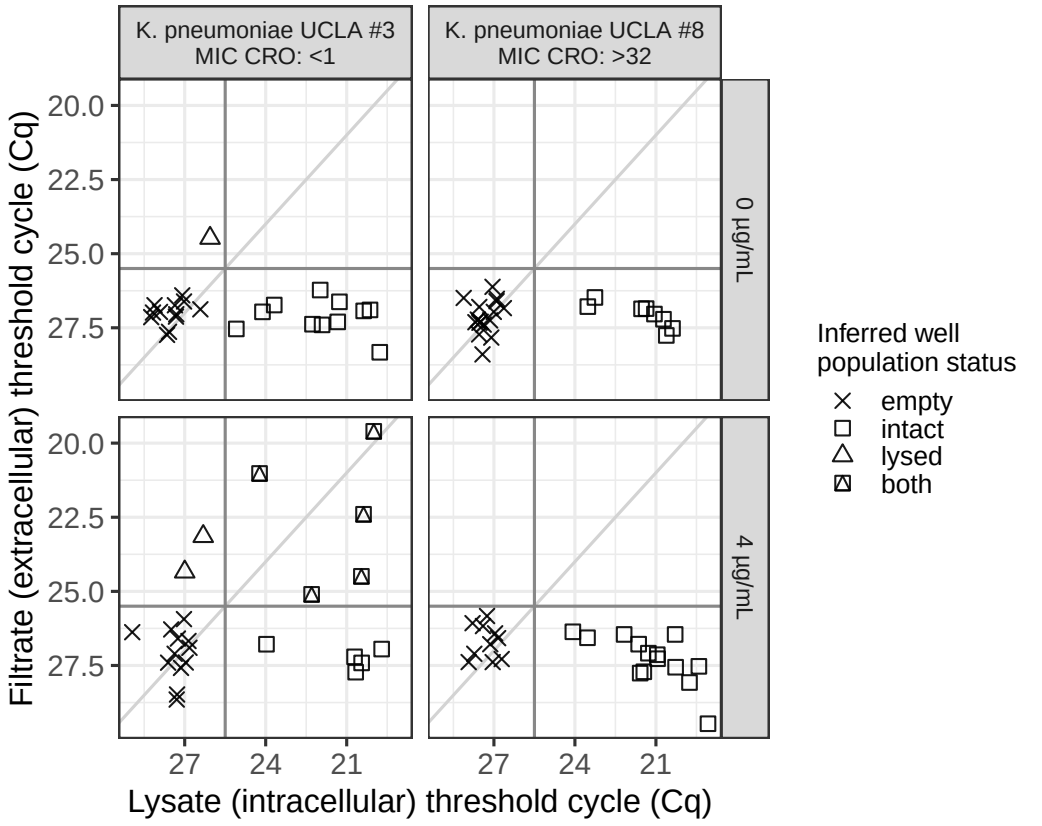


Figure 4.11: Digital filtration AST for *Klebsiella pneumoniae* clinical isolates, 110 minutes.

4.4 Pharmacodynamic models for discrete populations in digital filtration AST

In comparing the results of disparate digital filtration AST to the gold standard MIC values and bulk filtration AST results, it becomes desirable to more rigorously and uniformly define our predictions, so as to know how far the results are from our predictions. Modeling provides this quantitative precision. The motivation of this section is to better understand the mechanism of digital filtration AST by defining simple but usable discrete population and pharmacodynamic models for the underlying system, the antibiotic exposures.

In section 4.4.4, the adequacy of these models and our understanding of digital filtration AST are assessed by comparing fitted parameters with fits from the bulk filtration AST data and models. The ability to fit pharmacodynamic parameters from digital filtration AST results is also demonstrated. In contrast, section 4.5 explores how the same equations and pharmacodynamic models can be employed as classification algorithms for calling susceptibility, as if in clinical practice.

In this section, I define two relatively simple population models for discrete populations that differ by the omission or inclusion of new cell births. The population models are linked to pharmacodynamic models. I then derive equations that enable one to infer the parameters of the pharmacodynamic model from the well population statuses described in 4.2.3.5.

There are many other population models in the literature that could also be applied to the digital filtration AST protocol, but these models were not further explored in this work. They include age-structured stochastic processes such as branching processes or kinetic theory. Such classes of models are mathematically more complex than the Markov birth-death process. These models may have closed form expressions for their moments, which can be fitted to data, but also may not. It is also doubtful whether the parameters of these more complex models can be fit from the censored well population statuses instead of the richer nucleic acid concentration measurements.

4.4.1 Pharmacodynamics without cell division: the Markov death process

It is possible to assume a "no-births model" for the population and pharmacodynamics of one's antibiotic exposure system. In the no-births model, there is no further creation of bacterial cells after the start of the antibiotic exposure. In other words, no births and no cell division are assumed to occur, but bacterial cells may die and

transition to a dead state. This model is sometimes known as the simple Markov death process, when contrasted with the Markov birth-death process in the next section (4.4.2). The no-births assumption is at odds with the results of bulk filtration AST. However, keeping the total number of cells constant over time simplifies the mathematics needed to describe and fit one's models. The assumption of no births is reasonable for short exposures, especially those at the same scale as the expected doubling time, which for *E. coli* is between 20 and 40 minutes in rich broth. It is common in the AST literature for this population model to be implicitly assumed [88, 125].

When there are no new births, then the population of cells at the start of the exposure form a cohort, a group of individuals experiencing a common event (the antibiotic exposure). We ignore the age of cells by assuming they are of identical age that begins at the start of the exposure. One can define the survival probability for this cohort as the probability that a given cell has died by time t, as in section 3.3.1.

In the no-births model, I use the same survival probability function as in section 3.3.2, but with the background death rate and antibiotic-induced death rates combined. The events of death from background death rate (analogous to equation 3.17) and death from antibiotics (see equation 3.26) are independent, and to survive a cell must have neither event occur. Therefore,

$$\begin{aligned} S[t] &= e^{-kt} Q[\alpha, \beta t] \\ \beta &= \frac{\beta_{\max} A^{\gamma}}{\mathrm{EC}_{50}^{\gamma} + A^{\gamma}}, \end{aligned} \tag{4.25}$$

where parameters $\beta_{\max}$, EC_{50}, and γ are defined as in section 3.3.2, and A is the antibiotic dose. Equation 3.28 in section 3.3.2 contains the same overall survival function.

It is reasonable to assume that the bacterial cells in a nutrient-rich antibiotic exposure react independently to antibiotics, so that the probability of one cell dying does not influence the probability of another cell dying. Otherwise, one would be positing the existence of cell-to-cell communication or cooperative survival. Such mechanisms do exist in the form of biofilms and secreted beta-lactamases. But these effects were not obvious during bulk filtration AST, and they were not further explored for simplicity. Under the assumption of independent survival probabilities, and without information about the number of live cells at a prior time point, the number of live cells L at each time point becomes binomially distributed with a size parameter that is the total number of cells L_0 and a probability parameter that is the survival

probability at time t. The bacterial population in the "no-births" model therefore obeys the following equation:

$$L, D|L_0 \sim \text{Multinomial}\left(\text{size} = L_0, \text{prob} = (S[t], 1 - S[t])\right), \tag{4.26}$$

where $D = L_0 - L$ is the number of dead cells, $L, D|L_0$ is the joint distribution of L and D given L_0, and $S[t]$ is also parameterized by A, $\beta_{\max}$, EC_{50}, and γ. This equation is a discrete, stochastic model for population dynamics and pharmacodynamics.

In defining the survival probability, any contribution of cell age to the survival probability is subsumed by the form of our survival function. In more complex pharmacodynamic models, the distribution of cell lifetimes could be de-convolved into the distribution of cell ages at the start of the antibiotic exposure and the distribution of time to death as a function of antibiotic dosage and of cell age. The theory of fragmentation processes or kinetic theory offer suitable frameworks, but they were not further explored in detail. Some general arguments are made below.

If we were to model the hazard rate as a function of cell age, one reasonable assumption would be to allow the probability of cell death to increase toward the end of the cell lifespan when synthesis of the peptidoglycan septum separating the daughter cells occurs. This is because experiments have shown that distinct protein machinery, the MreB and FtsZ complexes, are responsible for peptidoglycan synthesis during elongation and during cytokinesis in most bacteria, and the rate of peptidoglycan synthesis is highest during cytokinesis [26]. Indeed, certain beta-lactams at certain doses have been observed by live microscopy to cause cell lysis specifically at the midpoint of the cell where cytokinesis is occurring [28, 124]. Another reasonable assumption to make is that the distribution of cell ages at the start of the exposure reflects the population at steady state, a consequence of assuming that the cells are in exponential growth and completely unsynchronized. An ideal age distribution, $\text{Prob}(a) = 2\mu e^{-\mu a}, 0 \leq a \leq \frac{\ln 2}{\mu}$, convoluted with a normally-distributed variation in age at division, was proposed by Koch in Chapter 103 of Neidhardt et al. 2001 [137]. Together, these two assumptions imply that the extent to which the probability of death concentrates around the time of septum formation would correlate with the value of the α parameter, analogous to the reasoning of section 3.3.2.

4.4.2 The Markov birth-death model is a mathematically tractable discrete model for digital filtration AST

The Markov birth-death (MBD) model [142–144], also known as the Kendall process, tracks the numbers of individuals (e.g., bacteria) in the population as the system state variable. The population number increases by one at a rate proportional to the number of individuals alive and to a birth constant μ. Similarly, the population decreases by one at a rate proportional to the number of individuals alive and to a death constant $\mathcal{B}$[1]. The set of differential equations, together often called the master equation, describing the MBD is below:

$$\frac{\partial}{\partial t}P_n[t] = \mu(n-1)P_{n-1}[t] - (\mu+\mathcal{B})nP_n[t] + \mathcal{B}(n+1)P_{n+1}[t], n \geq 1 \quad (4.27)$$

$$\frac{\partial}{\partial t}P_0[t] = \mathcal{B}P_1[t]. \quad (4.28)$$

Here, n is the number of individuals in the population, with $n \in \mathbb{Z}, 0 \leq n \leq \infty$[2]. $P_n[t]$ is the traditional notation for the probability of having n individuals in the population at time t; it is a function of both n and t.

The boundary condition of interest is when there are n_0 cells at time t:

$$\begin{aligned} P_{n_0}[0] &= 1 \\ P_n[0] &= 0 \text{ if } n \neq n_0. \end{aligned} \quad (4.29)$$

The master equation can be rewritten in terms of a generating function $G[z,t]$ by taking the Z-transform of $P_n[t]$, defined as $G[z,t] = \sum_{n=0}^{\infty} P_n[t]z^n$. Then, the master equation implies that

$$\frac{\partial G[z,t]}{\partial t} = \left(\mu z^2 - (\mu+\mathcal{B})z + \mathcal{B}\right)\frac{\partial G[z,t]}{\partial z}. \quad (4.30)$$

The initial condition for the initial number of cells is

$$G[z,0] = z^{n_0}. \quad (4.31)$$

When $\mathcal{B}$ is a constant function of time, then the closed form for $P_n[t]$ given n_0 can

[1] Note that my choice of μ and $\mathcal{B}$ as symbols are opposite the convention used in the stochastic process literature. Usually, the birth rate is denoted r or λ, and the death rate is denoted μ.

[2] In the previous section, we denoted n_0 as L_0, and we denoted n as L.

be written as follows:

$$P_n[t] = \begin{cases} \left(\frac{\mu}{\mathcal{B}}\right) \sum_{i=0}^{\min(L,L_0)} (-1)^i \binom{L+L_0-i-1}{L-i}\binom{L_0}{i} \\ \quad \times \left(\frac{1-e^{(\mu-\mathcal{B})t}}{1-\left(\frac{\mu}{\mathcal{B}}\right)e^{(\mu-\mathcal{B})t}}\right)^{L+L_0-i} \left(\frac{1-\left(\frac{\mathcal{B}}{\mu}\right)e^{(\mu-\mathcal{B})t}}{1-e^{(\mu-\mathcal{B})t}}\right)^{i} & \text{if } \mu \neq \mathcal{B} \\ \left(\frac{\mu t}{\mu t+1}\right)^{L+L_0} \sum_{i=0}^{\min(L,L_0)} \binom{L+L_0-i-1}{L-1}\binom{L_0}{i} \left(\frac{1-\mu^2t^2}{\mu^2t^2}\right)^{i} & \text{if } \mu = \mathcal{B}. \end{cases} \tag{4.32}$$

The derivation of this equation can be found in Bailey 1964 [143]. The right-hand side is the inverse Z transform of the generating function in equation 4.30.

For our purposes, in addition to the live cells n, we wish to also track another random variable, the number of dead cells[3] m. The generating function for the joint distribution of n and m is:

$$G[z,w,t] = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P_{n,m}[t] z^n w^m. \tag{4.33}$$

The differential equation governing $G[z,w,t]$ is:

$$\frac{\partial G[z,w,t]}{\partial t} = \left(\mu z^2 - (\mu+\mathcal{B})z + \mathcal{B}w\right)\frac{\partial G[z,w,t]}{\partial z}. \tag{4.34}$$

For $G[z,w,t]$, with the relevant initial condition $G[z,w,0] = z^{n_0}$, the solution to the partial differential equation can be found using the method of characteristics:

$$\begin{aligned} G[z,w,t] &= \left(\frac{(z-z_+)z_- - z_+(z-z_-)e^{-\gamma t}}{(z-z_+)-(z-z_-)e^{-\gamma t}}\right)^{n_0} \\ z_+ &= \frac{(\mu+\mathcal{B})}{2\mu}\left(1+\sqrt{1-\frac{4\mu\mathcal{B}w}{(\mu+\mathcal{B})^2}}\right) \\ z_- &= \frac{(\mu+\mathcal{B})}{2\mu}\left(1-\sqrt{1-\frac{4\mu\mathcal{B}w}{(\mu+\mathcal{B})^2}}\right) \\ \gamma &= \mu(z_+ - z_-). \end{aligned} \tag{4.35}$$

To relate the death rate $\mathcal{B}$ to pharmacodynamics (and to antibiotic dose and strain susceptibility), we can use the single-hit model dose-response curve from section 3.3.2. Recall that in the single-hit model, $h[t] = \beta = \frac{\beta_{max}[\text{ABX}]^\gamma}{\text{EC}_{50}^\gamma + [\text{ABX}]^\gamma}$. Then,

$$\mathcal{B} = \beta + k = \frac{\beta_{max}[\text{ABX}]^\gamma}{\text{EC}_{50}^\gamma + [\text{ABX}]^\gamma} + k, \tag{4.36}$$

[3] Denoted as D in the previous section.

where k is the antibiotic-independent, background death rate.

The multi-hit model of equation 3.25, with its delay in antibiotic killing α, was not explored to maintain algebraic simplicity. In the multi-hit model, $\mathcal{B}$ would be a non-constant function of overall elapsed time. A Markov birth-death process where the birth and death rates are arbitrary functions of overall time (but not of cell age) is called a time-inhomogeneous Markov birth-death process. The general solution for time-inhomogeneous Markov birth-death processes was explored in Kendall 1948 [142]. It is possible our particular gamma-distribution multi-hit model yields closed forms for quantities of interest, but this was not further explored in this thesis.

The following approximation enables one to compare the β fitted when the single-hit model is assumed to the β fitted from the same data with a multi-hit model assumed.

$$\beta_{\text{single}} = \frac{\beta_{\text{single}}}{1} \approx \frac{\beta_{\text{multiple}}}{\alpha} \tag{4.37}$$

This works because the mean cell lifetime when lifetimes are gamma-distributed with shape parameter α and rate parameter β is $\frac{\beta}{\alpha}$. The coefficient of variation of the lifetimes is $\alpha^{-\frac{1}{2}}$. Changing the value of α and β proportionally does not affect the average. It is only the spread of the lifetimes around the average lifetime that changes. By normalizing β_{multiple} by an α greater than one to obtain an equivalent β_{single}, the lifetimes become less concentrated about the same mean lifetime. See the curves in Figure 3.10 for a visualization.

The MBD's algebraic simplicity is its major draw. Unfortunately, there is no state variable for age in the model, and so the lineage and ages of the cells are not recorded, only the current total count at time t. The model does not keep track of which cell creates the next live or dead cell, and the chance of a cell to undergo a birth or death event is the same at every instant in time, even if the cell was born the instant prior. Such assumptions are not realistic for actual bacteria cells, which in liquid culture should exhibit a relatively narrow range of doubling times (centered around 20–25 minutes), as bacteria elongate with similar rate constants and only divide upon reaching a size falling within a narrow range [137]. Unfortunately, the short time scales and low inoculums examined in my digital AST experiments both argue against the appropriateness of the Markov assumption; the behavior would have converged after long times or with many cells per well.

Nonetheless, the casting of the assay state as a memoryless Markov process enables a very significant simplification in the algebra needed to describe the system state

over time, and so will be used here. Indeed, the MBD is likely the simplest stochastic model for discrete populations that is continuous in time.

4.4.3 Derivation of well population status probabilities for discrete pharmacodynamic models

To fit either of the discrete pharmacodynamic models above, the no-births or the Markov birth-death process, one needs to relate its parameters to output variables that one can measure. As discussed in section 4.2.3.5, the outputs of digital filtration AST are the nucleic acid concentrations measured for each fraction from each digital partition. The parameters of one's model could be fitted to the nucleic acid concentration matrices directly (see section 4.2.3.5), but I have not yet explored this approach. The approach I have taken instead fits the well population status tallies described in section 4.2.3.5.

As long as digital filtration AST is performed as intended without cross-contamination, then each well represents an independent instance of filtration AST, and the populations in each well progress stochastically and independently. I have derived the commonly-assumed Poisson-loading model for the loading of the wells in sections 4.1.3 and 4.2.2. We incorporate this model into our explanation of digital filtration AST below.

Let s_i^O, s_i^I, s_i^B, and s_i^E be the mutually exclusive events that well i has no cells, only intact cells, both killed and intact cells, or only killed cells, respectively. Let s_i be the status of well i, a categorical random variable taking on a value of s_i^O, s_i^I, s_i^B, or s_i^E with a certain probability for each outcome.

Let O, I, B, and E, as outputted by the procedures in section 4.2.3.5, be the observed counts of wells with the corresponding well population statuses indicated by their names. For these metrics, The total number of wells measured is denoted $N = I + B + E + O$. For example, $O = \sum_{i=1}^{N} \mathbf{1}(s_i = s_i^O)$, where $\mathbf{1}$ is an indicator random variable.

Let the vector $\mathbf{S} = [O, I, B, E] \in \mathbb{Z}^4$ represent the vector of well population status tallies. Because the number of wells N is fixed, and each s_i is an independent categorical random variable, $\mathbf{S}$ is then a multinomial random variable with a size parameter of N and a vector of probability parameters, the "well population status probabilities," as denoted below.

$$\mathbf{S} \sim \text{Multinomial}\left(N, \left[p_{\text{empty}}, p_{\text{no deaths}}, p_{\text{mixed}}, p_{\text{extinct}}\right]\right), \tag{4.38}$$

where $p_{\text{empty}} = \text{Prob}\left(s_i = s_i^O\right)$, $p_{\text{no deaths}} = \text{Prob}\left(s_i = s_i^I\right)$, $p_{\text{mixed}} = \text{Prob}\left(s_i = s_i^B\right)$, and $p_{\text{extinct}} = \text{Prob}\left(s_i = s_i^E\right)$.

The task now at hand is to define the four well population status probabilities as functions of relevant pharmacodynamic parameters and the parameters that determines our model of digital loading. These likelihood equations will link the assumed underlying pharmacodynamic parameters, including the EC50 susceptibility metric, to the available and observed (or more precisely, "inferred" by the methods of section 4.2.3.5) well loading statuses, allowing estimation or fitting of the underlying parameters.

I derive the well population status probabilities without births, then with births. For the Markov birth-death process, I derive the probabilities with and without a lysis efficiency parameter that represents the probability of a given intact cell not being collected and observed by the on-membrane lysis. Although the inclusion of the lysis efficiency parameter complicates the equations algebraically and computationally, it addresses a key source of measurement error seen during the prototype implementations of digital filtration AST herein.

Surprisingly, we can derive analytic expressions for the well status probabilities for all of these cases. These analytic expressions enable fitting to be implemented by basic statistical software.

4.4.3.1 Well population status probabilities without births

Let the density per well volume of bacteria be C, and let the probability that a given bacteria will die be p, with $q = 1 - p$ being the probability of a bacteria living. Let m_i be the number of cells loaded into well $i, i = 1 \ldots N$, there being N wells. Let M be the number of cells loaded into all N wells, with $M = \sum_{i=1}^{N} m_i$. Let $D = \sum_{i=1}^{N} d_i$ and $L = \sum_{i=1}^{N} l_i$ be the total number of dead and live cells across all N wells. We can assert that $M = D + L$.

According to our model of digital loading, $m_i \sim \text{Poisson}(\text{rate} = C)$. Let d_i and l_i be the number of dead and live cells in well i after the exposure. It follows that $m_i = d_i + l_i$.

Because of the assumption of no births, $d_i | m_i \sim \text{Binomial}(\text{n trials} = m_i, \text{prob} = p)$. That is, the number of dead cells d_i given a starting number of m_i cells is binomially distributed. Likewise, $l_i | m_i \sim \text{Binomial}(\text{n trials} = m_i, \text{prob} = q)$. (In a model

with births, one would generally need to distinguish between founder cells and new progeny when defining d_i and l_i.)

But by the multinomial-Poisson transformation, $d_i \sim \text{Poisson}(\text{rate} = Cp)$ and $l_i \sim \text{Poisson}(\text{rate} = Cq)$. Since the sum of Poisson random variables is also Poisson distributed, we have $M \sim \text{Poisson}(\text{rate} = NC)$, $D \sim \text{Poisson}(\text{rate} = NCp)$, and $L \sim \text{Poisson}(\text{rate} = NCq)$, which is consistent with our assertion that $M = D + L$. Thus, the no-births model is equivalent to a situation where the cells are balls of two colors drawn with replacement from an urn containing the two kinds of balls at a proportion p, then loaded into the N wells as a Poisson process. The loading of multiple cells into a well does not influence the propensity of each cell from dying or not.

In our digital filtration AST, we do not directly observe D, L, and M. Instead, we have only observed the well population statuses $\mathbf{S} = [O, E, I, B]$ defined in the preceding section (4.4.3). The well population status probabilities in no-births model are

$$\begin{aligned} p_{\text{empty}} &= \text{Prob}\left(s_i = s_i^O\right) \\ &= \text{Prob}\left(d_i = 0\right) \text{Prob}\left(l_i = 0\right) \\ &= e^{-Cp} e^{-Cq} \\ &= e^{-C}, \end{aligned} \tag{4.39}$$

$$\begin{aligned} p_{\text{no deaths}} &= \text{Prob}\left(s_i = s_i^I\right) \\ &= \text{Prob}\left(d_i = 0\right) \text{Prob}\left(l_i \geq 1\right) \\ &= e^{-Cp}\left(1 - e^{-Cq}\right), \end{aligned} \tag{4.40}$$

$$\begin{aligned} p_{\text{mixed}} &= \text{Prob}\left(s_i = s_i^B\right) \\ &= \text{Prob}\left(d_i \geq 1\right) \text{Prob}\left(l_i \geq 1\right) \\ &= \left(1 - e^{-Cp}\right)\left(1 - e^{-Cq}\right), \end{aligned} \tag{4.41}$$

$$\begin{aligned} p_{\text{extinct}} &= \text{Prob}\left(s_i = s_i^E\right) \\ &= \text{Prob}\left(d_i \geq 1\right) \text{Prob}\left(l_i = 0\right) \\ &= \left(1 - e^{-Cp}\right) e^{-Cq}. \end{aligned} \tag{4.42}$$

Since **S** is then a multinomial random variable, the probability of observing **S** given parameters p and C is

$$\begin{aligned}\text{Prob}(\mathbf{S}|C,p) = \frac{N!}{O!I!B!E!}\left(e^{-C}\right)^{O}\left(e^{-Cp}\left(1-e^{-Cq}\right)\right)^{I} \\ \times\left(\left(1-e^{-Cp}\right)\left(1-e^{-Cq}\right)\right)^{B}\left(\left(1-e^{-Cp}\right)e^{-Cq}\right)^{E}.\end{aligned} \tag{4.43}$$

4.4.3.2 Well population status probabilities with births and without lysis efficiency observation error

The chance of the well beginning with L_0 cells, denoted p_{L_0}, is Poisson-distributed with a mean parameter C that is the mean number of cells per well, so

$$p_{L_0} = P_{L_0}[0] = \frac{C^{L_0}e^{-C}}{L_0!}. \tag{4.44}$$

An empty well that does not receive any cells will always remain with 0 live and 0 dead cells, so

$$p_{\text{empty}} = P_0[0] = e^{-C}. \tag{4.45}$$

A well that only contains live cells must start with any number of non-zero live cells, and then subsequently never experience a death event. From the generating equation 4.35, we can derive the probability of no death events by time t, given L_0, since setting $z = 1$ and $w = 0$ generates the marginal probability (over all values of z) of having 0 dead cells.

$$\begin{aligned} z_+ &= \frac{\mu+\beta+k}{\mu} \\ z_- &= 0 \\ p_{\text{no deaths}|L_0} &= G[1,0,t] \\ &= \left[\frac{-z_+e^{-\mu z_+ t}}{1-z_+-e^{-\mu z_+ t}}\right]^{L_0} \\ &= \left[\frac{\mu+\beta+k}{(\beta+k)e^{(\mu+\beta+k)t}+\mu}\right]^{L_0} \end{aligned} \tag{4.46}$$

To get to $p_{\text{no deaths}}$, the probability of observing no deaths in a given well, one can marginalize over the different starting inoculums of the well, using equation 4.44 for p_{L_0}. (The case for $L_0 = 0$ is accounted for in p_{empty}.)

$$p_{\text{no deaths}} = \sum_{L_0=1}^{\infty} p_{\text{no deaths}|L_0}p_{L_0} \tag{4.47}$$

Similarly, we can find the probability of all the cells in a well dying, given L_0. By setting $z = 0$ and $w = 1$, the generating equation becomes the marginal probability (over all values of w dead cells) of having 0 live cells at time t. Then, we marginalize over L_0 in equation 4.49:

$$z_+ = \frac{\mu+\beta+k}{2\mu} + \sqrt{1 - \frac{4\mu(\beta+k)}{(\mu+\beta+k)^2}}$$

$$z_- = \frac{\mu+\beta+k}{2\mu} - \sqrt{1 - \frac{4\mu(\beta+k)}{(\mu+\beta+k)^2}}$$

$$\begin{aligned} p_{\text{extinct}|L_0} &= G[0,1,t] \\ &= \left[\frac{-z_-z_+ + z_-z_+e^{-\mu(z_+-z_-)t}}{-z_+ + z_-e^{-\mu(z_+-z_-)t}}\right]^{L_0} \\ &= \left[\frac{\beta+k}{\mu}\left(\frac{e^{(\mu-(\beta+k))t}-1}{e^{(\mu-(\beta+k))t} + \frac{\beta+k}{\mu}}\right)\right]^{L_0} \end{aligned} \tag{4.48}$$

$$p_{\text{extinct}} = \sum_{L_0=1}^{\infty} p_{\text{extinct}|L_0} p_{L_0}. \tag{4.49}$$

The last component of the well probability vector is p_{mixed}. It is the remaining probability not covered by empty, all alive, or extinct wells.

$$p_{\text{mixed}|L_0} = 1 - p_{\text{no deaths}|L_0} - p_{\text{extinct}|L_0} \tag{4.50}$$

$$p_{\text{mixed}} = \sum_{L_0=1}^{\infty} p_{\text{mixed}|L_0} p_{L_0} \tag{4.51}$$

4.4.3.3 Well population status probabilities with births and with lysis efficiency observation error

In many digital filtration AST experiments, an imbalance in the observation of intact cells versus killed cells is observed, and the null hypothesis of this imbalance being due to stochastic loading was rejected at a significance of 0.05. Therefore, it was decided to incorporate a lysis efficiency term to the Markov birth-death model. For simplicity, cells lysed by antibiotics were considered always observable. In contrast, intact cells were assumed to have an independent probability of lysing λ that could be less than 1. Denote as $p_{\text{missed}|L}$ the probability that none of a well intact cells were lysed by the lysis buffer and observed, given that the well contains L intact cells. Then,

$$p_{\text{missed}|L} = (1-\lambda)^L. \tag{4.52}$$

Due to the lysis inefficiency, two more well statuses are created, each indistinguishable from another status. First, if a well has not experienced any death events, and none of the intact cells are lysed and observed, then the well will appear to have been empty. Let the probability of this event be called $p_{\text{missed,no deaths}}$, and $p_{\neg\text{missed,no deaths}}$ to be the case that at least one intact cell is observed. Second, if a well contains both intact and dead cells, yet none of the intact cells are lysed and observed, then the well appears as one in which all the cells have died from antibiotics. The probability of this well status is $p_{\text{missed,mixed}}$, and its complement is $p_{\neg\text{missed,mixed}}$. The model for our observed well status tallies becomes

$$\mathbf{S} \sim \text{Multinomial}\left(N, \begin{bmatrix} p_{\text{empty}} + p_{\text{missed,no deaths}} \\ p_{\neg\text{missed,no deaths}} \\ p_{\neg\text{missed,mixed}} \\ p_{\text{extinct}} + p_{\text{missed,mixed}} \end{bmatrix}\right). \tag{4.53}$$

The formulas for $p_{\text{missed,no deaths}}$ and $p_{\text{missed,mixed}}$ are not immediately obvious, but by factorizing the probability into simpler conditional terms, an analytical expression is obtained that can be approximated by computational software.

First, we factor out the starting inoculum, skipping the term for an empty well. Then we factor out the true well status.

$$\begin{aligned} p_{\text{missed,no deaths}} &= \sum_{L_0=1}^{\infty} p_{\text{missed,no deaths}|L_0} p_{L_0} \\ &= \sum_{L_0=1}^{\infty} p_{\text{missed}|\text{no deaths},L_0} p_{\text{no deaths}|L_0} p_{L_0} \end{aligned} \tag{4.54}$$

Similarly,

$$\begin{aligned} p_{\neg\text{missed,no deaths}} &= \sum_{L_0=1}^{\infty} p_{\neg\text{missed,no deaths}|L_0} p_{L_0} \\ &= \sum_{L_0=1}^{\infty} p_{\neg\text{missed}|\text{no deaths},L_0} p_{\text{no deaths}|L_0} p_{L_0} \\ &= \sum_{L_0=1}^{\infty} (1 - p_{\text{missed}|\text{no deaths},L_0}) p_{\text{no deaths}|L_0} p_{L_0}. \end{aligned} \tag{4.55}$$

Repeating for mixed wells,

$$p_{\text{missed,mixed}} = \sum_{L_0=1}^{\infty} p_{\text{missed}|\text{mixed},L_0} p_{\text{mixed}|L_0} p_{L_0}, \tag{4.56}$$

$$p_{\neg\text{missed,mixed}} = \sum_{L_0=1}^{\infty} (1 - p_{\text{missed}|\text{mixed},L_0}) p_{\text{mixed}|L_0} p_{L_0}. \tag{4.57}$$

The expressions for $p_{\text{no deaths}|L_0}$ and $p_{\text{mixed}|L_0}$ were derived in equations 4.46 and 4.50.

Next, we factor $p_{\text{missed}|\text{no deaths},L_0}$ and $p_{\text{missed}|\text{mixed},L_0}$ by the number of intact cells L present before the lysis buffer is added.

$$\begin{aligned} p_{\text{missed}|\text{no deaths},L_0} &= \sum_{L=0}^{\infty} p_{\text{missed}|L,\text{no deaths},L_0} p_{L|\text{no deaths},L_0} \\ &= \sum_{L=0}^{\infty} p_{\text{missed}|L} p_{L|\text{no deaths},L_0} \end{aligned} \tag{4.58}$$

$$\begin{aligned} p_{\text{missed}|\text{mixed},L_0} &= \sum_{L=0}^{\infty} p_{\text{missed}|L,\text{mixed},L_0} p_{L|\text{mixed},L_0} \\ &= \sum_{L=0}^{\infty} p_{\text{missed}|L} p_{L|\text{mixed},L_0} \end{aligned} \tag{4.59}$$

Above, I used the fact that in our model, the probability of an intact cell being lysed by lysis buffer is independent of whether the well is in an "all alive" or "mixed" state. Therefore, we can use the expression for $p_{\text{missed}|L}$ in equation 4.52.

The quantities $p_{L|\text{no deaths},L_0}$ and $p_{L|\text{mixed},L_0}$ are still not easily defined. But they can be redefined using the definition of conditional probability.

$$p_{L|\text{no deaths},L_0} = \frac{p_{L,\text{no deaths}|L_0}}{p_{\text{no deaths}|L_0}} \tag{4.60}$$

$$p_{L|\text{mixed},L_0} = \frac{p_{L,\text{mixed}|L_0}}{p_{\text{mixed}|L_0}} \tag{4.61}$$

The denominators are known from equations 4.46 and 4.50. The numerators are found as follows. The quantity $p_{L,\text{no deaths}|L_0}$ is equivalent to finding the inverse Z-transform of the generating equation (equation 4.35) where $w = 0$, while z remains a free dummy variable. This yields:

$$p_{L,\text{no deaths}|L_0} = \begin{cases} \binom{L-1}{L_0-1} e^{-L(\mu+\beta+k)t} \left[\frac{\mu}{\mu+\beta+k} \left(e^{(\mu+\beta+k)t} - 1 \right) \right]^{L-L_0} & \text{if } L \geq L_0 \\ 0 & \text{if } L < L_0. \end{cases} \tag{4.62}$$

The quantity $p_{L,\text{mixed}|L_0}$ can be found by subtracting the probability of mutually exclusive sub-events from the probability of the event $L|L_0$ of having L cells at time

t with a starting inoculum of L_0.

$$p_{L,\text{mixed}|L_0} = p_{L|L_0} - p_{L,\text{no deaths}|L_0} - p_{L,\text{extinct}|L_0} \tag{4.63}$$

The $p_{L,\text{no deaths}|L_0}$ term was derived above. The closed form for $p_{L|L_0}$ was given in equation 4.32, in which $p_{L|L_0}$ was denoted $P_n[t]$. The remaining unspecified term is $p_{L,\text{extinct}|L_0}$.

Note that since a population is extinct only when $L = 0$ has been reached, the following equation holds:

$$p_{L,\text{extinct}|L_0} = \begin{cases} p_{\text{extinct}|L_0} & \text{if } L = 0 \\ 0 & \text{if } L > 0. \end{cases} \tag{4.64}$$

The expression for $p_{\text{extinct}|L_0}$ was derived earlier in equation 4.48.

The probabilities for the four possible well population statuses have now been defined. Graphs of these probabilities, and their component probability terms, are shown in Figure 4.12 for aid in understanding the model.

