A PIPELINE FOR HIGH-THROUGHPUT DRUG SCREENING

By

Elisa Shireen Yazdani

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biostatistics

December 16, 2023

Nashville, Tennessee

Approved:

Amir Asiaeetaheri, Ph.D.

Andrew Spieker Ph.D.

ACKNOWLEDGMENTS

I would first like to thank Dr. Amir Asiaeetaheri for the opportunity to work on this project and his patience and guidance throughout the process. Thank you for your support throughout the project, and for teaching me more efficient ways to code. I would also like to extend my gratitude to Dr. Andrew Spieker for being the second reader of my thesis, and providing valuable feedback throughout the process. To all the faculty that I have interacted with over the past two years, thank you for being some of the best professors I have ever had. Graduate school is never easy, but the professors throughout my time at Vanderbilt have helped me develop a passion and understanding for an amazing field. I am very thankful that this program accepted me, despite having a BA in only Chemistry and Psychology, and has helped teach me so many intricate details of Biostatistics.

To my classmates, thank you for your friendship and help throughout our courses together; I will always remember the late-night slack messages as we figured out homework problems together. I want to specifically thank a few students who were extremely helpful during my time at Vanderbilt and helped make it an enjoyable experience. Lisa, thank you for helping me understand any biological concepts I may have needed clarification about, and for the many laughs we have shared throughout our time in the program. Alexis, thank you for always letting me bounce ideas off you, talk through things, and explaining many difficult statistical concepts to me. Julia, thank you for being there for me from our very first days as biostatistic students. Your friendship throughout this process has helped get me through the coursework and also many life-changes the past two years. You are a constant inspiration to me and have shown me that anything is possible.

I would also like to extend my gratitude to my mentor from my undergraduate studies, who I am also lucky to call a friend, Dr. Jason Kilgore. Jason, thank you for igniting my love of biostatistics and constantly believing in me to achieve at the highest level possible. I would not be where I am today without your support and guidance throughout the past few years.

I would also like to extend an immense amount of gratitude to my family. To my mom and dad, thank you for your constant support and guidance for the past 25 years. Your unwavering belief in me to achieve at the highest level has helped me push through challenges and difficulties, and most importantly given me the ability to persevere and believe in myself. I hope you are proud of my accomplishments, but know that none of my accomplishments would be possible without you. To my sister, thank you for always making me laugh, believing in me, and for making me feel like a genius anytime I show you my code. To my dog, Bear, for being the absolute light of my life and patiently waiting for me to take him on walks as I finished studying. Bear attended every thesis meeting with me and sat on my lap during my defense. He has added a sense of purpose to my life and calmness that has helped me get through the end of this program. And lastly, I would like to acknowledge my faith for getting me through the program, and God's constant blessings in my life that have brought me to where I'm supposed to be.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The study of novel drug combinations is essential to improving patient care. Focus has shifted away from single drug use in favor combination drug therapy because there are more favorable therapeutic responses, reduced side effects, and delayed resistance Tallarida (2011); Yadav et al. (2015). However, identifying effective drug combinations is a complex statistical challenge. For decades, the standard has been to use synergy to assess these drug combinations. As this topic has been studied extensively researchers have continued to propose various definitions and statistical methods to determine if drug combination data exhibits synergy He et al. (2018); Meyer et al. (2020); Tallarida (2011); Yadav et al. (2015). Researchers want to collect robust evidence of synergistic drug combinations before proceeding to clinical trials. This has caused widespread use and proliferation of synergy across fields in STEM. The traditional view of using synergy scores to elucidate the most effective combination of drugs has encountered challenges in contemporary literature for several reasons:

- There now exist a large number of newly defined synergy scores Meyer et al. (2020).
- These scores often conflict with each other when applied to real-world data Hwangbo et al. (2023).
- Advances in high-throughput screening allow for extensive drug combination testing without specific optimization for synergy Hwangbo et al. (2023).
- Studies have revealed that in-vitro synergistic combination of drugs (typically observed when applied to cell lines) do not necessarily translate to clinical success, and additivity alone may be sufficient Hwangbo et al. (2023).

In light of these challenges, this thesis introduces a straightforward, non-parametric model to study the dose-response curve of drug combinations using empirical data. This offers researchers a simplified method to evaluate therapeutic drug performance and efficacy. We first describe the mechanics of a typical high-throughput combination drug screening and the resulting data. We then discuss standard quality control and data normalization techniques. Finally, we discuss the various definitions of synergy and the surrounding debate.

1.1 High Throughput Screening

High-throughput drug screening (HTS) is an adaptation of classic drug screening methods. In classic drug screening, the process of identifying active compounds, antibodies, or genes for a particular biomolecular

pathway was a labor-intensive and time-consuming task. Each potential drug candidate was tested individually under controlled conditions requiring significant resources and producing a limited amount of data.

HTS automates this process, allowing for the simultaneous testing of thousands to millions of compounds within a short time frame Attene-Ramos et al. (2014). With the help of modern technology, HTS employs robotic arms, liquid handling devices, and advanced detection systems to precisely and quickly analyze samples, Figure 1.1. This is invaluable when performing high-throughput combinatorial screenings, where multiple drug combinations are tested for synergy. During this process, robotic systems can mix and match drug combinations in various concentrations, generating vast amounts of data on how drugs interact with each other and the biological system they target. This approach accelerates the drug discovery process while providing a more comprehensive view of potential therapeutic compounds and their interactions.

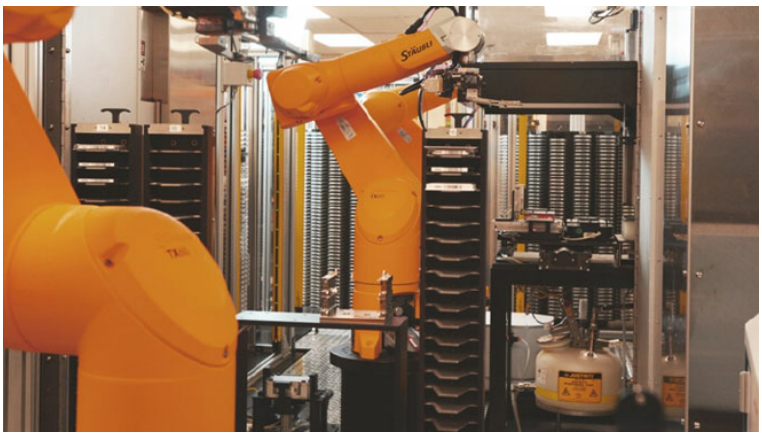


Figure 1.1: A High-Throughput Screening experimental setup with robotic arms, and well plates being put through automated testing Eder and Herrling (2016).

A typical HTS experiment begins with careful plate design, to ensure that samples, controls, and replicates are systematically arranged for accurate readouts and analysis. Experiments utilize multi-well plates, usually in formats of 96, 384, or 1,536 wells Attene-Ramos et al. (2014). The dimension of the plates come from the rows and columns of wells. For a 96-well plate, the format is 8 rows and 12 columns. For a 384-well plate, the format is 16 rows and 14 columns, as seen in Figure 1.2. Each plate has designated wells for positive controls (known compounds producing a desired effect) and negative controls (often a vehicle or base solution such as dimethyl sulfoxide [DMSO]) to establish assay parameters An and Tolliday (2010). Assay complexity, reagent choices, and detection methodologies all contribute to the experimental cost. A single well may cost several cents to several dollars, so when thousands to millions are being run (which is standard in an automated system), it can quickly become costly An and Tolliday (2010); Attene-Ramos et al. (2014); Doig et al. (2006). HTS becomes more cost-effective as more compounds are screened concurrently, making the process ideal for expansive drug discovery. Therefore, it is important to design techniques that leverage

information from as few plates as possible, while adjusting for outliers and errors using normalization.

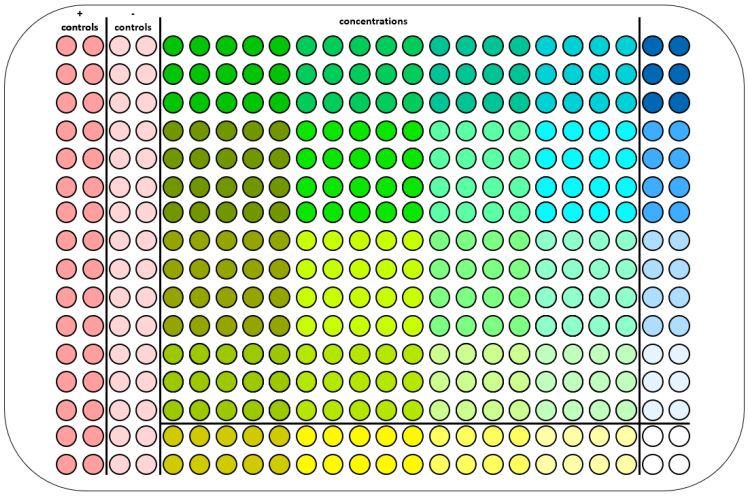


Figure 1.2: **Illustration of a 384-well plate designed for two-drug combination screening.** The first four columns are controls: columns 1 & 2 are positive controls known to inhibit cell activity (either stopping or slowing their growth), whereas columns 3 & 4 are negative controls expected to have no impact (containing either an ineffective drug or a solvent like DMSO). The white wells (bottom right) are reserved for cells in a vehicle, typically DMSO. The bottom two rows, transitioning from light to dark yellow, represent increasing concentrations of drug one, while the two rightmost columns, transitioning from light to dark blue, represent increasing concentrations of drug two. The central grid, ranging from cyan to green, represents combined treatments of both drugs; their concentrations are determined by the intersection of the respective rows and columns. Typically, this experimental setup is replicated three times, producing three identical plates.

In cell viability or cytotoxicity assays, drug combinations are placed in their plates and given a settling period of 48 hours. The cells are then stained with fluorescent dyes and imaged to assess viability Wei et al. (2021). Viability is defined as the proportion, or percentage, of cells that are living. Inhibition is the opposite, indicating the proportion or percentage of cells that have been killed. Different fluorescent dyes are used for different purposes, including specific dyes for high-density micro-plates with fluorescent signals in the background Wei et al. (2021). The duration of these experiments depends upon the cell line growth rate, drug mechanism, and specific assays, but a standard screening may span from one to several weeks from start to finish.

