A Non-linear Mixed Modeling Framework for Quantifying Drug Synergy in Patient-Derived Xenograft Data

Junho Yoon

Graduate School of Public Health

Yonsei University

Department of Health Informatics & Biostatistics

Table of Contents

List of Tables

I. Introduction

Combination therapy has emerged as a promising approach in cancer treatment, offering potential advantages over monotherapy by targeting multiple oncogenic pathways simultaneously and overcoming resistance mechanisms (Mokhtari et al., 2017). A key goal in developing combination therapies is to identify combinations that produce a synergistic effect, where the combined effect is greater than the additively predicted effect. Patient-Derived Xenograft (PDX) models, where patient tumors are implanted into immunodeficient mice, have become a vital preclinical data for evaluating such therapies. Their strength lies in retaining the histopathological features and genetic heterogeneity of the original human tumor, leading to a high correlation between drug responses in PDX models and clinical outcomes in patients (Gao et al., 2015).

The analysis of combination treatment in PDX experiments often involves comparing tumor volumes at a fixed time point or using summary metrics like tumor growth inhibition (TGI). While useful, these methods may not fully represent the longitudinal nature of tumor growth data. For example, combPDX, a statistical framework for assessing drug synergy from tumor growth curves, assesses the combination effect based on the observed tumor volumes, without assuming an underlying model of tumor growth (Huang et al., 2022).

Mathematical modeling of tumor growth provides a more mechanistic approach to understanding and predicting therapeutic effects. Models such as the exponential, logistic, and Gompertzian curves have been widely used to describe the dynamics of tumor progression (Ribba et al., 2014). By characterizing the baseline growth trajectory, one can model the effect of a drug as a perturbation of this growth.

This study proposes a novel framework for assessing the joint action of combination therapies by integrating nonlinear mixed-effects (NLME) models of tumor growth with dynamic drug effect models. The present method first characterizes the inherent growth dynamics of untreated tumors and then quantifies drug effects as a deviation from this baseline. This allows for a more direct and mechanistically interpretable assessment of how drugs, both alone and in combination, alter the tumor's natural trajectory. We introduce a methodology to: (1) select an appropriate tumor growth model from several candidates; (2) use the Stochastic Approximation Expectation-Maximisation (SAEM) algorithm to fit the NLME models; (3) using distance-based statistics to test the hypothesis of synergy; and (4) validate our approach using both actual and synthetically generated PDX datasets.

II. Materials and Methods

1. Dataset

1.1. Novartis PDX encyclopedia

The experimental data for this study were sourced from the Novartis Institutes for BioMedical Research PDX Encyclopedia (NIBR PDXE), a large-scale collection of approximately 1,000 patient-derived xenograft models developed to better predict clinical trial outcomes. This resource was used to perform extensive *in vivo* compound screening with a 'one animal per model per treatment' (1 × 1 × 1) experimental design.

We curated a dataset for this analysis from the publicly available Novartis Institutes for BioMedical Research PDX Encyclopedia (NIBR PDXE). The dataset comprises 1,951 unique model-treatment experiments. These experiments span six major cancer indications: breast cancer (BRCA), cutaneous melanoma (CM), colorectal cancer (CRC), gastric cancer (GC), and non-small cell lung cancer (NSCLC). The dataset is composed of 218 control arms, 999 monotherapy arms, and 734 combination therapy arms, providing a robust foundation for developing and testing our modeling framework (Melillo et al., 2023; Gao et al., 2015).

To meet the requirements of our NLME framework and to reduce inter-study variability, we selected PDX models (subjects) for which complete longitudinal tumor growth data were available for all four treatment groups: the

control, monotherapy A, monotherapy B, and the combination of A and B.

1.2. Data normalization

To standardize the starting tumor volume and account for inherent inter-subject variability, all raw volume measurements (*V(t)*) for each subject were normalized to their respective initial volume (*V(0)*). The log-transformed normalized tumor volume was used to model the tumor growth and drug effect.

$$y(t) = \ln\left(\frac{V(t)}{V(0)}\right)$$

This transformation standardizes the starting point for all subjects by setting the initial log-relative volume, *y(0)*, to 0, which can simplify the parameters of the growth models.