4.4.4 Estimates of pharmacodynamic parameter values from digital filtration AST using Bayesian Hamiltonian Monte Carlo

In this section, we wish to know whether digital filtration AST results recapitulate the *in vitro* pharmacodynamics observed in bulk filtration AST. Specifically, we wish to rule out the existence of processes (natural or man-made design flaws) during digital filtration AST that would cause the digital outputs to deviate significantly from those predicted in bulk experiments. To do this, we fit digital filtration AST data to the MBD model implied by the well population status probability equations above, then for the pharmacodynamic parameters shared with the bulk filtration AST's model in section 3.3, compare their values.

Additionally, fitting pharmacodynamic models may be useful in the future for finding values (or at least narrowing priors) of parameters not dependent on the clinical specimen, namely the measurement error of lysis efficiency. One might do this in clinical practice for quality control of one's device.

To fit data (to an overdetermined system of equations), we must assume a source of random error, which in some algorithms implies a loss function to minimize. In the compartment model of bulk filtration AST, the error model comprised PCR measurement errors modeled in section 3.3.3. For digital filtration AST, some of

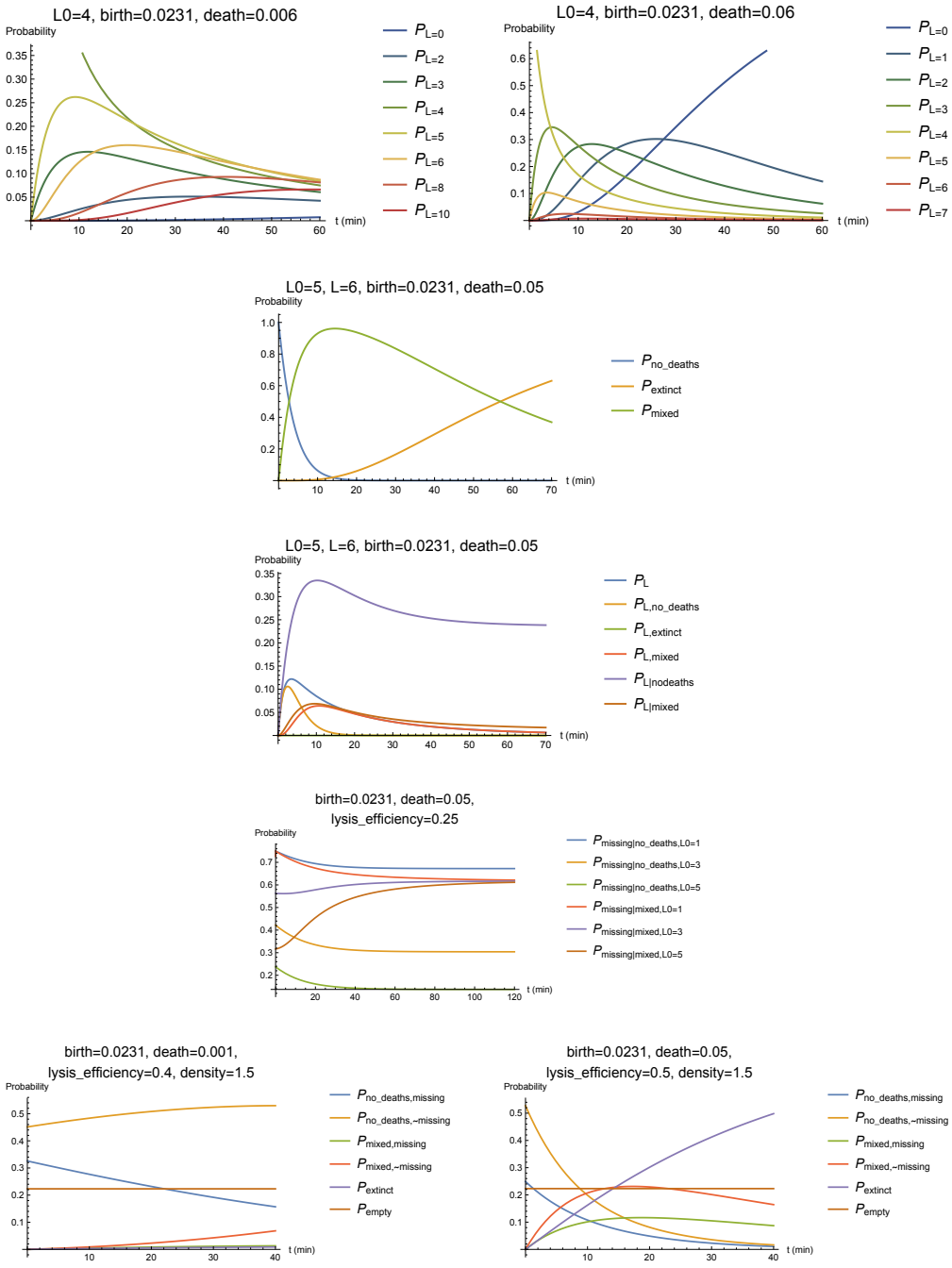


Figure 4.12: Probabilities in the Markov birth-death model related to the well population status. The plot titles show the values of fixed parameters used to generate the plots.

the stochasticity in any measured data is assumed to arise from the stochasticity of the underlying system. Specifically, the multinomial probability of equation 4.53 was assumed for the analysis below. Some of the error was also assumed to arise from batch effects introduced during the human-operated measurement process, as discussed in the next section. Unlike the bulk filtration AST data, PCR measurement error was not modeled for the digital filtration AST data. Instead, the PCR measurement error was implicitly assumed to have been corrected during well population status calling.

Using Bayesian Monte Carlo methods, each parameter in the model/well population status probability equations can be considered a random variables representing our uncertainty in their values. Variables which we believe are fixed constants are simply random variables with no spread and no uncertainty. Variables which we believe are not shared across groups of data can be modeled hierarchically.

4.4.4.1 Digital filtration AST measurement error model

For this section's purposes, the growth rate was assumed equal for all strains. The EC_{50} parameters were assumed to be unique to each strain, and the maximum antibiotic kill rate and Hill coefficient were assumed to be unique for each antibiotic. For the batch parameters, the target inoculum I_{target} was inputted as known data.

To account for variability beyond the multinomial stochasticity of the well statuses, two sets of noise parameters were introduced.

First, the deviations from the target inoculum were modeled as hierarchically, gamma distributed constants I with a mean of $\bar{I}$ and a standard deviation of σ_I. The mean $\bar{I}$ was given a mean of 1 and a coefficient of variation of 0.5. The variance of I, σ_I^2, was assumed to be an inverse gamma distribution with a mean of $9 = 3^2$ and a standard deviation of 9.

$$I \sim \text{Gamma}\left(\text{shape}=\frac{\bar{I}^2}{\sigma_I^2}, \text{rate}=\frac{\bar{I}}{\sigma_I^2}\right) \tag{4.65}$$

$$\bar{I} \sim \text{Gamma}\,(\text{shape}=2, \text{rate}=4) \tag{4.66}$$

$$\sigma_I^2 \sim \text{InvGamma}\,(\text{shape}=3, \text{rate}=18) \tag{4.67}$$

$$C = I \cdot I_{\text{target}} \tag{4.68}$$

Second, each AST experiment batch was given an independent lysis efficiency

parameter l with a beta distribution prior:

$$l \sim \text{Beta}\,(\text{mean}{=}0.5, \text{concentration}{=}3)\,. \tag{4.69}$$

To clarify, the inoculum densities C for each condition and the global lysis efficiency l were free (non-pharmacodynamic) parameters in equation 4.53. The prior distributions required for Bayesian analysis are now given in the above equations.

The priors for the rest of the pharmacodynamic parameters were kept the same as in the bulk filtration AST model of section 3.3.4. An exception was made for the maximum antibiotic kill rate $\beta_{\max}$. The assumed mean of the prior of $\beta_{\max}$ was divided by the assumed mean for the prior of α. The justification for this crude correction was explained in section 4.4.2.

4.4.4.2 Results and posterior predictive checks

All five Monte Carlo chains converged and sampled the posterior with low autocorrelation, as seen in Figures 4.13–4.14b.

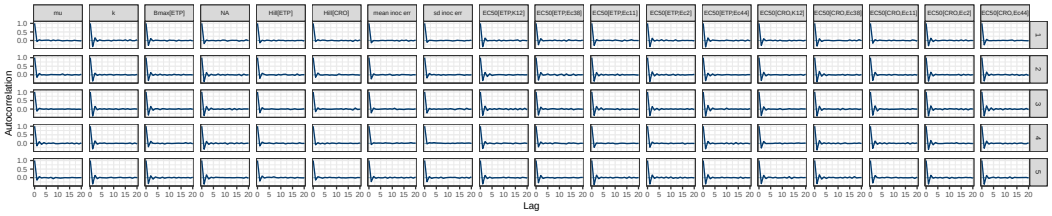
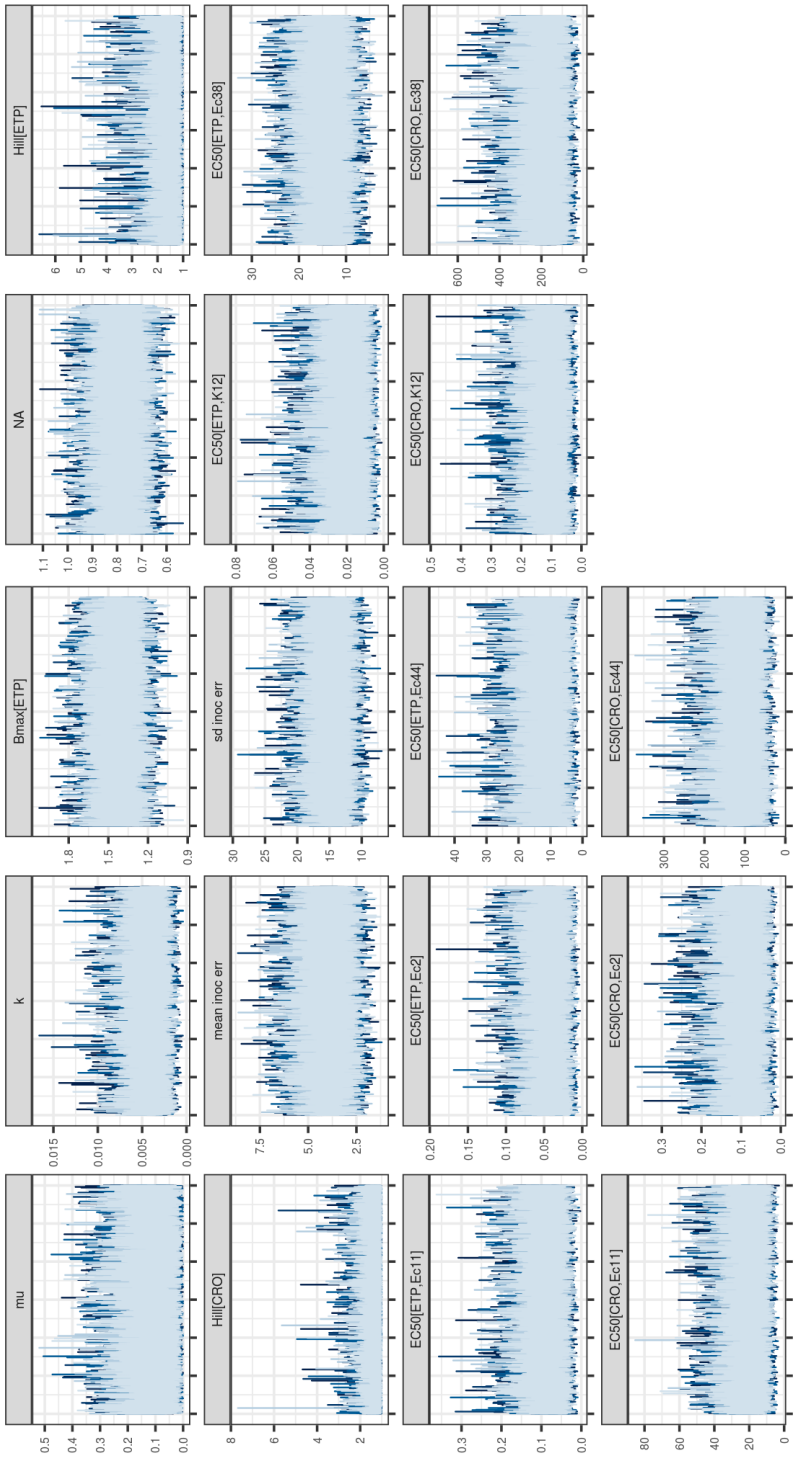


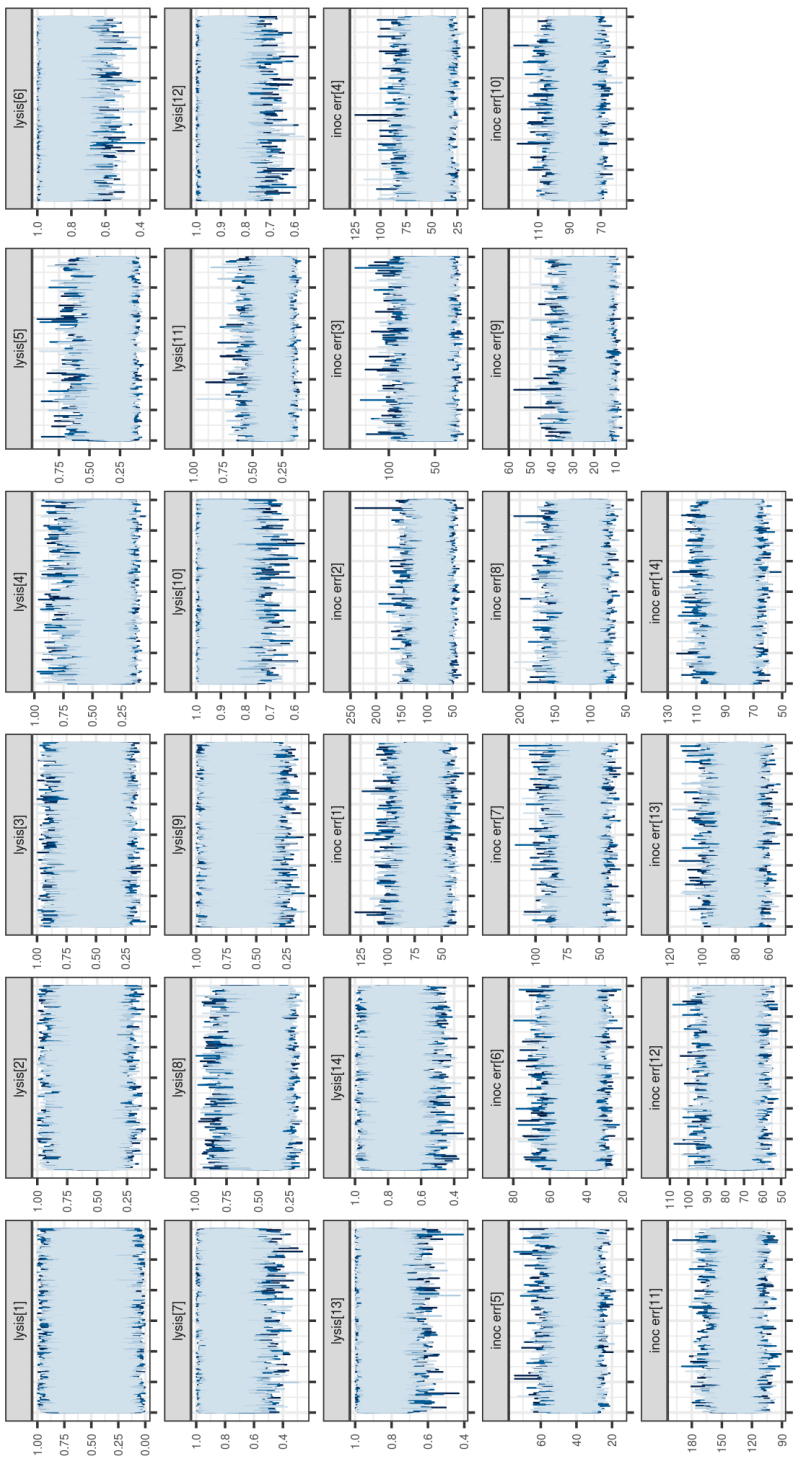
Figure 4.13: Autocorrelation of Monte Carlo chain sampling of the Markov birth-death pharmacodynamic model.

The marginal posterior means of the fitted parameters are listed below in Table 4.2 alongside the values from Table 3.3 in the bulk compartment model.



(a) Shared pharmacodynamic parameters

Figure 4.14: Monte Carlo chain traces of the Markov birth-death pharmacodynamic model.



(b) Experiment batch effect parameters

Figure 4.14: Monte Carlo chain traces of the Markov birth-death pharmacodynamic model (cont'd).

Parameter	interpretation	bulk value	digital value
μ	intrinsic growth rate (min^{-1})	0.0163	0.00183
k	intrinsic death rate (min^{-1})	0.0000341	0.0000705
B_{max}[ETP]	maximum kill rate for ertapenem (min^{-1})	0.0843	0.0237
B_{max}[CRO]	maximum kill rate for ceftriaxone (min^{-1})	0.0617	0.0131
α[ETP]	delay in ertapenem activity	4.65	NA
α[CRO]	delay in ceftriaxone activity	5.17	NA
hill[ETP]	Hill coefficient for ertapenem	1.11	1.58
hill[CRO]	Hill coefficient for ceftriaxone	1.08	1.37
EC_{50}[ETP, K12]	EC_{50} for ertapenem + E. coli K12 (S, µg/mL)	0.0908	0.0196
EC_{50}[CRO, K12]	EC_{50} for ceftriaxone + E. coli K12 (S, µg/mL)	0.822	0.108
EC_{50}[ETP, Ec38]	EC_{50} for ertapenem + E. coli #38 (R, µg/mL)	9.39	14.2
EC_{50}[CRO, Ec38]	EC_{50} for ceftriaxone + E. coli #38 (R, µg/mL)	298	182

Table 4.2: Fitted marginal posterior mean values of the digital filtration AST pharmacodynamic model.

The parameter with the biggest fold-change in value was the growth rate μ. The decrease may be related to the lysis inefficiency masking only live cells, not dead cells, but further experiments are needed. The other parameters remained within an order of magnitude of each other. The digital fitted value more resembles the prior distributions, indicating that the digital AST data set is less informative, as is expected given the fewer number of experiments. Although no new information is gained from the digital data alone, the absence of strong value changes indicates that the MBD model is sufficient for describing the observed results. No new mechanisms are needed.

4.5 Diagnostic performance of digital filtration AST

Digital filtration AST is a novel rapid phenotypic AST method, and the accuracy of the method across strains as a function of exposure duration and inoculum are of interest for future investigators. In this section, I demonstrate that digital filtration AST does infer susceptibilities accurately in a diagnostic setting resembling clinical practice. In a diagnostic setting, the identities and susceptibilities of the strains used are unknown. Furthermore, the goal of modeling shifts from accurate inference of

all parameters to the maximal accuracy in classifying strains as susceptible.

In particular, the inference performed earlier in section 4.4.4 (and reported in Table 4.2) pooled data across experiments performed with the same strain. Informative priors based on known strain MICs were assumed for the EC_{50} parameters. In this section, Bayesian estimates were instead obtained by pretending that each experiment tested a unique strain with unknown susceptibility. This is further detailed in section 4.5.2.

In section 4.4.4, the EC_{50} is the only parameter in the model reflecting strain susceptibility. However, the use of Bayesian Monte Carlo to estimate the EC_{50} is also not the only way susceptibility can be called from the outputs of digital filtration AST. In fact, there are infinite algorithms one can propose; only some of which will approach the best possible classification accuracy. Six algorithms are discussed in the next section (4.5.1), only one of which is the estimate of the EC_{50} by Bayesian HMC.

4.5.1 Six susceptibility metrics and susceptibility classification algorithms

Six algorithms for calculating susceptibility were performed in this section in parallel. Multiple algorithms were included to better summarize the performance of the tangible steps of digital filtration AST as a method, and also to enable comparison of the algorithms' performances.

Every classification algorithm for susceptibility calls implicitly assumes a pharmacodynamic system model, but only the algorithms that calculate the EC_{50} and β metrics use the duration of exposure as an input. The other metrics are equivalent to models in which the exposure comprises a single instant of time during which cells die or not (e.g., a coin flip).

All of the metrics calculated utilized well population statuses instead of the unprocessed nucleic acid concentration measurements, as discussed in section 4.2.3.5. The performance of algorithms directly using concentration measurements is left as future work.

A digital filtration AST experiment batch comprises a set of antibiotic exposures digitally loaded in parallel from the same 2 mL batch culture. An experimental condition in that experiment batch is a subset of the exposures sharing the same antibiotic dose, exposure duration, and inoculum. Two of the metrics (PE_{naive}, PE_{MLE}) were calculated for each experimental condition. Three (p_{binomial}, the RDI, and β) were calculated by comparing a treated experimental condition with

an untreated condition within the same digital experiment batch. The sixth metric (EC_{50}) was defined and calculated for each strain, and therefore once for each batch of experimental conditions.

The six algorithms herein all comprise calculating a real-valued susceptibility metric first, then comparing the metric against a threshold to obtain the categorical susceptibility classification. (An example of an algorithm that does not calculate a metric would be a k-nearest neighbor classifier.)

One of the metrics is actually the p-value of a hypothesis test. Hypothesis testing of universal truths is traditionally and theoretically a concept distinct from the classification of specific instances [145]. Clinical diagnostics falls under the latter framework, so I will cast hypothesis testing for susceptibility as a classification algorithm in this thesis.

The six algorithms for the well statuses are now defined below.

4.5.1.1 The proportion or probability of death

The simplest metric PE_{naive} is an approximation for the proportion or percentage of killed (extracellular) cells in a given experimental condition. In a crude approximation, it is assumed that all wells have 1 cell only, except that wells that have both extracellular and intracellular are assumed to have 1 intact and 1 killed cell.

$$PE_{\text{naive}} = \frac{E + B}{E + 2B + I} \tag{4.70}$$

The second metric PE_{MLE} is a better approximation of the probability of antibiotic killing than PE_{naive}. It is the frequentist maximum likelihood estimator $\hat{p}$ for the survival probability of the no-births model.

$$PE_{\text{MLE}} = \hat{p} = \begin{cases} \frac{1}{\hat{c}} \ln\left(\frac{(E-I)+\sqrt{(E-I)^2+4(B+I)(B+E)e^{-\hat{c}}}}{2(B+E)e^{-\hat{c}}}\right), & \text{if } E + B > 0 \\ 0, & \text{if } E + B = 0 \end{cases} \tag{4.71}$$

$$\hat{c} = -\ln\left(\frac{(O + E)(O + I)}{(O + I + B + E)^2}\right) \tag{4.72}$$

For the case where $E + B > 0$, one can substitute equation 4.72 into 4.71 to obtain:

$$PE_{\text{MLE}} = \frac{\ln\left(\frac{O+I}{N}\right)}{\ln\left(\frac{(O+E)(O+I)}{N^2}\right)} \tag{4.73}$$

$$= \log_{\frac{(O+E)(O+I)}{N^2}}\left[\frac{O+I}{N}\right] \tag{4.74}$$

$$= \frac{\ln\left(\frac{O+I}{N}\right)}{\ln\left(\frac{O+I}{N}\right) + \ln\left(\frac{O+E}{N}\right)}, \tag{4.75}$$

where $N = O + I + B + E$.

To derive equations 4.71 and 4.72, we find the values of p and c that maximize the likelihood $\mathcal{L} = \text{Prob}(\mathbf{S}|C, p)$ in equation 4.43. This is equivalent to finding the maximum of the log-likelihood $\ell = \ln(\mathcal{L})$. In other words,

$$\begin{aligned}\ell = &-OC - ICp + I\ln\left(1 - e^{-Cq}\right) - ECq + E\ln\left(1 - e^{-Cp}\right) \\ &+ B\ln\left(1 - e^{-Cp}\right) + B\ln\left(1 - e^{-Cq}\right) + \ln\left(\frac{N!}{O!I!B!E!}\right).\end{aligned} \tag{4.76}$$

The maximum occurs when $\frac{\partial \ell}{\partial p} = 0$, $\frac{\partial \ell}{\partial c} = 0$. First solving for p, we write:

$$\frac{\partial \ell}{\partial p} = EC + \frac{ECe^{-pC}}{1 - e^{-pC}} - IC - \frac{ICe^{-qC}}{1 - e^{-qC}} + \frac{BCe^{-pC}}{1 - e^{-pC}} - \frac{BCe^{-qC}}{1 - e^{-qC}}. \tag{4.77}$$

To solve for the MLE $\hat{p}$ (and defining $\hat{c}$ to be the MLE for C and $\hat{q} = 1 - \hat{p}$ for convenience), we set the equation above to 0, eliminate a solution $\hat{c} = 0$ since we are only interested in $C > 0$, and define the change of variables $u = e^{\hat{p}\hat{c}}$ and $v = e^{\hat{q}\hat{c}} = \frac{e^{\hat{c}}}{u}$.

$$0 = E + \frac{Eu^{-1}}{1 - u^{-1}} - I - \frac{Iv^{-1}}{1 - v^{-1}} + \frac{Bu^{-1}}{1 - u^{-1}} - \frac{Bv^{-1}}{1 - v^{-1}} \tag{4.78}$$

$$= E + \frac{E}{u - 1} - I - \frac{I}{v - 1} + \frac{B}{u - 1} - \frac{B}{v - 1} \tag{4.79}$$

$$= \frac{(E - I)(u - 1)(v - 1) + (E + B)(v - 1) - (B + I)(u - 1)}{(u - 1)(v - 1)} \tag{4.80}$$

$$= \frac{(E - I)(u - 1)(e^{\hat{c}} - u) + (E + B)(e^{\hat{c}} - u) - (B + I)(u - 1)u}{(u - 1)(e^{\hat{c}} - u)}. \tag{4.81}$$

$$0 = (E-I)(-u^2 + (e^{\hat{c}}+1)u - e^{\hat{c}}) + (E+B)(e^{\hat{c}} - u) - (B+I)(u^2 - u) \quad (4.82)$$

$$= -(B+E)u^2 + e^{\hat{c}}(E-I)u + e^{\hat{c}}(B+I). \quad (4.83)$$

$$u = \frac{e^{\hat{c}}(E-I) \pm \sqrt{e^{2\hat{c}}(E-I)^2 + 4(B+E)e^{\hat{c}}(B+I)}}{2(B+E)} \quad (4.84)$$

$$e^{\hat{c}\hat{p}} = \frac{(E-I) + \sqrt{(E-I)^2 + 4(B+E)(B+I)e^{-\hat{c}}}}{2(B+E)e^{-\hat{c}}} \quad (4.85)$$

Solving for $\hat{p}$ yields equation 4.71. We have discarded extraneous solutions created when we simplify the rational terms containing u, including the "-" solution in the last equation. In the case that $B + E = 0$, the solution is $\hat{p} = 0$.

To solve for $\hat{c}$, we define $w = e^{\hat{c}}$.

$$\frac{\partial \ell}{\partial c} = -O - Eq + \frac{Epe^{-cp}}{1-e^{-cp}} - Ip + \frac{Iqe^{-cq}}{1-e^{-cq}} + \frac{Bpe^{-cp}}{1-e^{-cp}} + \frac{Bqe^{-cq}}{1-e^{-cq}} \quad (4.86)$$

$$0 = -(E+O)\hat{q} - (I+O)\hat{p} + \frac{(E+B)\hat{p}}{w^{\hat{p}}-1} + \frac{(I+B)\hat{q}}{w^{\hat{q}}-1} \quad (4.87)$$

We then substitute in the solution for $\hat{p}$ in equation 4.85 and solve for w.

4.5.1.2 The relative difference index (RDI)

The third metric is the batch run-specific relative difference index RDI of any experimental condition-specific metric X between a treated and an untreated reference condition. For example, either of the PE metrics can be used for X; in this thesis, PE_{MLE} was used. The RDI is defined as

$$RDI = \frac{X_{\text{treated}} - X_{\text{reference}}}{X_{\text{treated}} + X_{\text{reference}}}. \quad (4.88)$$

RDI takes on a value between -1 and 1. When there is no difference in the X's, $RDI = 0$. The RDI is more interpretable between experimental batch runs than either the difference or the ratio of X_{treated} and $X_{\text{reference}}$.

4.5.1.3 Binomial hypothesis testing

The fourth metric p_{binomial} is the p-value from a version of the one-tailed Barnard's exact test for binomially-distributed binary outcomes. A crude assumption is made that each well represents 1 cell, except that cells with mixed populations have 1 intact and 1 killed cell each. Let E_T and E_U be the number of wells with extinct populations

in the treated and untreated conditions, and similarly for I_T, I_U, B_T, B_U, O_T, and O_U. Let N_T and N_U be the assumed total number of cells in each condition. We assume that each cell/well dies or survives with an independent, identical probability. The results in each condition is therefore binomially distributed, and Barnard's test can be applied to the 2x2 contingency table containing the numbers of live and dead cells in each condition. In the null hypothesis, the probability of death p is the same across both treated and untreated conditions. p is an unknown nuisance parameter. In Barnard's test, the value of p giving the maximum (least-significant) p-value is chosen. In calculating p_{binomial}, we instead choose p to be the PE_{naive} metric described above when considering the treated and untreated conditions together as one condition.

$$\begin{aligned} p = PE_{\text{naive,null}} &= \frac{E_T + E_U + B_T + B_U}{N_T + N_U} \\ &= \frac{E_T + E_U + B_T + B_U}{E_T + 2B_T + I_T + E_U + 2B_U + I_U} \end{aligned} \tag{4.89}$$

To test the null hypothesis, we calculate p_{binomial} to be the probability in each condition of a result equal or more extreme than the observed result. Such results are those values of S whose PE_{naive} is $\geq p$ for the treated condition, and $\leq p$ for the untreated condition. Because of our crude assumption that each well with a status of s^E or s^B has 1 killed cell each, such results are those values of the status vector S with a sum of E and B greater than observed.

$$\begin{aligned} p_{\text{binomial}} &= \text{Prob}\left(E + B \geq E_T + B_T; N_T, p\right) \text{Prob}\left(E + B \leq E_U + B_U; N_U, p\right) \\ &= \left(\sum_{x=E_T+B_T}^{N_T} \binom{N_T}{x} p^x \left(1-p\right)^{N_T - x} \right) \\ &\quad \times \left(\sum_{x=0}^{E_U+B_U} \binom{N_U}{x} p^x \left(1-p\right)^{N_U - x} \right) \end{aligned} \tag{4.90}$$

We reject the null hypothesis if the p-value of the observed results, or a more extreme result, is less than 0.05. Alternatively, we call a strain susceptible to the dose of antibiotic in the treated condition if p_{binomial} is less than a threshold that we must choose beforehand. The threshold does not have to be the traditional 0.05 significance threshold of hypothesis testing.

It may be possible to derive the analogue of p_{binomial} for PE_{MLE}. To do so, one needs to define what is a "more extreme" result of the multinomially distributed random

variable **S** than the observed results. Since the possible values of a multinomial distribution are partially ordered, unlike the binomial distribution, the situation constitutes a discrete optimization problem, similar to that discussed in Greenberg 1985 [146]. Unfortunately, I have not finished this investigation. Such a test would be more accurate. On the other hand, p_{binomial} is already a decent approximation for PE_{MLE} in the regime of the observed data, so the same may be true of the hypothesis testing p-value.

4.5.1.4 The antibiotic kill rate and EC_{50}

The fifth metric is the instantaneous antibiotic kill rate defined in the pharmacodynamic model (with the lysis efficiency parameter) in 4.4. To calculate this metric, one would need to either solve or fit the model equations for β. In this thesis, I used Stan to fit β for each experimental condition in all digital filtration AST experiment batches. However, in contrast to the fitting done in section 4.4.4, the starting inoculum, lysis efficiency, background lysis rate, and EC_{50} was estimated separately for each batch. Treating each batch as a separate strain (even if not) emulates the level of knowledge a clinical laboratory has when it receives a clinical specimen. Furthermore, the rest of the pharmacodynamic parameters that were not strain-specific were fixed at the value measured by the bulk experiments in section 3.4.3. This emulates a situation in the future when these values could plausibly be taken from a database compiled over the finite number of clinically useful antibiotics and known bacterial pathogens or common contaminant species. In this situation, the Hamiltonian Monte Carlo performed by Stan essentially functions as a numerical equation solver for the EC_{50} parameter, with the three other non-fixed nuisance parameters marginalized out.

The sixth metric is the EC_{50} as defined by the Hill equation and is found simultaneously with the above fitting of β. Because β_{max} and γ have been fixed, the values of the EC_{50} is deterministic given β. Both are shown separately, however, because β is defined for each treated condition, while the EC_{50} is defined for each strain.

The β and EC_{50} parameters of the no-births model and the MBD model without lysis efficiency could also be calculated using Stan. The results are not shown for brevity.