The measurement of cell viability in drug screening experiments is often achieved through optical methods. Many viability assays rely on colorimetric, fluorometric, or luminescent readouts. For instance, colorimetric assays, like the MTT assay, result in a color change when viable cells convert soluble tetrazolium salt into a visible formazan product. In fluorometric assays, viable cells metabolize certain compounds, producing fluorescence that is detected and quantified. In contrast, luminescent assays measure the presence of ATP as an indicator of cell viability and the live cell produce a luminescent signal. Advanced imaging techniques incorporating high-content screening systems can also be employed Riss et al. (2016).

1.2 Plate Quality Controls

One of the most important aspects of a drug screening experiment is ensuring the quality of data derived from the plates. Quality control can be performed by checking the consistency and reliability of results obtained from the plate assays. Variability can arise from numerous sources, such as pipetting errors, uneven cell distribution, or equipment malfunctions. To account for this, researchers rely on various statistical measures to assess plate quality before analyzing the data. It is standard to discard a low quality plate, even if a subset of wells are compromised.

1.2.1 Quality Control with Z' Factor

A popular metric used for plate quality control is the Z' factor. The Z' factor quantifies the difference between the distributions of positive and negative controls, providing a measure of assay robustness and is defined as An and Tolliday (2010):

$$Z' = 1 - \frac{3 \times (\sigma_p + \sigma_n)}{|\mu_p - \mu_n|} \tag{1.1}$$

where σ_p and σ_n are the standard deviations of the positive and negative controls, respectively, and μ_p and μ_n are their means. A Z' factor between 0.5 to 1 indicates an excellent assay, 0 to 0.5 suggests a marginal assay, and values below 0 suggest that the assay is unreliable, as seen in Table 1.1.

1.2.2 Quality Control with SSMD

Another important metric in plate quality control is the Strictly Standardized Mean Difference (SSMD). Unlike the Z' factor, which compares the distribution spread, SSMD focuses on the mean difference between positive and negative controls in relation to their standard deviations. The SSMD will always be positive because the signal from the instrument is positive. The SSMD is calculated as Huang et al. (2012):

$$\beta = \frac{\mu_p + \mu_n}{\sqrt{\sigma_p^2 + \sigma_n^2}} \tag{1.2}$$

A higher SSMD value typically indicates a more reliable assay. By utilizing both the Z' factor and SSMD together, researchers gain a comprehensive understanding of the quality and reliability of their plate assays.

Z'	Assay Quality
$Z' \geq 0.5$	excellent assay
$0 \leq Z' < 0.5$	marginal assay
$Z' < 0$	not suitable for HTS

β	Assay Quality
$\beta \geq 5$	excellent assay
$3 \leq \beta < 5$	good assay
$2 \leq \beta < 3$	inferior assay
$\beta < 2$	poor assay

Table 1.1: Comparison of Z' factor and SSMD values in relation to assay quality An and Tolliday (2010); Huang et al. (2012).

1.3 Plate Normalization

Despite the efficacy of HTS, inherent systematic errors, such as edge effects, plate-to-plate variation, and instrument drift, can occur. Systemic errors within plates are introduced through plate-specific bias, within-plate bias, and across plate-well location bias. These errors can lead to incorrect results, undermining the accuracy of the assay by producing false positives Murie et al. (2014). Plate normalization can be used to correct these errors, ensuring that the data is adjusted to a reference scale or normalized using specific controls. Percent of control, normalized percentage inhibition, and B-score are methods that account for these systematic errors.

1.3.1 Percent of Control (POC)

The POC offers a qualitative measure of a test compound's activity, defined by:

$$POC = \frac{x_i}{\mu_p} \times 100$$

where x_i denotes the raw measurement for the i^{th} compound and μ_p represents the mean of the measurements on the positive controls Malo et al. (2006).

1.3.2 Normalized Percentage Inhibition (NPI)

This method measures the inhibition effect of a compound by normalizing it relative to controls:

$$NPI = \frac{\mu_p - x_i}{\mu_p - \mu_n}$$

where x_i is the raw measurement for the i^{th} compound, and μ_p and μ_n denote the means of the measurements on the positive and negative controls, respectively Malo et al. (2006).

1.3.3 B-score

The B-score is designed to correct for spatial artifacts in the screening data by comparing the response of a well to the median response of its neighboring wells. It is defined as:

$$r_{ip} = y_{ip} - \left(\hat{\mu}_p + \hat{R}_i^p + \hat{C}_j^p\right)$$

The residual (r_{ip}) is defined by the difference between the observed value (y_{ip}) and the fitted value, which is the sum of the estimated average of the plate ($\hat{\mu}_p$), estimated systematic row offset ($\hat{R}_i^p$), and column offset

($\hat{C}_j^p$) Malo et al. (2006). The B-score is then calculated as follows:

$$Bscore = \frac{r_{ip}}{MAD_p}$$

MAD_p denotes the median absolute deviation for each plate, and is given by:

$$MAD = \text{median}\left|r_{ip} - \text{median}(r_{ip})\right|$$

1.4 Dose-Response Curve in Drug Screening

The dose-response curve is a fundamental concept in pharmacology and drug screening. This curve represents the relationship between varying concentrations of a drug (dose) and the observed effect on a biological system. Mathematically, the curve can be described by an equation relating the dose (typically on a logarithmic scale) to the measured response. The x-axis represents the drug concentration, often denoted in milligrams per liter (mg/L) or molar concentration (M). The y-axis quantifies a biological response such as cell viability, inhibition, enzyme activity, or any other relevant endpoint, and is usually expressed as a percentage or absolute value, as seen in Figure 1.3 Light et al. (2019). Multiple replicates of each drug concentration are performed to ensure data robustness and reproducibility. The curve can then be used to determine key pharmacological parameters such as drug potency and efficacy, aiding in the further development and evaluation of potential therapeutic agents.

Two important parameters derived from dose-response curves are *drug efficacy and potency*. Drug efficacy refers to the maximum response achievable from a drug, regardless of dose. It provides insight into the overall potential of a drug to elicit its intended biological effect. In contrast, drug potency refers to the concentration of a drug required to achieve a specified response. A drug with higher potency will elicit the same response at a lower concentration compared to a less potent drug, as seen in Figure 1.3 Light et al. (2019).

Metrics such as IC50 and EC50 can also be derived from dose-response curves. IC50 represents the concentration at which a drug inhibits a biological response by half, commonly used in studies of drug toxicity or inhibition (Figure 1.3). EC50 denotes the concentration of a drug that induces half of its maximal possible effect, typically employed when assessing a drug's therapeutic effect. These metrics are determined by fitting experimental data to models, which are often sigmoidal, such as the Hill or logistic function. IC50 and EC50 values can be extrapolated from these curves and provide pharmacodynamic information to inform further drug development and therapeutic application Light et al. (2019).

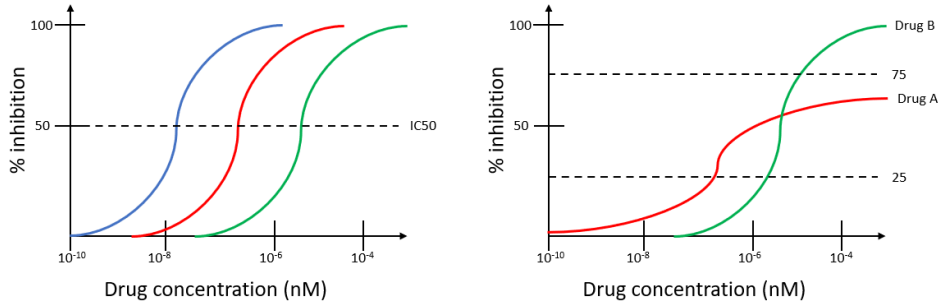


Figure 1.3: An in-depth exploration of dose-response relationships in drug screening. The figures illustrate the nuances of drug efficacy, potency, and the mathematical frameworks like the Hill equation used for analysis. Collectively, they emphasize the multidimensionality of drug comparison, underscoring that a singular value might not wholly represent a drug's effect. In the left Figure, you can see three separate drugs with the highest potency drug represented by the blue line and the lowest potency drug represented by the green line. The IC50 value indicates the concentration at which a drug inhibits a biological response by half. The right figure shows how to different drugs can be represented, with Drug B being more potent at a lower concentration, but as the concentration increases Drug A becomes more potent Light et al. (2019)

1.4.1 Dose Response Surfaces and Synergy in Drug Combination

In therapeutic interventions, particularly for multifaceted conditions like cancer, the use of multiple drugs concurrently may offer distinct advantages. Combining drugs may reduce individual drug doses, which in turn may minimize toxicity and associated side effects He et al. (2018); Tallarida (2011); Yadav et al. (2015). An upward trajectory of the toxicity curve with increasing drug dosages underscores the importance of dose optimization. Moreover, heterogeneous diseases such as cancer display multiple subpopulations or subclones, each with distinct characteristics. Administering a combination of drugs may better target these diverse subclones, minimizing the potential for resistance and ensuring a comprehensive therapeutic approach Dagogo-Jack and Shaw (2018). Furthermore, the integration of FDA-approved drugs in these combinations paves the way for cost-effective treatment strategies. Not only does this approach leverage the existing knowledge and safety profiles of these drugs, it also significantly shortens the arduous and costly drug development process. Together, these features present a compelling case for the exploration of dose-response surfaces and the synergy that may be found within drug combinations.

When drugs are administered in combination, their collective effect can manifest in three distinct ways: additive, synergistic, or antagonistic. It is essential to note that the nature of these interactions—whether additive, synergistic, or antagonistic—is dose-specific. An additive outcome arises when the combined effect is equivalent to the sum of each drug's individual effect at those doses. Synergy is characterized by the combined effect exceeding what would be anticipated from simple additivity at the given doses. Conversely, an antagonistic response is observed when the combined effect is lesser than the sum of the individual drug

effects at those doses Yadav et al. (2015). In the following section, we explore three prominent definitions of additivity and the challenges of finding a universal definition.

1.4.2 Definitions of Additivity

To comprehend and quantify synergy, we have to first outline the concept of additivity. Numerous models and mathematical formulations have been proposed to encapsulate this idea Yadav et al. (2015). In this section, we highlight the historically significant definitions of additivity. For a comprehensive overview of more contemporary definitions, readers are directed to Meyer et al. (2020). Here, the effects of drugs A and B, when used individually at specific doses, can be denoted as E_A and E_B, respectively. These represent the percentages of cells impacted by each drug (either inhibition or viability). It is imperative to emphasize that a drug should not demonstrate synergy when combined with itself, a tenet all additivity models uphold.