Table 1. Data obtained from the Novartis Institutes for BioMedical Research PDX Encyclopedia

Tumor type	Treatment type	Treatment	No. Patient Models (Replications)
Breast Cancer	Control	-	39
	Monotherapy	BYL719	39
		LEE011	39
		LJM716	39
		Trastuzumab	38
	Combination therapy	BYL719 + LEE011	39
		BYL719 + LJM716	39
		LJM716 + trastuzumab	38
Cutaneous Melanoma	Control	-	33
	Monotherapy	BMK120	33
		LEE011	33
		binimetinib	33
		Encorafenib	33
	Combination therapy	BKM120 + encorafenib	33
		LEE011 + binimetinib	18
		LEE011 + encorafenib	33
		encorafenib + binimetinib	33
Colorectal Cancer	Control	-	43
	Monotherapy	BKM120	41
		BYL719	42
		INC424	1
		LJC049	40
		binimetinib	43
		cetuximab	41
		encorafenib	42
		figitumumab	1
	Combination therapy	BKM120 + LJC049	40
		BKM120 + binimetinib	1
		BYL719 + binimetinib	42
		BYL719 + cetuximab	41
		BYL719 + encorafenib	41
		INC424 + binimetinib	1
		cetuximab + encorafenib	41
Gastric Cancer	Control	-	37
	Monotherapy	BYL719	37
		HSP990	37
		INC280	19

		LEE011	35
		LJM716	36
		everolimus	35
		trastuzumab	20
	Combination therapy	BYL719 + HSP990	37
		BYL719 + LJM716	33
		INC280 + trastuzumab	19
		LEE011 + everolimus	35
		LJM716 + trastuzumab	20
Non-Small Cell Lung Cancer	Control	-	29
	Monotherapy	BKM120	29
		BYL719	29
		LGH447	29
		binimetinib	29
	Combination therapy	BKM120 + binimetinib	29
		BYL719 + LGH447	29
Pancreatic Ductal Adenocarcinoma	Control	-	35
	Monotherapy	BKM120	30
		INC424	28
		binimetinib	35
		figitumumab	33
	Combination therapy	BKM120 + binimetinib	30
		INC424 + binimetinib	28
		figitumumab + binimetinib	33

2. Models

2.1. Growth and drug effect models

The structural model, $f(\psi_i, t)$, which defines the predicted tumor growth trajectory for an individual subject *i*, is formulated using a mechanistic framework based on ordinary differential equations (ODEs). This approach models the rate of change in tumor volume, allowing for a more direct interpretation of how drugs alter the tumor's natural growth trajectory.

2.1.1. General ODE framework

The foundational model assumes a first-order kill process where the drug's effect is proportional to the tumor volume at any given time. The rate of change for the tumor volume, $V(t)$, is a balance between the tumor's natural growth and the drug-induced killing effect:

$$\frac{dV(t)}{dt} = (Growth\ term) - (Drug\ effect\ rate) \cdot V(t)$$

Using the log-transformed normalized volume, this relationship can be expressed as a direct subtraction of rates, where the drug effect rate reduces the tumor's relative growth rate:

$$\frac{dy(t)}{dt} = f_{growth}(t, y) - Effect_g(t)$$

2.1.2. Baseline tumor growth models

We adopted three distinct mathematical models to describe the natural growth dynamics of tumor volume (Ribba et al., 2014).

1) Exponential growth: This model assumes an unrestricted, constant relative growth rate, *a*.

$$\frac{dy}{dt} = a$$

$$y(t) = at$$

2) Gompertz growth: This asymmetric sigmoidal model assumes an initially high relative growth rate that decelerates exponentially over time, reflecting biological constraints such as limited nutrient access. The model is defined by an initial growth velocity, α_1, and a rate of deceleration, α_2.

$$\frac{dy}{dt} = \alpha_1 - \alpha_2 y(t)$$

$$y(t) = \frac{\alpha_1}{\alpha_2}(1 - e^{-\alpha_2 t})$$

3) Logistic growth: This sigmoidal symmetric model assumes the growth rate decreases as a function of tumor size. It is defined by an relative intrinsic growth rate, *r*, and a carrying capacity, *K*.

$$\frac{dy}{dt} = r(1 - \frac{e^y}{K})$$

$$y(t) = \ln K - \ln\left[1 + \left(\frac{K}{V(0)} - 1\right)e^{-rt}\right]$$

2.1.3. Drug effect model

The therapeutic effect is modeled as a dynamic process that perturbs the baseline growth rate. For a given drug g, we consider two functional forms:

1) Constant effect: the simplest model where the drug exerts a constant kill rate, β_g. Here $g \in \{A, B\}$ indicates the drug identity. The same functional form is applied separately to each drug.

$$Constant\ drug\ effect\ rate: \beta_g$$

$$Effect_g(t) = \beta_g$$

2) Exponentially decaying effect: This model assumes the drug effect is maximal at treatment initiation (β_g) and decays exponentially over time with a rate constant, k_g. This structure can represent phenomena such as the development of drug resistance and metabolic breakdown.

$$Decaying\ drug\ effect\ rate: \beta_g e^{-k_g t}$$

$$Effect_g(t) = \beta_g(e^{-k_g t})$$

2.1.4. Structural models

By combining the three baseline growth functions, $f_{grwoth}(t)$, with the two dynamic drug effect functions $Effect_g(t)$, we generate a set of candidate structural models.

A single, unified structural model was constructed to analyze all treatment arms (Control, Monotherapy A, Monotherapy B, Combination) simultaneously. The model employs binary indicator variables, I_A, I_B, I_{AB}, to selectively apply drug effects corresponding to each subject's treatment group. The predicted log-normalized volume, y(t), for any subject is given by:

$$y(t) = f_{growth}(t) - (I_A + I_{AB}) \cdot Effect_A(t) - (I_B + I_{AB}) \cdot Effect_B(t)$$

Under this structure, the null hypothesis for the combination therapy arm (I_{AB}=1, I_A=0, I_B=0) is that the combined effect is the simple sum of the individual monotherapy effects:

$$y_{AB}(t) = f_{growth}(t) - Effect_A(t) - Effect_B(t)$$

The final structural model, $f(\psi_i, t_{ij})$, used in the subsequent NLME analysis is selected from these six options based on statistical goodness-of-fit criteria, ensuring the most appropriate model is chosen for any given dataset.

2.1.5. Equivalence to Bliss Independence

The additive null hypothesis used in our framework, where the combined

drug effect is the sum of the individual effects on the log-volume scale, is directly equivalent to the classical Bliss Independence model. The Bliss Independence principle is based on a probabilistic interpretation of drug action, where two drugs are said to act independently if the presence of one does not affect the probability of the other's effect on tumor cells (Huang et al., 2022).

This equivalence can be formally demonstrated. Let $V_g(t)$ be the raw tumor volume at time *t* for a subject in treatment group *g* (g = A or B), where *g* can be Control (C), Drug A, Drug B, or the combination (AB).

The Bliss Independence model is most directly expressed in terms of the surviving fraction of tumor volume relative to the control. The surviving fraction for the combination is the product of the surviving fractions for the individual monotherapies:

$$\frac{AB(t)}{V_C(t)} = \frac{V_A(t)}{V_C(t)} \cdot \frac{B(t)}{V_C(t)}$$

Taking the natural logarithm of both sides of the equation yields:

$$\ln(\frac{AB(t)}{V_C(t)}) = \ln\left(\frac{V_A(t)}{V_C(t)}\right) + \ln(\frac{V_B(t)}{V_C(t)})$$

This can be expanded to show the additive relationship on the raw log-volume scale:

$$\ln(V_{AB}(t)) - \ln\big(V_C(t)\big) = \big[\ln(V_A(t)) - \ln\big(V_C(t)\big)\big] + [\ln(V_B(t)) - \ln\big(V_C(t)\big)]$$

$$\ln(V_{AB}(t)) = \ln(V_A(t)) + \ln(V_B(t)) - \ln(V_C(t))$$

This relationship can be translated into the log-normalized volume, $y(t) = \ln(V(t)/V(0))$. By subtracting $\ln(V(0))$ from each term in the equation above, we get:

$$y_{AB}(t) = y_A(t) + y_B(t) - y_C(t)$$

Finally, we can substitute our structural model definitions into this equation. Recall that for any treatment group g, the predicted log-normalized volume is:

$$y_g(t) = f_{growth}(t) - Effect_g(t)$$

with the drug effect for the control group being zero ($Effect_C(t) = 0$) in the following equation,

$$y_{AB}(t) = [f_{growth}(t) - Effect_A(t)] + [f_{growth}(t) - Effect_B(t)] - [f_{growth}(t)]$$

Simplifying this expression reveals the null model for the combination therapy model used in our framework:

$$y_{AB}(t) = f_{growth}(t) - Effect_A(t) - Effect_B(t)$$

Therefore, the null hypothesis of our unified structural model, that the drug effects are additive in the log-normalized unit is mathematically equivalent to the widely accepted Bliss Independence principle of non-interaction.