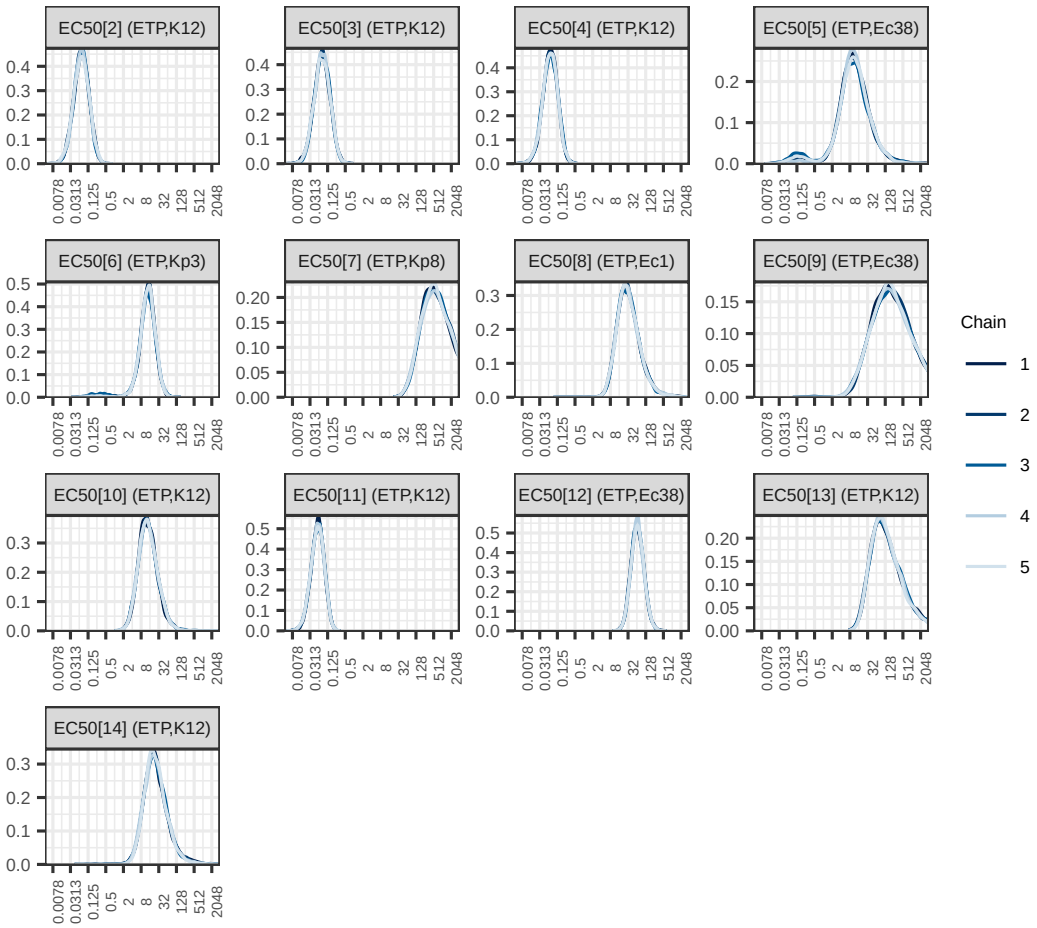
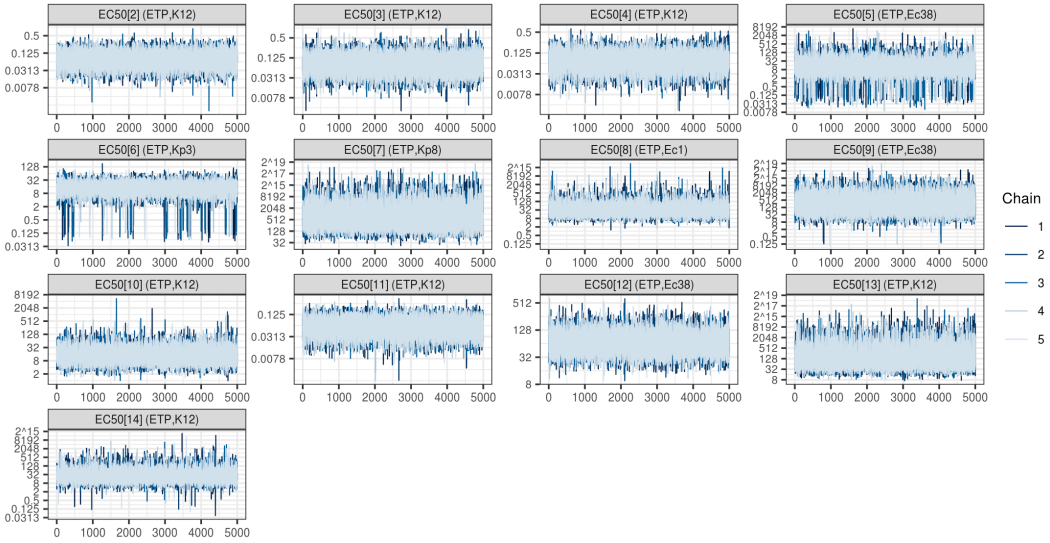


Figure 4.15: Estimates of EC_{50}, digital filtration diagnostic performance model.

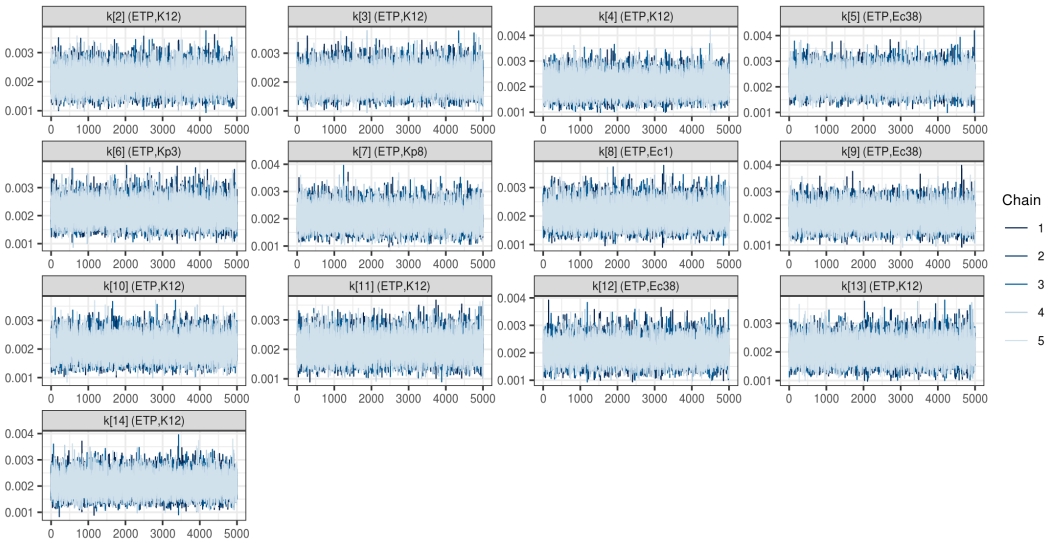
4.5.2 Bayesian Hamiltonian Monte Carlo estimates of the EC_{50}

The results of calculating the EC_{50}, one of the six metrics of section 4.5.1, are shown in Figures 4.15 and 4.16. The model was able to provide estimates for all parameters without any NUTS divergences or high autocorrelation (Figure 4.17).

Some multimodality was seen in the traces of the EC_{50} for the two experiments where background lysis was observed. This could indicate that the priors for the background lysis rate may be too constrained.

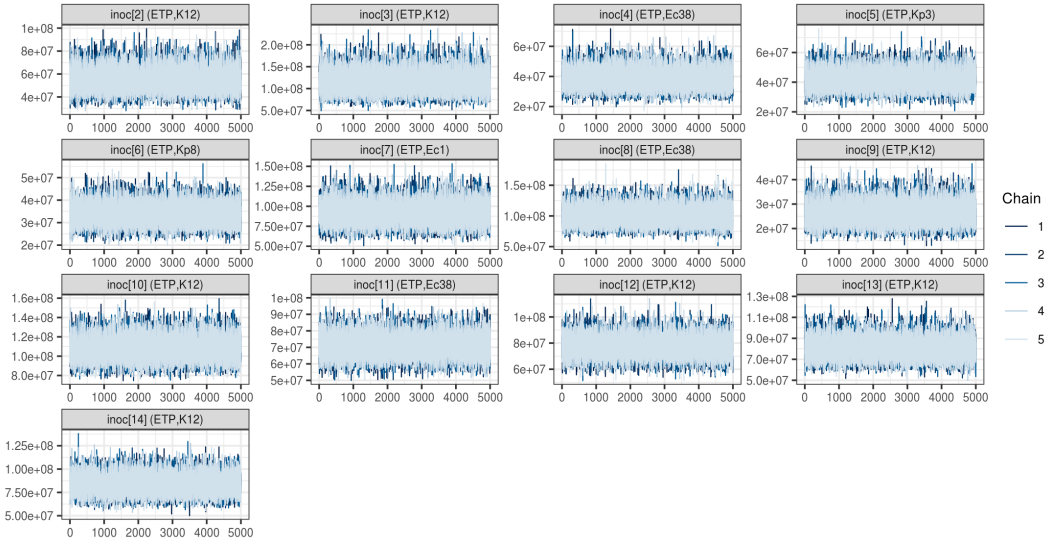


(a) Effective concentration 50. Note the log-scale y-axis

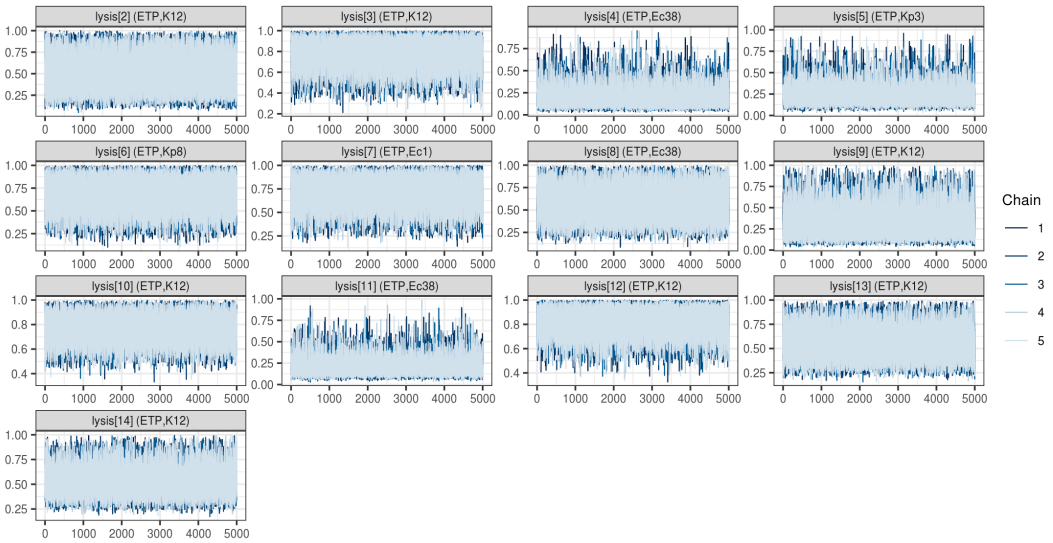


(b) Background lysis parameters

Figure 4.16: Markov chain traces during calculation of susceptibility metrics by HMC for unknown strains.

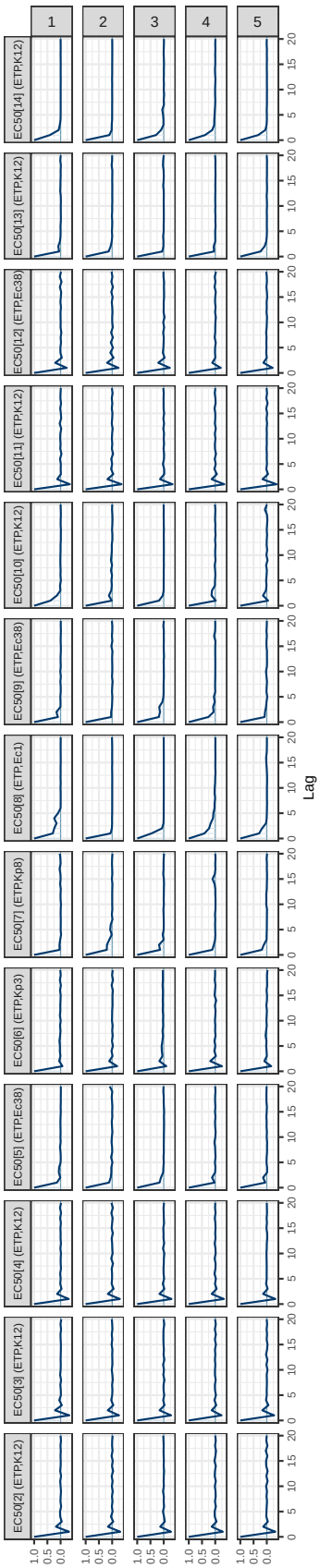


(c) Inoculum density in the batch culture

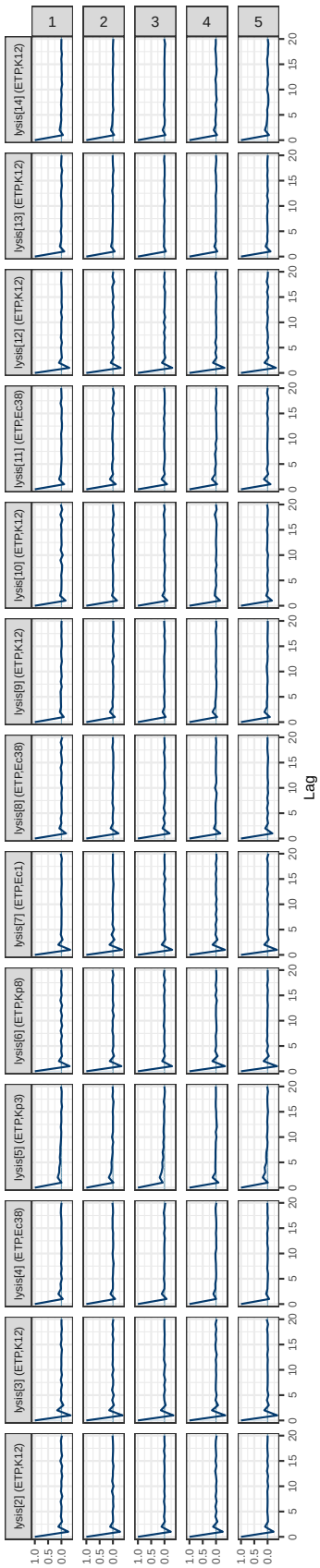


(d) Lysis efficiency observation error

Figure 4.16: Markov chain traces during calculation of susceptibility metrics by HMC for unknown strains (cont'd).

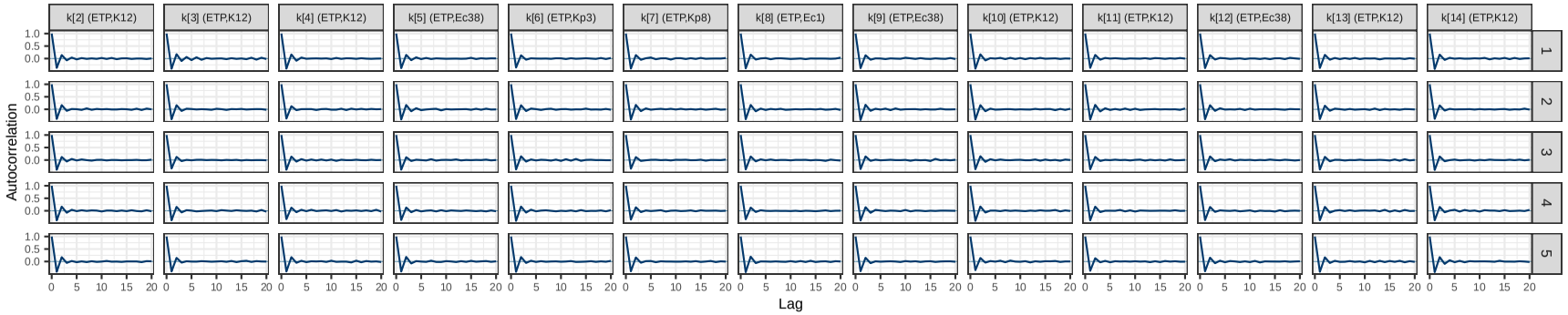


(a) EC_{50} parameters

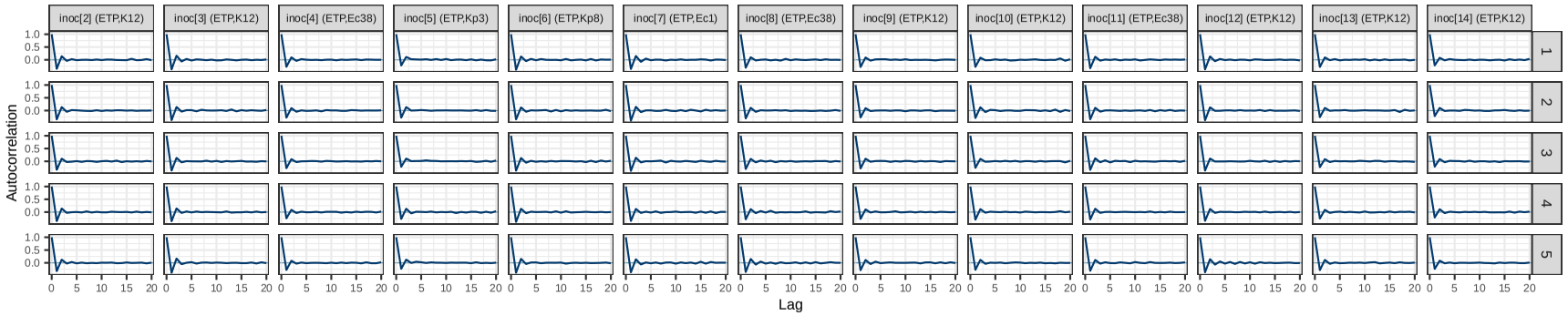


(b) Lysis efficiency (observation error) parameters

Figure 4.17: Autocorrelation of parameters during calculation of susceptibility metrics by HMC for unknown strains.



(c) Background lysis rate parameters



(d) Batch culture inoculum before the serial dilution.

Figure 4.17: Autocorrelation of parameters during calculation of susceptibility metrics by HMC for unknown strains (cont'd)

4.5.3 Results and diagnostic accuracy of digital filtration AST

The results of each susceptibility metric for all digital filtration ASTs performed are shown in Figure 4.18. As a matter of convention, in this section, a positive AST result occurs when a call of "resistant" is made.

There were only three experimental conditions where the phenotype was expected to be resistant (including if no antibiotic was present), yet lysis was seen in a well. All other resistant conditions had the maximum possible resistance signal of no lysis. Therefore, regardless of the metric chosen, one can already conclude that the two experiment batches (figures 4.7a and 4.4d) containing these three experimental conditions are the only ones that cause false negative calls. For the susceptible strains, there were two digital experiment batches (figures 4.8 and 4.4c) that had results lower in magnitude, for at least one metric, than either a resistant condition with the same antibiotic or an untreated condition. For the batch in 4.4c, the only metrics where this happened were the two PE metrics, in which the background lysis seen in 4.7a (part of a batch intended for ertapenem assessment, not ceftriaxone) and 4.4d are eligible for comparison because the PE metrics are condition-specific.

With this information, for each of the metrics of section 4.5.1, one can trace the effect of drawing a threshold to map the susceptibility metrics to a final susceptibility call. The receiver-operating characteristic (ROC) curves in Figure 4.19 are created by varying the threshold from its minimum to its maximum possible values.

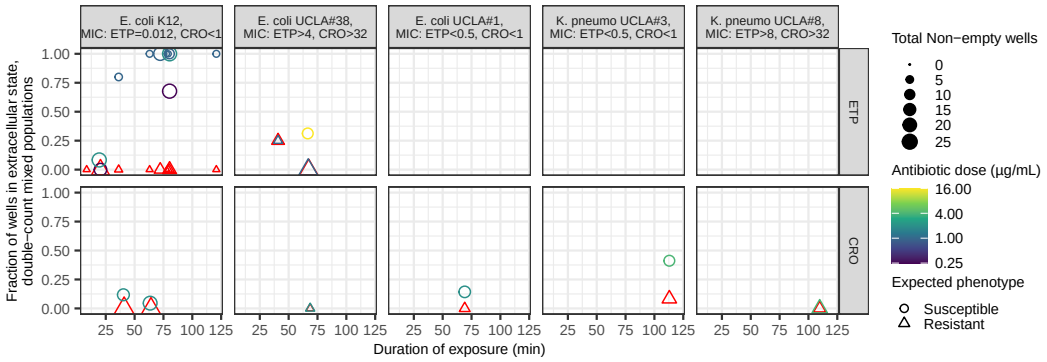
It is important not to compare the area under the curves (AUC) for metrics that are not defined over the same number of observations. The numbers of susceptible and resistant data for ertapenem and ceftriaxone conditions available for each metric are written in the captions of Figure 4.19.

Using the PE_{naive} and PE_{MLE} metrics, there are three errors for the ceftriaxone conditions and seven errors for the ertapenem conditions.

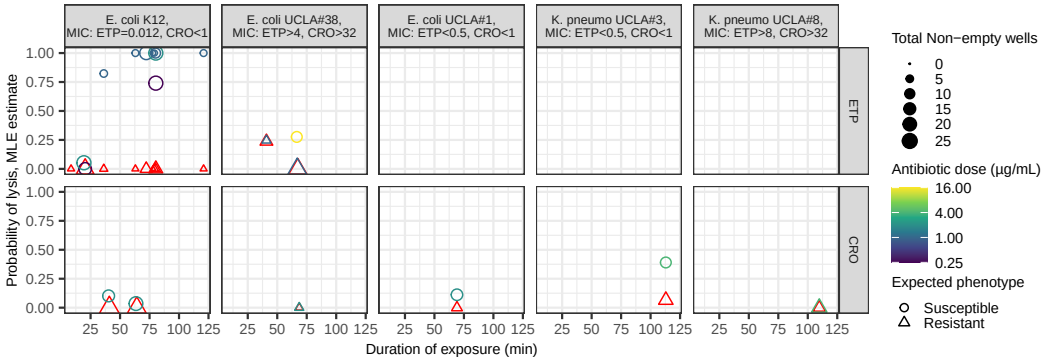
Interesting, the PE_{naive} yields values surprisingly close to the PE_{MLE}. However, for $PE_{\text{naive}} < 0.5$, PE_{naive} is too high, while for $PE_{\text{naive}} > 0.5$, PE_{naive} is too low—an intuitive result.

Using the RDI metric on the PE_{MLE} metric, one sees no errors for the ceftriaxone conditions and 1 error in the ertapenem conditions.

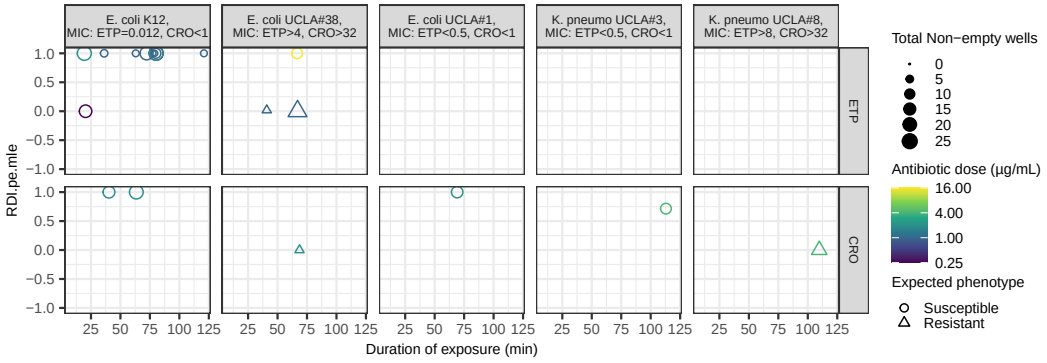
Figure 4.19d shows that there exists a significance threshold for which p_{binomial} only yields one error in the ertapenem condition. Indeed, any threshold between 0.46 and 0.25 would suffice. However, using the traditional a prior threshold of 0.05, there are



(a) Naive proportion of wells extracellular

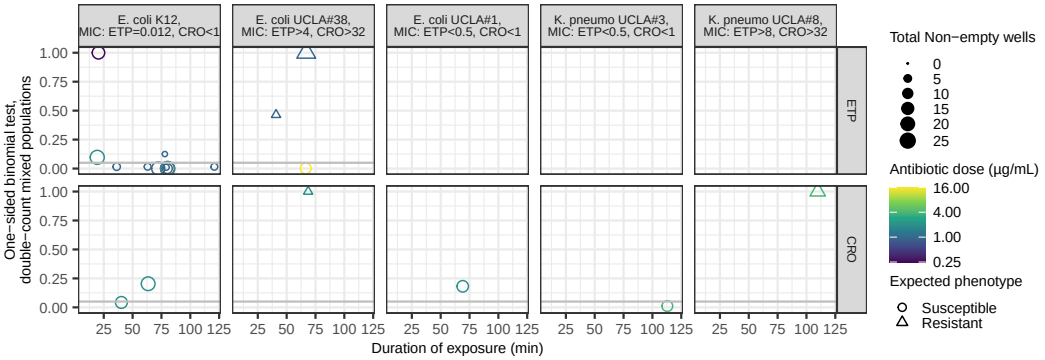


(b) MLE of the death probability

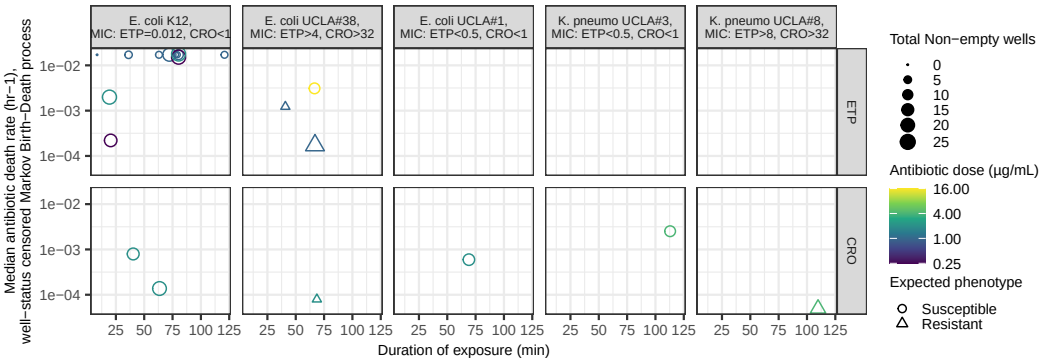


(c) Relative difference index of death probability MLE. $RDI = 0$ implies no change relative to the control.

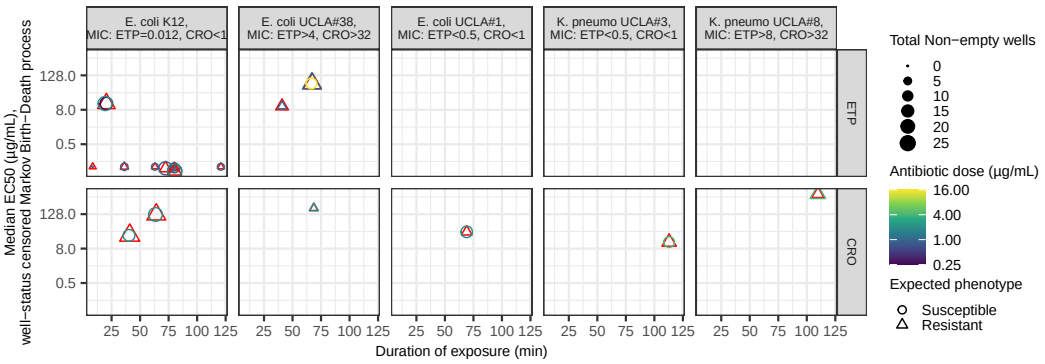
Figure 4.18: Beta-lactam susceptibility measured by digital filtration AST using six susceptibility metrics.



(d) One-sided binomial exact test using the naive proportion of wells extracellular

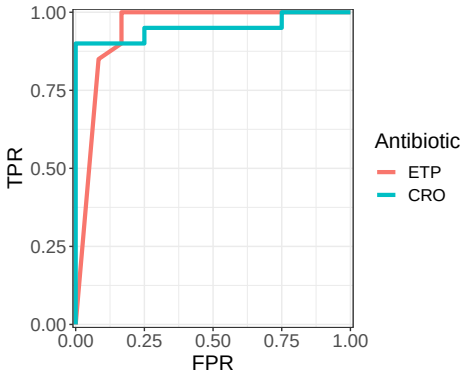


(e) Instantaneous antibiotic kill rate. Note the log-scale y-axis.

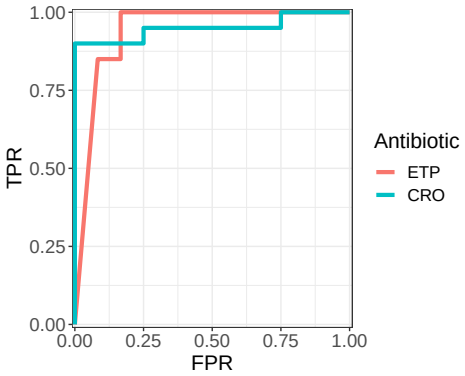


(f) Effective concentration 50. The expected phenotype legend can be ignored. All metrics for a given strain should be similar and should reflect the MIC in the column headers. Note the log-scaled y-axis.

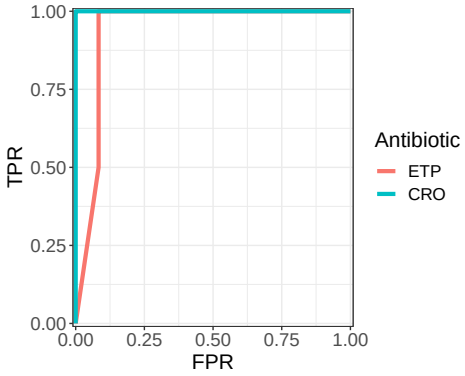
Figure 4.18: Beta-lactam susceptibility measured by digital filtration AST using six susceptibility metrics (cont'd).



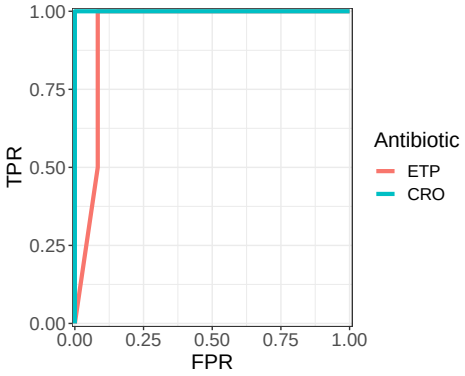
(a) Naive proportion of wells extracellular. ETP: 12 S, 20 R. CRO: 4 S, 20 R.



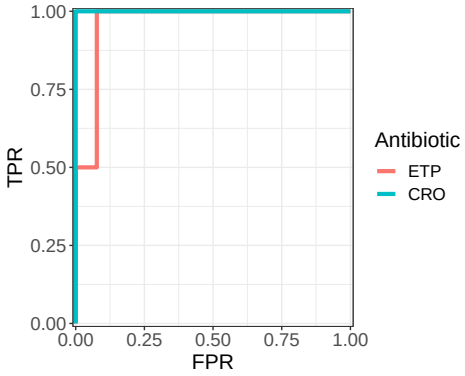
(b) MLE of the death probability. ETP: 12 S, 20 R. CRO: 4 S, 20 R.



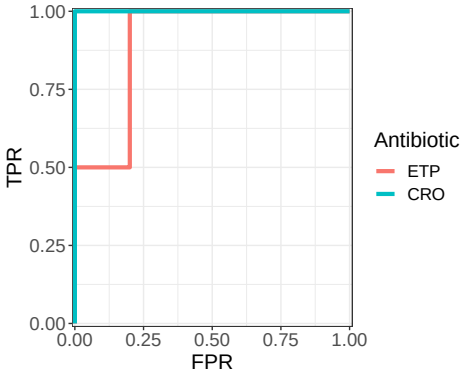
(c) Relative difference index of PE_{MLE}. ETP: 12 S, 2 R. CRO: 4 S, 2 R.



(d) One-sided binomial exact test using PE_{naive}. ETP: 12 S, 2 R. CRO: 4 S, 2 R.



(e) Instantaneous antibiotic kill rate. ETP: 13 S, 2 R. CRO: 4 S, 2 R.



(f) Effective concentration 50. ETP: 5 S, 2 R. CRO: 4 S, 2 R.

Figure 4.19: ROC curves of the metrics used to call digital filtration AST.

five errors. All five errors involve susceptible, treated conditions that failed to reach significance. (Multiple hypothesis correction is not needed because the p-value is being used as a real-valued classification metric, not as an actual probability of wrongly rejecting a null hypothesis about a single statement.)

Examining the β metric, perfect classification is so far possible for the ceftriaxone conditions. Among the ertapenem conditions, there are two errors: a pair of conditions (the 0.25 µg/mL ETP condition in Figure 4.8 and the 1 µg/mL ETP condition in Figure 4.7a) whose β would allow perfect classification if flipped.

Using the EC_{50} metric, it is still the case that one of the two strains contributing to misclassification with the β metric must be misclassified by any threshold. The difference in the EC_{50} has narrowed, however, because the signal in the other treated conditions of those experimental batches pulls the estimates of EC_{50} towards their respective true directions.

There is high categorical agreement between digital filtration AST and the gold-standard susceptibility call, as shown in Table 4.3. Only the EC_{50} had a maximum accuracy less than 90%, and the error was slight (see first two panels from top left in Figure 4.18f. With appropriate thresholds, no false positives occurred where a resistant strain was called susceptible.

For comparison, FDA approval of commercial AST devices requires that the false positive rate of susceptibility classification, also known as a very major errors (VME), be <1.5%. The rate of false negatives, or major errors, should be <3%. The overall categorical agreement should be >90%. (There are additional error categories for when the predicted MIC differs by >2 fold.)

It should be noted that the experimental batch in Figure 4.8, the source of false negatives outside of the PE metrics, was performed with the shortest exposure duration of the comparable experimental conditions. If one were to restrict performance of digital AST to greater than 30 minutes for ertapenem, then 100% agreement is seen.

Similarly, imposing a minimum exposure duration of 75 minutes for ceftriaxone is suggested by the lag in antibiotic activity seen in bulk filtration AST results in Figures 3.6h to 3.6l. This would technically result in 100% agreement for all metrics. However, the uncertainty in that duration threshold is large because only 2 data points are currently available.

To definitively answer whether digital filtration AST is accurate enough for clinical use, one would need to perform a proper clinical trial on one implementation of

Susceptibility metric	ETP	CRO
PE_{naive}	30/32 = 0.9375	22/24 = 0.9167
PE_{MLE}	30/32 = 0.9375	22/24 = 0.9167
RDI (PE_{MLE})	13/14 = 0.9286	6/6 = 1.000
p_{binomial} (PE_{naive})	13/14 = 0.9286	6/6 = 1.000
β	14/15 = 0.9333	6/6 = 1.000
EC_{50}	6/7 = 0.8571	6/6 = 1.000

Table 4.3: Maximum categorical agreement of digital filtration AST applied to unknown strains.

digital filtration AST, as done for FDA approval. However, the optimization of digital filtration AST has not progressed to the point where a clinical trial would be worth the time and resources. Nonetheless, as a proof of principle, digital filtration AST can distinguish susceptibilities from low numbers of cells, and the accuracy improves with longer exposure durations and increased inoculum.

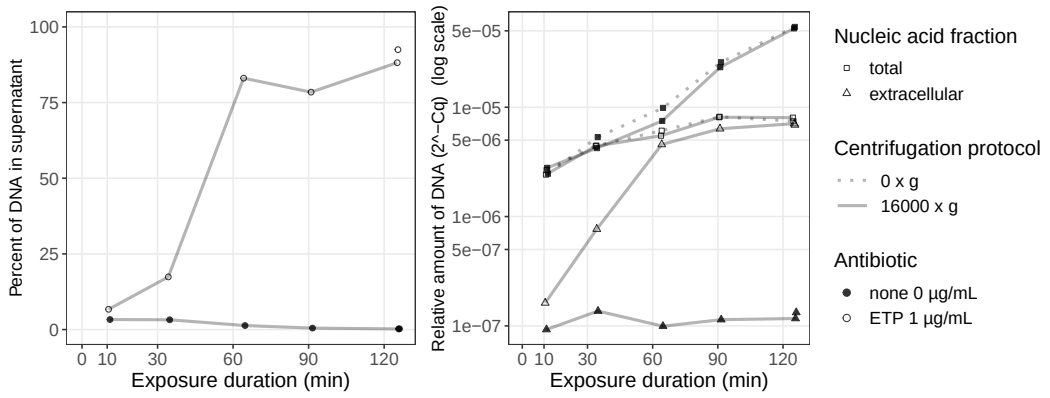
4.6 Future work

One pre-requisite of digital filtration AST is knowledge of which taxon is present, such that compatible NAA primers are selected. In future uses, taxon identification could be obtained separately. Alternatively, the design of the NAA primers can be chosen to yield information about taxon identity. The ability to identify taxa is particularly important to distinguish susceptibilities of polymicrobial infections when isolation during sub-culture is not performed. Theoretically, this could be achieved during digital filtration AST by the choice of primers. Primers could ignore known contaminant taxa (e.g., *Micrococcus*, coryneform bacteria) and only amplify the most prevalent pathogen groups. Multiplexed primers or probes could be used to measure the responses of several target taxa. Targeted taxa could include taxons that are usually contaminants but occasionally pathogens, like coagulase-negative *Staphylococcus* or *Cutibacterium*. In digital filtration AST, the well population status outputs of each primer set can be analyzed independently. Thus, species identification and polymicrobial infections do not pose an inherent limitation of digital filtration AST. Further primer optimization is needed to achieve multiplexed digital filtration AST, nonetheless.