1. **Highest Single Agent (HSA)**: The HSA model suggests that the combined effect of two drugs should not be more than that of the most potent drug used alone. Thus, we define:

$$E_{AB} = \max(E_A, E_B)$$

2. **Bliss Independence Model**: The Bliss independence assumption states that the combined effect of two drugs is the union of their individual effects when the two drugs are independent. Under this assumption, we can write their expected additive effect E_{AB} using the definition of union:

$$E_{AB} = E_A + E_B - E_A \times E_B$$

3. **Loewe Additivity**: If D_A and D_B are the doses of two drugs that produce the same effect E individually, then according to Loewe additivity:

$$\frac{D_A}{D_{A,comb}} + \frac{D_B}{D_{B,comb}} = 1$$

   where $D_{A,comb}$ and $D_{B,comb}$ are the doses of A and B in combination that achieve the same effect E.

These definitions may not always agree. The determined synergy or antagonism may depend on the chosen additivity model Meyer et al. (2020).

1.4.3 Challenges and Controversies Surrounding Synergy Metrics

The concept of drug synergy, as represented in current literature, varies across specific combination doses (x, y) rather than being a universal property of drug pairs. Interestingly, the global synergy assessment, which uses straightforward statistics like the mean or maximum, may not consistently align with the designated interactions (synergistic, antagonistic, additive) observed at specific combination doses. Differing methodologies applied to identical datasets have yielded conflicting synergy scores. Furthermore, even when using the same synergy metric, variations in normalization and data preprocessing can lead to conflicting synergy outcomes Meyer et al. (2020).

The clinical relevance of synergy is also debated. Despite clear evidence of synergy in laboratory settings, the transition to meaningful therapeutic results in clinical scenarios remains ambiguous. The complexities introduced by drug pharmacokinetics, potential adverse effects, and variations in patient responses challenge the reliability of synergy as a definitive measure in combination therapies Hwangbo et al. (2023).

Recent trends show an increased interest in harnessing machine learning techniques to predict drug synergy. This is driven, in part, by the broader availability of large-scale drug combination datasets. However, given the nuances and debates surrounding the definition of synergy, combined with its sensitivity to outliers, normalization, preprocessing, and questionable clinical utility, excessive focus on such predictions may lead to resource misallocation. We introduce a method anchored in isotonic regression to produce a more dependable method for assessing drug combinations. This method could serve as a reliable target for future machine learning predictions.

CHAPTER 2

METHODS

We introduce three original contributions to the field of pharmacological analytics. First, we redefine synergy scores within the framework of causal inference to clarify their interpretation. Then, we explore various methodologies for assessing the monotonicity of drug response surfaces, providing a comprehensive toolkit for analyzing drug effectiveness. Lastly, we advocate for the application of isotonic regression as a superior method for modeling the dose-response surface of drugs, addressing the limitations inherent in conventional synergy scores.

2.1 Causal Inference Notation for Defining Synergy

The study of synergy spans many fields and consequently does not have a cohesive or intuitive notation compatible with statistical perspectives. Our approach introduces a notation rooted in the principles of causal inference, and strives for a more comprehensive understanding of synergy scores. The intention is not to present novel breakthrough in causal inference literature; we simply utilize the notation to simplify the understanding of synergy scores. To the best of our knowledge, the application of potential outcome notation in the realm of drug screening literature is unprecedented and remains unfamiliar territory for those within synergy research. Our proposition stems from the belief that this notation augments conceptual clarity, thereby enhancing the understanding of synergy equations.

In causal inference, the concept of potential outcomes plays a pivotal role. An individual's potential outcome under a specific treatment condition is often symbolized as $Y(t)$, where t represents the treatment received. For instance, $Y(0)$ denotes the potential outcome if the individual does not receive treatment, while $Y(1)$ signifies the outcome if the individual does receive treatment Hernán and Robins (2010).

To relate this to a single dose-response curve: imagine a drug A. For any given dose d, the potential outcome can be depicted as $Y_A(d)$, which describes the anticipated response of an individual to that specific dose of drug A.

Extending this idea to a scenario involving two drugs, drug A and drug B, we have: the dose-response surface as a function $Y_{AB}(d_A, d_B)$. Here, d_A and d_B represent the doses of drug A and drug B, respectively, and $Y_{AB}(d_A, d_B)$ quantifies the combined response of a cell to that particular combination of drug doses.

2.1.1 Causal Notation Applied to Synergy Equations

Synergy equations quantify the combined effect of multiple drugs in relation to their individual impacts. By using the causal framework, these equations can be presented in a more structured and intuitive format. In the context of drug combinations, let $Y_{AB}(d_A, 0)$ represent the potential outcome for dosage d_A of drug A, and $Y_{AB}(0, d_B)$ denote the potential outcome for dosage d_B of drug B. Let $Y_{AB}(d_A, d_B)$ be the potential outcome when both drugs are administered together.

1. **Highest Single Agent (HSA):** HSA estimates the combined effect as the greater effect of the individual drugs. In causal terms:

$$Y_{AB}^{HSA}(d_A, d_B) = \max\{Y_{AB}(d_A, 0), Y_{AB}(0, d_B)\}$$

If the factual (observed) $Y_{AB}(d_A, d_B) > Y_{AB}^{HSA}(d_A, d_B)$ two drugs are synergistic.

2. **Bliss Independence:** The Bliss independence model posits that the combined effect of two agents is equivalent to the product of their separate effects. Translated into causal notation, the combined effect under Bliss independence can be expressed as:

$$Y_{AB}^{Bliss}(d_A, d_B) = Y_{AB}(d_A, 0) + Y_{AB}(0, d_B) - Y_{AB}(d_A, 0) \times Y_{AB}(0, d_B)$$

If the factual (observed) $Y_{AB}(d_A, d_B) > Y_{AB}^{Bliss}(d_A, d_B)$ two drugs are synergistic.

3. **Loewe Additivity:** Loewe's principle asserts that drug A alone at dose D_A achieves an effect equivalent to drug B alone at dose D_B, then their doses should replicate this effect. Using causal notation:

$$\frac{d_A}{D_A} + \frac{d_B}{D_B} = 1$$

where $Y_{AB}^{Loewe}(d_A, d_B) \triangleq Y_{AB}(D_A, 0) = Y_{AB}(0, D_B)$, signifying that drug A and drug B have the same efficacy at dosages D_A and D_B respectively. Any pair (d_A, d_B) satisfying this equation (representing a line in the dosage space) indicates additivity.

If the factual (observed) $Y_{AB}(d_A, d_B) > Y_{AB}^{Loewe}(d_A, d_B)$, then the drugs exhibit synergy.

Utilizing causal notation for synergy equations offers a structured approach and improves understanding, especially when examining the intricate interactions of drug combination effects.

2.2 Directly Modeling Dose-Response Surface: A Robust Approach

Synergy analysis, as traditionally approached, aims to assess how two drugs interact with each other, often using reference models or metrics. However, a direct modeling of the dose-response surface offers several inherent advantages, especially when considered through the lens of non-parametric methods like isotonic

regression.

2.2.1 The Need for Direct Modeling

Comprehensiveness: Modeling the dose-response surface directly gives a complete picture of how combined drug dosages affect the outcome. This approach avoids the need for assumptions inherent in synergy scores or reference models.

Flexibility: Traditional synergy models, like Bliss, HSA, and Loewe, are based on specific assumptions about drug interactions. Direct modeling allows for the drug interaction landscape to reveal itself without being forced into predefined patterns.

Improved Interpretability: By directly capturing the landscape of drug interactions, it becomes more intuitive to visualize and understand the joint effects of two drugs across a range of dosages, rather than interpreting scores that indicate synergy or antagonism.

2.2.2 Why Non-Parametric? The Case for Isotonic Regression

Isotonic regression is a non-parametric technique that models a response variable as a monotonically increasing or decreasing function of the predictor variable(s). The primary benefit of this method is its ability to find the best-fitting monotonic curve to the data without making restrictive assumptions about its form.

Preserving Natural Order: Biological processes, including responses to drugs, often exhibit monotonic behavior. As the dose of a drug increases, it is natural to expect the response to either consistently increase or decrease, without sudden reversals. Isotonic regression, by ensuring monotonicity, captures this inherent trait.

Flexibility in Complexity: Unlike parametric methods, which might require choosing the correct function form in advance, isotonic regression adapts to the complexity of the data. Whether the true relationship is nearly linear or more intricate, isotonic regression will attempt to capture it.

Avoiding Overfitting: Non-parametric methods, while flexible, are prone to overfitting. The monotonic constraint of isotonic regression naturally regularizes the model, reducing the risk of fitting noise in the data.

2.2.3 Monotonicity in Dose-Response Surfaces

In pharmacological contexts, it is reasonable to expect dose-response surfaces to be monotonic van Wijk et al. (2015). As the dose of drugs increases, biological systems typically respond in a consistent direction, whether that be an increase in efficacy, saturation, or even the onset of adverse effects. Several reasons underscore this expectation:

Physiological Constraints: Organisms tend to have upper and lower bounds to how they can respond to

external stimuli, including drugs. This naturally enforces monotonic behavior on the response as drug doses increase.

Receptor Dynamics: At the molecular level, as the concentration of a drug increases, more receptors might be occupied or modulated, leading to a consistent directional change in response.

In summary, direct modeling of the dose-response surface, especially through non-parametric methods like isotonic regression, offers a more comprehensive, flexible, and intuitive understanding of drug interactions. The inherent monotonic behavior of biological responses further validates the choice of isotonic regression in this domain.

2.2.4 Isotonic Regression

Isotonic regression is a non-parametric technique used for fitting a sequence of observations with a function that is monotonically increasing or decreasing. We use the 'isotone' package in R to fit our data. In the context of the 'isotone' package in R, isotonic regression is implemented by finding a monotonic fit to a set of data points where the monotonicity constraints are provided as partial order using a Directed Acyclic Graph (DAG). This method is particularly useful when the relationship between the independent and dependent variables is assumed to be non-decreasing or non-increasing but not necessarily linear. The algorithm works by solving an optimization problem that minimizes the sum of squared errors from the data points, subject to the monotonicity constraint. The result is a step function that increases (or decreases) monotonically, capturing the trend of the data without assuming a specific parametric form, such as a straight line or a polynomial curve Dai et al. (2020). This makes isotonic regression a flexible tool for modeling dose-response relationships in pharmacology, where the response is expected to increase with the dose without the necessity of a linear relationship.