2.2. General NLME model framework

To analyze the longitudinal tumor growth data from a heterogeneous population of PDX models, we employ a Nonlinear Mixed-Effects (NLME) modeling framework. This hierarchical approach is essential for simultaneously characterizing the typical population response (fixed effects) and the subject-to-subject variability (random effects). The complete mathematical structure, as implemented in the R *saemix* package (ver.3.3), is expressed as follows:

$$y_{ij} = f(\psi_i, t_{ij}) + \epsilon_{ij}$$

where y_{ij} is the observed log-normalized tumor volume for subject *i* at time *j*. $f(\psi_i, t_{ij})$ is the nonlinear model that predicts the tumor volume based on the individual parameter vector ψ_i for subject *i*. ψ_i is the vector of individual parameters for subject *i* (e.g., growth rates, drug sensitivity). ϵ_{ij} is the residual error term, which accounts for measurement error and intra-subject variability. We assume the errors are independent and identically distributed, following a normal distribution with mean 0 and variance σ^2.

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

2.3. Hierarchical model selection

A unified NLME model is fitted simultaneously to the complete longitudinal dataset, encompassing all four treatment arms (Control, Drug A,

Drug B, and Combination). This approach allows the model to estimate a single, consistent set of subject-specific growth parameters informed by the data from all treatment conditions for that subject.

2.3.1. Individual parameter model and transformation

Each individual parameter in the vector ψ_i is modeled as a function of a fixed effect and a random effect. To ensure that biological parameters such as growth rates and drug effect magnitude are constrained to be positive, a log-normal distribution is assumed.

$$\psi_i = \exp(\phi_i)$$

where ϕ_i is the vector of transformed parameters for subject *i*, defined by the linear model:

$$\phi_i = \mu + \eta_i$$

In the model, μ is the vector of fixed effects, which is the logarithm of the population parameter, and η_i is the vector of random effects for subject *i*, which is the deviation of that subject's transformed parameters from the population mean. The random effects are assumed to follow a multivariate normal distribution, $\eta_i \sim N(0, \Omega)$.

2.3.2. Covariance structure

The covariance matrix of the random effects, Ω, defines the magnitude of the inter-subject variability. For this framework, a diagonal covariance structure is used. This structure assumes that the random effects for different parameters are independent, with each parameter having its own variance component, ω_k^2.

$$\Omega = \begin{pmatrix} \omega_1^2 & 0 & 0 & 0 \\ 0 & \omega_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \omega_p^2 \end{pmatrix}$$

2.3.3. Individual parameters

This framework is applied to all parameters in the selected structural model. The specific formulation for each parameter *k* for subject *i* is $\psi_{ik} = \exp(\mu_k + \eta_{ik})$ (Drikvandi 2017). Below are the explicit definitions for each model component.

1) Exponential growth: for the growth rate a:

$$a_i = \exp(\mu_a + \eta_{a,i})$$

2) Gompertz growth: for the initial growth velocity α_1, and the deceleration rate α_2:

$$\alpha_{1i} = \exp(\mu_{\alpha 1} + \eta_{\alpha 1,i})$$

$$\alpha_{2i} = \exp(\mu_{\alpha 2} + \eta_{\alpha 2,i})$$

3) Logistic growth: for the carrying capacity *K* and intrinsic growth rate *r*:

$$K_i = \exp(\mu_K + \eta_{K,i})$$

$$r_i = \exp(\mu_r + \eta_{r,i})$$

4) Constant drug effect: for the effect magnitude β_g:

$$\beta_g = \exp(\mu_{\beta g} + \eta_{\beta g,i})$$

5) Decaying drug effect: for the effect magnitude β_g and the decay rate k_g:

$$\beta_g = \exp(\mu_{\beta g} + \eta_{\beta g,i})$$

$$k_{gi} = \exp(\mu_{kg} + \eta_{kg,i})$$

2.3.4. Random effect selection

To identify the most parsimonious model that adequately describes the data without over-parametrization, a stepwise selection procedure was employed. This process performs a joint selection of both fixed-effect covariates and the random effect structure. The selection algorithm is guided by Bayesian Information Criterion (BIC), a penalized likelihood criterion used for model

comparison. The goal is to find the model that minimizes this value. The general formula for BIC is:

$$BIC = -2 \cdot \ln(\hat{L}) + k \cdot \ln(N)$$

where $\hat{L}$ is the maximized likelihood of the model, k is the number of estimated parameters, and N is the number of subjects. The *saemix* package uses specific versions of BIC that are adapted for covariate and random effect selection in mixed-effect models.