Lysis efficiency is as low as 25% in certain experiment runs. While this lysis efficiency can be estimated in future experiments (if the experiment run contains one susceptible and treated condition and one untreated condition), the variability should be a target of optimization.

Once the digital filtration AST method is optimized, perhaps automated, exploration of more pathogen taxa and antibiotic compounds can be done. There are likely phenomena that interfere with filtration AST in other taxa not seen in the strains and antibiotics so far tested. When more data is available, the susceptibility metrics' performance can be evaluated with separate test and training sets.

Filtration is not the only separation that can be performed to achieve accessibility AST. Centrifugation in particular is a simple method that will separate out both host tissue and intact bacteria from freely-dissolved extracellular nucleic acids. Centrifugation has two benefits over filtration: it is not limited by the buildup of a filter cake, and the separation is easily reversible. Filtering a sample containing diluted whole blood would quickly lead to clogged filters. The filtrate would be easily interpretable, but limiting in quantity. Resuspension and clearing of the filter cake is difficult to perform completely, while resuspension and sedimentation of a centrifuge pellet can be repeated many times. Washing of the pellet to improve separation efficiency is possible as it was in filtration. The drawback of centrifugation is that physical manipulation of the supernatant may dislodge single cells that are sedimented and not anchored to the bottom of the tube. In filtration, the single cells were sometimes lost but never would appear in the filtrate to cause a background signal (which is why filtration was pursued in the first place).



(a) Percent of extracellular DNA

(b) Extracellular DNA

Figure 4.20: Bulk centrifugation AST pharmacodynamics and background signal is equal to that of bulk filtration AST. The gDNA was quantified using qPCR, then converted to relative amount assuming a PCR efficiency of 2.

Figure 4.20 shows a bulk centrifugation AST where the same antibiotic exposure was separated, then re-suspended to continue the exposure as a time sequence. A replicate exposure that did not undergo centrifugation was incubated in parallel.

The difference in feed fraction measurements was not significant. A small amount of fresh media was added to the sequential exposure to keep the total volume the same. Similar population dynamics are seen as measured in Chapter 3.

Chapter 5

THE IDEAL PHENOTYPIC ASSAY

A motivation of this thesis was to investigate the possibility (or impossibility) of performing phenotypic AST on blood clinical specimens without culturing, with respect to the number of bacteria present. It was posited by some that below a certain number of cells, a sufficiently accurate phenotypic AST would be "impossible," in some way. The discussion of impossibilities makes relevant the concept of an "ideal assay." Using the colloquial definition of "ideal," one would conclude that an ideal assay would be an assay that no other assay can perform better than, according to some metric of performance or accuracy.

As one can see, the above question of finding the minimum necessary number of cells contains much ambiguity in its framing, and the definition of an ideal assay remains vague. Unfortunately, the existing literature contains neither a precise definition of an ideal assay nor a relevant discussion of the theoretical limits of assays, including those not yet designed. There are decades of literature from the fields of clinical chemistry, laboratory medicine, analytical chemistry, and metrology on how to empirically quantify or calibrate the performance of a given diagnostic test or instrument, either in the hands of a specific operator or more generally for the average operator in clinical settings. These existing concepts, frameworks, and statistical techniques include the concepts as diagnostic sensitivity and specificity, lower limits of quantitation and analytic sensitivity, and traditional statistics (e.g., confidence intervals, regression, propagation of uncertainty, and hypothesis testing). The above techniques are generally applied to the results of clinical trials, and each such clinical trial only examines a single specific, static, and finalized implementation of a diagnostic assay [59, 117]. I am not aware of published examples of quantitative predictions or optimizations being made for the performance of a class of clinical diagnostic assay through statistical modeling (with or without empirically measured random error). These kinds of forward predictions and theoretical benchmarks would be useful to engineers during the earlier stages of assay design, prior to committing to a costly clinical trial.

This chapter first describes a definition of an ideal phenotypic assay that also serves as a theoretical framework for analyzing assays. Although an attempt has been made

to be precise, the definition will remain outlined mathematically at an informal level. A comprehensive mathematical definition, derived from axiomatic set theory and covering all conceivable phenotypic assays, lies outside the current scope of this thesis.

Next, this chapter will apply the above theoretical framework definition to filtration AST. The results of Chapter 3 are plugged into the model of filtration AST to predict the performance of an ideal filtration AST as a function of the number of bacterial cells initially present in the assay system.

Lastly, this chapter will compare the real results of digital filtration AST shown in Chapter 4 to the predicted ideal results. Because the experimental results and predictions are compatible, I conclude that the predicted ideal results provide a relevant estimate for the minimum number of cells needed for a phenotypic filtration AST, and that the deductive approach of predicting ideal assay performance can provide a useful performance benchmark for engineers. The amount of data is not sufficiently powerful to prove that my implementation of filtration AST has indeed reached its ideal performance, but thus far, the data is compatible with such a conclusion.

5.1 A mathematical framework for describing phenotypic assays

5.1.1 Defining assay performance as accuracy

An ideal phenotypic assay is one with the best performance. More specifically, the invocation of a "performance" implies the existence of one or more metrics of performance, also known as "figures of merit," such as speed, cost, shelf life, safety, and accuracy. The existence of metrics of performance imply that each assay can be located at a point in the space of the metrics of performance. One of the points in the space may possess the highest desirable value for all of the metrics. Let us call it the point of maximum metrics. If that point can be occupied by an assay, then that assay is the ideal assay. But there may also be a region of the space, surrounding and including the point of maximum metrics, that no assay can occupy, except perhaps asymptotically. The boundary of this inaccessible region delineates the limit of what is possible or impossible. This boundary, which is a set of related assays differing likely by some trade-off in its parameters, then constitutes the set of ideal assays.

The defining purpose of an assay is to provide information about a specimen's physical properties. Concerns such as speed or cost, though important in actual

clinical practice, do not affect the scientific consideration of whether a human-made assay is at all possible. Thus, in this thesis, only accuracy (decomposed into diagnostic specificity and sensitivity) is used as a metric of performance.

Accuracy, as used above, refers to the ability to infer the physical property from the assay results, while inaccuracy refers to the total amount of error accumulated during the measurement process. This definition of accuracy is compatible with the definition put forth by the International Organization for Standardization (ISO) [147].

Other definitions of accuracy exist in the literature. In metrology, which is concerned with the measurement of physical properties that have numerical (e.g., real number) values, "accuracy" traditionally refers to the closeness of real-valued measurements to a true value and related to systematic error. Accuracy is contrasted with the "precision," which measures the spread within a set of measurements and is related to random error. In the ISO definition, "precision" retains the definition above. However, a different term, "trueness," refers to closeness to a true value, while "accuracy" instead refers to the combination of trueness and precision. In clinical diagnostics, "accuracy" of binary diagnostic tests refers to the total fraction of calls made correctly. Specifically,

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} = SN \times Prev + SP \times (1 - Prev), \tag{5.1}$$

with ACC, TP, TN, FP, FN, SN, SP, and $Prev$ standing for accuracy, true positives, true negatives, false positives, false negatives, sensitivity, specificity, and prevalence, respectively.

In the existing literature, the ideal performance of an assay has been described in terms of the above definitions. First, in the metrological definition, the phrase "ideal assay" is typically understood to refer to an assay with perfect accuracy and precision; that is, always returning the exact true value of a variable in a non-dynamic system with each repeated measurement, and with systematic and random errors of 0. In clinical diagnostics, the phrase "ideal assay" is typically understood to refer to an assay with perfect sensitivity and specificity (and thus an ACC of 1), with the assay somehow maintaining that perfect performance in every imaginable real-world context in which the assay is performed, such as for all possible prevalences of the positive and negative cases. This is because the most common occurrence of the phrase "ideal diagnostic" arises when discussing a receiver operator characteristic curve that lies along the x and y-axes. Although the above definitions are viable

definitions of an ideal assay, I believe they are not as useful as the more detailed framework I am proposing herein when serving as benchmarks for assay designers. Simply requiring the assay to possess perfect, oracular powers sets a benchmark that is trivial, does not provide much insight into how to improve an assay, and does not enable one to discuss different sources of error.

In this thesis, I adopt the ISO definition of accuracy and state that an assay relays information about unknowable, theoretically-defined system variables (real-valued or categorical) into a format understood by humans. An assay has higher accuracy if it relays information with less ambiguity or uncertainty in the system variables' values.

The Guide to Uncertainty in Measurement 2008 does list, in prose, ten different potential sources of uncertainty in a measurement, but leaves the details to individual scenarios and does not attempt to group these sources of uncertainty into a mathematical framework as I have [148].

5.1.2 The use of ideal models and deductive reasoning

Philosophically, an ideal assay can exist only as an abstract concept. Any empiric, experimental approach that evaluates assays by inductive reasoning cannot not define a performance limit because the possibility will always exist for the assay to be improved in some yet-undiscovered way. Conversely, the inability of anyone to achieve a better assay accuracy in practice does not prove that a better assay will not be attained in the future.

In contrast, deductive reasoning does offer a framework to discuss ideals. In deductive reasoning, one must assume a set of axioms and assumptions about the universe or system in which the relevant assays operate. The assumptions and axioms (for what is in the system and how they behave) constitute a model of the assay. Deductive reasoning leads one to conclusion of what is possible or impossible given the model. An assay can then be called ideal among the conceivable assays that satisfy the given assumptions. The model of the assay defines the scope of one's inquiry, so the modeler should make the minimally necessary assumptions to be useful. The real world may often be more complex than our assumed system. If experimental results deemed impossible by one's model are subsequently demonstrated, it will be because one of the assumptions of the model has been transcended.

For comparison, there are many examples of theoretical limits in the physical sciences or engineering, such as the diffraction limit of microscopy, the Carnot effi-

ciency for heat engines, and the time complexity of integer factorization by conventional computers. All are defined (or conjectured, for the last example) by deduction from assumed axioms (including values of physical constants), rather than by empirical experiments. A given physical implementation of a microscopy, engine, or computer may not reach the theoretical limit due to suboptimal design or additional limitations of implementation not modeled. For example, most man-made combustion engines—which unlike the Carnot engine need to provide high power—do not approach their Carnot efficiencies. On the other hand, super-resolution microscopy and quantum computers have or may surpass their above respective limits. They do so by circumventing assumptions defining the operation of a traditional microscope or a computer.

Conditioning on assumptions makes it necessary to discuss ideal assays only for a given type of assay. For example, it is hard to define what is a universally ideal assay for every imaginable real-world scenario, but it is possible to define an ideal accessibility AST assay, an ideal digital filtration AST assay, an ideal live-imaging AST assay, or an ideal bioassay for pharmacologic potency. Each ideal type of assay is defined by the assumptions comprising one's definition of that type of assay. As an analogous example, engineers often compare the important classes of combustion engine designs so far invented (e.g., spark-ignition, diesel, and gas turbine) to corresponding models of their ideal thermodynamic cycles (e.g., the Otto, Diesel, and Brayton cycles).

5.1.3 What is a phenotypic assay?

In this section, I will propose a unified, general definition for a phenotypic assay. This definition naturally leads to a mathematical framework introduced in the next section. It was my hope that this general definition and mathematical framework would offer some clarity in the classification and invention of new diagnostics assays.

An "assay" is a procedure performed on a particular specimen that yields information about the quantity, or presence, of some property or quality of that specimen. Meanwhile, the dictionary definition of a "phenotype" is "the sum total of the observable characteristics of an individual," the individual usually being a biological organism whose observable traits are being contrasted with an underlying, causal genotype [149].

In contrast with the usage of "phenotype" in genetics, the term "phenotypic assay" does not yet have an established, unified definition in the scientific literature, even

though the term is widely used and generally understood by those in the life sciences. A quick search of any publication database will show that the phrase "phenotypic assay" or "phenotypic test" has been applied to several kinds of assays. The three most established usages of the phrase are the phenotypic ASTs, phenotypic screenings for drug discovery, and tests for human genetic disorders. In each of these three cases, the term "phenotypic assay" was coined to contrast with an alternative, existing "non-phenotypic" assay.

The non-phenotypic alternative to phenotypic AST is a genotypic AST, where the possession of a resistance gene in an organism's genome is used to conclude that a microorganism is resistant to an antimicrobial compound. (See Chapter 1.)

In the field of drug discovery, a phenotypic screening examines a large library of small molecules or natural products for a desired physiologic effect on cultured cells or whole organisms [150–152]. The non-phenotypic alternative to the phenotypic screening is a target-based screening [153]. A target-based screening measures the effect of the candidate chemicals on (often purified) components of a specific molecular pathway (e.g., binding affinity to an protein, affect on enzyme activity, or even translation of a reporter construct) known to be relevant to the desired physiologic effect.

Lastly, a phenotypic test for a genetic disorder generally involves measuring a deficiency in an enzyme's activity in a clinical specimen (e.g., circulating blood cells or liver biopsy) or a human individual (e.g., the sweat test for cystic fibrosis [154]). The non-phenotypic alternative for diagnosing a genetic disorder would be the sequencing of the presumed causal gene—from any somatic genome-containing clinical sample—to identify a known or predicted disease-causing allele.

Some careful deliberation reveals at least five distinct conceptual themes running through the above three usages of the term "phenotypic." Not all themes are shared by all cases, however. Only the themes that are shared in all cases should be part of the unified definition of "phenotypic," unless there is also a compelling distinction shared by the cases without the theme that can justify their being excluded.

5.1.3.1 The phenotype as genotype expression

First, notice that the original meaning of the word "phenotype" from the field of genetics is present in the phenotypic AST and phenotypic tests for genetic disorders. In this sense, the phenotype is the word describing all traits of an organism that

arise from the expression of genetic information—that is, via transcription and translation of genes, and subsequent interaction of the gene products. It would be tempting to define phenotypic assays using the Central Dogma of molecular biology and using the contrast with sequencing-based diagnostic assays. However, this first attempt at a definition would exclude the phenotypic screenings for drug discovery because genetic variation is not the cause of the phenotype of interest measured in a pharmacological assay, and because target-based screening does not use sequencing. Yet, there are still many aspects of drug discovery assays that seem "phenotypic," like their being agnostic to mechanism or their focus on function, and a better definition would capture these aspects.

5.1.3.2 The phenotype as the properties of biological life

A second theme, broadening the first, is that the above examples all include living organisms as potential specimens: bacteria in AST, animals in drug screening, and humans in certain diagnostics tests. A second attempt at defining a phenotypic assay would thus require a phenotypic assay to be performed on a living organism. This is the definition offered in Swinney 2013 [150]. The definition of biological life most widely used recognizes several criteria that life fulfills, namely 1) maintenance of homeostasis through the metabolic extraction of environmental energy, 2) responsiveness to the environment, and 3) reproduction, with any heritable traits subject to evolution. Additionally, biological life achieves these criteria by being highly organized and complex, being organized as cells with a plasma membrane, being comprised of a small finite number of macromolecular polymer classes (traditionally the carbohydrates, lipids, proteins, nucleic acids, and various secondary metabolites), and using only nucleic acids as their genetic material.

However, specimens that are not living organisms can still be analyzed by a phenotypic test. For example, certain phenotypic tests for genetic disorders are performed on extracts derived from biological specimens. As specific examples, glucose-6-phosphate dehydrogenase (G6PD) deficiency is diagnosed by enzyme activity in red blood cell lysates, and routine newborn screening for lysosomal storage diseases assesses enzyme activity in dried blood spots. The extract specimens were once living but now are non-living[1]. Yet these tests have been rightfully deemed more phenotypic than the contrasting genotypic test, at least in the way they are agnostic to mechanism.

[1] Should the definition of biological life be stretched to include extracts as living? In general we

Furthermore, using biological life to define a phenotype conflates all of the properties of life with the concept of the phenotype, but not all properties of life are necessary to define the concept. For the above examples, I believe the only[2] relevant part of their specimens' biological nature is biological life's responsiveness to environmental perturbations.

5.1.3.3 The phenotype as system behavior

From the above discussion, we arrive at another attempt at a definition. All of the above phenotypic assays involve measuring some sort of response to a perturbation or behavior[3]. In AST, the response is death from antibiotics. In genetic testing, the response can be the catalysis of an added substrate. In drug discovery, the perturbation is the addition of the candidate compound, and the response is any of the diverse read out technologies employed so far.

We can further refine what we mean by the terms response and behavior by renaming them as a transformation. There must be a state the system is in. If after a perturbation, the system remains in the same state, it has not responded. Taking the contrapositive, if the system does respond, it must be in a new state. Thus, the system must occupy at least two states, one before and one after the perturbation. Occupying a sequence of different states is a transformation. Furthermore, because there is an order to the states, one being before the transformation and one being after, the transformation can be said to be dynamic, or changing over some notion of time.

Intuitively, we would like our definition of a phenotypic assay to exclude all assays that only measure the quantity (or presence) of an analyte at one point in time

consider cellular extracts to be non-living because they do not maintain homeostasis or reproduce without interventions that are indistinguishable from attempting to resurrect the original cells, which would be considered living.

[2] The maintenance of homeostasis is generally not required; as an example, bacteria that die during antibiotic exposure in a phenotypic AST have failed to maintain homeostasis by definition. Neither is population growth required for many phenotypic assays. The concept of being more complex or more organized are subjective terms that would require definitions of their own. Meanwhile, the particular structure and composition of extant biological life may turn out to be a coincidence rather than fundamental physical law—a hypothesis testable by bio-orthogonal synthetic polymers or by any discovery of alien life.

[3] The Merriam-Webster Dictionary similarly defines behavior either as "anything that an organism does involving action and response to stimulation" or "the way in which something functions or operates." Here, I use "behavior" interchangeably with any response to the environment. Similar terms include "function" and "physiology." The Oxford English Dictionary defines "physiology" as "the functional processes of an organism, organ, or system."

because these assays do not measure a non-trivial "behavior." In such assays, the specimen is assumed to be stable over time and structurally homogeneous. Clinical examples would include the majority of assays measuring serum ions, metabolites, or proteins, often through immunoassays or other photometric assays using dyes. Non-clinical examples would include all of the elementary measurements in analytical chemistry or physics, such as gravimetric mass balances, volumetric pipettes, photometry and spectroscopy, and measuring physical dimensions with calipers. The majority of the examples of measurements given in the GUM, and in the literature on analytical chemistry and clinical chemistry, focus on assays making these simplifying assumptions.

However, we must be careful in which transformations we consider to be a behavior. For example, performing a titration to measure the amount of acid in a solution is not a phenotypic assay, but the specimen of acid is transformed by the addition of known amounts of base, and the indicator dye must transform in color. As another example, consider the enzyme-linked immunosorbent assays or the lateral flow assay for urine human chorionic gonadotropin (hCG), a well-known clinical diagnostic test for pregnancy. These antibody-based tests transform the specimen by binding the target protein to antibodies, yet we do not generally consider the bound analyte to have exhibited an interesting behavior. The example of protein binding is particularly relevant because it is sometimes the modality used in target-based drug screenings.

In these assays lacking specimen transformation, the lack of dynamics in the model does not impact assay accuracy, other than any specimen degradation from improper storage or excessive delays. In contrast, phenotypic assays require, or at least benefit from, acknowledging and modeling the dynamical changes in the assayed specimen during the assay.

5.1.3.4 The phenotype as a dynamic transformation

As a refinement of our definition of a phenotypic assay, let us only include those assays whose specimen's transformation is influenced by the specimen system's internal dynamics, and where these internal dynamics are captured in the measurement. A volumetric titration for acid does not measure the kinetics of the reaction between the base and the acid, only the endpoint equilibrium concentrations. For that (and other reasons discussed later), we tend not to call the acid-base titration a phenotypic assay. Meanwhile, a phenotypic broth microdilution AST usually does

not involve an explicit time series measurement, but the fact that all wells begin without visible turbidity acts as a de facto time series measurement. Thus, the internal dynamics of the system, the population growth of the bacteria, is observed by the assay.

Unfortunately our definition is too broad, even with this requirement for an observed transformation. As a specific example, note that some target-based drug screens involve measuring the activity of an enzyme over time in response to the perturbation of substrate addition. Since the system is transformed, and the dynamics of the transformation are observed, this assay would be phenotypic. Thus, the existence of a behavior or observed dynamic transformation should be part of an eventual definition of a phenotypic assay, but not the sole criterion.

As a even more problematic example, some target-based drug screens include live cells in their specimens. These cells may not be the correct target tissue but have incorporated a recombinant genetic construct allowing the activity of the target molecule to be quantified while the cells is alive (e.g., fluorescent protein expression) or after the cells are killed (e.g., luciferase activity). The activity of the target molecule is compared to pre-perturbation baseline levels (and/or to negative controls). This use of cells fulfills all the criteria we have examined thus far: possessing a phenotype caused by gene expression, being biologically alive, and responding to a perturbation in a way that is measured. However, there is something less phenotypic than a phenotypic drug screen on a whole animal. In a cell-based, target-based screen, the living cells are not the system of interest, but are part of the readout method. Furthermore, the mechanism of the phenotype is known, because the underlying genetic construct was designed by humans to be orthogonal to any preexisting natural process in the living cells. Two additional themes are suggested by this example and some previous examples.

5.1.3.5 The phenotype as an empirical measurement

A fourth theme of phenotypic assays is their empirical design. In this thesis, we use "empirical" to describe a measurement that is made in a manner agnostic to any theory or model, apart from models defining the property itself, about the specimen's inner workings, nature, or behavior, especially where the inner workings are expected to differ between specimens[4]. The definition of phenotypic screens

[4]The Oxford English Dictionary defines "empirical" as describing something "that pursues knowledge by means of direct observation, investigation, or experiment (as distinct from deductive

for drug discovery given by Moffat et al. 2017 also hinges on being agnostic to mechanism [151].

The alternative to an empirical measurement of a phenotype would be to deductively infer the phenotype via a mechanistic model after making measurements of other properties that are not the objective of the query and therefore not part the phenotype, even if those other properties influence or cause the phenotype.

When an assay is designed to be empirical, it is implied that the designer's available mechanistic models are expected beforehand not to adequately predict the specimen's behavior. The anticipated surprises in one's specimen's behaviors arise either from an incomplete or incorrect model of the system, from the existence of multiple plausible distinct mechanisms for the system, or both.

For systems where there is only one true mechanism in existence, more knowledge about the system will eventually and asymptotically reduce the probability of unknown phenomena toward zero. Very few biological specimens have completely known ontologies—that is, knowledge about all components, their properties, and their interactions. Even though "multiomics" studies and systems biology models have become popular, it is still common practice in scientific research to always assume that unexpected phenomena are present in the system one studies. In other words, in the life sciences, the remaining hyperbolic doubt about the completeness of the system's ontology is treated as non-zero by convention, requiring one to control for any such factor when performing assays on biological systems. As of 2023, no living organism I am aware of has been fully explained to the point where the community believes its ontology has been completed. This explains why it is intuitive to associate phenotypic tests with those performed on living specimens.

For some systems, it is further known (or just assumed) that there exists more than one true mechanisms that could be present in the set of relevant specimens and which can produce the phenotype in question, in addition to the possibility of yet unknown mechanisms. Further studies of this class of specimens will not reduce to zero the uncertainty of which mechanism will be present in a given specimen.

---

reasoning, abstract theorizing, or speculation); that relates to or derives from this method of pursuing knowledge. Later also: relating to or espousing empiricism as a methodology" [155].

The use of "empiric" is synonymous with the later form "empirical." The meaning of "empiric" in empiric antibiotic therapy, used in Chapter 1, is a separate, more specific meaning: "of medical practice or a medical treatment: based on experience of the outcome of previous cases; based on clinical judgement or diagnosis; (in later use) not dependent upon the results of laboratory investigations or formal clinical trials" [155].

In all three examples of AST, drug discovery, and genetic testing, more than one true mechanism was anticipated. In AST, the same species of bacterium can achieve beta-lactam resistance through acquisition of beta-lactamases, acquisition of efflux pumps, mutations in porins, or mutations in their penicillin-binding proteins. A phenotypic AST would detect resistance regardless of which genes and which alleles are responsible for the resistance.

Similarly, the range of what we consider a single genetic syndrome can be caused by a diversity of mutations that cause different levels of severity. The mutations may be in different genes whose products acting at different steps in the same biochemical pathway. There may also be alleles in other pathways that interact with the main causal pathway to modulate the overall phenotype. A phenotypic test, instead of genotyping, may be required in clinical practice in order to make a firm diagnosis.

An additional example of an assay that is not empirical is the lateral-flow hCG pregnancy test. This test can give false positives due to ovarian or testicular tumors and molar pregnancies instead of true, viable pregnancies, indicating that it is not a phenotypic test[5].

The agnostic nature of the empirical assay, considered in isolation, increases the accuracy of an empirical assay in detecting phenotypes. If only one mechanism is being detected by a non-empirical assay, but there exists a diversity of plausible internal mechanisms that can produce the phenotype desired, then some specimens possessing the phenotype will be missed by the non-empirical assay. In contrast, the empirical assay would have a non-zero chance at detecting all specimens with the phenotype. However, whether an empirical assay is overall more accurate than a corresponding non-empirical assay depends on other reasons for inaccuracy. For example, if an empirical measurement incurs higher instrumental measurement noise than a non-empirical measurement, and the non-empirical measurement detects a mechanism that is highly predominant, the non-empirical measurement may be more accurate.

[5]There is no phenotypic test for pregnancy. The operational/definitive test for pregnancy is the delivery of a living child. An alternative test is a prenatal ultrasound. This test fulfills the criteria of relevance and empirical measurement because it directly addresses the question of whether a pregnancy is present, and because the principle of ultrasound is agnostic to the different variations of fetal implantation and development that still constitute a pregnancy. Whether the third criterion is met is debatable. There is an inherent use of time, in that the specimen—the mother and fetus—is required to transform—by the fetus growing in size—for at least 7 weeks so that the ultrasound interpretation is unambiguous. However, if one narrows the scope of what is considered the ultrasound test to only the operation of the ultrasound machine, then, like most non-functional imaging studies, the ultrasound is not phenotypic.

There is a special case in the above comparison of empirical and non-empirical assay accuracy. If the phenotype is defined operationally[6] by the assay, then by definition, the chance of detecting all specimens becomes one. One can then conclude generally that the empirical phenotypic assay is always more accurate than a non-empirical assay. For example, the MIC, breakpoint MICs, and susceptibility phenotypes of bacteria are defined operationally by the gold standard phenotypic ASTs, so these phenotypic ASTs are by definition always correct, even though the shortcomings of *in vitro* phenotypic ASTs has been debated for decades [156, 157].

In summary, I define a phenotypic assay as one fulfilling the following criteria.

- *Relevant:* The quantity (or nominal) value [159] of the property of a specimen measured by the assay exactly answers the question asked by the assay's user. The measured property is not a proxy or a marker for another property of more fundamental interest to the user.
- *Empirical:* The property can be measured empirically—that is, agnostic to any model about the specimen's behavior.
- *Dynamic:* The specimen undergoes a dynamic transformation, such as by changing over time or by responding to a perturbation. Though not required by my definition, most specimens considered to possess a phenotype undergo a transformation that is complex in some way. The transformation may even vary stochastically.

Clinical examples of phenotypic assays thus defined, beyond those discussed above, include the interferon-gamma release assay for tuberculosis exposure and recently reported assays that detect sepsis by neutrophil mobility [160]. Diagnostic procedures performed on intact human individuals to measure their metabolic activity or other biological phenomena, like functional imaging studies (e.g., PET and fMRI), drug pharmacokinetic studies, and hormone challenges (e.g., oral glucose tolerance and cortisol tests), also count as phenotypic assays, even though we do not usually call them assays. More examples of phenotypic assays can be found in research

[6] A physical property is defined operationally when its definition is based on a standard physical operation. The idea comes from a school of thought called operationalism, "a form of positivism which defines scientific concepts in terms of the operations used to determine or prove them" [158]. An operationalist would not define temperature as the theoretical average kinetic energy of atoms in a specimen. Instead, he or she would define temperature by the mechanisms of different types of thermometers that could be constructed.

settings, sometimes under the term "bioassay." These include the LD_{50} assays in toxicology, EC_{50} assays in pharmacology, assays for metabolic activity, and other laboratory experiments one can perform on cultured tissues or organisms.

Interestingly, our definition of a phenotypic test now extends to entities that are not biological, like man-made machines. Our discussion may have been motivated by the life sciences, but it is now generally accepted that no fundamental physical property, other than the particular organization assumed by biological life, separates biological life and organic matter from dead, non-living, or inorganic substances[7]. Therefore, activities such as measuring the real-life performance of a computer's processor, the charge capacity of a lithium ion battery over its lifetime, or even measuring the Young's modulus of a bar of metal alloy can all be considered phenotypic assays.

5.1.4 All phenotypic assays can be modeled as four mappings

Modeling constitutes an inherent aspect of understanding a natural phenomenon. Some level of mathematical modeling also anchors any reliable analysis of an assay's results in real-world practice. It is thus valuable to further define the phenotypic assay in terms of mathematical entities beyond the semantics of the previous section.

The definition of a phenotypic assay suggests that any conceptual model of a phenotypic assay should comprise at least four mappings: a transformation mapping F between system variables; a measurement mapping G from system variables (and measurement parameters) to measurement outcomes, a input error mapping J from system inputs (a subset of the system variables) to measured inputs, and a classifier mapping H from measurement outcomes to assay conclusions. These mappings will be described in the following sections.

Figure 5.1 depicts this framework of four mappings applied to two models used in Chapters 3 and 4. Table 5.1 shows how all models used in this thesis can be organized according to this framework.

[7]The alternative theory of vitalism, which stated that living organisms and organic matter possessed a special quality, would require that separate physical laws govern organic matter. Yet in all cases examined, the existing laws for inorganic matter have been found to be sufficient. Thus, the acceptance of only one set of physical laws is demanded by the principle of parsimony. The observation that life only arises from previous life is merely a temporary cosmic circumstance that synthetic biology may one day refute.

Table 5.1: Taxonomy of models for phenotypic filtration AST.

Model name	Section	System variables			Measurement & input error variables				Classifier
		System state variables	System inputs	System parameters	Input error parameters	Measured inputs	Measurement parameters	Measurement outputs	Classifier metric; classification parameters
empirical treated (T) vs. reference (R)	2	L, D	T, L_0, A	μ_T, μ_R	none	NA[a]	m, b, σ implied.	Y per pair of conditions (i.e., $Y_{L,T}, Y_{L,R}$, $Y_{D,T}, Y_{D,R}$).	TTPD_{CT} or TTPD_{LT} or p_{naive}[c]; ξ[b]
empirical constant net growth model	3.2.5	L or T, D	t_1, t_i, L_0	μ^*	none	NA[a]	m, b, and σ implied.	Y per pair of conditions (pretending no replicates in actual data)	μ^*; $\xi = 0$
continuous ODE model, hierarchical error	3.3	L, D	t, A, L_0	μ, k, α, $\beta_{\max}$, EC_{50}, γ	none, error in L_0 subsumed by m_e	NA[a]	c, m_e, b_e, σ_{ϵ_p}, m_p^*, $\sigma_{m_p}^*$, b_p, σ_{b_p}	Y per experiment batch	β[c] or EC_{50}[c] or p[c]; ξ[b]
discrete, no births, no digitally-loaded batches, ideal assay model	5.2	L, D	T, A, L_0	k, α, $\beta_{\max}$, $\text{EC}_{50}[c]$, γ, c	none	exact by definition	none	L, D	L; λ, w_{FN}, w_{FN}, $\Pr(c)$
discrete, no births, digitally-loaded batches. See 4.4.1, 4.4.3.1	4.5.1.1, 5.3	L, D per well	T, A	L_0 per well, C, k, α, $\beta_{\max}$, EC_{50}, γ	none	NA[a, d]	ddPCR positive droplet thresholds, well population status copy number or C_q thresholds.	$[O, I, B, E]$ per condition	PE_{naive} or PE_{MLE}; $\xi_{PE_{\text{MLE}}}$
	4.5.1.2, 4.5.1.3				none	[a]		$[O, I, B, E]$ per pair of conditions	RDI or p-value; ξ[b]
					prior for C	C_{target}. T, A assumed exact.		$[O, I, B, E]$ per experiment batch	β[c] or EC_{50}[c] or p[c]; ξ[b]
discrete, Markov birth-death, digitally-loaded batches	4.4.2, 4.4.3.2			L_0 per well, C, μ, k, α, $\beta_{\max}$, EC_{50}, γ	$\bar{I}$, σ_I	I_{target}		$[O, I, B, E]$ per experiment batch	β[c] or EC_{50}[c] or p[c]; ξ[b]
discrete, Markov birth-death, digitally-loaded batches, lysis inefficiency	4.4.2, 4.4.3.3, 4.5.1.4				$\bar{I}$, σ_I	I_{target}	above thresholds + l	$[O, I, B, E]$ per experiment batch	β or EC_{50} or p[c]; ξ[b]

[1] all inputs assumed exact. [2] threshold not given a specific name in the text, so here it is given the symbol ξ. [3] not calculated or discussed in the thesis.
[4] C_{target} was controlled, but it is not a variable, and it does not affect the non-Bayesian analysis.

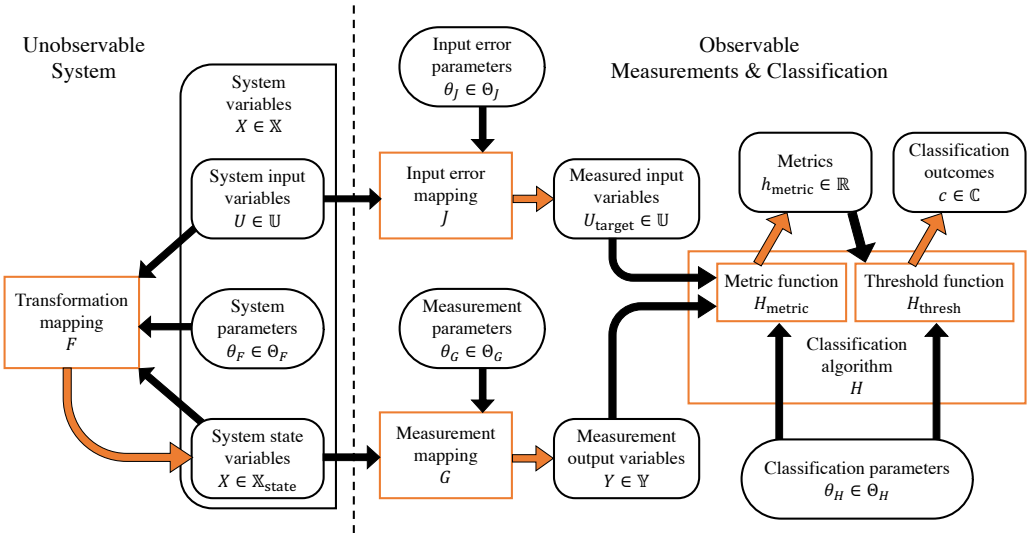


Figure 5.1: Phenotypic assays can be modeled as four mappings.