2.3 Statistical Tests for Monotonicity

To assess statistical tests for monotonicity, we performed a literature review for various approaches. Patton and Timmermann (2010) assessed monotonic trends in an economics paper with asset data that utilized several ways to test data for a monotone trend. They proposed a Monotonic Relation (MR) test that focuses on the minimum delta, which is the smallest increment change in stock assets from time-point A to time-point B. This can be applied to drug data by looking at the smallest increment change, the minimum delta, in the response value at the next dose level. It is useful in that it looks for a general monotonic increase, even if individual data points do not follow a strictly monotone trend. This can occur due to noise in the data that may create some violations of monotonicity, even if the data overall does follow a monotone trend. The robustness of the MR test is achieved by the use of bootstrapping to calculate the p-value. By generating

numerous resampled datasets, the MR test assesses the likelihood of observing the minimum delta by chance, thus allowing for a more nuanced interpretation of the data's underlying trend.

We present two statistical tests of our own to help capture if the data has a true monotonicity trend. We have operationalized these tests, along with the previously mentioned methods, within the R programming environment for both one-dimensional and two-dimensional (grid) datasets. Specific attention was paid to accommodating data replicates to mimic drug screening experiments. To evaluate the efficacy of these two tests, we generate synthetic datasets based on linear models and Hill equation models, allowing us to evaluate and compare their ability to detect monotonic trends. We then apply the tests to real-life data to assess if the assumption for monotonicity is met to then be able to proceed with isotonic regression.

2.3.1 Monotonic Relation (MR) Test

The monotonicity constraint can be imposed through a Directed Acyclic Graph (DAG). Assume the value of the function over the neighboring nodes i and j of the DAG are y_i and y_j, and there is a directed edge $i \to j$, then monotonicity requires that $y_j \geq y_i$. For such neighboring nodes, we define $\delta_{ij} = y_i - y_j$. We collect all such δ_{ij} from the DAG into one vector and denote it as Δ. The null hypothesis and the alternative hypothesis (asserting monotonicity) can be formulated as:

$$H_0 : \Delta \leq 0 \quad \text{vs.} \quad H_1 : \Delta > 0$$

The alternative hypothesis can also be expressed as:

$$H_1 : J = \min_{ij} \delta_{ij} > 0$$

We define the test statistic as the empirical minimum $\hat{J} = \min_{ij} \hat{\delta}_{ij}$ and use bootstrapping to compute the p-value associated with it.

2.3.2 Monotonic Steps (MS) Test

As an alternative to the MR methods, we propose a straightforward approach in which we define the statistic C as the normalized count of positive δ_{ij} values in the DAG:

$$C = \frac{1}{N} \sum_{ij} 1(\delta_{ij} > 0)$$

where N is the total number of edges in the DAG and $1(\cdot)$ is the indicator function that takes the value 1 if the condition inside is true and 0 otherwise. This was adapted from the methods and concepts of property tests

Goldreich (2017). The p-value for this statistic is computed using a permutation test.

CHAPTER 3

EXPERIMENT

In our experiments, we generated monotone data to assess the robustness of the proposed monotonicity tests. First, we chose a linear model of data due to its simplicity and interpretability when assessing the monotone tests. Second, we chose the Hill equation model, which is prevalent in drug screening literature as this model represents the trends that drug responses show.

The linear model is inherently monotonic, providing a clear baseline for assessing the performance of our tests. Conversely, the Hill equation model captures the saturation effect commonly observed in biological systems, where the response tends to plateau at the beginning doses and as the dose increases. This non-linearity poses a more challenging test for our monotonicity assessment.

Drug screening experiments typically include triplicate measurements to account for biological variability, and our simulation followed this standard. To realistically simulate an experiment, we included Gaussian noise in the dataset to represent random fluctuations that can occur due to experimental error or biological variation.

For the linear model, we selected a concentration range from 1 to 10. For the Hill equation model, we chose a more biologically relevant range, using logarithmically spaced concentrations from 10^{-8} to 10^{-2}.

The Hill equation in 1D can be described by the following function:

$$E(d) = E_{\max} + \frac{E_0 - E_{\max}}{1 + \left(\frac{d}{C}\right)^h}$$

where $E(d)$ is the effect at dose d, $E_{\max}$ is the maximal effect, E_0 is the baseline effect, C is the concentration of drug resulting in 50% of the maximal effect, and h is the Hill coefficient, which describes the steepness of the dose-response curve.

In the 2D case, the Hill equation is extended to account for the effects of two drugs in combination, and is given by:

$$E_d = \frac{C_1^{h1} C_2^{h2} E_0 + d_1^{h1} C_2^{h2} E_1 + C_1^{h1} d_2^{h2} E_2 + (\alpha_2 d_1)^{h1} (\alpha_1 d_2)^{h2} E_3}{C_1^{h1} C_2^{h2} + d_1^{h1} C_2^{h2} + C_1^{h1} d_2^{h2} + (\alpha_2 d_1)^{h1} (\alpha_1 d_2)^{h2}}$$

where the parameters are defined as in Table 3.1 Meyer et al. (2020).

Table 3.1: Annotation of parameters for the 2D Hill equation Meyer et al. (2020).

Parameter	Description
d_1, d_2	Drug concentrations for drug pair
C_1, C_2	EC_{50} for drugs 1 and 2 in isolation
h_1, h_2	Hill coefficients for dose response curves of drug 1 and 2 in isolation
E_0	The basal rate of proliferation in drug naive condition
E_1, E_2	$E_{\max}$ of drug 1 and 2 in isolation
E_3	$E_{\max}$ of the combination of drugs 1 and 2
α_1	Measure of how d_1 modulates the effective dose of d_2
α_2	Measure of how d_2 modulates the effective dose of d_1

Our simulation framework involves generating dose-response surfaces from these models and subsequently applying the proposed monotonicity tests to determine their power in detecting violations of monotonicity under various noise conditions.

3.1 1-D Data

In the one-dimensional (1-D) data simulations, we generate dose-response curves based on either a linear or Hill function model. Each simulated dataset incorporates three replicates to represent repeated experimental measurements at the same drug concentration. Additionally, Gaussian noise is introduced to the simulated data to mimic the deviations from a perfect monotone pattern, as is often observed in both linear and Hill 1-D dose-response scenarios.

Figure 3.1 showcases four examples of simulated one-dimensional dose-response data. Figure 3.2 displays the dose-response relationships with a Loess curve fitted to each replicate, illustrating the individual replicate trends. The dashed line represents the Loess curve fitted to the average of the replicates, providing a visual summary of the overall response pattern across the replicates.

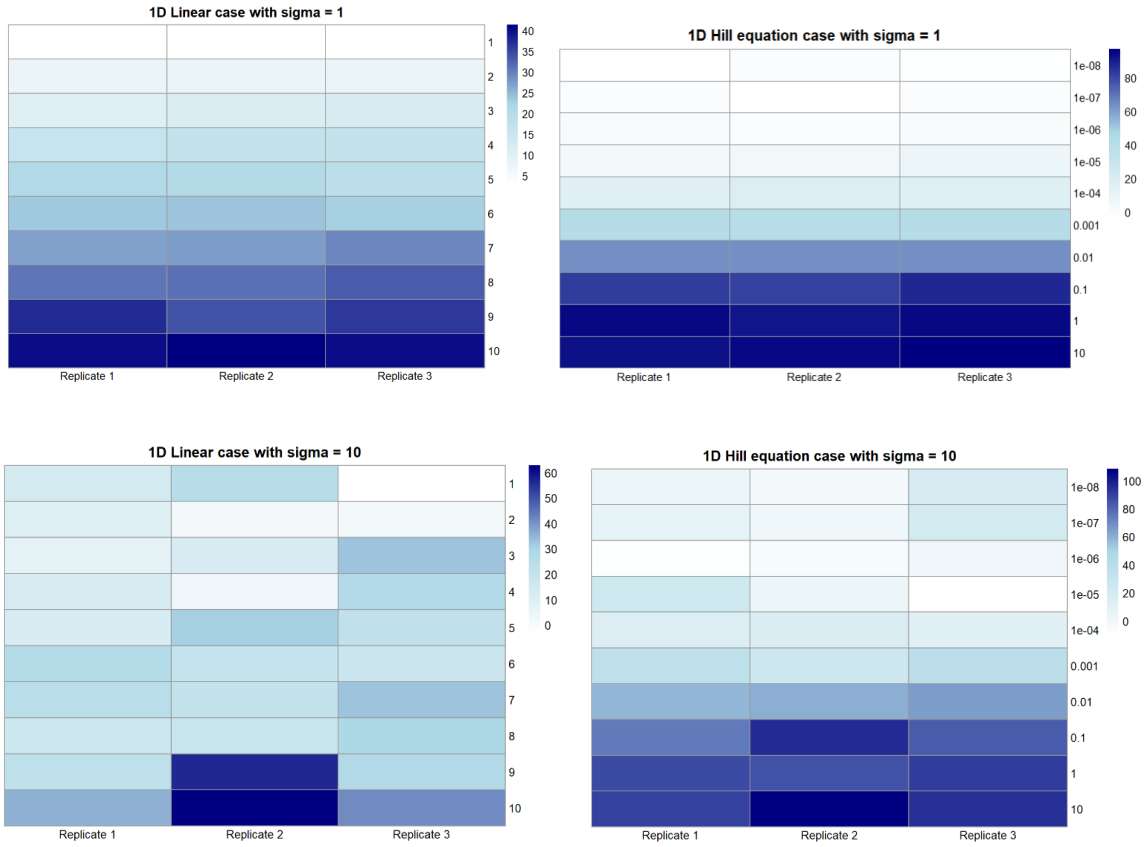


Figure 3.1: Simulated 1-D dose-response scenarios (from left to right): Linear model with small noise, $y = 4x$ and $\sigma = 1$; Hill model with small noise $C = 2 \times 10^{-3}, h = 1/2, E_0 = 0, E_{max} = 95$; Linear model with noise $\sigma = 10$; Hill model with noise $\sigma = 10$.