The backward elimination procedure begins with a full model, which includes all candidate covariates and a random effect on every parameter. From this starting point, the algorithm iteratively removes one term at a time.

At each step, the algorithm evaluates the effect of removing each term by testing:

1) Covariance relationships: for each parameter ψ_k with a covariance effect β_{kj} in the model, the algorithm calculates the BIC of the model without that specific effect. The underlying statistical model for the log-transformed parameter ϕ_{ik} for subject *i* is:

$$\phi_{ik} = \mu_k + \sum_{j=1}^{q} \beta_{kj} C_{ij} + \eta_{ik}$$

The algorithm identifies the single covariate effect β_{kj} whose removal results in the largest decrease in BIC.

2) Random effect: Simultaneously, for each parameter with a random effect component η_{ik} the algorithm tests the effect of removing it. This is equivalent to setting the corresponding variance ω_k^2 to zero.

The algorithm then removes the single term, either a covariate effect or a random effect, that results in the greatest improvement (lowest value) for the model's BIC. This elimination process continues until no further terms can be removed without increasing the BIC. The final output is the optimal model structure reached through this subtractive process.

The final mathematical form of an individual subject's parameters depends on the outcome of the random effect selection process. For parameters where a random effect was retained, the value for subject *i* is a combination of fixed and random effect:

$$\psi_{ik} = \exp(\mu_k + \eta_{ik}), \qquad \eta_{ik} \sim N(0, \omega_k^2)$$

for parameters where the random effect was removed, the parameter is described only by the fixed effect. Its value is therefore constant across all subjects in the population:

$$\psi_{ik} = \exp(\mu_k)$$

2.4. Parameter estimation using SAEM algorithm

The population parameters of the unified NLME model, denoted by the

vector $\theta = (\mu, \Omega, \sigma^2)$ are estimated by maximizing the marginal log-likelihood of the observed data. For a population of *N* subjects, the likelihood function is given by:

$$l(\theta; y) = p(y; \theta) = \prod_{i=1}^{N} p(y_i; \theta) = \prod_{i=1}^{N} \int p(y_i|\psi_i; \theta) p(\psi_i; \theta) d\psi_i$$

where the integral is taken over the distribution of the individual random effects. For nonlinear models, this integral has no analytical or closed-form solution, making direct maximization of the likelihood intractable. To overcome this challenge, we employ the Stochastic Approximation Expectation-Maximization (SAEM) algorithm, a robust iterative method for maximum likelihood estimation in NLME models, as implemented in the R package *saemix* (ver. 3.3) (Comets et al., 2017).

The SAEM algorithm is an extension of the standard Expectation-Maximization (EM) algorithm, which treats the unobserved individual parameters, ψ_i, as missing data. The standard EM algorithm iterates between two steps: an Expectation (E) step, where the expectation of the complete data log-likelihood is computed, and a Maximization (M) step, where the population parameters are updated to maximize this expectation. However, in the context of NLME models, the E-step is itself an intractable integral.

Therefore, SAEM replaces the intractable E-step with a stochastic procedure, breaking each iteration *k* into three steps: Simulation (S-step), Stochastic Approximation (SA-step), and Maximization (M-step). This three-step

cycle is repeated until the parameter estimates converge.

In S-step, instead of calculating the conditional expectation analytically, we simulate it. A set of individual parameters, $\psi_i^{(k)}$, is drawn for each subject from the current estimate of its conditional distribution, $p\left(\psi_i \middle| y_i; \theta^{(k-1)}\right)$. This is performed using a Markov Chain Monte Carlo (MCMC) procedure in the package.

Then, in SA-step, the conditional expectation of the complete data log-likelihood, $Q_k(\theta)$, is updated using the simulated parameters from the S-step. This is not a full recalculation but a gradual update based on a decreasing step size, γ_k:

$$Q_k(\theta) = Q_{k-1}(\theta) + \gamma_k(\log(p(y, \psi^{(k)}; \theta)) - Q_{k-1}(\theta))$$

The sequence of step sizes, γ_k, is crucial for convergence. It is typically set to 1 for an initial number of iterations to allow the algorithm to explore the parameter space, and then it decreases towards zero to ensure the algorithm converges to a stable solution.