5.1.4.1 The transformation mapping F: definition

The transformation mapping $F : \mathbb{T}\times\mathbb{X} \to \mathbb{X}$ takes as input a time variable $t \in \mathbb{T}$ and a set $\mathcal{S}$, with size $|\mathcal{S}|$, of system variables $X_1 \in \mathbb{X}_1, X_2 \in \mathbb{X}_2, \ldots, X_i \in \mathbb{X}_i, \ldots, X_{|\mathcal{S}|} \in \mathbb{X}_{|\mathcal{S}|}$. The set $\mathbb{X} = \prod_{i=1}^{|\mathcal{S}|} X_i = \mathbb{X}_1 \times \mathbb{X}_2 \times \cdots \times \mathbb{X}_{|\mathcal{S}|}$ is called the system space, and it is the collection (here specified as an ordered tuple) of the sets of all the possible values of the system variables X_i. The output of the mapping is a new set of values of each X_i in $\mathbb{X}$.

At this general level, the variables X_i may each be discrete, continuous, categorical. Likewise, the time variable $t \in \mathbb{T}$ can be discrete ($\mathbb{T} = \mathbb{Z}^+$) or continuous ($\mathbb{T} = \mathbb{R}^+$)—whichever is appropriate for the phenotypic assay being discussed.

The system variables can be grouped into categories—namely, the system input variables, the system parameters, and the system state variables—based on their role in the model.

Some system variables have values not determined by the system state but rather by an external force, the experimenter. Their value may (or may not) vary over time. We will call the set U of these variables the system input variables, or system inputs. The Cartesian product of the domains $\mathbb{X}_i$ of the variables in U is $\mathbb{U}$, and a vector in $\mathbb{U}$ will be denoted u. If all the variables in U are real-valued, then we can simply say $u \in \mathbb{R}^{|U|}$.

Some system variables have constant values that do not change over time, and are not set by an experimenter. We will call the set of these system variables the system

parameters θ_F. The Cartesian product of the domains $\mathbb{X}_i$ of the variables in θ_F is named Θ_F. If all the variables in θ_F are real-valued, then we can say $\Theta_F = \mathbb{R}^{|\theta_F|}$.

Some system variables describe an aspect of the system that changes over time, but are not input variables because they are not controlled externally. Rather, they only change in response to other system variables, including the system input variables. As the system changes over time, different combinations of these changing system variable values are assumed. We call these combinations of values a state. More specifically, a state is one combination of values of a minimum set of system variables required to predict future states by summarizing sufficient information about the past history of the system and its inputs. The subset $\mathcal{S}_{\text{state}}$ of the system variables in $\mathcal{S}$ whose values form the system states are the system state variables[8]. The subset $\mathbb{X}_{\text{state}} = \prod_{X_i \in \mathcal{S}_{\text{state}}} X_i$ of the system space $\mathbb{X}$ is called the system state space.

In phenotypic AST, relevant system state variables would include the number of types of cells (e.g., live, dead, arrested) or molecules; and the attributes of each cell, like age, size, ancestry, expression of genes, and spatial location in the specimen. Or, as another example, the state variables could just be the total biomass.

The categories of variables θ_F, U, and $\mathcal{S}_{\text{state}}$ are now three disjoint subsets of $\mathcal{S}$. In the literature on dynamic systems, these three categories are usually defined separately. I have given a name to them altogether as "system variables" in order to emphasize how an entity in one category (e.g., system parameter) in one system may move to a different category (e.g., system input variable or system state variable) in another system. For example, if antibiotics were added during the antibiotic exposure by the experimenter, then the added antibiotic concentration would become a system input variable. If the bacteria create enzymes to degrade antibiotics, then the total antibiotic concentration would need to be a system state variable.

In my definition, F is the update rule of a dynamic system[9] tracing a trajectory through the state space $\mathbb{X}$ because of the presence of the time variable. A dy-

[8] Any remaining system variables would be those that change over time, are not input variables, but are not necessary to predict future states. We could call these variables "alternative system state variables," and we could swap them with some system state variables to create a reparameterized model. Or, we could just ignore them and not include them in our model.

[9] Specifically, F as we have defined implies a Markovian dynamic system, as the next state X_{t_0} does not depend on *all* previous states, only the previous state. In practice, I believe the Markov assumption will not be noticeably restrictive because non-Markovian dependence on a *finite* number of multiple previous states $X_{t_{-1}}, X_{t_{-2}}, \ldots$ can be redefined as Markovian dependence on a single previous state $X'_{t=-1}$ that is a higher dimensional combination of the $X_{t_{-1}}, X_{t_{-2}}, \ldots$. If the system state is non-Markovian, with its evolution depending on all past states, further generalization of this definition can be done. In this thesis, I am only covering the high-level framework.

namic mapping is useful for modeling the transformation of a phenotypic assay because most biological specimens exist in space-time, and so their behaviors and transformations take place over time, not instantaneously.

F may be deterministic, where each system state maps to one subsequent system state, or F may be stochastic, where the system state maps to one of several possible new system states according to a probability distribution. Whether one uses ordinary differential equations or stochastic differential equations/processes to model the evolution of the system state is up to the assay in consideration and may require additional restrictions on the nature of X_i or F (e.g., measurability) to be rigorously defined. A vast literature exists detailing how to model dynamical systems across many academic fields, and it should be consulted for determining the appropriate form of F.

The most familiar dynamic systems are those involving ordinary differential equations. In these cases, the mapping F is a smooth flow, with $X(t) = F(X(0))$. Differentiating the flow with respect to time, and separating $\mathcal{S}$ into θ_F, U, and $\mathcal{S}_{\text{state}}$, yields an ordinary differential equation $\frac{dX}{dt} = F'(t, x, u, \theta_F)$. When F does not change over time, then we can write F as the familiar time-invariant state space model $\frac{dX}{dt} = F'(x, u, \theta_F)$.

Another useful dynamic system are those modeled as discrete time difference equations. In these cases, the mapping F can be the transition to the next state, or F can be the overall transition after several time steps. The latter case essentially becomes an ordinary differential equation as the time step tends toward 0.

The takeaway of this rather general overview of systems modeling is that it will be good practice for anyone studying diagnostics to model any dynamic behaviors exhibited by the specimens of their assay.

5.1.4.2 The input error mapping J: definition

It is common to use information about the system inputs U in calculations during the analysis of the outcomes Y (defined in the next section 5.1.4.3) of a measurement. In the literature, U is typically assumed to take on the exact values intended by the assay user. However, a more general approach is to assume some amount of error in the control of the inputs to the assay. In keeping with the philosophy of section 5.1.4.4, and with the convention in statistics literature[10], we assume that the system

[10] An alternative would be to model the true input values as the outcome of a function of the target input values. In real life, this framework appeals because we conceptualize the target value

receives a true input value, which are the values of the system input variables in U, but that we only know and record the values of "measured input variables" (or just "measured inputs") $U_{\text{target}} \in \mathbb{U}$ after the intervention of an input error mapping $J : \mathbb{U} \times \Theta_J \rightarrow \mathbb{U}$, such that $U_{\text{target}} = J(U, \theta_J)$. The symbol θ_J represents a set of variables called the input error parameters. Each variable θ_{Ji} has a domain Θ_{Ji}, and the combination of their domains I have denoted as Θ_J: $\Theta_J = \prod_{i=1}^{|\theta_J|} \Theta_{Ji}$.

Models whose independent or input variables are assumed to be recorded with error are known in the statistics literature as "measurement error models" or "errors-in-variable models." In this thesis, to avoid confusion with the measurement mapping defined in section 5.1.4.3, I refer to errors in the system input variable as input errors.

5.1.4.3 The measurement mapping G: definition

The measurement mapping G is a model for the measurement process accomplished by instruments and sensors. M can be deterministic or stochastic.

A deterministic measurement mapping can be defined as a function $G : \mathbb{X} \times \Theta_G \rightarrow \mathbb{Y}$. This function takes as input the system variables X_i in $\mathcal{S}$ and a set M of measurement parameters $\theta_{G1} \in \Theta_{G1}, \ldots, \theta_{Gk} \in \Theta_{Gk}, \ldots, \theta_{G|M|} \in \Theta_{G|M|}$, where $|M|$ is the number of variables (a.k.a. parameters) in M. For brevity, we let $\Theta_G = \prod_{k=1}^{|O|} \Theta_{Gk} = \Theta_{G1} \times \Theta_{G2} \times \cdots \times \Theta_{G|O|}$ be the measurement output space of all possible combinations of measurement output variable values. G outputs a set O of measurement output variables $Y_1 \in \mathbb{Y}_1, \ldots, Y_j \in \mathbb{Y}_j, \ldots, Y_{|O|} \in \mathbb{Y}_{|O|}$, where $\mathbb{Y}_j$ is the set of all possible values variable Y_j can assume, and $|O|$ is the number of measurement output variables. For brevity, I denote $\mathbb{Y} = \prod_{j=1}^{|O|} Y_j = \mathbb{Y}_1 \times \mathbb{Y}_2 \times \cdots \times \mathbb{Y}_{|O|}$ to be the measurement output space of all combinations of all the possible measurement output variable values. The vector $Y \in \mathbb{Y}$ with $|O|$ components represents one particular assay outcome. Similarly, $\theta_G \in \Theta_G$ are the parameter values of one assay outcome.

A stochastic measurement mapping can be defined as a probability distribution over $\mathbb{Y}$ parameterized by inputs $\mathcal{S}$ and M, denoted as $Y \sim G\left(X_1, \ldots, X_{|\mathcal{S}|} | \theta_1, \ldots, \theta_{|M|}\right)$. In this more general stochastic case, the inputs can potentially map to multiple members of $\mathbb{Y}$. To clarify, the current state will map to one particular measurement output, and the probability of mapping to each possible output has a probability according to G. It is common to assume that random measurement errors account

as existing before the true value is realized. However, I have chosen to use the opposite direction, where the target values are a function of the true value.

for much of an assay's stochasticity. Thus, the stochastic form for G is appropriate in many cases. For example, the most frequently seen form of G for continuous outputs, in my experience, is a linear regression with normal error:

$$Y \sim \text{Normal}\,(\text{mean} = \beta_0 + \beta_1 X_1, \text{sd} = \sigma) \tag{5.2}$$

In this case, $Y = \mathbb{R}$, $\Theta_G = \mathbb{R}^3$, $M = \{\beta_0, \beta_1, \sigma\}$, $\theta_{G1} = \beta_0$, $\theta_{G2} = \beta_1$, and $\theta_{G3} = \sigma$.

The measurement outcome variables represent the directly measured quantities produced by a measurement technology. In practice, it is often the case that the physical quantity measured is not literally the values of X_i, but a correlated quantity from which an estimate of the true value of X_i will eventually be inferred. In different wording, the outcomes in $\mathbb{Y}$ are generally in different units of measurement than are natural for describing variables in $\mathbb{X}$. For example, in microscopy-based AST assays, the system state variables X_i in $\mathcal{S}_{\text{state}}$ may include the numbers, sizes, ages, and ancestry of cells in an antibiotic exposure. The system input variables U would include the time each frame is imaged. The measured input variables U_{target} are the recorded timestamps of the images, while the measurement output parameters θ_G will include microscope settings like gain and total magnification. The measurement outcome variables Y_j in M are a matrix of pixel intensities. Alternatively (see section 5.1.4.6), the measurement output variables can be some subset of the numbers, areas, ages, and ancestry of cells inferred by an image processing software.

In a fluorimetric assay measuring the production of a fluorescent metabolic byproduct such as resorufin (from resazurin), the system input variables U may include the elapsed time, the volume of the exposure, the initial living biomass, and the initial concentration of resazurin. The system parameters θ_F would include the rate of resazurin conversion per unit biomass and other pharmacodynamic or population dynamic parameters. The system state variables X_i in $\mathcal{S}_{\text{state}}$ would include the total living biomass at a given time and the concentration of resorufin at a given time. The measured input variables U_{target} are the target values of exposure duration, exposure volume, initial resazurin concentration, and (unless the clinical specimen's biomass is not measurable) initial biomass. The measurement output parameters θ_G would include the fluorescence per unit of resorufin and the autofluorescence of the chosen growth media. The measurement outcome variable Y_j in M would be a relative fluorescence unit measured by a camera.

5.1.4.4 The system state is unobservable, while the measurement outputs and measured inputs are observable

In some cases, it is hard to decide whether a variable belongs to the system variables or is part of the measurement. I propose that the criteria for membership in the system are 1) being shared across all measurement technologies in use and not particular to one type of measurement technology, and 2) being an quantity defined by a theoretical model and purported to exist in reality, as opposed to a value that is objectively manifest in a measurement device.

For example, in filtration AST, the nucleic acid amplification reaction could be considered part of the system. However, the purity of reagents and the primer designs can be altered at will. PCR reactions can be repeated on the same sample later on to create replicates. Furthermore the same antibiotic exposure could be analyzed by a different measurement technology than PCR. Thus, it is more useful to model phenomena seen in PCR measurements as part of the measurement mapping. In contrast, the antibiotic exposure of cells in media prior to filtration is shared by all phenotypic AST measurement technologies, being the optimal conditions for any phenotypic AST. So, any entity present in that exposure is a system variable. The bacteria assayed are by definition obtained from the wild, and so their characteristics cannot be engineered beforehand.

Autofluorescence of media is a property of a chemical substance, and so it could be considered a parameter of nature. However, it is a quantity that appears only when an antibiotic exposure is subjected to an artificial excitation beam during the measurement process. Furthermore, it is not a relevant quantity when, say, a microscale cantilever is used as the measurement technology. Therefore, media autofluorescence is a measurement output parameter rather than a system parameter. On the other hand, the amount of resazurin and resorufin are not relevant in other measurement technologies, so they could be considered measurement output parameters as well. However, the substrate molecules were physically introduced into the exposure, where they exist and dynamically interact with the cells. Regardless of whether a fluorescent measurement is made or what excitation wavelength is used, any physical effect of the redox dyes on the growth of the cells will exist. Any such effect would also influence non-fluorescent measurements of the system. I would therefore model the amount of resazurin and resorufin as system variables.

The second criteria reveals an important property separating the system variables from the measurement outputs and measured inputs: observability. In this frame-

work for phenotypic assays, humans can only observe and record the measured outcome variables Y_j, never the system variables X_i. By definition, any attempt to observe the system state simply generates measured outcomes. The true values of the system state are inherently unknowable and only defined by a chosen model for the system.

Though the nature of measurements and their relationship to theory is an active philosophical topic, the stance used above (that the true value of a physical quality is inherently unknowable, and only approximately estimable by measurements) is the prevailing view in practice [159]. It is adopted by the BIPM (International Bureau of Weights and Measures), the international governing metrology body, and it underlies the traditional statistical tools we use to quantify uncertainty [161].

I believe it is important to further recognize and separately consider the stochasticity caused by the system from the stochasticity of the measurement technology. Doing so requires defining the system transformation mapping separately from the measurement mapping, as we have done here. The alternatives to this choice would be to consider the measurement stochasticity as caused by either the system only or by the measurement only. In my experience, it is more common in the literature to assume that the system evolves deterministically, if dynamics are even considered, and then to model all variability as being due to measurement error (e.g., 5.2).

There are precedents in the literature for separating the system and measurement stochasticity. An example of separating system and measurement stochasticity is found in the state-space model paradigm [162] for analyzing time series data, commonly used in control theory, economics, and ecology. State-space models are hierarchical statistical models that take on the following form:

$$\begin{aligned} \frac{\partial \mathbf{x}(t)}{\partial t} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{w} \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{v} \end{aligned} \tag{5.3}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector; $\mathbf{u} \in \mathbb{R}^p$ is the input or control vector; $\mathbf{y} \in \mathbb{R}^q$ is the output vector; and $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$, and $D \in \mathbb{R}^{q \times p}$ are the state (or system), input, output, and feedthrough matrices. To model stochasticity, Weiner or Gaussian processes $\mathbf{w}$ and $\mathbf{v}$ are included, making the equations into stochastic differential equations.

State-space models are a special case of the general phenotypic assay framework described herein. The term "state-space model" traditionally implies the use of

continuous stochastic dynamical systems like Gaussian or Weiner processes, for which elegant algorithms for inference, like the Kalman filter, have been derived. Phenotypic assays may be nonlinear or non-Gaussian however, which is the motivation to generalize the framework describing them in this Chapter. Each part of the traditional state-space model can be understood to correspond to a part of the general phenotypic assay framework. Specifically, the elements of **A**, **B**, and **w** are system parameters; the components of **x** are the system state variables; and the components of **u** are the system input variables. Together, **A**, **B**, **x**, **w**, and **u** are the system variables X_i in the set of system variables $\mathcal{S}$. The state-space model output vector **y** corresponds to the output variables Y_j in the general framework. The elements of **C** and **v** are the measurement output parameters θ_G in M. The feedthrough matrix **D** if often assumed to be zero and was not included in my framework. If included, they would best be part of θ_G. Although **D** and J both transform the system inputs, they are not equivalents because the measured inputs should not be added to the measurement outputs.

As an aside, there is a concept, definitional uncertainty[11], which should not be confused with the above concept of the system's unobservability. The current BIPM framework (since the 3rd edition of the International Vocabulary of Metrology[12]) uses an "Uncertainty Approach," where error and uncertainty represent different concepts [159]. Within the Uncertainty Approach, it is assumed that a range[13]of true quantity values exists due to uncertainty in one's definition or model of the quantity. For example, in measuring the diameter of a sample cylinder, the sample may not be perfectly cylindrical, and the diameter may be defined as a range of values between the maximum inscribing and minimum circumscribing circles [161]. Thus, multiple diameter values will arise during measurement that fit the definition of the true value. The width of the range of true values is a definitional uncertainty. The definitional uncertainty may be further compounded by additional uncertainties during the measurement process. In my framework, the inherent stochasticity of the natural system is an additional uncertainty compounded to the uncertainty of

[11] Definitional uncertainy is defined as a "component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand" and "the practical minimum measurement uncertainty achievable in any measurement of a given measurand" in the BIPM's International Vocabulary of Metrology [159]

[12] The preceding "Error Approach" assumed that a single true quantity value existed, and that systematic measurement error could be defined as the difference between one's (mean) measurements and the one true quantity value.

[13] An exception is that fundamental constants of nature, like the Planck constant, can be assumed to have a single true quantity value.

the true value. Definitional uncertainty of a specimen system is not part of my framework. The confusion may arise because both concepts are used to discuss forms of uncertainty that cannot be reduced by improved measurement technologies. (See section 5.1.5.)

5.1.4.5 The classification mapping H: definition

The classification algorithm $H : \mathbb{Y} \times \mathbb{U} \times \Theta_H \rightarrow \mathbb{C}$, also called[14] a classifier, a hypothesis, or a prediction rule, is a function that takes in the measurement output variables and returns a classification outcome or assay conclusion. The measurement outputs $\mathbb{Y}$ and measured inputs $\mathbb{U}$ from the measurement mapping above now serve as an input space containing all possible examples or instances that can be classified. Let θ_H be a set of classification parameters, and let Θ_H stand for the combined domain of the parameters in θ_H, as done earlier for θ_F, θ_G, and θ_J.

$\mathbb{C}$ is a set of discrete classification values such as "positive" and "negative," either unordered or ordered, that represents all possible classification outcomes. We will denote a particular output of H as the variable C. The value of C is denoted $c \in \mathbb{C}$. While often binary, $\mathbb{C}$ can have more categories, such as "susceptible," "intermediate," "resistant," and "error" for phenotypic AST. If $\mathbb{C}$ is real-valued, then H is called a regression, especially when C represents a property of a real entity. Because clinicians are most concerned with taking discrete actions based on test results, most real-valued diagnostic test results (e.g., units of a substance in serum) end up being interpreted using thresholds or other guidelines to form categorical outcomes (e.g., "within normal limits," "abnormal.") Therefore, I will assume that $\mathbb{C}$ is categorical in the remainder of this thesis.

Most of the classifiers used in this thesis involve the calculation of a real-valued metric or score. This metric is then compared to a threshold λ to make the final classification call. Thus, for convenience, I will define a intermediate classification metric $h_{\text{metric}} \in \mathbb{R}$ that is the output of a metric function $H_{\text{metric}} : \mathbb{Y} \times \mathbb{U} \times \Theta_{H\text{metric}} \rightarrow \mathbb{R}$. The metric h_{metric} is then transformed by the thresholding function $H_{\text{thresh}} : \mathbb{R} \times \Theta_{H\text{thresh}} \rightarrow \mathbb{C}$ to the final classification call $C = c$. θ_H has been split into $\theta_{H\text{metric}}$ and $\theta_{H\text{thresh}}$, with respective domains of $\Theta_{H\text{metric}}$ and $\Theta_{H\text{thresh}}$. In the

[14] The discussion herein is compatible with the framework and terminology of decision theory and the Probably Approximately Correct (PAC) learning model [163, 164], specifically the Agnostic PAC model with generalized loss functions [165]. Note that the PAC framework is one of several overlapping mathematical formulations of the learning task, of which classification is a subclass, that have been put forth in the machine learning literature.

examples in this thesis, $\theta_{H\text{metric}}$ are all the classification parameters except for the threshold λ, and $\theta_{H\text{thresh}}$ is simply $\{\lambda\}$.

For the remainder of this section, the measured input variables U will be understood to be part of the measurement outputs Y. Both sets of variables are treated the same by H, and this notational shortcut will improve readability.

In our context of supervised classification, we assume that every specimen assayed by a phenotypic assay possesses a correct classification that the assay user wishes to know. Let $y \in \mathbb{Y}$ be one instance of the measurements from one assayed specimen, and let $c \in \mathbb{C}$ be the correct classification of y. To adopt the more general Agnostic PAC model, we allow that two instances y_1 and y_2 have the same value in $\mathbb{Y}$ (i.e., $y_1 = y_2$), yet their correct classifications may differ (i.e., $c_1 \neq c_2$).

Let the sample or test set $\mathcal{S}_{\text{test}} = \{(y_1, c_1), \ldots, (y_j, c_j), \ldots, (y_{N_{\text{test}}}, c_{N_{\text{test}}})\}$ be a set of N_{test} labeled specimen measurement instances in $\mathbb{Y} \times \mathbb{C}$. The test set $\mathcal{S}_{\text{test}}$ is drawn from a domain set $\mathbb{Y} \times \mathbb{C}$ according to a distribution $\mathcal{D}$. Formally, $\mathcal{D}$ is a mapping from A, a σ-algebra of $\mathbb{Y}\times\mathbb{C}$ representing events, to the interval $[0, 1] \subset \mathbb{R}$. The joint distribution $\mathcal{D}$ describes the probability $\Pr(y)$ of y arising from $\mathbb{Y}$ (i.e., being drawn into $\mathcal{S}_{\text{test}}$) times the probability $\Pr(c|y)$ that y carries label c, since $\Pr(c, y) = \Pr(c|y)\Pr(y)$ by the laws of probability. The distribution $\mathcal{D}$ is unknown in practice to the learning algorithm and the assay user.

The purpose of the function H is to output classifications $H(y) \in \mathbb{C}$ of y that match the correct classification c as often as possible, in some way. A classifier H that makes correct classifications often is said to perform well. Mathematically, the performance of H on a single instance to be classified is defined by a loss function $\ell : \mathbb{H} \times \mathbb{Y} \times \mathbb{C} \to \mathbb{R}^+$ that reflects the (non-negative) difference between the true classification and the predicted classification. Better performing classifiers exhibit smaller losses. $\mathbb{H}$ is the "hypothesis class," the domain of the possible classifiers, a set which is chosen beforehand by the practitioner in the PAC framework. The choice of $\mathbb{H}$ determines the extent of overfitting, and whether learning is even possible (as described by PAC theory).

The performance of H on a set of instances to be classified is measured by a risk function, also called an error function [164, 166]. When considering the performance of H in classifying the theoretical true population of instances $\mathbb{Y}$, the risk function is called the generalization risk or generalization error $R(H) : \mathbb{H} \to \mathbb{R}^+$,

defined as

$$R(H) = \mathop{\mathrm{E}}_{(y,c)\sim\mathcal{D}}[\ell(H, y, c)]. \tag{5.4}$$

When considering the performance of H against a randomly drawn test set $\mathcal{S}_{\text{test}}$, the risk function is called the empirical risk or empirical error $\hat{R}(H) : \mathbb{H} \to \mathbb{R}$, defined as

$$\hat{R}(H) = \frac{1}{N_{\text{test}}} \sum_{j=1}^{N_{\text{test}}} \ell(H, y_j, c_j) \tag{5.5}$$

If the instances of $\mathcal{S}_{\text{test}}$ is drawn independently and identically distributed by $\mathcal{D}$, then the generalization risk is the expectation of the empirical risk:

$$R(H) = \mathop{\mathrm{E}}_{\mathcal{S}_{\text{test}}}[\hat{R}(H)] \tag{5.6}$$

Furthermore, it is sometimes useful to define the conditional (generalization) risk $R(H|y)$ for a particular observation y as follows:

$$R(H|y) = \mathop{\mathrm{E}}_{(c|y)\sim\mathcal{D}}[\ell(H, y, c)] \tag{5.7}$$

$$= \sum_{c} \ell(H, y, c)\Pr(c|y). \tag{5.8}$$

The overall generalization risk can be written as the conditional risk averaged over all possible observations: $R(H) = \int_{\mathbb{Y}} R(H|y)\Pr(y)dy$.

As with loss functions, better performing classifiers exhibit smaller risks. One can therefore define an "ideal classifier," a concept discussed later in the next section 5.1.5, that has the minimal possible risk over $\mathbb{Y}$. In fact, for our definition of the phenotypic assay, the accuracy performance metric discussed in section 5.1.1 should be synonymous with the risk functions defined in this section. Whether the generalization or empirical risk function is used as a performance metric depends on whether one is discussing an ideal assay or an actual assay, respectively.

There are a number of logical choices for a loss function depending on one's application. The squared error loss function $\ell_{\text{SE}}(H, y, c) = (H(y) - c)^2$ is the most commonly used loss function for regression (and sometimes classification), thanks to its mathematical convenience. The hinge loss is another loss function used in machine learning. For classification tasks, the simplest and an intuitive loss function is the 0-1 loss function that records the number of errors of any kind. The 0-1 loss function is

$$\ell_{0\text{-}1}(H, y, c) = \mathbf{1}(H((y) \neq c) = \begin{cases} 0, & \text{if } h(y) = c \\ 1, & \text{if } h(y) \neq c. \end{cases} \tag{5.9}$$

The corresponding generalization and empirical risks are, respectively,

$$\begin{aligned} R_{0\text{-}1}(H) &= \mathop{\mathrm{E}}_{(y,c)\sim\mathcal{D}}[\mathbf{1}(H((y) \neq c)] \\ &= \Pr_{(y,c)\sim\mathcal{D}}[H(y) \neq c] \\ &= \mathcal{D}((y,c) : H(y) \neq c) \end{aligned} \tag{5.10}$$

and

$$\hat{R}_{0\text{-}1}(H) = \frac{1}{N_{\text{test}}} \sum_{j=1}^{N_{\text{test}}} \mathbf{1}(H((y) \neq c) \tag{5.11}$$

In many decisions in clinical practice, making an incorrect call of a false negative (or a false positive) causes worse consequences than making the complementary call. Thus, one can consider a "weighted 0-1 loss function":

$$\ell_{0\text{-}1,\text{weight}}(H,y,c) = \begin{cases} 0, & \text{if } h(y) = c \\ w_{FN}, & \text{if } h(y) = \text{Neg and } c = \text{Pos} \\ w_{FP}, & \text{if } h(y) = \text{Pos and } c = \text{Neg}, \end{cases} \tag{5.12}$$

where "Pos" and "Neg" are the elements of $\mathbb{C} = \{\text{Pos}, \text{Neg}\}$ representing positive and negative classifications, and $w_{FN}, w_{FP} \in \mathbb{R}^+$ so that ℓ remains non-negative.

To choose H, the engineer can either assume a class of functions from first principles and/or use a learning algorithm to find improved classifiers, possibly after training on a set $\mathcal{S}_{\text{training}} \in \mathbb{Y} \times \mathbb{C}$ of known instances. How to further characterize, implement, assess, and choose learning algorithms is a topic addressed by the large body of literature on statistical machine learning [167], which I will not cover here.

For the purposes of this thesis, I wish to point out the following conclusions of how the classification algorithm fits into the proposed framework for phenotypic assays. First, all phenotypic assays inherently include a classification algorithm as the final step of the assay, up to the trivial case where the measurement outcome space $\mathbb{Y}$ is equal (or isomorphic) to the classification space $\mathbb{C}$. (See the next section 5.1.4.6 for an explanation.) Secondly, the performance of the classification algorithm is necessarily a component of the overall assay's performance and must be considered by anyone involved in assay development. I believe that future engineers will find it helpful to conceptually modularize the contribution of assay errors due to data analysis apart from the physical, "hardware" limitations from the system or the measurement technology. The hardware and analysis components of an assay are optimized, after all, by separate sets of techniques and literature.

5.1.4.6 The boundary between the measurement mapping and the classification algorithm is subjective

Data obtained by measurement technologies often undergoes processing to reduce dimensionality and/or to smooth out random noise. It is the task of the modeler to decide how much data processing is modeled as part of the measurement or the final analysis algorithm.

For example, in digital filtration AST, the first transduction of the measurement technology are the electronic signals from the sensors of the BioRad QX200 ddPCR droplet reader device I used. Although proprietary and not seen by humans, if these signals were digitized, they would be the first human-readable numbers during the measurement process. These voltage changes are segmented and converted at least to two-channel fluorescence intensities of individual droplet detections, as displayed by the proprietary software. Droplets outside of pre-set tolerances are discarded. Then, a thresholding is performed to denote positive and negative droplets. Next, the counts of positive and negative droplets are used to calculate a nucleic acid concentration. The outcomes of any of these steps could be considered the measurement outcome variable Y, and the subsequent steps considered part of the classification algorithm. In Chapter 3, I modeled the measurement outcome variable to be the nucleic acid concentration. Modeling the droplet counts would also have been a conveniently simple choice for Y. In Chapter 4, I modeled the measurement outcome variable as the well population status tallies, but could have also modeled nucleic acid concentrations or droplet counts. The alternative choices for Y would be more complex to model due to their higher dimensionality. Continuing with the filtration AST example, $\mathbb{C} = \{\text{Susceptible}, \text{Resistant}\}$, and the choices of F include the metrics of section 4.5.1.

Although there is subjectivity in separating the measurement mapping and the classification algorithm, I believe modeling and discussing the two mappings as two modules is the most useful generalization as the engineering optimization of the measurement hardware and the statistical analysis software are performed in different contexts (e.g., stages of bringing an assay to market) using different skill sets.

5.1.5 The ideal phenotypic assay is defined by minimizing all controllable sources of stochasticity

As stated at the beginning of this chapter, an ideal phenotypic assay is one with the highest accuracy. Since an assay comprises only the four mappings F, J, G, and H, the overall inaccuracy of a phenotypic assay arises from the ambiguities introduced by these four mappings[15]. In other words, the assay error is some composition of the stochasticity of the system evolution in F, the stochasticity of input error from J, the stochasticity of measurement error during G, and the accuracy (in the binary diagnostic sense) of the classification algorithm H. Specifically, inaccuracies or errors arise when multiple system variable values can map to the same measured input or measurement output, or when measurement outputs can map to multiple final classifications.

Whether the inaccuracy created during a mapping can be minimized depends on whether the mapping is controllable by humans. In the measurement mapping G, all the components of the measurement technology are controlled by humans. We design and manufacture the measurement technology to reduce uncertainty in its specification, and we set the measurement input parameters that set the controlled system variables. In practice, there is always stochasticity and ambiguity in measurements due to random errors. But the notion of an ideal assay asks us to make the unrealistic assumption that with infinite human resources and ingenuity, the measurement uncertainty will tends towards zero. When the measurement technology has no uncertainty, the measurement technology will be one-to-one mappable to each system variable, such that the measurement space and the system space become the same entity up to isomorphism. In other words, in an ideal assay, the measurement space is simply the system space, as if the latter were magically knowable to humans. Similarly, the input error mapping J can theoretically be reduced to the identity function in an ideal assay.

In transformation mapping F, the assayed system contains some variables in $\mathcal{S}$ that humans can influence, namely the system input variables U. By our definition, U is not perfectly set by humans, but as the input errors asymptotically approach zero, U will approach U_{target}. But other components, such as the biological properties of the bacteria and their stochastic response to antibiotics, are derived from the "wild" and/or cannot logically be controlled by humans. Thus, biological stochasticity and

[15] As a reminder, I am ignoring definitional uncertainty (see section 5.1.4.4) created by the finite amount of detail in the definition of the system variables.

any ambiguity it generates cannot be engineered away, even in an ideal assay.

For the classification algorithm mapping H, there are two benchmarks relevant to ideal performance. In a perfect classification, all instances are correctly classified. However, if the same instance of measurement outputs can be labeled with more than one true classifications, as in the Agnostic PAC model, then the measurement outputs are not perfectly separable, and a perfect classification is not possible on average when using the same classifier. Instead, classifiers must find a balance between bias (underfitting) and variance (overfitting). Balancing bias and variance is indeed a primary pre-occupation of anyone implementing a classier in the real world. Nonetheless there is a benchmark performance for an ideal classifier that can be defined, the Bayes optimal classifier, even when a perfect classifier cannot be realized.

The Bayes optimal classifier H^* is the classifier that possesses the minimal possible (technically, the infimum) expected risk given the distribution $\mathcal{D}$, which can be proven to be of the form shown in equation 5.14. In practice, one cannot know whether one's classifier is a Bayes classifier because one does not know the nature of $\mathcal{D}$. Only samples of the distribution can be observed as training data. Nonetheless, the concept serves as a general benchmark and is important in deriving bounds on the total error [168]. Furthermore, in the ideal case, the domain of Y and how it is generated is presupposed to exist, so the ideal classifier performance can be calculated explicitly. For example, we do so in the next section, 5.2.