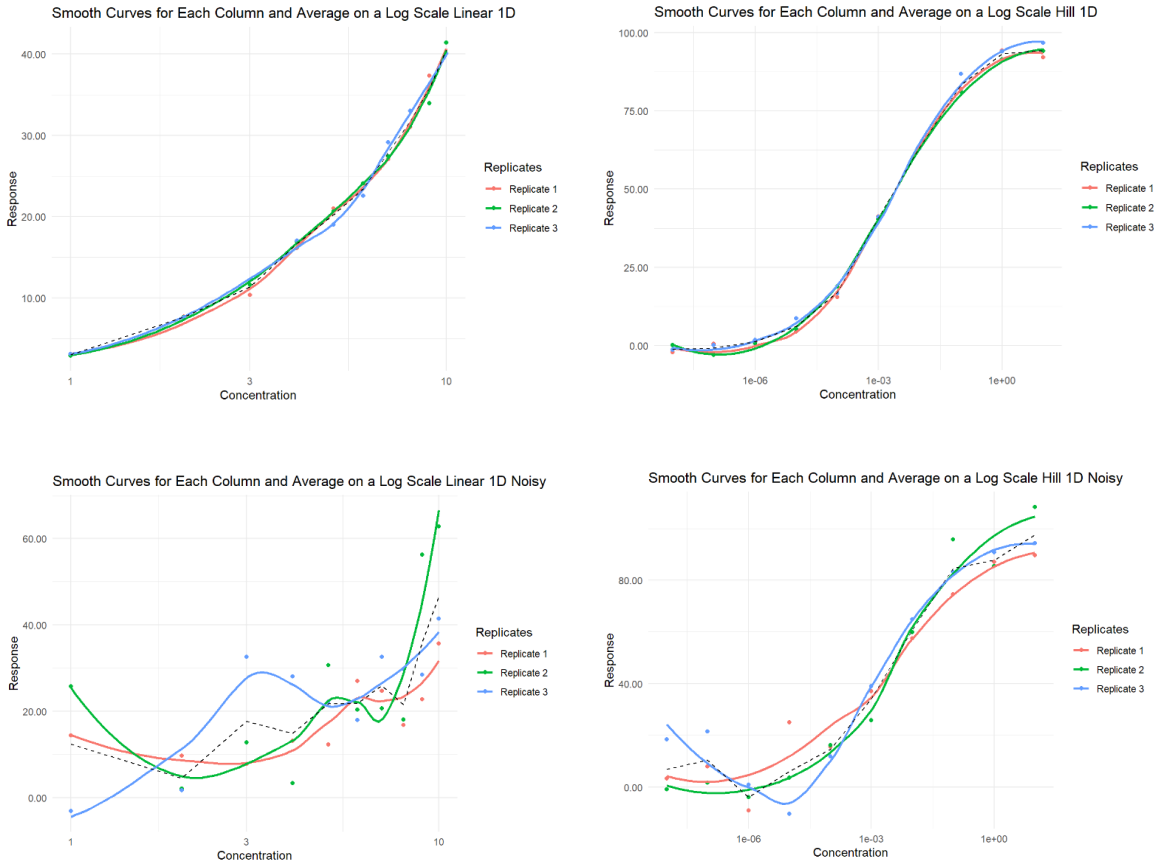


Figure 3.2: Dose-response relationships with Loess curve fitting of the same data of Figure 3.1.

3.1.1 Statistical Tests for Monotonicity

To evaluate the monotonicity of the dose-response relationship, we employ statistical tests that utilize resampling methods—bootstrap for the Monotonic Relation (MR) test, and permutation for the Monotonic Violation (MS) tests—to calculate p-values. Each of these resampling procedures is repeated 1,000 times to ensure a robust estimation of the p-values.

3.1.1.1 MR Test

For the bootstrap resampling used in the MR test, a new replicate dataset is generated by randomly selecting one of the three available replicates at each concentration level. This approach allows us to create a variety of unique replicates from a limited number of actual samples, allowing us to understand how the model behaves asymptotically as seen in Figure 3.3.

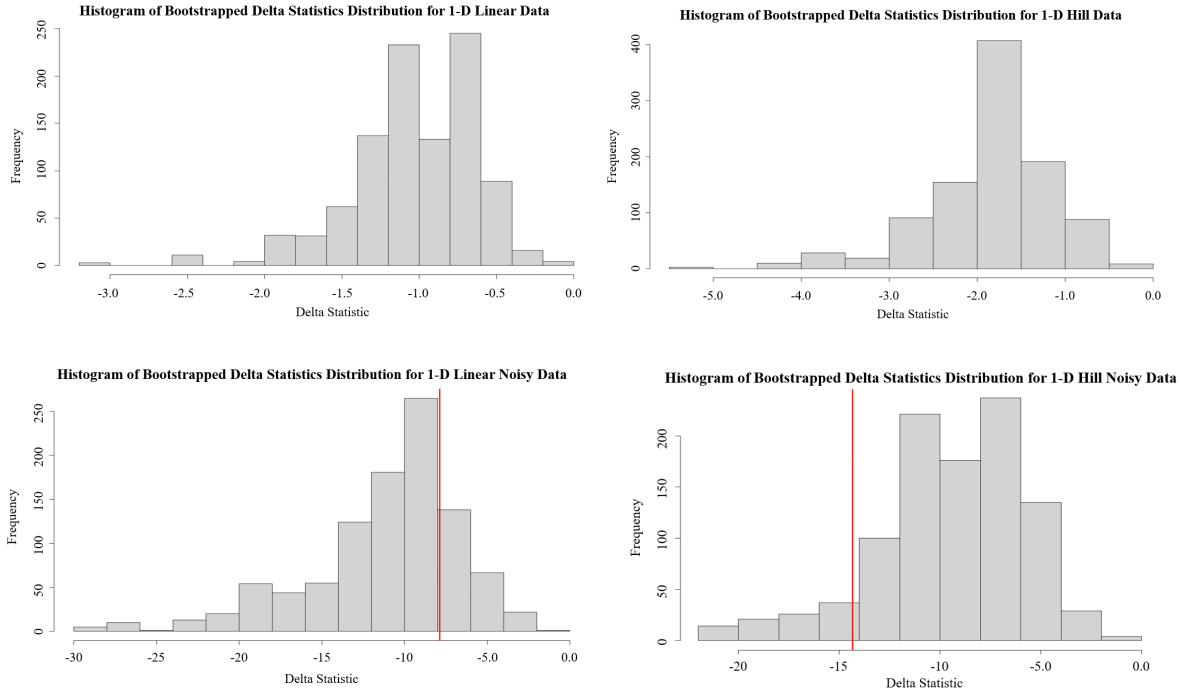


Figure 3.3: Bootstrap distributions of the Monotonic Relation (MR) statistic for simulated 1-D dose-response data. Figures (a) and (b) show the distributions for noise-free linear and Hill models, respectively, where the original data's statistics are not visible due to extremely low p-values ($<$0.001). In contrast, figures (c) and (d) depict scenarios with added Gaussian noise, where the original data's statistics fall within the bootstrap distributions, resulting in higher p-values of 0.217 and 0.904, respectively.

To evaluate the robustness of our statistical test under varying conditions, we conducted a power analysis to determine how noise affects the ability of our test to detect true monotonic relationships. This involved generating ten distinct datasets for each specified level of noise, computing the bootstrap distribution of the δ statistics for each to compute the p-value, and recording the frequency with which the test yielded a significant result.

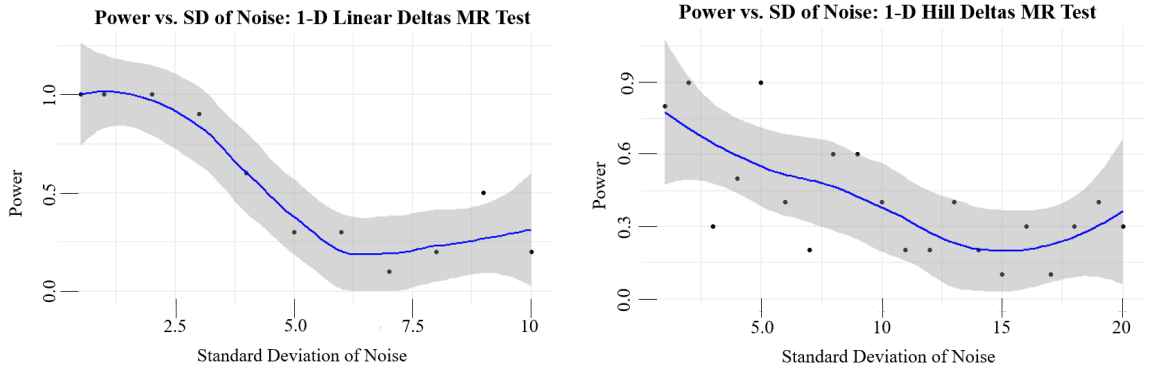


Figure 3.4: The power of the MR Test as a function of noise level for 1-D data. Each plot illustrates the proportion of simulations in which the MR test successfully identified a monotonic trend at varying noise intensities. A Loess curve has been fitted to the data points to highlight the overall trend.

Figure 3.4 depicts the power of the MR test as a function of noise for two different model types. In each subplot, the power is represented by the proportion of tests that successfully identified monotonicity, with noise levels ranging from low to high. The smooth curve fitted to these data points represents a loess regression, providing a clear visualization of the trend. As expected, the power of the MR test diminishes with increasing noise levels, indicating the impact of variability on test sensitivity.

3.1.1.2 MS Test

MR power seems to be noise sensitive, as for moderated amount of noise its power drops. We rerun the same experiment using our introduced MS test which shows less sensitivity to noise.

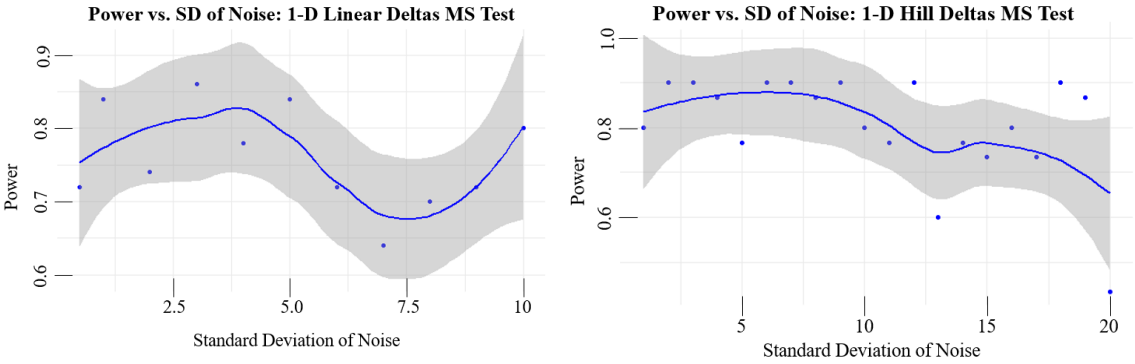


Figure 3.5: Power of MS Test as a function of noise level for 1-D data. The plot shows the percentage of experiments where the MS test detects monotonicity for various levels of noise. The curve is a Loess-fitted curve to illustrate the trend.