Finally, in M-step, the population parameters are updated by finding the value, $\theta^{(k)}$, that maximizes the updated log-likelihood function, $Q_k(\theta)$. To ensure a thorough exploration of the parameter space and to reduce the risk of converging to a local maximum, the *saemix* package utilizes the simulated annealing algorithm, which is a type of meta-heuristic algorithms.

3. Hypothesis testing

The selection of the most appropriate growth model for a given dataset is a critical step to avoid model misspecification (Mould and Upton 2013). Therefore, before the formal hypothesis test is conducted, a model selection procedure is performed. We evaluated the six potential structural models, formed by combining the three baseline growth functions with the two drug effect models. The model with the best fit to the data, as determined by the Bayesian Information Criterion (BIC), is selected and carried forward for the hypothesis test of synergy.

3.1. Null hypothesis

The assessment of drug synergy is devised as a formal statistical hypothesis test. The null hypothesis (H_0) assumes that the two drugs act independently, which, as established in section 2.1.5, is equivalent to the classical Bliss Independence model of non-interaction. The core of the test is to compare the observed longitudinal tumor growth data from the combination therapy arm to the predicted mean trajectory generated under this null hypothesis.

3.2. Distance-based statistics

To quantify the deviation between the observed data and the model's prediction under additivity, we employ a distance-based statistic. This statistic

provides a standardized measure of how far the observed data for a single subject deviates from the expected behavior under the null hypothesis. For each subject *i* in the combination drug group, the distance D_i, is calculated as:

$$D_i = \left[y_i - L(t, \hat{\theta}_0)\right]^T [Cov(y_i)]^{-1} \left[y_i - L(t, \hat{\theta}_0)\right]$$

This formula is the squared Mahalanobis distance, a multivariate statistic that standardizes the residual vector by the full variance-covariance matrix of the observations. The components of this equation are defined as follows:

1) y_i is the vector of n_i observed log-normalized tumor volumes for subject *i*. This is the actual experimental data collected for that subject.

2) $L(t, \hat{\theta}_0)$ is the predicted mean tumor growth trajectory under the null hypothesis of additivity. This prediction is generated using the population fixed-effect parameters ($\hat{\theta}$) that were estimated from the unified NLME model described in the preceding sections. It represents the model's best prediction of how a typical subject's tumor would grow if the two drugs acted together with no interaction.

3) $\left[y_i - L(t, \hat{\theta}_0)\right]$ is the vector of raw residuals for subject *i*, representing the difference between the observed data and the null prediction at each time point.

4) $Cov(y_i)$ is the estimated variance-covariance matrix of the observations for subject *i*. This component captures the total expected variability

in the data from two sources, as derived from the NLME model parameters. The intra-subject variability, which arises from the random effects, is calculated as $\left[\nabla_{\hat{\theta}_0} L(t, \hat{\theta}_0)\right] \hat{\Omega} \left[\nabla_{\hat{\theta}_0} L(t, \hat{\theta}_0)\right]^T$, where $\hat{\Omega}$ is the estimated variance-covariance matrix of the random effects and $\nabla_{\hat{\theta}_0} L(t, \hat{\theta}_0)$ is the Jacobian matrix of the model function with respect to the parameters. The intra-subject residual error, which represents measurement error and other unexplained variability is captured by the term $\hat{\sigma}^2 I$, where $\hat{\sigma}^2$ is the estimated residual error variance and I is the identity matrix. Therefore, the full expression of the covariance matrix is:

$$Cov(y_i) = \left[\nabla_{\hat{\theta}_0} L(t, \hat{\theta}_0)\right] \hat{\Omega} \left[\nabla_{\hat{\theta}_0} L(t, \hat{\theta}_0)\right]^T + \hat{\sigma}^2 I$$

5) $[Cov(y_i)]^{-1}$ is the inverse of the variance-covariance matrix. Its inclusion in the formula is to apply weights in the residuals, giving less weight to observations that are expected to be more variable and accounting for the correlation between repeated measurements on the same subject over time.

3.3. Testing methods

3.3.1. Method 1: Asymptotic Chi-squared test

The aggregated distance-based statistics, $\sum_{i=1}^{M} D_i$ (where M is the number of the subjects in the combination therapy group), is assumed to asymptotically follow a chi-squared (χ^2) distribution with degrees of freedom equal to the total number of observations. This is based on the properties of the Mahalanobis distance drawn under multivariate normal distribution

(Mukhopadhyay 2008).