When one uses the 0-1 loss function, the corresponding Bayes optimal classifier simply chooses the most likely classification. In the vocabulary of binary clinical diagnostics, the Bayes optimal classifier with 0-1 loss maximizes the accuracy ACC and the positive predictive value (PPV) metrics.

$$H^*_{0\text{-}1}(y) = \arg\max_{c \in \mathbb{C}} \Pr(c|y) \tag{5.13}$$

In the general case with arbitrary loss function ℓ, the Bayes optimal classifier (also known as the Bayes decision rule in decision theory) will choose the classification that incurs the lowest conditional generalization risk $(R|y)$ for each encountered y, thereby minimizing the overall generalization risk $R(H)$ of the classifier:

$$H^*(y) = \arg\min_{H(y) \in \mathbb{C}} \sum_c \ell(H, y, c) \Pr(c|y) \tag{5.14}$$

To summarize, a phenotypic assay is an ideal phenotypic assay if and only if all the following properties apply:

- The number of uncontrolled system state variables is minimized to the extent that can be logically justified, but some variables will remain uncontrolled.
- The system input variables are identical to the measured input variables, indicating perfect control over the inputs.
- The measurement space is equal (up to isomorphism) to the system state space. In other words, there is no measurement error and no loss of information about the system state.
- The classification algorithm is a Bayes optimal classifier for the system state/measurement space, with respect to some reasonable loss function.
- The assay's generalization risk, with respect to the chosen loss function, is minimized.

5.2 Calculation of the ideal performance for the accessibility AST assay

We now can apply our definition of the ideal phenotypic assay to answer the question of what the ideal performance of accessibility AST should be.

We define an ideal accessibility AST assay and its component variables, fix the constant system variables to our best estimates of their values for relevant strains, and arrive at a function whose level set in the state space represents the ideal assay performance.

5.2.1 Defining the assay system and model scope

For simplicity, we will assume the no-births model for the system described in section 4.4.1. The system comprises the time variable $t \in \mathbb{T}, \mathbb{T} = \mathbb{R}^+$; two non-negative integer state variables, D and L, in the state space $\mathbb{X} = Z^+ \times Z^+$; seven constant real-valued parameters L_0, k, α, $\beta_{\max}$, γ, EC_{50}, and antibiotic dose A; and an indicator parameter c for the true susceptibility phenotype class. (The parameter A was previously denoted $[Abx]$.) $D[t]$ is the number of dead cells at time t, and $L[t]$ is the number of live cells at time t. The parameter $L_0 \in \mathbb{Z}^+$, the inoculum, is the number of live cells at time 0, and also the total number of cells at any time t: $L_0 = L[t] + D[[t] = L[0]$. We assume no dead cells are present at time 0. Thus, L_0 sets the initial conditions for the dynamics of the state variables.

Only one end-point measurement is made at time $T \in \mathbb{R}^+$ to reflect the protocol used for the available digital filtration AST data. Analysis of time-series filtration AST,

in which information about population dynamics is available, will be left for future work.

The variables T, A, and L_0 are controlled or assumed given, although L_0 is not actually controllable in real life filtration AST. The three variables are the system input variables. In this ideal assay system, the system input variables are all constant. Furthermore, no live or dead cells are re-introduced from outside the system after the exposure has begun, so there is no need for system input variables representing such immigration or emigration processes.

The phenotypic class c takes on the values in the set $\mathbb{C} = \{\text{Pos}, \text{Neg}\}$, where "Pos" denotes "Resistant" and "Neg" denotes "Resistant." For each value of c, there is a value of the EC_{50} denoted by $EC_{50}[c]$, with $EC_{50}[\text{Neg}] < EC_{50}[\text{Pos}]$. In this section we will assume the two values of $EC_{50}[c]$ are a given constants. The remaining parameters are defined the same way as in Chapter 3.

For convenience, I will often group together the nine non-state system parameters (L_0, A, T, k, α, $\beta_{\max}$, γ, EC_{50}), or a subset of them depending on context, together under the symbol θ.

The transformation mapping F is thus a stochastic mapping where the joint distribution of D and L is multinomially distributed, and equation 4.25 (derived from ODEs in section 3.3) and equation 4.26 govern the state evolution to time T. (The equations are repeated below for clarity.)

$$L, D|\theta \sim \text{Multinomial}\,(\text{size} = L_0, \text{prob} = (S[\theta], 1 - S[\theta])) \tag{5.15}$$

$$S[\theta] = e^{-kT} Q[\alpha, \beta T] \tag{5.16}$$

$$\beta = \frac{\beta_{\max} A^{\gamma}}{EC_{50}[c]^{\gamma} + A^{\gamma}}. \tag{5.17}$$

The marginal distribution of L (and of D) is binomial. For convenience we will only discuss L in the remainder of this chapter. Let us denote the cumulative probability density function (CDF) of L as a function F_L, and the complementary CDF as $\overline{F_L}$:

$$F_L[L; \theta] = \sum_{l=0}^{L} \binom{L_0}{l} S[\theta]^l \, (1 - S[\theta])^{L_0 - l} \tag{5.18}$$

$$\overline{F_L}[L; \theta] = 1 - F_L[L; \theta], \tag{5.19}$$

The probability mass function (PMF) of L will be denoted

$$f_L[L; \theta] = \binom{L_0}{l} S[\theta]^l \, (1 - S[\theta])^{L_0 - l} \tag{5.20}$$

5.2.2 Deriving the ideal classifier

Since this is an ideal assay, the measurement output space $\mathbb{Y}$ is the same as the state space $\mathbb{X}$, the measurement mapping G can be ignored as it is trivially the identity function, and the classification mapping could be written $H : \mathbb{X} \to \mathbb{C}$. For consistency, however, we will continue to denote H in terms of $\mathbb{Y}$.

To complete the ideal assay, the choice of the classification function H must be an ideal classifier. When using the weighted 0-1 loss function, the ideal classifier can be shown to take the form of a scalar threshold. Starting from equations 5.12 and 5.14, we see that

$$H^*(y) = \underset{H(y)\in\mathbb{C}}{\arg\min} \sum_c \ell_{\text{0-1,weighted}}(H, y, c)\Pr(c|y) \tag{5.21}$$

$$= \underset{H(y)\in\mathbb{C}}{\arg\min} \begin{cases} 0 \cdot \Pr(c = \text{Neg}|y) + w_{FN} \cdot \Pr(c = \text{Pos}|y), & \text{if } H(y) = \text{Neg} \\ 0 \cdot \Pr(c = \text{Pos}|y) + w_{FP} \cdot \Pr(c = \text{Neg}|y), & \text{if } H(y) = \text{Pos} \end{cases} \tag{5.22}$$

$$= \begin{cases} \text{Neg}, & \text{if } w_{FN} \cdot \Pr(c = \text{Pos}|y) < w_{FP} \cdot \Pr(c = \text{Neg}|y) \\ \text{Pos}, & \text{if } w_{FN} \cdot \Pr(c = \text{Pos}|y) \geq w_{FP} \cdot \Pr(c = \text{Neg}|y) \end{cases} \tag{5.23}$$

$$= \begin{cases} \text{Neg}, & \text{if } \frac{w_{FN}}{w_{FP}} < \frac{\Pr(c=\text{Neg}|y)}{\Pr(c=\text{Pos}|y)} \\ \text{Pos}, & \text{if } \frac{w_{FN}}{w_{FP}} \geq \frac{\Pr(c=\text{Neg}|y)}{\Pr(c=\text{Pos}|y)} \end{cases} \tag{5.24}$$

where $y \in \mathbb{Y}$ will be replaced[16] by the random variable $L|\theta$ in our specific model, and where we have arbitrarily chosen "Pos" to be the output in the case of equality[17]. Note how $c \in \mathbb{C}$ denotes both a system variable and the true classification label. Meanwhile, $H(y) \in \mathbb{C}$ denotes the final assay output.

Equation 5.24 states that the ideal classifier for the weighted 0-1 loss function simply checks if the likelihood ratio $\frac{\Pr(c=\text{Neg}|y)}{\Pr(c=\text{Pos}|y)}$ of the classes is above or below the threshold of $\frac{w_{FN}}{w_{FP}}$, which varies from 0 to ∞^+. It is unknown, however, whether there are multiple thresholds that are crossed as $y = L|\theta$ varies, because the terms for $\Pr(c|y)$ have not yet been defined from this ideal assay model. The term for $\Pr(y|c) = \Pr(L|c, \theta)$, however, is defined in equations 5.15–5.17, and it is the binomial PMF of equation 5.20. Therefore, the $\Pr(c|y)$ terms can be factored using Bayes' Theorem and simplified to one unknown parameter—the prior distribution

[16] Notation not used in order to ease page formatting.

[17] An "Inconclusive" classifier output is also possible in practice, but setting boundaries for such an output introduces details not necessary for the theoretical aims of this thesis.

$\Pr(c)$, also called the prevalence, for the frequency of the labels.

$$\Pr(c|L,\theta) = \frac{\Pr(L|c,\theta)\Pr(c)}{\sum_c \Pr(L|c,\theta)\Pr(c)} \tag{5.25}$$

$$= \frac{\frac{L_0!}{L!(L_0-L)!} S[\theta]^L \left(1 - S[\theta]\right)^{L_0-L} \Pr\left(c\right)}{\sum_c \frac{L_0!}{L!(L_0-L)!} S[\theta]^L \left(1 - S[\theta]\right)^{L_0-L} \Pr\left(c\right)}. \tag{5.26}$$

Alternative factorizations of $\Pr(c|L,\theta)$ would not circumvent the fact that $\Pr(c)$ is not known. Note that $\sum_c \Pr(c) = 1$ by Kolmogorov's second axiom of probability [169].

The algebraic form for the argument of equation 5.26 is that of a $|\mathbb{C}|$-dimensional logistic function of L, except only at the points where L is an integer. For clarity, consider the equation when $c = \text{Pos}$:

$$\Pr(c = \text{Pos}|L,\theta) = \frac{\left(\frac{S[\text{EC}_{50}[\text{Pos}],\theta]}{1-S[\text{EC}_{50}[\text{Pos}],\theta]}\right)^L \kappa_{c=\text{Pos}}}{\left(\frac{S[\text{EC}_{50}[\text{Pos}],\theta]}{1-S[\text{EC}_{50}[\text{Pos}],\theta]}\right)^L \kappa_{c=\text{Pos}} + \left(\frac{S[\text{EC}_{50}[\text{Neg}],\theta]}{1-S[\text{EC}_{50}[\text{Neg}],\theta]}\right)^L \kappa_{c=\text{Neg}}} \tag{5.27}$$

where $\kappa_c = \Pr(c) \cdot (1 - S[\text{EC}_{50}[c],\theta])^{L_0}$ are terms constant with respect to L. Since the survival likelihood ratio terms (the bases of the exponentiation by L) are non-negative, $H^*(L|\theta)$ is always a monotonic function with respect to L. Similarly, the likelihood ratio $\frac{\Pr(c=\text{Neg}|y)}{\Pr(c=\text{Pos}|y)}$ is also a monotonic exponential function with respect to L:

$$\frac{\Pr(c = \text{Neg}|y)}{\Pr(c = \text{Pos}|y)} = \left(\frac{S[\text{EC}_{50}[\text{Neg}],\theta]}{S[\text{EC}_{50}[\text{Pos}],\theta]} \times \frac{1 - S[\text{EC}_{50}[\text{Pos}],\theta]}{1 - S[\text{EC}_{50}[\text{Neg}],\theta]}\right)^L \times \left(\frac{1 - S[\text{EC}_{50}[\text{Neg}],\theta]}{1 - S[\text{EC}_{50}[\text{Pos}],\theta]}\right)^{L_0} \times \frac{\Pr\left(c = \text{Neg}\right)}{1 - \Pr\left(c = \text{Neg}\right)} \tag{5.28}$$

Thus, our ideal model's classification algorithm always take the form of a indicator function with a single threshold λ for L. We can calculate, in terms of the parameters θ, the optimal threshold value $\lambda^* \in \mathbb{R}$ for L for our ideal classifier by substituting the integer L with the real-valued λ^* in equations 5.24 and 5.28, then solving for λ^* and only keeping real solutions.

$$\lambda^* = \frac{\log\left(\frac{w_{FN}}{w_{FP}} \times \frac{1-\Pr(c=\text{Neg})}{\Pr(c=\text{Neg})} \times \left(\frac{1-S[\text{EC}_{50}[\text{Pos}],\theta]}{1-S[\text{EC}_{50}[\text{Neg}],\theta]}\right)^{L_0}\right)}{\log\left(\frac{S[\text{EC}_{50}[\text{Neg}],\theta]}{S[\text{EC}_{50}[\text{Pos}],\theta]} \times \frac{1-S[\text{EC}_{50}[\text{Pos}],\theta]}{1-S[\text{EC}_{50}[\text{Neg}],\theta]}\right)}. \tag{5.29}$$

The optimal classifier for our ideal assay model, with the weighted 0-1 loss function, can now be written as:

$$H^*(L|\theta) = \begin{cases} \text{Neg}, & \text{if } L > \lambda^* \\ \text{Pos}, & \text{if } L \leq \lambda^*. \end{cases} \tag{5.30}$$

It is tempting to envision thresholds for L as elements of $\mathbb{Y} = \{0, 1, \ldots, L_0 - 1, L_0\} \subset \mathbb{Z}^+$, like L. However, note that the domain of λ, as I prefer to define it, is continuous and unbounded, while L is discrete. So, L will rarely exactly equal λ, and the behavior of $H(L|\theta)$ will contain discontinuous jumps caused by L's discrete nature. Encoding these discontinuities into the definition of the threshold λ is possible by using the floor function so that $\lambda \in \mathbb{Y}$, but the complexity does not seem to yield additional insights than the formulation above.

Furthermore, in some regions of its parameter space, λ^* lies outside of the interval $[0, L_0]$ and cannot be reached by L in our system, where reaching λ means to be able to assume a value that is greater than and a value that is less than λ since L do. Specifically, there are asymptotes when the two survival probabilities become equal to each other, or either one approaches 0 or 1. Extreme values of $\frac{w_{FN}}{w_{FP}}$ or $\frac{\Pr(c=\text{Pos})}{\Pr(c=\text{Neg})}$ can also send λ^* out of the range of L. The optimum threshold in these cases should be capped at either 0 or L_0, respectively, if λ were to be defined with a domain $\mathbb{Y}$. Lastly, When $S[\text{EC}_{50}[\text{Neg}], \theta] = S[\text{EC}_{50}[\text{Pos}], \theta]$, the optimum threshold is degenerate as any threshold would yield the same poor assay performance.

5.2.3 Summarizing free parameters

At this point, $H^*(L|\theta)$ remains a multidimensional function with at least thirteen arguments: three system input variables of interest (T, A, L_0); seven pharmacodynamic parameters (k, α, $\beta_{\max}$, γ, $\text{EC}_{50}[\text{Pos}]$, $\text{EC}_{50}[\text{Neg}]$, c); the loss function weights (w_{FN}, w_{FP}); and the prevalence $\Pr(c)$.

Showing all possible instances of $H^*(L|\theta)$ with multiple free parameters could be achieved with an interactive visualization software, capturing all the trade-offs made necessary by the model's mathematical structure. However, to more easily discuss and visualize H^*, and to make the static figures featured in subsequent sections, it is insightful to examine the function behavior only along a small subset of dimensions at a time. To reduce the dimensions or our analysis, we summarize and collapse certain parameters to interpretable constant values.

The pharmacodynamic parameters in $S[\theta]$ that are not assumed to be specific to the

strain are α, β, γ, and k. These are held constant at the fitted values from the bulk filtration experiments in Table 3.3.

In choosing the strain EC_{50} values, the two values must be chosen to yield interpretable conclusions. There exist several reasonable choices for the values of $EC_{50}[c]$ to represent the susceptible and resistant classes. First, one could choose the two closest values to challenge the ideal assay with the worst case scenario. In the currently accepted paradigm of the minimum inhibitory concentration, a measured MIC needs to be one multiplicative factor of 2 away from the breakpoint concentration A_{break} to be considered susceptible or resistant. Thus, one could choose $EC_{50}[\text{Neg}] = \frac{1}{2}A_{\text{break}}$ and $EC_{50}[\text{Pos}] = 2A_{\text{break}}$. However, the factor of 2 is an arbitrary constant close to the current measurement error of the standard broth microdilution assay. The truly worst case scenario would be two concentrations infinitesimally lower and higher than A_{break}, which would be impossible and pointless to classify. A second option for defining $EC_{50}[c]$ would be to find a summary statistic (e.g., arithmetic or geometric mean, median, or mode) of the observed susceptible and resistant MICs from a large sampling of healthcare facilities.

In this analysis, the values of the susceptible and resistant strain used most frequently in this thesis (*E. coli* K12 and UCLA #38) were chosen on the premise that their EC_{50} values were the most accurately estimated, and their MICs appeared sufficiently representative (i.e., within an order of magnitude) of typical clinical isolates. There is no reason why two strains with a smaller (and more challenging) difference in EC_{50} could have been analyzed. Such an analysis could be performed in the future.

The three system input parameters—the exposure duration T, the dose A, and the inoculum L_0—are the variables of highest interest to designers of AST assays because they can be controlled at will by the designer. In particular, the inoculum L_0 is the parameter that most limits phenotypic ASTs from whole blood, as discussed in section 1.1.2. The inoculum is also the focus of this thesis. Thus, we do not need to summarize these three variables.

Nonetheless, in the below analysis (Figure 5.2), the dose A was fixed at the breakpoint MIC concentration. By definition, the breakpoint MIC concentration is the concentration chosen by standards organizations as the separation between susceptible from resistant isolate MICs. The breakpoint MIC may not be the concentration that yields the fastest signal, but it is interpretable by current physicians. The analysis can be repeated to examine the effect of A on ideal assay performance, as done in the next section in Figure 5.3.

5.2.4 Plotting the expected accuracy

As mentioned in section 5.1.4.5, the generalization risk is the selfsame accuracy performance metric (see section 5.1.1) we wish to examine with our ideal assay model. Plotting the ideal assay's generalization risk as a single value is not meaningful, however. Even when the above pharmacodynamic parameters and system input variables are held constant, the optimal classifier still comprises a continuous set or family of thresholds parameterized by the loss function weights and the prevalence. Each optimal classifier in this set may possess a different generalization risk than another optimal classifier with different values for these parameters.

Certain conventional values exist for these remaining parameters: the FDA has set guidelines for acceptable accuracy for antibiotic susceptibility tests, and one can estimate resistance prevalence for various populations from epidemiological data. However, the full ranges of these parameters are relevant to the question of this thesis because they factor into the definition of the assay's generalization risk.

To visualize these parameter ranges, and to link with the conventional terminology for binary clinical diagnostic tests, we can reparameterize the relationship between w_{FN}, w_{FP}, and $\Pr(c)$ using quantities I call the expected diagnostic sensitivity and specificity[18]. The expected diagnostic sensitivity, $\mathrm{Sn}[\lambda]$, of an assay is the the expected proportion of instances, or probability, that an assay makes a correct positive (resistant) call given that the strain is resistant ($c = \mathrm{Pos}$), while the expected diagnostic specificity, $\mathrm{Sp}[\lambda]$, is the probability that the assay makes a correct negative (susceptible) call given that the strain is susceptible ($c = \mathrm{Neg}$).

$$\mathrm{Sn}[\lambda;\theta] = \Pr(L \leq \lambda | c = \mathrm{Pos}, \theta) = F_L[\ \lfloor\lambda\rfloor\ ; \mathrm{EC}_{50}[\mathrm{Pos}], \theta] \tag{5.31}$$

$$\mathrm{Sp}[\lambda;\theta] = \Pr(L > \lambda | c = \mathrm{Neg}, \theta) = \overline{F_L}[\ \lfloor\lambda\rfloor\ ; \mathrm{EC}_{50}[\mathrm{Neg}], \theta]. \tag{5.32}$$

$\mathrm{Sn}[\lambda]$ and $\mathrm{Sp}[\lambda]$ are functions of the threshold $\lambda \in \mathbb{R}^+$. The trade-off between SN and $1 - SP$ as λ varies is what is depicted in the traditional receiver operating characteristic (ROC) curve.

[18] For clarity, note that one can also define quantities that are best called the empirical or observed diagnostic sensitivity $SN = \frac{TP}{FN+TP}$, observed diagnostic specificity $SN = \frac{TP}{FN+TP}$, and observed accuracy $ACCU$ (see Equation 5.1 for definition.) These quantities, used in Chapter 4 and section 5.1.1, are defined using observed results and used only in contexts where a set of assay performance data is present. In this chapter's context, the observed diagnostic sensitivity, diagnostic specificity, and accuracy of an ideal assay would be random variables whose expected values are the expected diagnostic sensitivity, diagnostic specificity, and accuracy of that ideal assay. Also note that $\mathrm{Sn}[\lambda]$, $\mathrm{Sp}[\lambda]$, $\mathrm{Accu}[\lambda]$, SN, SP, and $ACCU$, as written above, are all only defined for binary classification algorithms.

The relationship between $\text{Sn}[\lambda]$, $\text{Sp}[\lambda]$, and the generalization risk for the weighted 0-1 loss function is derived below.

$$R(H(L;\lambda,\theta)) = \mathop{\text{E}}_{(L,c)\sim D}[\ell_{\text{0-1,weight}}(H,L,c)] \tag{5.33}$$

$$= \sum_{(L,c)} \ell_{\text{0-1,weight}}(H,L,c)\Pr(L|c,\theta)\Pr(c) \tag{5.34}$$

$$\begin{aligned} = &\sum_{L\le\lambda} w_{FP}\cdot f_L[L;c=\text{Neg},\theta]\Pr(c=\text{Neg}) \\ &+ \sum_{L>\lambda} 0\cdot f_L[L;c=\text{Neg},\theta]\Pr(c=\text{Neg}) \\ &+ \sum_{L\le\lambda} 0\cdot f_L[L;c=\text{Pos},\theta]\Pr(c=\text{Pos}) \\ &+ \sum_{L>\lambda} w_{FN}\cdot f_L[L;c=\text{Pos},\theta]\Pr(c=\text{Pos}) \end{aligned} \tag{5.35}$$

$$\begin{aligned} = &\, w_{FP}\cdot F_L[\ \lfloor\lambda\rfloor\ ;\text{EC}_{50}[\text{Neg}],\theta]\cdot\Pr(c=\text{Neg}) \\ &+ w_{FN}\cdot \overline{F_L}[\ \lfloor\lambda\rfloor\ ;\text{EC}_{50}[\text{Pos}],\theta]\cdot\Pr(c=\text{Pos}) \end{aligned} \tag{5.36}$$

$$\begin{aligned} = &\, w_{FP}\cdot(1-\text{Sp}[\lambda;\theta])\cdot(1-\Pr(c=\text{Pos})) \\ &+ w_{FN}\cdot(1-\text{Sn}[\lambda;\theta])\cdot\Pr(c=\text{Pos}). \end{aligned} \tag{5.37}$$

Let the expected accuracy, $\text{Accu}[\lambda]$, of an assay be the expected proportion of correct calls made by an assay.

$$\text{Accu}[\lambda;\theta] = \text{Sn}[\lambda;\theta]\Pr(c=\text{Pos}) + \text{Sp}[\lambda;\theta]\Pr(c=\text{Neg}) \tag{5.38}$$

By comparing equation 5.37 and 5.38, one can see that $-R(H)$ and $\text{Accu}[\lambda]$ are similar but not equivalent quantities, even though it is tempting to define an ideal assay as maximizing $\text{Accu}[\lambda]$. Our ideal assay does not maximize $\text{Accu}[\lambda]$ unless the unweighted 0-1 loss function is assumed, since equation 5.37 yields equations 5.1 and 5.38 when one chooses $w_{FN} = w_{FN} = 1$ and then defines $\text{Accu}[\lambda] = \text{E}[ACCU] = 1 - R(H)$.

Equation 5.37 shows that the generalization risk can be fully specified by the parameters w_{FN}, w_{FP}, $\Pr(c=\text{Neg})$, $\text{Sn}[\lambda]$, $\text{Sp}[\lambda]$. The sensitivity and specificity are not affected by the prevalence, since they are conditional on c. Thus, to visualize $R(H)$ as a function of all 5 of these parameters, it was decided that the $\text{Sn}[\lambda]$ and $\text{Sp}[\lambda]$ would be fixed subjectively (e.g., according to FDA guidelines). The viewer would be expected to vary the prevalence mentally. The choice of $\text{Sn}[\lambda]$, $\text{Sp}[\lambda]$, and $\Pr(c=\text{Neg})$ would then imply unique values of w_{FN} and w_{FP} that would minimize $R(H)$ and create H^*.

5.2.5 Numerical computation of the ideal assay isosurface

Of the possible depictions of the ideal assay's optimal classifier, the one most related to the original query of this thesis is to show, for each starting inoculum L_0, the minimum exposure duration needed for the classifier to achieve a given assay performance at a given antibiotic dose. If a minimum exposure duration did not exist, then achieving that assay performance is impossible when starting from that inoculum.

Figure 5.2 shows, for each value of L_0, the minimum value of T for which at least one threshold can be drawn that classifies the system state/measurement output with at least the desired diagnostic sensitivity (SN) and specificity (SP).

Finding T subject to these constraints constitutes a discrete optimization problem. The following algorithm was devised to calculate the isosurface of H^* in Figure 5.2.

First, the threshold λ was reparameterized as a scaled threshold ξ that varies from 0 to 1, such that $\xi L_0 = \lambda$. As ξ varies from 0 to 1, the expected sensitivity and specificity trace out a convex curve C on an ROC plot[19].

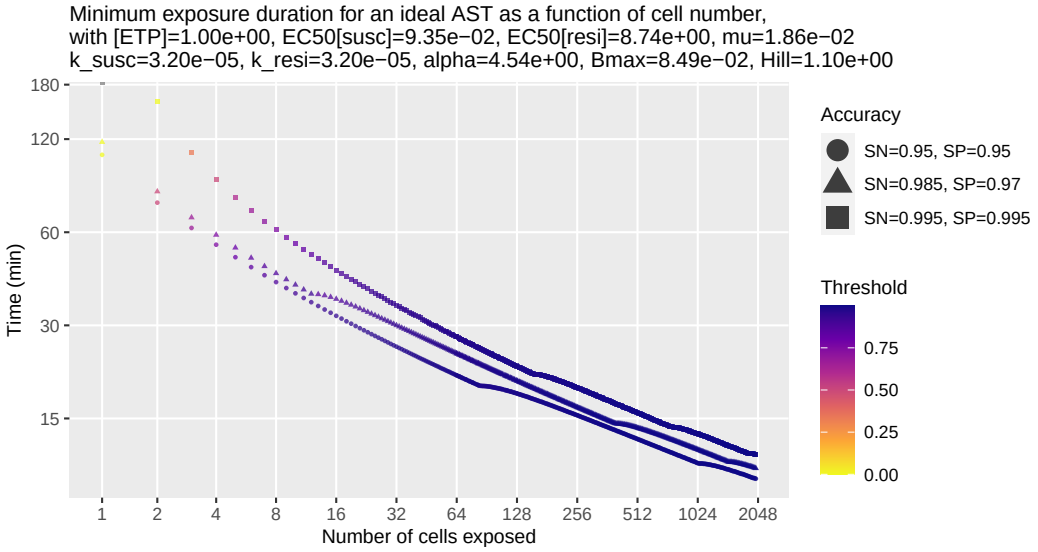
$$C(\xi, \theta) = \left(F_L[\lfloor \xi L_0 \rfloor; \mathrm{EC}_{50}[\mathrm{Pos}], \theta],\ \overline{F_L}[\lfloor \xi L_0 \rfloor; \mathrm{EC}_{50}[\mathrm{Neg}], \theta] \right) \tag{5.39}$$

This curve is not smooth, but rather advances to a new "achievable accuracy point" P_l in ROC space every time ξL_0 crosses an integer value. There are $L_0 + 1$ integer values l between 0 and L_0, each associated with a unique P_l. $C(\xi, \theta)$, and the location of the P_l, shifts as T is varied. ($C(\xi, \theta)$ also shifts when any of the other parameters in θ are varied.)

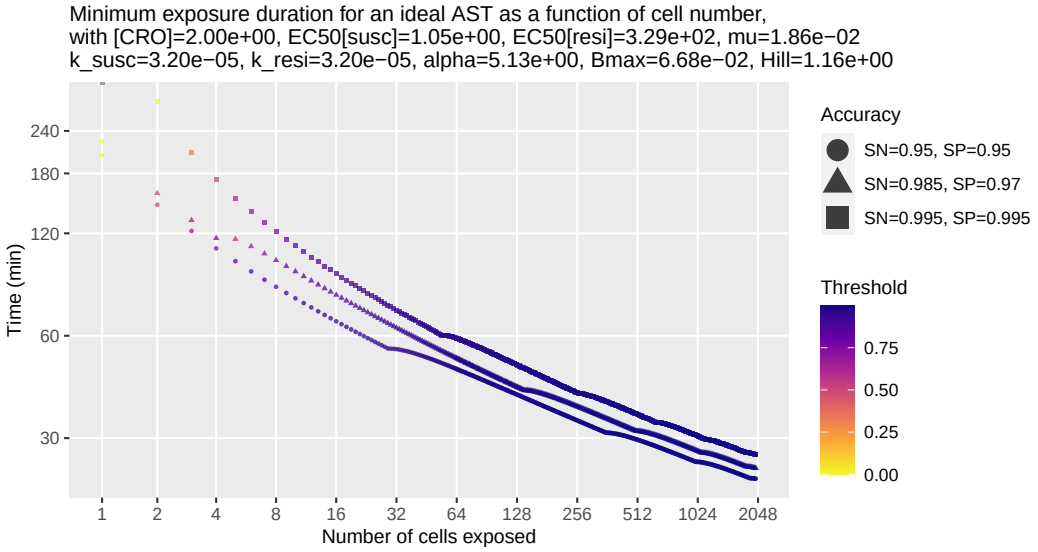
The given target assay performance was specified as a tuple $\hat{P} = (\hat{SN}, \hat{SP})$ of a particular expected sensitivity value $\hat{SN}$ and expected sensitivity value $\hat{SP}$. At a certain T, the point ($\hat{SN}$, $\hat{SP}$) may (or may never) lie along the convex hull of the $L_0 + 1$ points. This $T_{\min}$ is the minimum exposure duration at which it is possible to classify the two strains. The threshold value $\xi_{T_{\min}}$ at $T_{\min}$ is not the only threshold value that can satisfy the assay performance $\hat{P}$, nor does it achieve the highest assay performance as measured by $-R(H)$ or $\mathrm{Accu}[\lambda; \theta]$.

The solutions for each L_0 were found computationally as follows: For a given L_0 and T, use a binary search algorithm to find the closest pair of points P_{lower} and P_{upper} that straddled $\hat{SN}$. Then, the cross-product between the vectors $\overline{P_{\text{lower}}\ \hat{P}}$ and $\overline{P_{\text{lower}}\ P_{\text{upper}}}$ determines whether $\hat{P}$ lies within the convex hull of $\mathbb{C}$. For a given L_0,

[19] The convention of plotting the false positive rate $FPR = 1 - SP$ as the x-axis was not followed.



(a) Ertapenem



(b) Ceftriaxone

Figure 5.2: Minimum exposure duration for phenotypic AST as a function of initial inoculum. Note that the contour lines shown in Figure 5.3 depict the minimum number of cells as a function of exposure duration.

the inclusion of $\hat{P}$ in the convex hull is not a monotonic function (see Figure 5.3). Therefore, to find the minimum T where $\hat{P}$ is reached, the time axis was searched linearly starting with $T = 0$ and progressing by a small time step[20]. Repeating this linear search of the time axis for each L_0, for three different $\hat{P}$, and for two antibiotics (each at one dosage only) generated Figure 5.2.

In Figure 5.2, the SN=0.95, SP=0.95 isocline represents the FDA benchmark for an overall accuracy of 90%, assuming 50% prevalence for resistant strains. The SN=0.985, SP=0.97 isocline represents the FDA benchmark for no more than 1.5% very major errors and 3% major errors.

An alternative depiction of the ideal assay isosurface is to vary the exposure duration T, then find the minimum inoculum L_0 required to reach the given accuracy at the given dose. A similar algorithm was followed, differing from Figure 5.2 in that the search along the inoculum axis was nested inside the loop over the time axis.

Table 5.2 shows values of L_0 found by this second algorithm for selected T, $\hat{SN}$, and $\hat{SP}$, conditioned on the fitted parameters from Table 3.3. The level sets found by the second algorithm for various antibiotic dosages are also visible in Figure 5.3.

5.3 Biological stochasticity is so far sufficient to explain the limits of digital filtration AST

The preceding sections of this chapter have built up to Figure 5.2, in which the ideal performance of a filtration AST assay with one endpoint measurement was defined and quantitatively predicted. All parameter values used in the predictions were fitted to bulk filtration AST data only.

The main question of academic interest now asks whether the deductive reasoning that led to the model in Figure 5.2 is sufficiently useful for designing future assays. Since I have not undergone many rounds of commercial/industrial assay development in this thesis, I cannot definitively answer whether modeling is truly useful in practice. However, in the spirit of Popperian falsifiability, I can at least assess whether my model is insufficiently accurate for future assays.

Since digital filtration AST gives more information than bulk filtration AST, and because no digital filtration AST data was used to predict assay performance, I will compare the digital filtration AST results to the predictions of the model of Figure 5.2.

[20] Theoretically, the correctness of this brute-force approach is not guaranteed if the time step is too large. A more rigorous algorithm is left as future work.

Antibiotic	Time (min)	Sensitivity	Specificity	N
ETP	15	0.95	0.95	304
ETP	15	0.985	0.97	442
ETP	15	0.995	0.995	714
ETP	30	0.95	0.95	32
ETP	30	0.985	0.97	56
ETP	30	0.995	0.995	92
ETP	60	0.95	0.95	5
ETP	60	0.985	0.97	8
ETP	60	0.995	0.995	12
ETP	120	0.95	0.95	1
ETP	120	0.985	0.97	1
ETP	120	0.995	0.995	3
CRO	30	0.95	0.95	355
CRO	30	0.985	0.97	528
CRO	30	0.995	0.995	926
CRO	60	0.95	0.95	42
CRO	60	0.985	0.97	57
CRO	60	0.995	0.995	97
CRO	120	0.95	0.95	5
CRO	120	0.985	0.97	7
CRO	120	0.995	0.995	11

Table 5.2: Minimum number of cells needed for phenotypic AST.

Figure 5.3 shows the results of the comparison between the actual and ideal filtration AST performance.

The PE_{MLE} metric was chosen to be the final assay result because it is defined for each experimental condition, and its performance metric space is therefore more populated with the available data than other metrics' spaces.

The fill color of each point is the PE_{MLE} value. An experimental condition was called as susceptible if the PE_{MLE} metric was above 0.07 for either antibiotic. Note however that only the shape of the point is affected by the 0.07 PE_{MLE} threshold.