3.1.2 2-D Data

The experimentation is extended to 2-D data, where monotonicity is conceptualized through a grid-shaped Directed Acyclic Graph (DAG). In this DAG, each node (point) on the grid should have a value less than or equal to the values of all nodes to its right and above, as seen in Figure 3.6. The MR and MS test statistics are computed in accordance with this monotonicity constraint. We simulate both linear (planar) surfaces and Hill equation-based surfaces, incorporating varying levels of noise and including three replicates for each condition.

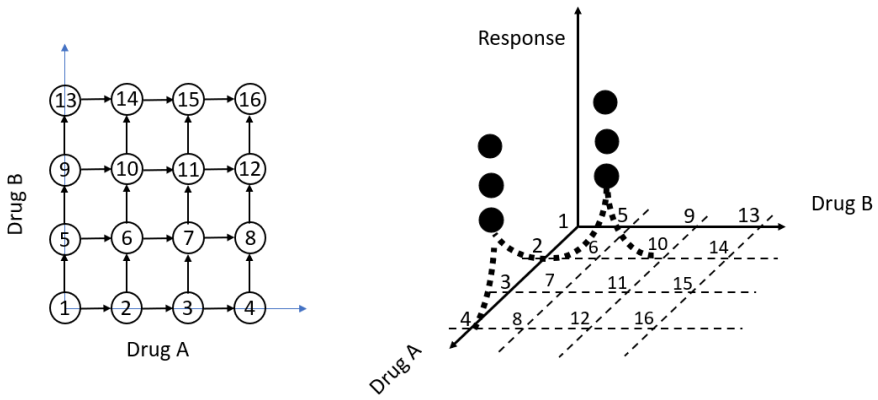


Figure 3.6: On the left side is an example of a DAG for 2-D data, where the 3-D graph represents the DAG on the axis of the drug concentrations and the black dots represent the response replicates for the data.

Figure 3.7 presents heat maps of plate responses for the linear model with minimal noise across three replicates and their averaged results. Similarly, Figure 3.8 shows the responses for the same model under conditions of large noise.

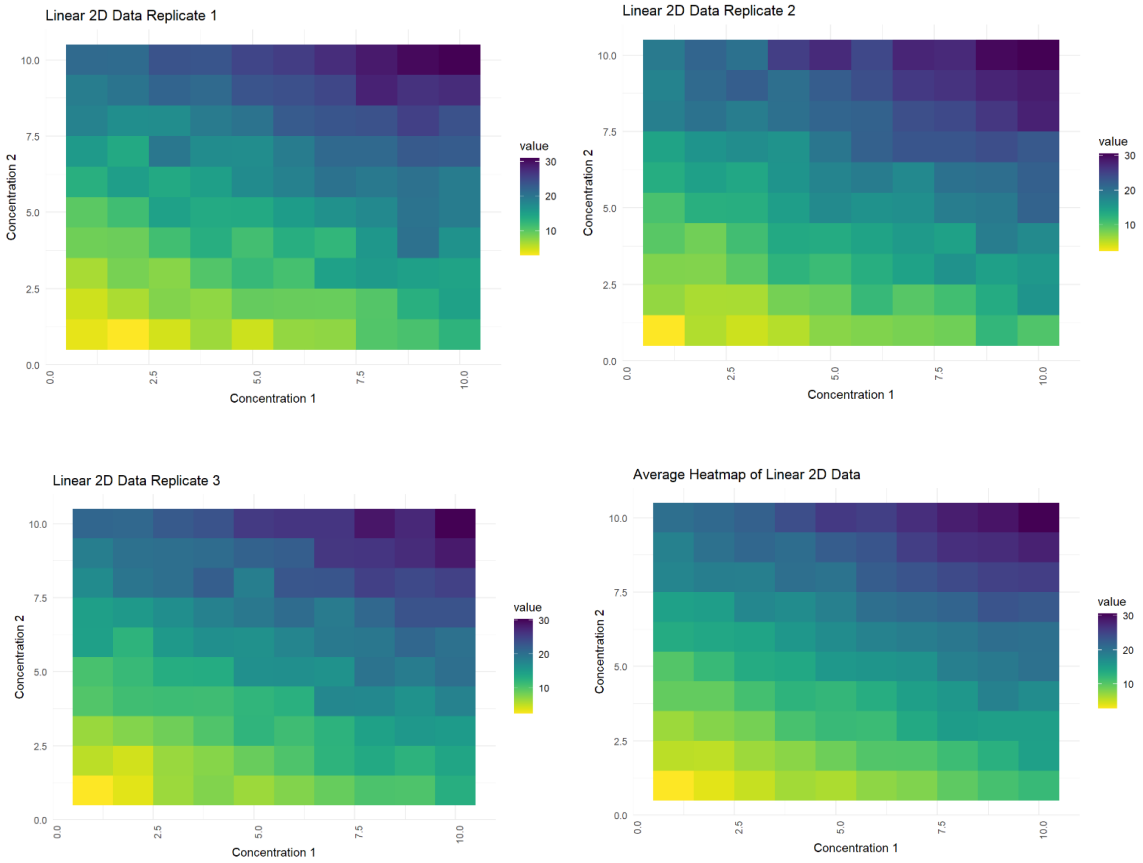


Figure 3.7: Heat maps of plate responses for three replicates and their average for the linear model with minimal noise.

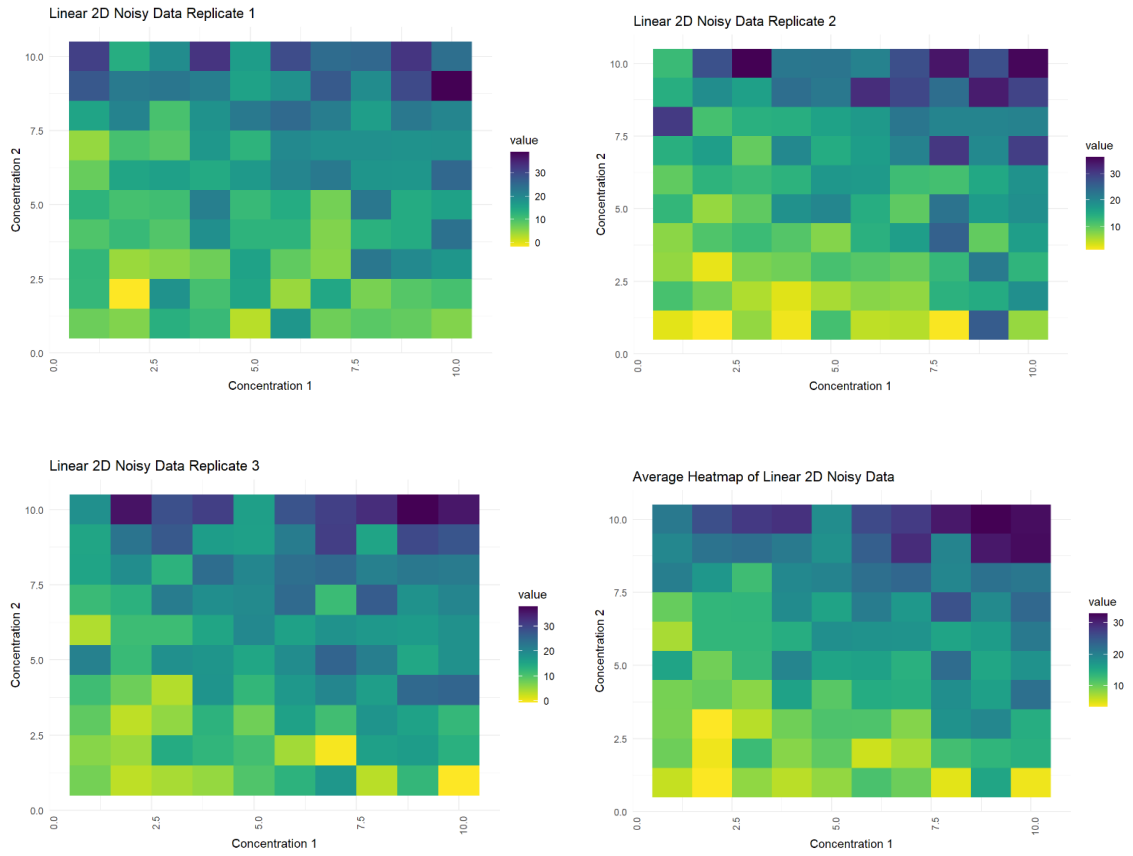


Figure 3.8: Heat maps of plate responses for three replicates and their average for the linear model with substantial noise.

For the 2-D Hill equation-based surfaces, Figures 3.9 and 3.10 display the outcomes with small and large noise levels, respectively.