However, practically, the true population parameters are unknown. We use estimates ($\hat{\theta}_0, \hat{\Omega}, \hat{\sigma}^2$) derived from the NLME model. When sample estimates are used, the exact distribution of the statistic is a Hotelling's T-squared (T^2) distribution, which can be asymptotically approximated when the sample size is large (Williams et al., 2018).

Under this classical framework, the aggregate test statistics $\sum_{i=1}^{M} D_i n_i$, is compared to a χ^2 distribution with $\sum_{i=1}^{M} n_i$ degrees of freedom. A p-value is calculated from this distribution to test for a significant deviation from additivity. A statistically significant result leads to the rejection of the additivity hypothesis, providing evidence for a synergistic or antagonistic interaction. The significance was tested at $p < 0.05$ level. Accordingly, the Type-I error rate (α) is fixed at 0.05.

3.3.2. Method 2: Bootstrap pivotal confidence interval

The χ^2 test described above relies on an asymptotic approximation that may not be reliable for the small sample sizes often encountered in PDX studies (Oberg et al., 2021). To perform a more robust and statistically valid hypothesis test, we also construct a confidence interval for the true distance statistic, D_i, using a bootstrap procedure to empirically derive its sampling distribution.

The pivotal method is based on finding a 'pivotal quantity', a function of the data and the unknown parameter whose distribution does not depend on the

parameter itself. For our case, let D_i be the true, unknown distance for subject *i*, and let $\widehat{D}_i$ be its estimate calculated from our original dataset. The pivotal quantity is the estimation error:

$$R_i = \widehat{D}_i - D_i$$

If we knew the distribution of R_i, we could find its $\alpha/2\,(q_{\alpha/2})$ and $1-\alpha/2\ (q_{1-\alpha/2})$ quantiles to construct a $100(1-\alpha)\%$ confidence interval for D_i as follows:

$$1-\alpha = P\left(q_{\alpha/2} \le \widehat{D}_i - D_i \le q_{1-\alpha/2}\right) = P\left(\widehat{D}_i - q_{1-\alpha/2} \le D_i \le \widehat{D}_i - q_{\alpha/2}\right)$$

However, R_i is unknown, so the bootstrap is used to approximate this unknown distribution. The bootstrap analogue, which is the difference between a bootstrap estimate $D^*_{i,b}$ and the original estimate $\widehat{D}_i$ is denoted as R^*_i.

$$R^*_i = D^*_{i,b} - \widehat{D}_i$$

We can generate a large number of these bootstrap replicates, $R^*_{i,b}$, to create an empirical distribution that serves as an approximation for the true distribution of R_i.

The empirical distribution of $R^*_{i,b}$ is generated via the following procedure. First, the unified NMLE model is fit to the original, complete dataset. Using the resulting population parameter estimates, the observed distance statistic, $\widehat{D}_i$, is calculated for each subject *i* in the combination therapy using the

formula from section 3.2. Then, a large number of bootstrap samples (e.g., B = 100 or 1,000) are generated. For each bootstrap iteration b = 1,...,B, a bootstrap dataset is created by resampling subjects with replacement from the original dataset. The entire unified NLME model is fit to this new bootstrap dataset, and using the parameters estimated from the bootstrap dataset, the bootstrap replicate of the distance statistic, $D_{i,b}^{*}$ is calculated for each subject i.

After the resampling loop, we have the bootstrap distribution of the statistic, $\{D_{i,1}^{*},\ D_{i,2}^{*},\cdots,D_{i,B}^{*}\}$ which is used to construct the pivotal confidence interval. Let $\hat{q}_{\alpha}^{*}$ be the empirical α-quantile of the bootstrap distribution of the statistic $D_{i,b}^{*}$. We use these quantiles to approximate the unknown quantiles, q_{α}, of the pivotal quantity's distribution. The $100(1-\alpha)\%$ pivotal confidence interval for the true distance, D_i, is then given by the formulas:

$$Lower\ Bound\colon L_i = 2\widehat{D}_i - \hat{q}_{1-\alpha/2}^{*}$$

$$Upper\ Bound\colon U_i = 2\widehat{D}_i - \hat{q}_{\alpha/2}^{*}$$

For a 95% confidence interval, the 97.5th and 2.5th percentile of the bootstrap distribution of $D_{i,b}^{*}$ was used.