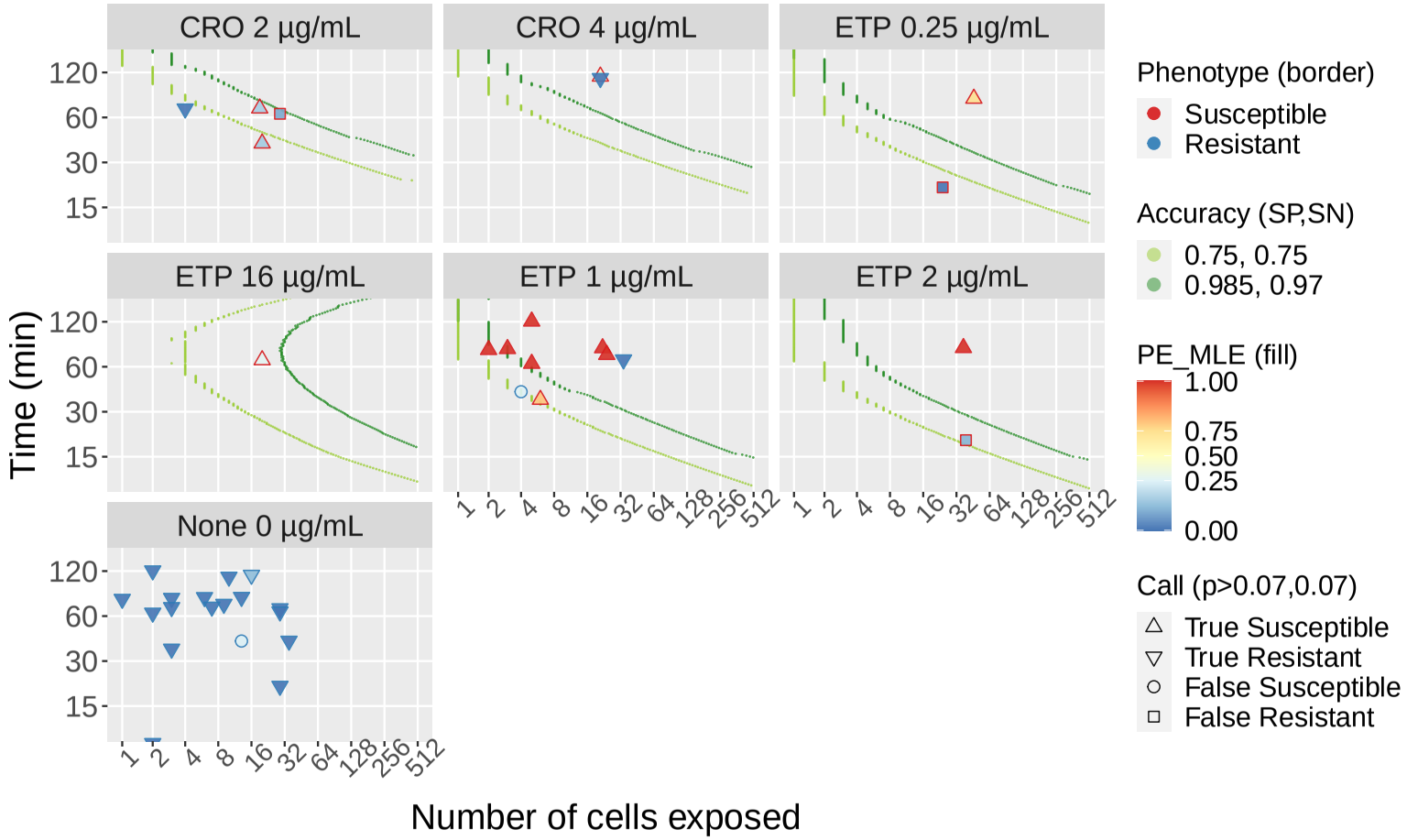


Figure 5.3: Comparison of actual digital filtration AST results to fitted ideal assay accuracy. See text for further details.

The use of different shapes was meant to ease interpretation. The effect of alternative thresholds can be inferred from the degree of color mismatch between the border and fill of each point. (Note that it is not necessary for antibiotics to share the same PE_{MLE} threshold. However, in this figure, the optimal thresholds were coincidentally both close to 0.07.)

Since the exact number of bacteria exposed in each digital filtration experimental condition cannot be exactly controlled by a serial dilution, the depicted x-axis position of each actual data point is the maximum likelihood estimate for the number of bacteria exposed from the digital filtration AST results, without taking into account the concurrent growth control. In contrast with the actual data points, the accuracy contour lines, representing the ideal performance, are calculated assuming an exact number of cells exposed.

Of the five errors incurred when using the optimal thresholds of 0.07 and 0.07 (circle and square shapes), all occurred when the specimen assayed lay to the left or below the SN=0.985, SP=0.97 isocline representing the FDA requirements for AST assays. There are seven potential errors (triangles with fill colors that do not match the shape border) that were still correctly called using the 0.07 thresholds, of which two occurred to the right of the SN=0.985, SP=0.97 isocline. Most of these errors (9 of 12) were between the SN=0.75, SP=0.75 isocline depicted for scale and the SN=0.985, SP=0.97 isocline.

The match between the model and data can be quantified by calculating a likelihood $\mathcal{L}$ of the observed data in Figure 5.3. I will treat this likelihood as a p-value of a hypothesis test, although this may violate some philosophical definitions of a p-value favored by traditional statisticians. Specifically, I am testing the hypothesis that the ideal assay model could have given rise to all 39 observed classification results, at the target/estimated antibiotic dose, inoculum, exposure duration of each of the observed points. To summarize the unspecified optimal classification algorithm threshold of an ideal assay, I will define $\mathcal{L}$ to be the likelihood if the threshold $\xi_{\hat{p}_{MLE}} \in [0, 1]$ used for the observed data were applied to the ideal assay as well. It is assumed that the fitted pharmacodynamic parameter values are part of the model.

For the i-th data point, the likelihood $\mathcal{L}_i$ of making the observed call is

$$\begin{aligned}\mathcal{L}_i &= \Pr(c_{i,\text{obs}}|c_{i,\text{true}};\theta_i) \\ &= \begin{cases} \overline{F_L}[\lfloor \xi_{\hat{p}_{MLE}} \hat{c}_{\text{MLE}} N_i \rfloor; \text{EC}_{50}[\text{Susc}], \theta_i] & \text{if } c_{i,\text{obs}} = \text{Susc}, c_{i,\text{true}} = \text{Susc} \\ F_L[\lfloor \xi_{\hat{p}_{MLE}} \hat{c}_{\text{MLE}} N_i \rfloor; \text{EC}_{50}[\text{Susc}], \theta_i] & \text{if } c_{i,\text{obs}} = \text{Resi}, c_{i,\text{true}} = \text{Susc} \\ \overline{F_L}[\lfloor \xi_{\hat{p}_{MLE}} \hat{c}_{\text{MLE}} N_i \rfloor; \text{EC}_{50}[\text{Resi}], \theta_i] & \text{if } c_{i,\text{obs}} = \text{Susc}, c_{i,\text{true}} = \text{Resi} \\ F_L[\lfloor \xi_{\hat{p}_{MLE}} \hat{c}_{\text{MLE}} N_i \rfloor; \text{EC}_{50}[\text{Resi}], \theta_i] & \text{if } c_{i,\text{obs}} = \text{Resi}, c_{i,\text{true}} = \text{Resi}, \end{cases} \end{aligned} \tag{5.40}$$

where $\xi_{\hat{p}_{MLE}} = 0.07$ is a threshold for PE_{MLE} values; $\hat{c}_{\text{MLE}}$ is the maximum likelihood estimate (equation 4.72) for the cell concentration for the i-th datum (representing the i-th condition); N_i is the number of wells in the i-th condition; $c_{\text{obs}} \in \mathbb{C} = \{\text{Susc}, \text{Resi}\}$ is the susceptibility phenotype observed by digital filtration AST; $c_{\text{true}} \in \mathbb{C}$ is the known susceptibility phenotype called by CLSI-compliant broth microdilution at UCLA; θ_i stands for all the model parameters for the i-th condition, including the classification threshold $\lambda = 0.07$, the i-th inoculum L_{0i}, the i-th exposure duration t_i, the i-th dose A_i, and $\text{EC}_{50}[c_{i,\text{true}}]$; the magnitude of resistance $\text{EC}_{50}[c_{i,\text{true}}]$ changes value depending on $c_{i,\text{true}}$; and F_L is defined in equation 5.18.

Lastly, I also assume for simplicity that the data are independent and identically distributed, so that $\mathcal{L} = \prod_i \mathcal{L}_i$.

For all 39 conditions (or PE_{MLE} outputs) of digital filtration AST included in this thesis, $\mathcal{L} = 0.0000010855$. This is a low likelihood, less than 0.001282, the traditional arbitrary cutoff of 0.05 after the Bonferroni correction for multiple hypothesis testing. The multiple hypothesis correction is justified since 39 distributions were tested to make one conclusion.

Looking more carefully at the data, however, one notices that one early experiment (Figure 4.7a) contributed two outliers that pushed the overall likelihood low, with individual likelihoods of 0.01800 and 0.005675, compared to the median individual likelihood of 0.9930. In these two experiments, killed resistant cells or untreated cells were observed at short durations and doses, when no extracellular nucleic acids were expected under the assumed model. This extracellular signal could have easily arises from an effect not modeled. Possible nuisance effects include cross contamination of wells during extraction or PCR, or the culture conditions of the batch culture may have promoted cell disruption not related to antibiotic activity.

Upon removing the outlier experiment of Figure 4.7a, $\mathcal{L}_1 = 0.010626$. This likelihood is relatively high considering that the geometric mean of the likelihood for each datum is $0.010626^{\frac{1}{37}} = 0.884$. Specifically, $\mathcal{L}$ is higher than 0.001351, the conventional significance cutoff of 0.05 after the Bonferroni correction for 37 hypotheses.

Thus, I conclude that the observed digital filtration AST results, with the exception of two conditions of one experiment, are consistent with the ideal performance for any phenotypic filtration AST on the systems so far examined. When the two outlier conditions, and the factors causing their outlier behavior, are included as part of the assay, then I conclude that the observed digital filtration AST results are not compatible with the ideal performance, specifically that the observed assay performs worse than its ideal performance due to those factors causing outlier behaviors. It should be noted, however, that in each of the above two situations (concluding that the actual performance is consistent with the ideal performance or concluding that the actual performance is worse than ideal), there is an alternative conclusion that still needs to be considered.

In the situation where the observed data are considered compatible with the ideal predictions, especially if the number of data points is low, it remains possible that the actual performance differs and that the concordance of data and theory arose by chance. After more observations, the observed actual performance would regress to the mean, and the hidden difference would be revealed. On the other hand, as the number of data points for observed performance increases to infinity, any difference, no matter how small, from the ideal performance would eventually become discernible and statistically significant. The difference may not be practically relevant, though. To balance between these two possible conclusions, a performance gap that is practically meaningful should be chosen *a priori*, and then a statistical power analysis performed to estimate the minimum number of data points required.

The power of a hypothesis test is the probability that the null hypothesis is rejected when the alternative hypothesis is in fact true. In the typical statistical power analysis, the hypothesis test asks if two population means differ by a certain amount. There are an infinite number of alternative hypotheses that could be true, but all the alternative hypotheses are parameterized by one position parameter and often parameters for normal distribution standard deviations. In such contexts, the question the hypothesis testing needs to answer does not involve higher-level questions such as whether the normal distribution is a good model for the populations sam-

pled. Under these circumstances, the statistical power, effect size, sample size, and significance threshold become related in a deterministic equation.

Applying the two-population power analysis framework to our hypothesis is difficult, however. The questions with which this thesis grapples do involve questions about whether fixed pharmacological parameters were properly fitted and whether the specific ideal assay model is a close representation of reality. First, there is no single position parameter that we can vary to distinguish our null and alternative hypotheses. Instead, the ideal assay model is a function of at least 13 parameters. One could choose a significant difference for each parameter of our ideal model, and each parameter would either be shifted by this difference or not. A power analysis would then need to cover every possible combination of parameter value shifts. Secondly, we are also interested in alternative hypotheses where the structure of the model changes, beyond shifts in parameter values. For example, if any of the equations used to define our ideal assay model were altered, that model mismatch would be a meaningful explanation for a mismatch between the ideal and actual performances. There is no way, at least reported in the literature, to easily parameterize an infinite collection of all plausible mathematical models. A third complication is that our experiments test multiple hypotheses, that is, multiple points in the parameters space, to answer one overall hypothesis about model adequacy. The definition of statistical power needs to be generalized across these disparate conditions. Unfortunately, deriving a new type of power analysis was left as future work due to time constraints.

My null hypothesis states that a match between the ideal accessibility AST model and actual digital filtration AST data exists. Because the data was consistent with the match, I have failed to reject this null hypothesis. The power analysis would have altered my belief that certain alternative hypothesis are false, decreasing it if the power analysis did suggest the sample size was sufficiently large, and increasing it otherwise at the expense of my belief in the null hypothesis. The more my belief in alternative hypotheses decreases, ideally due to more data supporting the null hypothesis, the more the null hypothesis appears to be "necessary" to explain the data, even if it can never be completely accepted. But regardless of the power analysis results, the main conclusion still holds that the ideal assay model is sufficient to explain the currently-known performance of manually-performed digital filtration AST.

In the situation where the observed performance of the digital filtration AST is

considered worse than the predicted performance, one would like to conclude that the realized protocol has limitations (known and/or unknown). However, an alternative conclusion would be that the model is unsuitably specified. The difference between these conclusions lies in whether one intellectually "blames" the performance gap on the actual experiment or on the ideal model. In the case where the model's validity is rejected, any level of modeling could be the source of the unrealistic predictions. For example, inclusion of originally omitted noise variables or influence quantities may give rise to a lower ideal performance. The resulting more complicated model would be a more relevant benchmark, but whether it is more useful than the current simpler benchmark depends on how readily the additional variables can be fitted to data. The pharmacodynamic model (equation 5.15) and the fitted parameters constitute a part of the overall model which has a reasonable likelihood for future improvement due to their justification by empirical observation instead of deduction. Any underfitting or poor fitting of the pharmacodynamic parameters from the bulk data would need to be overcome either with additional or cleaner bulk filtration AST data or by a more complicated model. Differing pharmacodynamics of single cells compared to bulk populations could also give rise to an irrelevant model, but the lack of an inoculum effect in the short term bulk filtration AST data and the lack of evidence for sub-populations with different pharmacodynamics makes this interpretation less likely.

In the case where the model of the ideal assay is deemed reasonably specified, then the gap between the actual and ideal performance indicates a flaw in the assay design, present as unmodeled variables or "influence quantities." Knowing the magnitude of these flaws is inherently useful to an engineer. In fact, as prior sections have shown, the following specific factors are known to cause non-ideal performance in the filtration AST protocol: 1) not being able to guarantee the observation of single cells due to lack of control of the number of initial cells per partition and the inability to create further subpartitions after the incubation, and 2) not being able to infer the numbers or mass of the cells after incubation due to high RT-PCR noise, 3) having lysis inefficiency observation error, and 4) being manually performed and therefore inconsistent compared to robotic machines due to unknown factors. The lysis inefficiency observation error (reason 3) in particular confounds the ability to assess model adequacy because higher probability of missing cells decreases the inferred number of cells present, shifting the observed points to the left in Figure 5.3. Counteracting this tendency is the fact that the lysis inefficiency observation error biases toward over-estimating the probability of cell death, which would lead

to more false susceptible calls. Yet, a preponderance of false susceptible calls is not observed in the actual data set, so we fail to reject the existence of a strong effect from the lysis inefficiency observation error.

The parsimony of my current model, the known existence of variability during execution, and the lack of strong evidence for potential flaws in the current ideal model therefore lead me to interpret any gap between the actual and ideal performance, where the actual performance is worse, as the consequence of variable execution of my protocol for digital filtration AST rather than a repudiation of the usefulness of modeling.

Lastly, in a situation where the observed performance of the digital filtration AST had been better than predictions, then the data would not be compatible with the predictions. Somewhere in the long chain of justifications for model choices and assumptions of the preceding sections, an assumed obstacle did not exist in reality. One could have interpreted the faulty assumption either as a failure of the modeler's deductive reasoning or as a windfall to the engineer in exceeding a valid deductive prediction.

5.4 Future work

Of the four caveats[21] to section 5.3's conclusion for worse actual performances, caveats 1 and 3 can be addressed by further experiments. Although the analyses that would address caveat 3 is present in section 4.18, they were not included in this thesis partly because the metrics defined across multiple conditions outputted fewer resulting points, and they offer no further insight beyond what can be gleaned from Figure 5.3. Caveat 2 and 4 require richer models of the specimen system. For caveat 4, the Markov birth-death model or an age-structured kinetic theory model would be preferred. These more complicated models would add incremental improvements to the truthfulness of the ideal performance predictions. However, deriving their optimal classifiers would probably involve significant mathematical labor, and so these models of ideal assays were not pursued in this thesis. Lastly, although I have made efforts to make my framework of phenotypic assay modeling as general and comprehensive as possible, the practice of mathematical modeling inherently involves subjective choices. Applying the four mapping framework to other phenotypic assays beyond AST will likely require further critical thinking and possibly further generalizations or revisions of the framework herein.

[21] 1) the low number of data points, 2) not modeling lysis inefficiency observation error when inferring inoculum, 3) the choice of metric, and 4) the lack of growth in the ideal system

Bibliography

[1] C. J. L. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. Kashef Hamadani, E. A. P. Kumaran, B. McManigal, S. Achalapong, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F.-X. Babin, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, J. A. Berkley, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, R. Clotaire Donatien, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. J. Day, M. De Luca, K. Dokova, A. Dramowski, S. J. Dunachie, T. Duong Bich, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, C. Garcia, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, Y. Hsia, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, A. W. J. Jenney, M. Khorana, S. Khusuwan, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, K. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, A. Manoharan, F. Marks, J. May, M. Mayxay, N. Mturi, T. Munera-Huertas, P. Musicha, L. A. Musila, M. M. Mussi-Pinhata, R. N. Naidu, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, T. J. Ochoa, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, P. Ounchanum, G. D. Pak, J. L. Paredes, A. Y. Peleg, C. Perrone, T. Phe, K. Phommasone, N. Plakkal, A. Ponce-de Leon, M. Raad, T. Ramdin, S. Rattanavong, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, V. D. Rosenthal, K. E. Rudd, N. Russell, H. S. Sader, W. Saengchan, J. Schnall, J. A. G. Scott, S. Seekaew, M. Sharland, M. Shivamallappa, J. Sifuentes-Osornio, A. J. Simpson, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Tigoi, C. Turner, P. Turner, H. R. Van Doorn, S. Velaphi, A. Vongpradith, M. Vongsouvath, H. Vu, T. Walsh, J. L. Walson, S. Waner, T. Wangrangsimakul, P. Wannapinij, T. Wozniak, T. E. M. W. Young Sharma, K. C. Yu, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek, and M. Naghavi, "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis," *The Lancet*, vol. 399, pp. 629–655, Feb. 2022.

[2] A. Dierig, C. Berger, P. K. A. Agyeman, S. Bernhard-Stirnemann, E. Giannoni, M. Stocker, K. M. Posfay-Barbe, A. Niederer-Loher, C. R. Kahlert, A. Donas, P. Hasters, C. Relly, T. Riedel, C. Aebi, L. J. Schlapbach, U. Heininger, and Swiss Pediatric Sepsis Study (last), "Time-to-positivity of blood cultures in children with sepsis," *Frontiers in Pediatrics*, vol. 6, 2018.

[3] E. M. Ransom, Z. Alipour, M. A. Wallace, and C.-A. D. Burnham, "Evalua-

tion of optimal blood culture incubation time to maximize clinically relevant results from a contemporary blood culture instrument and media system," *Journal of Clinical Microbiology*, vol. 59, pp. e02459–20, Feb. 2021.

[4] Karen C. Carroll, Michael A. Pfaller, Marie Louise Landry, Alexander J. McAdam, Robin Patel, Sandra S. Richter, and David W. Warnock, *Manual of Clinical Microbiology*. ASM Books, [N.p.]: ASM Press, 2019.

[5] Y. P. Tabak, L. Vankeepuram, G. Ye, K. Jeffers, V. Gupta, and P. R. Murray, "Blood culture turnaround time in U.S. acute care hospitals and implications for laboratory process optimization," *Journal of Clinical Microbiology*, vol. 56, pp. e00500–18, Nov. 2018.

[6] Centers for Disease Control and Prevention (U.S.), "Antibiotic resistance threats in the United States, 2019," tech. rep., US Department of Health and Human Services, Centres for Disease Control and Prevention, Atlanta, GA, Dec. 2019.

[7] S. Chandrasekaran, A. Abbott, S. Campeau, B. L. Zimmer, M. Weinstein, L. Thrupp, J. Hejna, L. Walker, T. Ammann, T. Kirn, R. Patel, and R. M. Humphries, "Direct-from-blood-culture disk diffusion to determine antimicrobial susceptibility of Gram-negative bacteria: Preliminary report from the Clinical and Laboratory Standards Institute Methods Development and Standardization Working Group," *Journal of Clinical Microbiology*, vol. 56, pp. e01678–17, Mar. 2018.

[8] P. P. Bourbeau and J. K. Pohlman, "Three days of incubation may be sufficient for routine blood cultures with BacT/Alert FAN blood culture bottles," *Journal of Clinical Microbiology*, vol. 39, pp. 2079–2082, June 2001.

[9] J. C. L. B. d. Quirós, C. Sánchez-Carrillo, E. Cercenado, C. Fron, P. Catalán, and E. Bouza, "Clinical impact of a 5-day versus 7-day blood culture incubation period," *Clinical Microbiology and Infection*, vol. 2, pp. 143–145, Oct. 1996.

[10] M. R. Jacobs, J. D. Colson, and D. D. Rhoads, "Recent advances in rapid antimicrobial susceptibility testing systems," *Expert Review of Molecular Diagnostics*, vol. 21, pp. 563–578, June 2021.

[11] A. van Belkum, C.-A. D. Burnham, J. W. A. Rossen, F. Mallard, O. Rochas, and W. M. Dunne, "Innovative and rapid antimicrobial susceptibility testing systems," *Nature Reviews Microbiology*, vol. 18, pp. 299–311, May 2020.

[12] W. G. Pitt, M. Alizadeh, G. A. Husseini, D. S. McClellan, C. M. Buchanan, C. G. Bledsoe, R. A. Robison, R. Blanco, B. L. Roeder, M. Melville, and A. K. Hunter, "Rapid separation of bacteria from blood—review and outlook," *Biotechnology Progress*, vol. 32, pp. 823–839, July 2016.

[13] R. Datar, S. Orenga, R. Pogorelcnik, O. Rochas, P. J. Simner, and A. van Belkum, "Recent advances in rapid antimicrobial susceptibility testing," *Clinical Chemistry*, vol. 68, pp. 91–98, Jan. 2022.

[14] J. Q. Boedicker, L. Li, T. R. Kline, and R. F. Ismagilov, "Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics," *Lab on a Chip*, vol. 8, no. 8, pp. 1265–1272, 2008.

[15] F. Zhang, J. Jiang, M. McBride, Y. Yang, M. Mo, R. Iriya, J. Peterman, W. Jing, T. Grys, S. E. Haydel, N. Tao, and S. Wang, "Direct antimicrobial susceptibility testing on clinical urine samples by optical tracking of single cell division events," *Small*, vol. 16, no. 52, p. 2004148, 2020.

[16] World Health Organization, *Blood Donor Selection: Guidelines on Assessing Donor Suitability for Blood Donation*. Geneva: World Health Organization, 2012.

[17] J. H. Kang, M. Super, C. W. Yung, R. M. Cooper, K. Domansky, A. R. Graveline, T. Mammoto, J. B. Berthet, H. Tobin, M. J. Cartwright, A. L. Watters, M. Rottman, A. Waterhouse, A. Mammoto, N. Gamini, M. J. Rodas, A. Kole, A. Jiang, T. M. Valentin, A. Diaz, K. Takahashi, and D. E. Ingber, "An extracorporeal blood-cleansing device for sepsis therapy," *Nature Medicine*, vol. 20, p. 1211, Oct. 2014.

[18] M. Arpi, M. W. Bentzon, J. Jensen, and W. Frederiksen, "Importance of blood volume cultured in the detection of bacteremia," *European Journal of Clinical Microbiology and Infectious Diseases*, vol. 8, pp. 838–842, Sept. 1989.

[19] B. Lamy, S. Dargère, M. C. Arendrup, J.-J. Parienti, and P. Tattevin, "How to optimize the use of blood cultures for the diagnosis of bloodstream infections? A state-of-the art," *Frontiers in Microbiology*, vol. 7, 2016.

[20] B. Jonsson, A. Nyberg, and C. Henning, "Theoretical aspects of detection of bacteraemia as a function of the volume of blood cultured," *APMIS*, vol. 101, pp. 595–601, Aug. 1993.

[21] N. G. Schoepp, T. S. Schlappi, M. S. Curtis, S. S. Butkovich, S. Miller, R. M. Humphries, and R. F. Ismagilov, "Rapid pathogen-specific phenotypic antibiotic susceptibility testing using digital LAMP quantification in clinical samples," *Science Translational Medicine*, vol. 9, p. eaal3693, Oct. 2017.

[22] T. Khazaei, J. T. Barlow, N. G. Schoepp, and R. F. Ismagilov, "RNA markers enable phenotypic test of antibiotic susceptibility in Neisseria gonorrhoeae after 10 minutes of ciprofloxacin exposure," *Scientific Reports*, vol. 8, p. 11606, Aug. 2018.

[23] E. S. Savela, N. G. Schoepp, M. M. Cooper, J. C. Rolando, J. D. Klausner, O. O. Soge, and R. F. Ismagilov, "Surfactant-enhanced DNA accessibility to nuclease accelerates phenotypic β-lactam antibiotic susceptibility testing of Neisseria gonorrhoeae," *PLoS biology*, vol. 18, p. e3000651, Mar. 2020.

[24] N. G. Schoepp, E. J. Liaw, A. Winnett, E. S. Savela, O. B. Garner, and R. F. Ismagilov, "Differential DNA accessibility to polymerase enables 30-minute phenotypic β-lactam antibiotic susceptibility testing of carbapenem-resistant Enterobacteriaceae," *PLoS Biology*, vol. 18, p. e3000652, Mar. 2020.

[25] J. T. Park and J. L. Strominger, "Mode of action of penicillin. Biochemical basis for the mechanism of action of penicillin and for its selective toxicity," *Science*, vol. 125, no. 3238, pp. 99–101, 1957.

[26] A. J. F. Egan, J. Errington, and W. Vollmer, "Regulation of peptidoglycan synthesis and remodelling," *Nature Reviews Microbiology*, vol. 18, pp. 446–460, Aug. 2020.

[27] A. Tomasz, "The mechanism of the irreversible antimicrobial effects of penicillins: How the beta-lactam antibiotics kill and lyse bacteria," *Annual Review of Microbiology*, vol. 33, no. 1, pp. 113–137, 1979.

[28] Z. Yao, D. Kahne, and R. Kishony, "Distinct single-cell morphological dynamics under beta-lactam antibiotics," *Molecular Cell*, vol. 48, pp. 705–712, Dec. 2012.

[29] T. M. File, Jr., "Treatment of community-acquired pneumonia in adults who require hospitalization," in *UpToDate* (J. A. Ramirez, S. Bond, and P. Dieffenbach, eds.), Waltham, MA: UpToDate, Apr. 2023.

[30] R. Moehring and D. J. Anderson, "Gram-negative bacillary bacteremia in adults," in *UpToDate* (S. B. Calderwood and K. K. Hall, eds.), Waltham, MA: UpToDate, Apr. 2023.

[31] V. G. Fowler, Jr. and T. L. Holland, "Clinical approach to Staphylococcus aureus bacteremia in adults," in *UpToDate* (F. D. Lowy and E. L. Baron, eds.), Waltham, MA: UpToDate, Apr. 2023.

[32] J. Garnacho-Montero, A. Escoresca-Ortega, and E. Fernández-Delgado, "Antibiotic de-escalation in the ICU: How is it best done?," *Current Opinion in Infectious Diseases*, vol. 28, pp. 193–198, Apr. 2015.

[33] T. P. Van Boeckel, S. Gandra, A. Ashok, Q. Caudron, B. T. Grenfell, S. A. Levin, and R. Laxminarayan, "Global antibiotic consumption 2000 to 2010: An analysis of national pharmaceutical sales data," *The Lancet Infectious Diseases*, vol. 14, pp. 742–750, Aug. 2014.

[34] A. Versporten, P. Zarb, I. Caniaux, M.-F. Gros, N. Drapier, M. Miller, V. Jarlier, D. Nathwani, H. Goossens, A. Koraqi, I. Hoxha, S. Tafaj, D. Lacej, M. Hojman, R. E. Quiros, L. Ghazaryan, K. A. Cairns, A. Cheng, K. C. Horne, F. F. Doukas, T. Gottlieb, J. Alsalman, K. Magerman, G. Y. Marielle, A. D. Ljubovic, A. A. M. Coelho, A. C. Gales, E. Keuleyan, D. Sabuda, J. L. Boswell, J. M. Conly, A. Rojas, C. Carvajal, J. Labarca, A. Solano, C. R. Valverde, J. M. Villalobos-Vindas, I. Pristas, V. Plecko, N. Paphitou, E. Shaqiri, M.-L. Rummukainen, K. Pagava, I. Korinteli, T. Brandt, S. Messler, A. Enimil, E. Iosifidis, E. Roilides, M. S. Sow, S. Sengupta, J. V. George, A. Poojary, P. Patil, J. Soltani, Z. Jafarpour, H. Ameen, D. Fitzgerald, Y. Maor, M. Chowers, E. Temkin, S. Esposito, L. Arnoldo, S. Brusaferro, Y. Gu, F. D. El-Hajji, N. J. Kim, B. Kambaralieva, J. Pavare, L. Zarakauska, V. Usonis, S. Burokiene, I. Ivaskeviciene, G. Mijovic, N. Duborija-Kovacevic, K. Bondesio, K. Iregbu, O. Oduyebo, D. Raka, L. Raka, S. Rachina, M. A. Enani, M. A. Shehri, B. Carevic, G. Dragovac, D. Obradovic, A. Stojadinovic, L. Radulovic, J. E. Wu, G. W. T. Chung, H. H. Chen, P. A. Tambyah, D. Lye, S. H. Tan, T. M. Ng, H. L. Tay, M. L. Ling, M. P. Chlebicki, A. L. Kwa, W. Lee, B. Beović, A. Dramowski, H. Finlayson, J. Taljaard, G. Ojeda-Burgos, P. Retamar, J. Lucas, W. Pot, C. Verduin, J. Kluytmans, M. Scott, M. A. Aldeyab, B. McCullagh, C. Gormley, D. Sharpe, M. Gilchrist, L. Whitney, M. Laundy, D. Lockwood, S. B. Drysdale, J. Boudreaux, E. J. Septimus, N. Greer, G. Gawrys, E. Rios, and S. May, "Antimicrobial consumption and resistance in adult hospital inpatients in 53 countries: Results of an Internet-based global point prevalence survey," *The Lancet Global Health*, vol. 6, pp. e619–e629, June 2018.

[35] S. S. Magill, J. R. Edwards, Z. G. Beldavs, G. Dumyati, S. J. Janelle, M. A. Kainer, R. Lynfield, J. Nadle, M. M. Neuhauser, S. M. Ray, K. Richards, R. Rodriguez, D. L. Thompson, S. K. Fridkin, and for the Emerging Infections Program Healthcare-Associated Infections and Antimicrobial Use Prevalence Survey Team, "Prevalence of antimicrobial use in US acute care hospitals, May-September 2011," *JAMA*, vol. 312, pp. 1438–1446, Oct. 2014.

[36] Y. Zhang, Q. Wang, Y. Yin, H. Chen, L. Jin, B. Gu, L. Xie, C. Yang, X. Ma, H. Li, W. Li, X. Zhang, K. Liao, S. Man, S. Wang, H. Wen, B. Li, Z. Guo, J. Tian, F. Pei, L. Liu, L. Zhang, C. Zou, T. Hu, J. Cai, H. Yang, J. Huang, X. Jia, W. Huang, B. Cao, and H. Wang, "Epidemiology of carbapenem-resistant Enterobacteriaceae infections: Report from the China CRE Network," *Antimicrobial Agents and Chemotherapy*, vol. 62, pp. e01882–17, Jan. 2018.

[37] M. J. Satlin, L. Chen, G. Patel, A. Gomez-Simmonds, G. Weston, A. C. Kim, S. K. Seo, M. E. Rosenthal, S. J. Sperber, S. G. Jenkins, C. L. Hamula, A.-C. Uhlemann, M. H. Levi, B. C. Fries, Y.-W. Tang, S. Juretschko, A. D. Rojtman, T. Hong, B. Mathema, M. R. Jacobs, T. J. Walsh, R. A. Bonomo, and B. N. Kreiswirth, "Multicenter clinical and molecular epidemiologi-

cal analysis of bacteremia due to carbapenem-resistant Enterobacteriaceae (CRE) in the CRE epicenter of the United States," *Antimicrobial Agents and Chemotherapy*, vol. 61, pp. e02349–16, Mar. 2017.

[38] A. Iovleva and Y. Doi, "Carbapenem-resistant Enterobacteriaceae," *Clinics in Laboratory Medicine*, vol. 37, pp. 303–315, June 2017.

[39] World Health Organization, "Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis," Tech. Rep. WHO/EMP/IAU/2017.12, World Health Organization, Geneva, Sept. 2017.

[40] Centers for Disease Control and Prevention (U.S.), "Antibiotic resistance threats in the United States, 2013," tech. rep., US Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, GA, 2013. CS239559-B.

[41] H. Gelband, M. Miller-Petrie, S. Pant, S. Gandra, J. Levinson, D. Barter, A. White, and R. Laxminarayan, "The state of the world's antibiotics, 2015," Tech. Rep. 20153355488, Center for Disease Dynamics, Economics & Policy, Washington, DC, 2015.

[42] A. Y. Guh, S. N. Bulens, Y. Mu, J. T. Jacob, J. Reno, J. Scott, L. E. Wilson, E. Vaeth, R. Lynfield, K. M. Shaw, P. M. S. Vagnone, W. M. Bamberg, S. J. Janelle, G. Dumyati, C. Concannon, Z. Beldavs, M. Cunningham, P. M. Cassidy, E. C. Phipps, N. Kenslow, T. Travis, D. Lonsway, J. K. Rasheed, B. M. Limbago, and A. J. Kallen, "Epidemiology of carbapenem-resistant Enterobacteriaceae in 7 US communities, 2012–2013," *JAMA*, vol. 314, pp. 1479–1487, Oct. 2015.

[43] G. C. Cerqueira, A. M. Earl, C. M. Ernst, Y. H. Grad, J. P. Dekker, M. Feldgarden, S. B. Chapman, J. L. Reis-Cunha, T. P. Shea, S. Young, Q. Zeng, M. L. Delaney, D. Kim, E. M. Peterson, T. F. O'Brien, M. J. Ferraro, D. C. Hooper, S. S. Huang, J. E. Kirby, A. B. Onderdonk, B. W. Birren, D. T. Hung, L. A. Cosimi, J. R. Wortman, C. I. Murphy, and W. P. Hanage, "Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks," *Proceedings of the National Academy of Sciences*, vol. 114, pp. 1135–1140, Jan. 2017.

[44] L. M. Weiner, A. K. Webb, B. Limbago, M. A. Dudeck, J. Patel, A. J. Kallen, J. R. Edwards, and D. M. Sievert, "Antimicrobial-resistant pathogens associated with healthcare-associated infections: Summary of data reported to the national healthcare safety network at the centers for disease control and prevention, 2011–2014," *Infection Control & Hospital Epidemiology*, vol. 37, pp. 1288–1301, Nov. 2016.