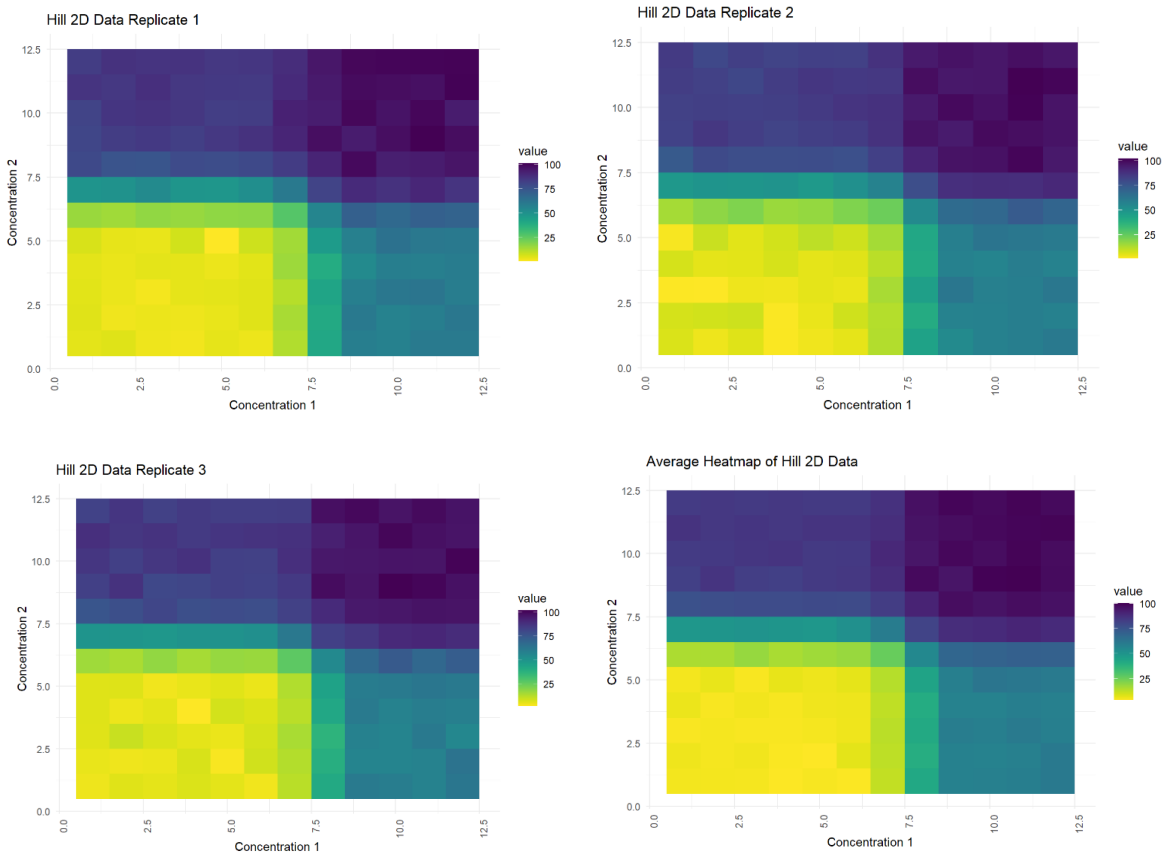


Figure 3.9: Heat maps of responses for replicates and average for the 2-D Hill model with minimal noise.

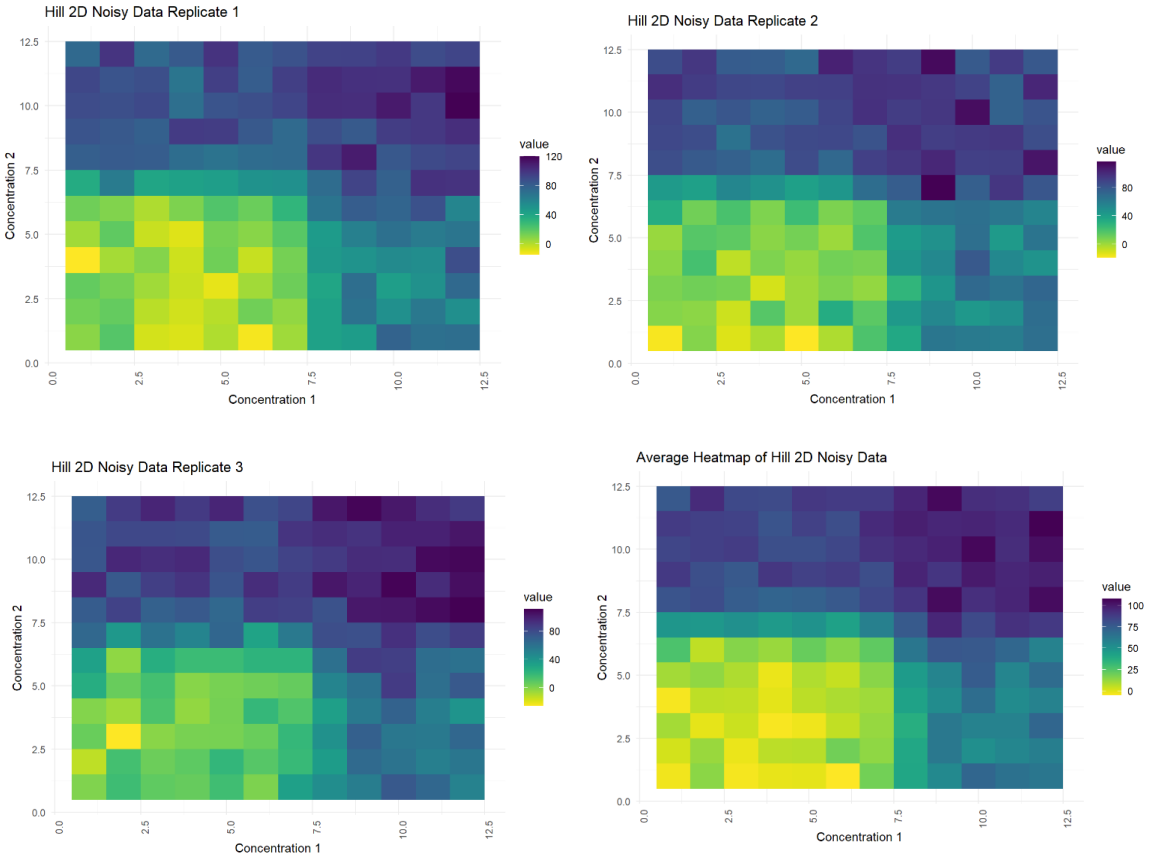


Figure 3.10: Heat maps of responses for replicates and average for the 2-D Hill model with substantial noise.

Similarly, for the 2-D case, we generated data using bootstrap methods for the Monotonic Relation test and permutation methods for the Monotonicity Score test to assess their power in detecting monotonicity. The results are illustrated in Figure 3.11, which demonstrates that the MS test exhibits greater robustness to noise compared to the MR test.

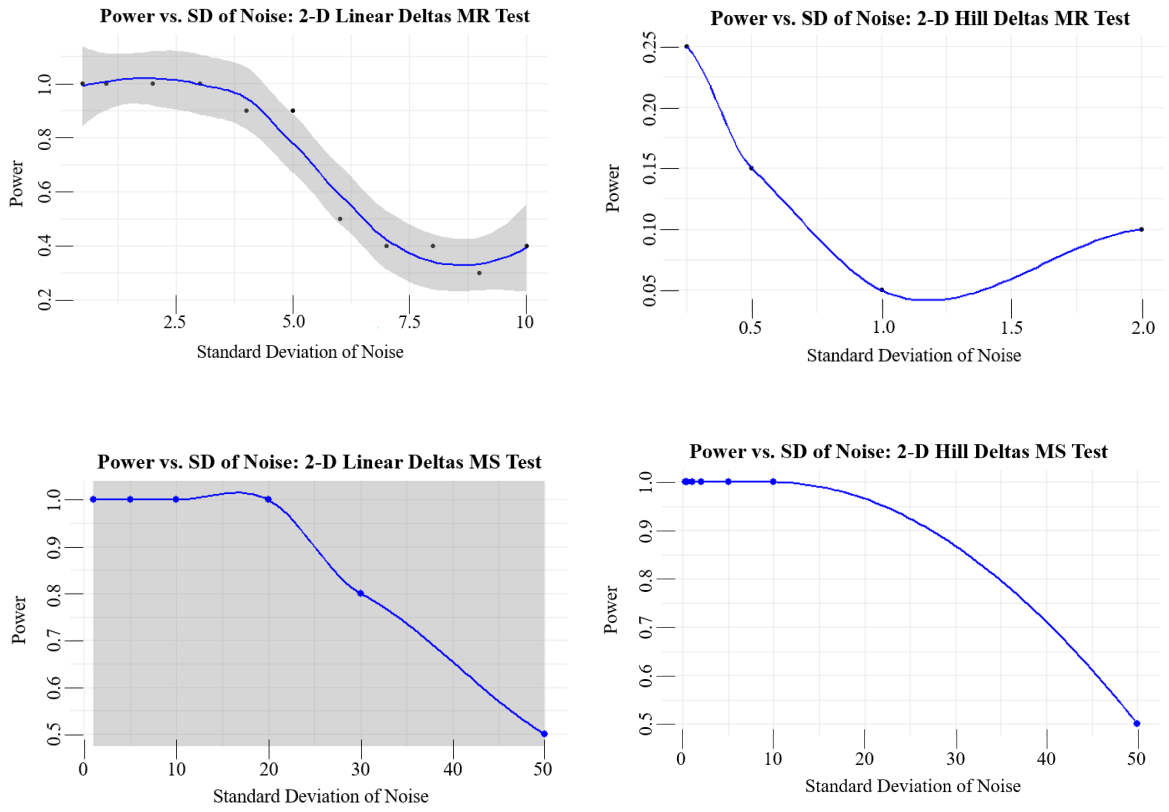
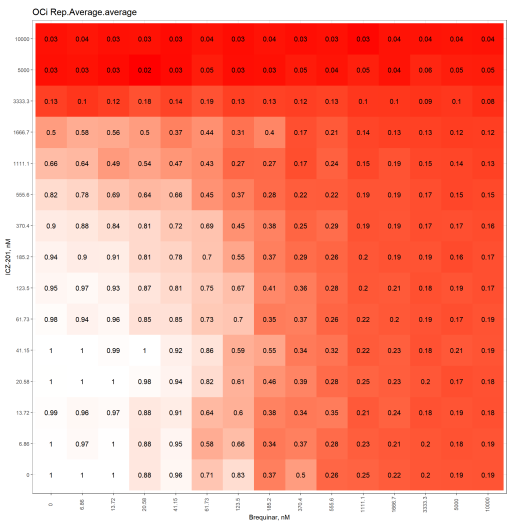


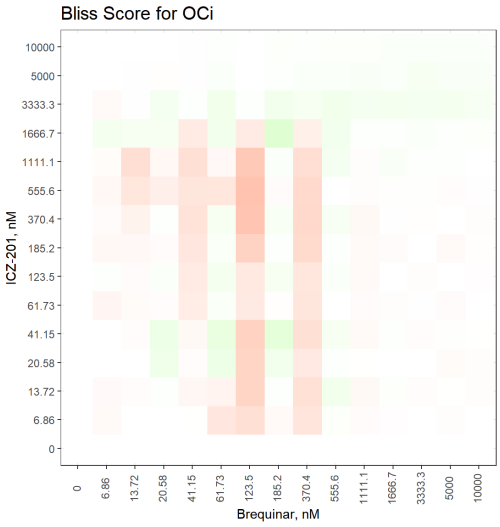
Figure 3.11: Comparison of test powers: The first row shows the power of the MS test, while the second row shows the power of the MR test as a function of noise levels, for both the 2D linear surface (first column) and Hill equation surface (second column).

3.2 Real Data

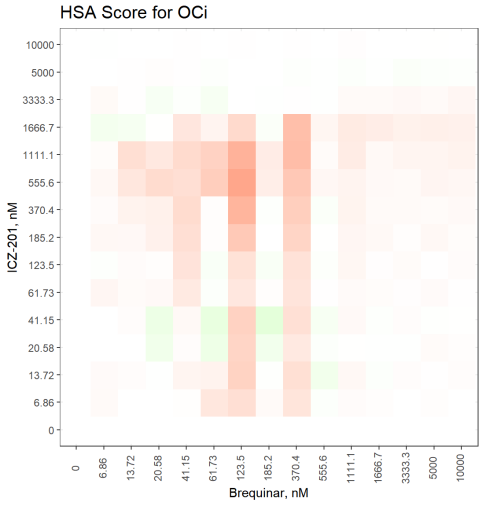
The dataset contains observations from a study on the combined effect of Itraconazole (ICZ) and Brequinar (known cancer drug treatments) on the OCi cancer cell line. Four replicates were performed for each combination, and the mean response is seen in Figure 3.12, which also includes two plots representing synergy scores. The results exhibit a discernible monotonic trend. To quantify this observation, the MR and MS tests were conducted, yielding p-values of 0.012 and 0.005, respectively, affirming the presence of monotonicity in the response. We can also see in Figure 3.12 that there is not a strong consensus in synergy scores, indicating that isotonic regression may be more helpful at elucidating effective combinations of drugs.



(a) Average response heatmap of the OCi cell line to ICZ and Brequinar.



(b) Synergy score plot using Bliss method.



(c) Synergy score plot using HSA method.

Figure 3.12: Visual representation of drug efficacy and synergy: (a) Average response to the drug combination across four replicates, (b) Synergy score calculated using the Bliss method, and (c) Synergy score calculated using the HSA method.

3.2.1 Isotonic Regression on Real Data

Isotonic regression was applied to the real-world dataset to identify the optimal drug combination for maximal therapeutic efficacy. The results, depicted in Figure 3.13, indicate promising outcomes. In the graphical representation, each colored patch corresponds to a specific combination of drugs, with the yellow square highlighting the lowest dose, most effective combination within that dose area. This information is partic-

ularly valuable in clinical practice where dosage constraints are often imposed due to safety concerns and potential side effects. Thus, the presented analysis provides a visual tool to select the most effective drug combination within the permissible dosage limits. The left panel of the figure uniquely displays the response values.

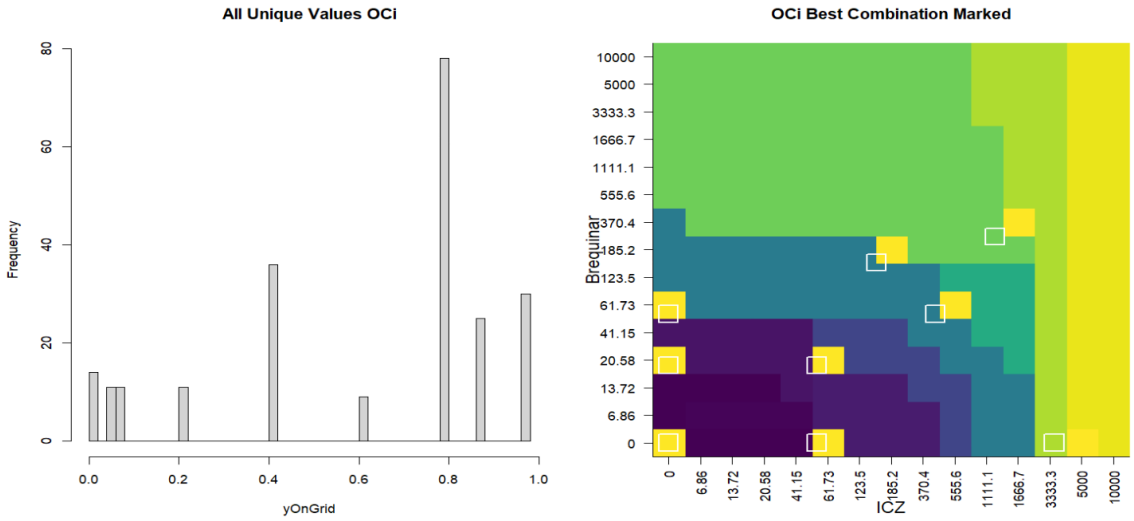


Figure 3.13: Isotonic regression analysis of the OCi cell line response to drug combinations. The left panel shows the unique response values, and the right panel illustrates the results of isotonic regression, with the yellow square indicating the optimal drug combination for maximum efficacy.

CHAPTER 4

DISCUSSION

As the scientific field continues to expand research findings related to drug combinations for therapies, it is important that we can properly assess drug combinations. By creating a pipeline to efficiently and accurately normalize plates, we reduce costs and improve efficiency of the experiments. However, when it comes to understanding true drug responses, we want to be sure we are as accurate as possible when determining the effect of two drugs. This is important because by finding effective drug pairings, we can improve efficacy of treatments while reducing side effects from medication. We propose a novel method in addition to synergy that vastly reduces the assumptions that need to be made about how the drugs behave and interact with each other.

In Chapter 2, we presented two ways to assess monotonic trends in drug response data. In Chapter 3, we simulated data to follow trends of monotonicity, included violations of monotonicity, and replicated what occurs in real-life experiments with drug combinations. We assessed the monotonicity of the data and also determined the sensitivity of the tests by looking at the power as the noise of the data increases. We found that the MS test method leads to more accurate conclusions about the data and is less sensitive to the noise within the data. We then used the tests to assess real-life data and confirm that it was monotone. Since the monotone assumption was met, we applied isotonic regression to the real-life data to help determine the most effective drug combinations within certain limitations of the drug doses.

There are obvious limitations of the test if the data is too noisy as was seen in Chapter 3. While we addressed the plateaus of the Hill equation within the tests, we did not account for magnitudes of monotone trends that may affect the statistics. Future tests can help characterize the magnitude of the monotone data and take that into account when testing for monotonicity. Future experiments can also look at other real-world data that has been documented to have synergy scores in disagreement (ie, Bliss and Loewe come to an opposite consensus) and assess if isotonic regression is helpful in determining the true effects of the drug combinations.

In summary, the monotone test for data will be helpful in future drug combination experiments. Since this is the only assumption for isotonic regression, if it is met, then isotonic regression can be run on the data. This can help elucidate the true effects of the drug combinations within varying dosage concentrations, and help clinicians determine the optimal dosage to maintain the most therapeutic effects while reducing side effects of the drugs for patients.

References

An, W. F. and Tolliday, N. (2010). Cell-based assays for high-throughput screening. *Molecular biotechnology*, 45:180–186.

Attene-Ramos, M. S., Austin, C., and Xia, M. (2014). High throughput screening.

Dagogo-Jack, I. and Shaw, A. T. (2018). Tumour heterogeneity and resistance to cancer therapies. *Nature reviews Clinical oncology*, 15(2):81–94.

Dai, R., Song, H., Barber, R. F., and Raskutti, G. (2020). The bias of isotonic regression. *Electronic journal of statistics*, 14(1):801.

Doig, S. D., Baganz, F., and Lye, G. J. (2006). High throughput screening and process optimisation. *Basic biotechnology*, 3.

Eder, J. and Herrling, P. L. (2016). Trends in modern drug discovery. *New approaches to drug discovery*, pages 3–22.

Goldreich, O. (2017). *Introduction to property testing*. Cambridge University Press.

He, L., Kulesskiy, E., Saarela, J., Turunen, L., Wennerberg, K., Aittokallio, T., and Tang, J. (2018). Methods for high-throughput drug combination screening and synergy scoring. *Cancer systems biology: methods and protocols*, pages 351–398.

Hernán, M. A. and Robins, J. M. (2010). *Causal inference*. CRC Boca Raton, FL.

Huang, W., Zhang, X. D., Li, Y., Wang, W. W., and Soper, K. (2012). Standardized median difference for quality control in high-throughput screening. In *2012 International Symposium on Information Technologies in Medicine and Education*, volume 1, pages 515–518. IEEE.

Hwangbo, H., Patterson, S., Dai, A., Plana, D., and Palmer, A. C. (2023). Additivity predicts the efficacy of most approved combination therapies for advanced cancer. *medRxiv*.

Light, J. G., Haidari, W., and Feldman, S. R. (2019). Assessing efficacy and the speed of response in psoriasis treatment.

Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J., and Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. *Nature biotechnology*, 24(2):167–175.

Meyer, C. T., Wooten, D. J., Lopez, C. F., and Quaranta, V. (2020). Charting the fragmented landscape of drug synergy. *Trends in pharmacological sciences*, 41(4):266–280.

Murie, C., Barette, C., Lafanechère, L., and Nadon, R. (2014). Control-plate regression (cpr) normalization for high-throughput screens with many active features. *Journal of biomolecular screening*, 19(5):661–671.

Patton, A. J. and Timmermann, A. (2010). Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics*, 98(3):605–625.

Riss, T. L., Moravec, R. A., Niles, A. L., Duellman, S., Benink, H. A., Worzella, T. J., and Minor, L. (2016). Cell viability assays. *Assay Guidance Manual [Internet]*.

Tallarida, R. J. (2011). Quantitative methods for assessing drug synergism. *Genes & cancer*, 2(11):1003–1008.

van Wijk, R., Tans, S. J., Wolde, P. R. t., and Mashaghi, A. (2015). Non-monotonic dynamics and crosstalk in signaling pathways and their implications for pharmacology. *Scientific reports*, 5(1):11376.

Wei, F., Wang, S., and Gou, X. (2021). A review for cell-based screening methods in drug discovery. *Biophysics Reports*, 7(6):504.

Yadav, B., Wennerberg, K., Aittokallio, T., and Tang, J. (2015). Searching for drug synergy in complex dose–response landscapes using an interaction potency model. *Computational and structural biotechnology journal*, 13:504–513.