[45] Task Force for Combating Antibiotic-Resistant Bacteria, "National action plan for combating antibiotic-resistant bacteria," tech. rep., U.S. Department

of Health and Human Services, U.S. Department of Defense, U.S. Department of Agriculture, Mar. 2015.

[46] World Health Organization, "Global action plan on antimicrobial resistance," tech. rep., World Health Organization, Geneva, 2015. WHA68/2015/REC/1, Annex 3.

[47] J. O'Neill, "Tackling drug-resistant infections globally: Final report and recommendations," tech. rep., Wellcome Trust, UK Department of Health, London, May 2016.

[48] L. D. Renner, J. Zan, L. I. Hu, M. Martinez, P. J. Resto, A. C. Siegel, C. Torres, S. B. Hall, T. R. Slezak, T. H. Nguyen, and D. B. Weibel, "Detection of ESKAPE bacterial pathogens at the point of care using isothermal DNA-based assays in a portable degas-actuated microfluidic diagnostic assay platform," *Applied and Environmental Microbiology*, vol. 83, pp. e02449–16, Feb. 2017.

[49] M. A. Poritz, A. J. Blaschke, C. L. Byington, L. Meyers, K. Nilsson, D. E. Jones, S. A. Thatcher, T. Robbins, B. Lingenfelter, E. Amiott, A. Herbener, J. Daly, S. F. Dobrowolski, D. H.-F. Teng, and K. M. Ririe, "FilmArray, an automated nested multiplex PCR system for multi-pathogen detection: Development and application to respiratory tract infection," *PLoS ONE*, vol. 6, p. e26047, Oct. 2011.

[50] J. S. Gootenberg, O. O. Abudayyeh, M. J. Kellner, J. Joung, J. J. Collins, and F. Zhang, "Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6," *Science*, vol. 360, pp. 439–444, Apr. 2018.

[51] C. R. Phaneuf, B. Mangadu, M. E. Piccini, A. K. Singh, and C.-Y. Koh, "Rapid, portable, multiplexed detection of bacterial pathogens directly from clinical sample matrices," *Biosensors*, vol. 6, p. 49, Dec. 2016.

[52] H. Ahn, B. S. Batule, Y. Seok, and M.-G. Kim, "Single-step recombinase polymerase amplification assay based on a paper chip for simultaneous detection of multiple foodborne pathogens," *Analytical Chemistry*, vol. 90, pp. 10211–10216, Sept. 2018.

[53] B. Raja, H. Goux, A. Marapadaga, S. Rajagopalan, K. Kourentzi, and R. Willson, "Development of a panel of recombinase polymerase amplification assays for detection of common bacterial urinary tract infection pathogens," *Journal of Applied Microbiology*, vol. 123, pp. 544–555, Aug. 2017.

[54] Clinical and Laboratory Standards Institute (CLSI), "Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically; Approved standard - Tenth Edition. M07-A10," Jan. 2015.

[55] L. B. Reller, M. Weinstein, J. H. Jorgensen, and M. J. Ferraro, "Antimicrobial susceptibility testing: A review of general principles and contemporary practices," *Clinical Infectious Diseases*, vol. 49, pp. 1749–1755, Dec. 2009.

[56] S. Blank and D. C. Daskalakis, "Neisseria gonorrhoeae—Rising infection rates, dwindling treatment options," *New England Journal of Medicine*, vol. 379, pp. 1795–1797, Nov. 2018.

[57] L. J. V. Piddock, "Assess drug-resistance phenotypes, not just genotypes," *Nature Microbiology*, vol. 1, pp. 1–2, July 2016.

[58] J. D. Bard and F. Lee, "Why can't we just use PCR? the role of genotypic versus phenotypic testing for antimicrobial resistance testing," *Clinical Microbiology Newsletter*, vol. 40, pp. 87–95, June 2018.

[59] A. van Belkum, T. T. Bachmann, G. Lüdke, J. G. Lisby, G. Kahlmeter, A. Mohess, K. Becker, J. P. Hays, N. Woodford, K. Mitsakakis, J. Moran-Gilad, J. Vila, H. Peter, J. H. Rex, and W. M. Dunne, "Developmental roadmap for antimicrobial susceptibility testing systems," *Nature Reviews Microbiology*, vol. 17, pp. 51–62, Oct. 2018.

[60] T. Kostić, M. Ellis, M. R. Williams, T. M. Stedtfeld, J. B. Kaneene, R. D. Stedtfeld, and S. A. Hashsham, "Thirty-minute screening of antibiotic resistance genes in bacterial isolates with minimal sample preparation in static self-dispensing 64 and 384 assay cards," *Applied Microbiology and Biotechnology*, vol. 99, pp. 7711–7722, Sept. 2015.

[61] S. Kalsi, M. Valiadi, M.-N. Tsaloglou, L. Parry-Jones, A. Jacobs, R. Watson, C. Turner, R. Amos, B. Hadwen, J. Buse, C. Brown, M. Sutton, and H. Morgan, "Rapid and sensitive detection of antibiotic resistance on a programmable digital microfluidic platform," *Lab on a Chip*, vol. 15, pp. 3065–3075, July 2015.

[62] R. Nakano, A. Nakano, Y. Ishii, T. Ubagai, T. Kikuchi-Ueda, H. Kikuchi, S. Tansho-Nagakawa, G. Kamoshida, X. Mu, and Y. Ono, "Rapid detection of the Klebsiella pneumoniae carbapenemase (KPC) gene by loop-mediated isothermal amplification (LAMP)," *Journal of Infection and Chemotherapy*, vol. 21, pp. 202–206, Mar. 2015.

[63] X. Mu, R. Nakano, A. Nakano, T. Ubagai, T. Kikuchi-Ueda, S. Tansho-Nagakawa, H. Kikuchi, G. Kamoshida, S. Endo, H. Yano, and Y. Ono, "Loop-mediated isothermal amplification: Rapid and sensitive detection of the antibiotic resistance gene ISAba1-blaOXA-51-like in Acinetobacter baumannii," *Journal of Microbiological Methods*, vol. 121, pp. 36–40, Feb. 2016.

[64] C. Hu, S. Kalsi, I. Zeimpekis, K. Sun, P. Ashburn, C. Turner, J. M. Sutton, and H. Morgan, "Ultra-fast electronic detection of antimicrobial resistance genes using isothermal amplification and Thin Film Transistor sensors," *Biosensors and Bioelectronics*, vol. 96, pp. 281–287, Oct. 2017.

[65] A. Cortegiani, V. Russotto, G. Graziano, D. Geraci, L. Saporito, G. Cocorullo, S. M. Raineri, C. Mammina, and A. Giarratano, "Use of Cepheid Xpert Carba-R® for rapid detection of carbapenemase-producing bacteria in abdominal septic patients admitted to intensive care unit," *PLoS One*, vol. 11, p. e0160643, Aug. 2016.

[66] BioFire Diagnostics (bioMérieux), "The BioFire® FilmArray® System." https://www.biofiredx.com/products/filmarray/.

[67] Food and Drug Administration, "510(k) Premarket Notification: (Xpert Carba-R) K160901," June 2016.

[68] C. Lasserre, L. De Saint Martin, G. Cuzon, P. Bogaerts, E. Lamar, Y. Glupczynski, T. Naas, and D. Tandé, "Efficient detection of carbapenemase activity in Enterobacteriaceae by matrix-assisted laser desorption ionization-time of flight mass spectrometry in less than 30 minutes," *Journal of Clinical Microbiology*, vol. 53, pp. 2163–2171, July 2015.

[69] P. Bogaerts, S. Yunus, M. Massart, T.-D. Huang, and Y. Glupczynski, "Evaluation of the BYG Carba test, a new electrochemical assay for rapid laboratory detection of carbapenemase-producing Enterobacteriaceae," *Journal of Clinical Microbiology*, vol. 54, pp. 349–358, Feb. 2016.

[70] A. M. Robinson, N. J. Medlicott, and J. E. Ussher, "The rapid detection of cefotaxime-resistant Enterobacteriaceae by HPLC," *Future Science OA*, vol. 2, p. FSO142, Dec. 2016.

[71] S. Bernabeu, L. Dortet, and T. Naas, "Evaluation of the β-CARBA™ test, a colorimetric test for the rapid detection of carbapenemase activity in Gram-negative bacilli," *Journal of Antimicrobial Chemotherapy*, vol. 72, pp. 1646–1658, June 2017.

[72] Y. Zhang, J.-E. Lei, Y. He, J. Yang, W. Wang, A. Wasey, J. Xu, Y. Lin, H. Fan, G. Jing, C. Zhang, and Y. Jin, "Label-free visualization of carbapenemase activity in living bacteria," *Angewandte Chemie (International Ed. in English)*, vol. 57, pp. 17120–17124, Dec. 2018.

[73] R. Santiso, M. Tamayo, J. Gosálvez, G. Bou, M. d. C. Fernández, and J. L. Fernández, "A rapid in situ procedure for determination of bacterial susceptibility or resistance to antibiotics that inhibit peptidoglycan biosynthesis," *BMC Microbiology*, vol. 11, p. 191, Aug. 2011.

[74] G. Bou, F. M. Otero, R. Santiso, M. Tamayo, M. d. C. Fernández, M. Tomás, J. Gosálvez, and J. L. Fernández, "Fast assessment of resistance to carbapenems and ciprofloxacin of clinical strains of Acinetobacter baumannii," *Journal of Clinical Microbiology*, vol. 50, pp. 3609–3613, Nov. 2012.

[75] K. Park, J. Jeong, S. Y. Yi, W. S. Lee, and Y. B. Shin, "FRET probe-based antibacterial susceptibility testing (F-AST) by detection of bacterial nucleases released by antibiotic-induced lysis," *Biosensors and Bioelectronics*, vol. 130, pp. 225–229, Apr. 2019.

[76] I. Faria-Ramos, M. J. Espinar, R. Rocha, J. Santos-Antunes, A. G. Rodrigues, R. Cantón, and C. Pina-Vaz, "A novel flow cytometric assay for rapid detection of extended-spectrum beta-lactamases," *Clinical Microbiology and Infection*, vol. 19, pp. E8–E15, Jan. 2013.

[77] C.-A. D. Burnham, R. A. Frobel, M. L. Herrera, and B. L. Wickes, "Rapid ertapenem susceptibility testing and Klebsiella pneumoniae carbapenemase phenotype detection in Klebsiella pneumoniae isolates by use of automated microscopy of immobilized live bacterial cells," *Journal of Clinical Microbiology*, vol. 52, pp. 982–986, Mar. 2014.

[78] I.-H. Su, W.-C. Ko, C.-H. Shih, F.-H. Yeh, Y.-N. Sun, J.-C. Chen, P.-L. Chen, and H.-C. Chang, "Dielectrophoresis system for testing antimicrobial susceptibility of Gram-negative bacteria to β-lactam antibiotics," *Analytical Chemistry*, vol. 89, pp. 4635–4641, Apr. 2017.

[79] W. Kang, S. Sarkar, Z. S. Lin, S. McKenney, and T. Konry, "Ultrafast parallelized microfluidic platform for antimicrobial susceptibility testing of Gram positive and negative bacteria," *Analytical Chemistry*, vol. 91, pp. 6242–6249, May 2019.

[80] M. F. Cansizoglu, Y. T. Tamer, M. Farid, A. Y. Koh, and E. Toprak, "Rapid ultrasensitive detection platform for antimicrobial susceptibility testing," *PLoS biology*, vol. 17, p. e3000291, May 2019.

[81] M. Mo, Y. Yang, F. Zhang, W. Jing, R. Iriya, J. Popovich, S. Wang, T. Grys, S. E. Haydel, and N. Tao, "Rapid antimicrobial susceptibility testing of patient urine samples using large volume free-solution light scattering microscopy," *Analytical Chemistry*, vol. 91, pp. 10164–10171, Aug. 2019.

[82] C. Lange, S. Schubert, J. Jung, M. Kostrzewa, and K. Sparbier, "Quantitative matrix-assisted laser desorption ionization–time of flight mass spectrometry for rapid resistance detection," *Journal of Clinical Microbiology*, vol. 52, pp. 4155–4162, Dec. 2014.

[83] J. M. Rolain, M. N. Mallet, P. E. Fournier, and D. Raoult, "Real-time PCR for universal antibiotic susceptibility testing," *Journal of Antimicrobial Chemotherapy*, vol. 54, pp. 538–541, Aug. 2004.

[84] K. E. Mach, R. Mohan, E. J. Baron, M.-C. Shih, V. Gau, P. K. Wong, and J. C. Liao, "A biosensor platform for rapid antimicrobial susceptibility testing directly from clinical samples," *Journal of Urology*, vol. 185, pp. 148–153, Jan. 2011.

[85] C. Halford, R. Gonzalez, S. Campuzano, B. Hu, J. T. Babbitt, J. Liu, J. Wang, B. M. Churchill, and D. A. Haake, "Rapid antimicrobial susceptibility testing by sensitive detection of precursor rRNA using a novel electrochemical biosensing platform," *Antimicrobial Agents and Chemotherapy*, vol. 57, pp. 936–943, Feb. 2013.

[86] A. Mezger, E. Gullberg, J. Göransson, A. Zorzet, D. Herthnek, E. Tano, M. Nilsson, and D. I. Andersson, "A general method for rapid determination of antibiotic susceptibility and species in bacterial infections," *Journal of Clinical Microbiology*, vol. 53, pp. 425–432, Feb. 2015.

[87] A. K. Barczak, J. E. Gomez, B. B. Kaufmann, E. R. Hinson, L. Cosimi, M. L. Borowsky, A. B. Onderdonk, S. A. Stanley, D. Kaur, K. F. Bryant, D. M. Knipe, A. Sloutsky, and D. T. Hung, "RNA signatures allow rapid identification of pathogens and antibiotic susceptibilities," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 6217–6222, Apr. 2012.

[88] R. P. Bhattacharyya, N. Bandyopadhyay, P. Ma, S. S. Son, J. Liu, L. L. He, L. Wu, R. Khafizov, R. Boykin, G. C. Cerqueira, A. Pironti, R. F. Rudy, M. M. Patel, R. Yang, J. Skerry, E. Nazarian, K. A. Musser, J. Taylor, V. M. Pierce, A. M. Earl, L. A. Cosimi, N. Shoresh, J. Beechem, J. Livny, and D. T. Hung, "Simultaneous detection of genotype and phenotype enables rapid and accurate antibiotic susceptibility determination," *Nature Medicine*, vol. 25, pp. 1858–1864, Dec. 2019.

[89] N. G. Schoepp, E. M. Khorosheva, T. S. Schlappi, M. S. Curtis, R. M. Humphries, J. A. Hindler, and R. F. Ismagilov, "Digital quantification of DNA replication and chromosome segregation enables determination of antimicrobial susceptibility after only 15 minutes of antibiotic exposure," *Angewandte Chemie (International Ed. in English)*, vol. 55, no. 33, pp. 9557–9561, 2016.

[90] Y. Ota, K. Furuhashi, T. Nanba, K. Yamanaka, J. Ishikawa, O. Nagura, E. Hamada, and M. Maekawa, "A rapid and simple detection method for phenotypic antimicrobial resistance in Escherichia coli by loop-mediated isothermal amplification," *Journal of Medical Microbiology*, vol. 68, no. 2, pp. 169–177, 2019.

[91] H. D. Marston, D. M. Dixon, J. M. Knisely, T. N. Palmore, and A. S. Fauci, "Antimicrobial resistance," *JAMA*, vol. 316, pp. 1193–1204, Sept. 2016.

[92] J. M. Hicks, R. Haeckel, C. P. Price, K. Lewandrowski, and A. H. B. Wu, "Recommendations and opinions for the use of point-of-care testing for hospitals and primary care: Summary of a 1999 symposium," *Clinica Chimica Acta*, vol. 303, pp. 1–17, Jan. 2001.

[93] M. Valiadi, S. Kalsi, I. G. F. Jones, C. Turner, J. M. Sutton, and H. Morgan, "Simple and rapid sample preparation system for the molecular detection

of antibiotic resistant pathogens in human urine," *Biomedical Microdevices*, vol. 18, p. 18, Feb. 2016.

[94] Centers for Disease Control and Prevention (CDC), "CDC & FDA Antibiotic Resistance Isolate Bank: Enterobacterales Carbapenem Breakpoint." https://wwwn.cdc.gov/ARIsolateBank/Panel/PanelDetail?ID=7.

[95] S. Pollett, S. Miller, J. Hindler, D. Uslan, M. Carvalho, and R. M. Humphries, "Phenotypic and molecular characteristics of carbapenem-resistant Enterobacteriaceae in a health care system in Los Angeles, California, from 2011 to 2013," *Journal of Clinical Microbiology*, vol. 52, pp. 4003–4009, Nov. 2014.

[96] Cepheid, "Xpert Carba-R." https://www.cepheid.com/en_US/tests/Healthcare-Associated-Infections/Xpert-Carba-R.

[97] P. M. Hawkey and D. M. Livermore, "Carbapenem antibiotics for serious infections," *BMJ*, vol. 344, p. e3236, May 2012.

[98] UCLA Clinical Microbiology, "Antimicrobial Susceptibility Summary 2019," tech. rep., UCLA Health System, Department of Pathology & Laboratory Medicine, Westwood, CA, 2019.

[99] E. Zankari, H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, F. M. Aarestrup, and M. V. Larsen, "Identification of acquired antimicrobial resistance genes," *Journal of Antimicrobial Chemotherapy*, vol. 67, pp. 2640–2644, Nov. 2012.

[100] Y.-Y. Chang, Y.-C. Chuang, L. K. Siu, T.-L. Wu, J.-C. Lin, P.-L. Lu, J.-T. Wang, L.-S. Wang, Y.-T. Lin, L.-J. Huang, and C.-P. Fung, "Clinical features of patients with carbapenem nonsusceptible Klebsiella pneumoniae and Escherichia coli in intensive care units: A nationwide multicenter study in Taiwan," *Journal of Microbiology, Immunology and Infection*, vol. 48, pp. 219–225, Apr. 2015.

[101] P. D. Tamma, K. E. Goodman, A. D. Harris, T. Tekle, A. Roberts, A. Taiwo, and P. J. Simner, "Comparing the outcomes of patients with carbapenemase-producing and non-carbapenemase-producing carbapenem-resistant Enterobacteriaceae bacteremia," *Clinical Infectious Diseases*, vol. 64, pp. 257–264, Feb. 2017.

[102] P. Rui and T. Okeyode, "National Ambulatory Medical Care Survey: 2015 state and national summary tables," tech. rep., National Center for Health Statistics, Centers for Disease Control and Control (U.S.), US Department of Health and Human Services, 2015.

[103] N. P. Braykov, M. R. Eber, E. Y. Klein, D. J. Morgan, and R. Laxminarayan, "Trends in resistance to carbapenems and third-generation cephalosporins

among clinical isolates of Klebsiella pneumoniae in the United States, 1999-2010," *Infection Control & Hospital Epidemiology*, vol. 34, pp. 259–268, Mar. 2013.

[104] A. C. Musta, K. Riederer, S. Shemes, P. Chase, J. Jose, L. B. Johnson, and R. Khatib, "Vancomycin MIC plus heteroresistance and outcome of methicillin-resistant Staphylococcus aureus bacteremia: Trends over 11 years," *Journal of Clinical Microbiology*, vol. 47, pp. 1640–1644, June 2009.

[105] Ö. Baltekin, A. Boucharin, E. Tano, D. I. Andersson, and J. Elf, "Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging," *Proceedings of the National Academy of Sciences*, vol. 114, pp. 9170–9175, Aug. 2017.

[106] J. Choi, J. Yoo, M. Lee, E.-G. Kim, J. S. Lee, S. Lee, S. Joo, S. H. Song, E.-C. Kim, J. C. Lee, H. C. Kim, Y.-G. Jung, and S. Kwon, "A rapid antimicrobial susceptibility test based on single-cell morphological analysis," *Science Translational Medicine*, vol. 6, pp. 267ra174–267ra174, Dec. 2014.

[107] H. Etayash, M. F. Khan, K. Kaur, and T. Thundat, "Microfluidic cantilever detects bacteria and measures their susceptibility to antibiotics in small confined volumes," *Nature Communications*, vol. 7, p. 12947, Oct. 2016.

[108] G. Longo, L. Alonso-Sarduy, L. M. Rio, A. Bizzini, A. Trampuz, J. Notz, G. Dietler, and S. Kasas, "Rapid detection of bacterial resistance to antibiotics using AFM cantilevers as nanomechanical sensors," *Nature Nanotechnology*, vol. 8, pp. 522–526, July 2013.

[109] S. E. Pidgeon and M. M. Pires, "Vancomycin-dependent response in live drug-resistant bacteria by metabolic labeling," *Angewandte Chemie (International Ed. in English)*, vol. 56, pp. 8839–8843, July 2017.

[110] M. R. L. Stone, M. S. Butler, W. Phetsang, M. A. Cooper, and M. A. T. Blaskovich, "Fluorescent antibiotics: New research tools to fight antibiotic resistance," *Trends in Biotechnology*, vol. 36, pp. 523–536, May 2018.

[111] S. Deshayes, W. Xian, N. W. Schmidt, S. Kordbacheh, J. Lieng, J. Wang, S. Zarmer, S. S. Germain, L. Voyen, J. Thulin, G. C. L. Wong, and A. M. Kasko, "Designing hybrid antibiotic peptide conjugates to cross bacterial membranes," *Bioconjugate Chemistry*, vol. 28, pp. 793–804, Mar. 2017.

[112] E. A. Phillips, T. J. Moehling, S. Bhadra, A. D. Ellington, and J. C. Linnes, "Strand displacement probes combined with isothermal nucleic acid amplification for instrument-free detection from complex samples," *Analytical Chemistry*, vol. 90, pp. 6580–6586, June 2018.

[113] S. Cai, C. Jung, S. Bhadra, and A. D. Ellington, "Phosphorothioated primers lead to loop-mediated isothermal amplification at low temperatures," *Analytical Chemistry*, vol. 90, pp. 8290–8294, July 2018.

[114] R. K. Daher, G. Stewart, M. Boissinot, and M. G. Bergeron, "Recombinase polymerase amplification for diagnostic applications," *Clinical Chemistry*, vol. 62, pp. 947–958, July 2016.

[115] B. J. Toley, I. Covelli, Y. Belousov, S. Ramachandran, E. Kline, N. Scarr, N. Vermeulen, W. Mahoney, B. R. Lutz, and P. Yager, "Isothermal strand displacement amplification (iSDA): A rapid and sensitive method of nucleic acid amplification for point-of-care diagnosis," *Analyst*, vol. 140, pp. 7540–7549, Nov. 2015.

[116] M. S. Reid, X. C. Le, and H. Zhang, "Exponential isothermal amplification of nucleic acids and assays for proteins, cells, small molecules, and enzyme activities: An EXPAR example," *Angewandte Chemie (International Ed. in English)*, vol. 57, pp. 11856–11866, Sept. 2018.

[117] S. Banoo, D. Bell, P. Bossuyt, A. Herring, D. Mabey, F. Poole, P. G. Smith, N. Sriram, C. Wongsrichanalai, R. Linke, R. O'Brien, M. Perkins, J. Cunningham, P. Matsoso, C. M. Nathanson, P. Olliaro, R. W. Peeling, and A. Ramsay, "Evaluation of diagnostic tests for infectious diseases: General principles," *Nature Reviews Microbiology*, vol. 4, pp. S21–S31, Sept. 2006.

[118] K. Matsuda, H. Tsuji, T. Asahara, Y. Kado, and K. Nomoto, "Sensitive quantitative detection of commensal bacteria by rRNA-targeted reverse transcription-PCR," *Applied and Environmental Microbiology*, vol. 73, pp. 32–39, Jan. 2007.

[119] E. Chern, S. Siefring, J. Paar, M. Doolittle, and R. Haugland, "Comparison of quantitative PCR assays for Escherichia coli targeting ribosomal RNA and single copy genes," *Letters in Applied Microbiology*, vol. 52, pp. 298–306, Mar. 2011.

[120] M. Adler, M. Anjum, D. I. Andersson, and L. Sandegren, "Influence of acquired β-lactamases on the evolution of spontaneous carbapenem resistance in Escherichia coli," *Journal of Antimicrobial Chemotherapy*, vol. 68, pp. 51–59, Jan. 2013.

[121] J. M. Zaengle-Barone, A. C. Jackson, D. M. Besse, B. Becken, M. Arshad, P. C. Seed, and K. J. Franz, "Copper influences the antibacterial outcomes of a β-lactamase-activated prochelator against drug-resistant bacteria," *ACS infectious diseases*, vol. 4, pp. 1019–1029, June 2018.

[122] R. Milo, R. Phillips, and N. Orme, *Cell Biology by the Numbers*. Garland Science, 2016.

[123] F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao, "The complete genome sequence of Escherichia coli K-12," *Science (New York, N.Y.)*, vol. 277, pp. 1453–1462, Sept. 1997.

[124] K. C. Huang, R. Mukhopadhyay, B. Wen, Z. Gitai, and N. S. Wingreen, "Cell shape and cell-wall organization in Gram-negative bacteria," *Proceedings of the National Academy of Sciences*, p. pnas.0805309105, Dec. 2008.

[125] M. Kalashnikov, J. C. Lee, J. Campbell, A. Sharon, and A. F. Sauer-Budge, "A microfluidic platform for rapid, stress-induced antibiotic susceptibility testing of Staphylococcus aureus," *Lab on a Chip*, vol. 12, pp. 4523–4532, Oct. 2012.

[126] C. B. Whitchurch, T. Tolker-Nielsen, P. C. Ragas, and J. S. Mattick, "Extracellular DNA required for bacterial biofilm formation," *Science*, vol. 295, pp. 1487–1487, Feb. 2002.

[127] C. Beloin, A. Roux, and J.-M. Ghigo, "Escherichia coli biofilms," *Current Topics in Microbiology and Immunology*, vol. 322, pp. 249–289, 2008.

[128] M. Kimmel and D. E. Axelrod, *Branching Processes in Biology*, vol. 19 of *Interdisciplinary Applied Mathematics*. New York, NY: Springer New York, 2015.

[129] R. R. Regoes, C. Wiuff, R. M. Zappala, K. N. Garner, F. Baquero, and B. R. Levin, "Pharmacodynamic functions: A multiparameter approach to the design of antibiotic treatment regimens," *Antimicrobial Agents and Chemotherapy*, vol. 48, pp. 3670–3676, Oct. 2004.

[130] J. W. Mouton and A. A. Vinks, "Pharmacokinetic/pharmacodynamic modelling of antibacterials in vitro and in vivo using bacterial growth and kill kinetics," *Clinical Pharmacokinetics*, vol. 44, pp. 201–210, Feb. 2005.

[131] A. Gelman, H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. Chapman & Hall/CRC texts in statistical science, Chapman and Hall/CRC, third edition ed., 2014.

[132] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch," *Science*, vol. 305, pp. 1622–1625, Sept. 2004.

[133] A. Nolting, T. D. Costa, K. H. Rand, and H. Derendorf, "Pharmacokinetic-pharmacodynamic modeling of the antibiotic effect of piperacillin in vitro," *Pharmaceutical Research*, vol. 13, pp. 91–96, Jan. 1996.

[134] S. Foerster, M. Unemo, L. J. Hathaway, N. Low, and C. L. Althaus, "Time-kill curve analysis and pharmacodynamic modelling for in vitro evaluation of antimicrobials against Neisseria gonorrhoeae," *BMC Microbiology*, vol. 16, p. 216, Sept. 2016.

[135] R. C. Li, "Simultaneous pharmacodynamic analysis of the lag and bactericidal phases exhibited by beta-lactams against Escherichia coli.," *Antimicrobial Agents and Chemotherapy*, vol. 40, pp. 2306–2310, Oct. 1996.

[136] M. Godin, F. F. Delgado, S. Son, W. H. Grover, A. K. Bryan, A. Tzur, P. Jorgensen, K. Payer, A. D. Grossman, M. W. Kirschner, and S. R. Manalis, "Using buoyant mass to measure the growth of single cells," *Nature Methods*, vol. 7, pp. 387–390, May 2010.

[137] F. C. Neidhardt, *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology.* American Society for Microbiology, 2nd ed ed., 2001.

[138] M. Kalashnikov, M. Mueller, C. McBeth, J. C. Lee, J. Campbell, A. Sharon, and A. F. Sauer-Budge, "Rapid phenotypic stress-based microfluidic antibiotic susceptibility testing of Gram-negative clinical isolates," *Scientific Reports*, vol. 7, p. 8031, Aug. 2017.

[139] Y. Zhang and H. Noji, "Digital bioassays: Theory, applications, and perspectives," *Analytical Chemistry*, vol. 89, pp. 92–101, Jan. 2017.

[140] R. A. Fisher, "On the interpretation of χ^2 from contingency tables, and the calculation of p," *Journal of the Royal Statistical Society*, vol. 85, no. 1, pp. 87–94, 1922.

[141] LGC Biosearch, "RapiDxFire thermostable reverse transcriptase manual," 2018. https://biosearchtech.a.bigcontent.io/v1/static/manual_307_RapiDxFire-reverse-transcriptase-manual.

[142] D. G. Kendall, "On the generalized "birth-and-death" process," *The Annals of Mathematical Statistics*, vol. 19, pp. 1–15, Mar. 1948.

[143] N. T. J. Bailey, *The elements of stochastic processes with applications to the natural sciences.* Wiley publication in applied statistics, Wiley, 1964.

[144] J. Coates, B. R. Park, D. Le, E. Şimşek, W. Chaudhry, and M. Kim, "Antibiotic-induced population fluctuations and stochastic clearance of bacteria," *eLife*, vol. 7, p. e32976, Mar. 2018.

[145] J. J. Li and X. Tong, "Statistical hypothesis testing versus machine learning binary classification: Distinctions and guidelines," *Patterns*, vol. 1, p. 100115, Oct. 2020.

[146] I. Greenberg, "A one-sided goodness-of-fit test for a multinomial population," *Journal of the American Statistical Association*, vol. 80, no. 391, pp. 558–562, 1985.

[147] International Organization for Standardization, *ISO 5725-1: 1994: Accuracy (Trueness and Precision) of Measurement Methods and Results-Part 1: General Principles and Definitions*. International Organization for Standardization, 1994.

[148] Joint Committee for Guides in Metrology, *Evaluation of measurement data—Guide to the expression of uncertainty in measurement. 100:2008E.* International Bureau of Weights, Measures (BIPM), International Electrotechnical Commission (IEC), International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), International Laboratory Accreditation Cooperation (ILAC), International Organization for Standardization (ISO), International Union of Pure and Applied Chemistry (IUPAC), International Union of Pure and Applied Physics (IUPAP), International Organization of Legal Metrology (OIML), Sept. 2008.

[149] Oxford English Dictionary, *"phenotype, n."*. Oxford University Press, 2022.

[150] D. C. Swinney, "Phenotypic vs. target-based drug discovery for first-in-class medicines," *Clinical Pharmacology & Therapeutics*, vol. 93, no. 4, pp. 299–301, 2013.

[151] J. G. Moffat, F. Vincent, J. A. Lee, J. Eder, and M. Prunotto, "Opportunities and challenges in phenotypic drug discovery: An industry perspective," *Nature Reviews Drug Discovery*, vol. 16, pp. 531–543, Aug. 2017.

[152] O. M. Lage, M. C. Ramos, R. Calisto, E. Almeida, V. Vasconcelos, and F. Vicente, "Current screening methodologies in drug discovery for selected human diseases," *Marine Drugs*, vol. 16, p. 279, Aug. 2018.

[153] G. E. Croston, "The utility of target-based discovery," *Expert Opinion on Drug Discovery*, vol. 12, pp. 427–429, May 2017.

[154] P. M. Farrell, T. B. White, C. L. Ren, S. E. Hempstead, F. Accurso, N. Derichs, M. Howenstine, S. A. McColley, M. Rock, M. Rosenfeld, I. Sermet-Gaudelus, K. W. Southern, B. C. Marshall, and P. R. Sosnay, "Diagnosis of cystic fibrosis: Consensus guidelines from the Cystic Fibrosis Foundation," *The Journal of Pediatrics*, vol. 181, pp. S4–S15.e1, Feb. 2017.

[155] Oxford English Dictionary, *"empirical, adj. and n."*. Oxford University Press, 2023.

[156] M. Mueller, A. d. l. Peña, and H. Derendorf, "Issues in pharmacokinetics and pharmacodynamics of anti-infective agents: Kill curves versus mic," *Antimicrobial Agents and Chemotherapy*, vol. 48, pp. 369–377, Feb. 2004.

[157] B. Kowalska-Krochmal and R. Dudek-Wicher, "The minimum inhibitory concentration of antibiotics: Methods, interpretation, clinical relevance," *Pathogens*, vol. 10, p. 165, Feb. 2021.

[158] Oxford English Dictionary, *"operationalism, n."*. Oxford University Press, 2022.

[159] Joint Committee for Guides in Metrology, *International vocabulary of metrology—Basic and general concepts and associated terms (VIM). 200:2012E*. International Bureau of Weights and Measures (BIPM), International Electrotechnical Commission (IEC), International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), International Laboratory Accreditation Cooperation (ILAC), International Organization for Standardization (ISO), International Union of Pure and Applied Chemistry (IUPAC), International Union of Pure and Applied Physics (IUPAP), International Organization of Legal Metrology (OIML), 2012.

[160] F. Ellett, J. Jorgensen, A. L. Marand, Y. M. Liu, M. M. Martinez, V. Sein, K. L. Butler, J. Lee, and D. Irimia, "Diagnosis of sepsis from a drop of blood by measurement of spontaneous neutrophil motility in a microfluidic assay," *Nature Biomedical Engineering*, vol. 2, pp. 207–214, Apr. 2018.

[161] Joint Committee for Guides in Metrology, *Guide to the expression of uncertainty in measurement—Part 6: Developing and using measurement models*. International Bureau of Weights and Measures (BIPM), International Electrotechnical Commission (IEC), International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), International Laboratory Accreditation Cooperation (ILAC), International Organization for Standardization (ISO), International Union of Pure and Applied Chemistry (IUPAC), International Union of Pure and Applied Physics (IUPAP), International Organization of Legal Metrology (OIML), 2020.

[162] K. J. Åström and R. M. Murray, *Feedback systems: An introduction for scientists and engineers*. Princeton university press, 2021.

[163] L. G. Valiant, "A theory of the learnable," *Communications of the ACM*, vol. 27, pp. 1134–1142, Nov. 1984.

[164] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press, 2014.

[165] V. N. Vapnik, *Statistical learning theory*. Adaptive and learning systems for signal processing, communications, and control, New York: Wiley, 1998.

[166] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistics, New York, NY: Springer New York, 1985.

[167] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer series in statistics, New York, NY: Springer, 2nd ed ed., 2009.

[168] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*. Cambridge, MA: MIT Press, 2012.

[169] J. B. Thomas, *Introduction to Probability*. Springer Texts in Electrical Engineering, New York, NY: Springer New York, 1986.