The resulting confidence interval of D_i, (L_i, U_i), is used to test the null hypothesis of additivity. By using a 95% confidence interval, the Type-I error rate (α) for the test is fixed at 0.05, identical to the previous method using asymptotic χ^2 test. The null hypothesis of additivity corresponds to a true distance value of

D_i = 0. The null hypothesis is rejected if the interval does not contain 0.

Since the test statistic D_i is a squared value and is always non-negative, the upper bound, U_i, will be positive. The test therefore can be simplified to examining the lower bound of the interval, L_i. If $L_i > 0$, the null hypothesis is rejected and conclude that a statistically significant interaction (synergy or antagonism) is present. Otherwise, if $L_i \leq 0$, the null hypothesis is not rejected and there is insufficient statistical evidence at the 5% significance level to conclude that the there is interaction between drugs.

Following the rejection of the null hypothesis, which interaction between synergy and antagonism occurs must be determined by examining the direction of the deviation. This can be done by comparing the Area Under Curve (AUC), of the observed combination data to the AUC of the predicted additive curve. It is synergy If the null hypothesis is rejected and the observed AUC is significantly less than the predicted additive AUC. Oppositely, it is antagonism if the null hypothesis is rejected and the observed AUC is significantly greater than the predicted additive AUC.

4. Evaluation of the framework

4.1. Monte Carlo simulation

To rigorously evaluate the statistical properties of our proposed framework, we conducted a series of Monte Carlo simulation studies. The primary objectives of this evaluation were to (1) assess the Type-I error rate (α) of both the asymptotic χ^2 test and the bootstrap pivotal confidence interval

methods, (2) evaluate statistical power (1-β) of both methods, and (3) Compare the performance of the robust bootstrap method against the classical asymptotic approach, particularly in the context of sample sizes relevant to preclinical PDX studies.

Synthetic longitudinal datasets were generated from a unified NLME model with a known, pre-specified 'true' structure (e.g., a Gompertz growth model with a decaying effect). To ensure the simulations were realistic, the values for these population parameters were based on the estimates obtained from the analysis of the real Novartis PDX dataset (Table 2).

Table 2. Parameters used to generate synthetic logistic growth/drug effect dataset

Parameter	Description	Population mean	Standard Deviation
K	Carrying capacity		
r	Intrinsic growth rate		
β_A	Magnitude of effect of drug A		
k_A	Decay rate of drug A		
β_B	Magnitude of effect of drug B		
k_B	Decay rate of drug B		
σ	Residual standard deviation	0.1	-

4.2. Evaluation of Type-I error rate

To evaluate the actual Type-I error rate, we simulated datasets under the null hypothesis of perfect additivity. Followed by hypothesis testing for synthetic dataset generated, Type-I error rate was calculated as the proportion of

simulations in which the null hypothesis was incorrectly rejected at a significance level of α = 0.05. A valid test should have a Type-I error rate close to the nominal 0.05 level.

4.3. Evaluation of Type-II error rate

To evaluate the statistical power of our framework, we simulated datasets under various alternative hypotheses, representing different degrees of synergy and antagonism. This was achieved by introducing an interaction parameter, δ_{int}, into the drug effect model for the combination therapy during data generation.

$$f(\psi_i, t_{ij}) = f_{growth}(t_{ij}, \theta_{growth,i}) - \delta_{int} \cdot (Effect_A(t_{ij}, \theta_{drugA,i}) + Effect_B(t_{ij}, \theta_{drugB,i}))$$

Synergy was simulated by setting $\delta_{int} > 1$, representing the combined effect greater than the null additivity. On the other hand, antagonism was simulated by setting $0 \leq \delta_{int} < 1$, which indicates the combined effect is less than the null additivity. The null hypothesis of additivity itself corresponds to $\delta_{int} = 1$.

For each simulated dataset generated under these synergistic or antagonistic conditions, we applied our full analytical framework. The statistical power for each scenario was then calculated as the proportion of simulations in which the null hypothesis was correctly rejected at a significance level of α = 0.05. Furthermore, for the simulations where the null hypothesis was correctly

rejected, we also assessed the accuracy of classifying the interaction type. This was done by comparing the Area Under the Curve (AUC) of the observed combination data to the AUC of the predicted additive curve.

III. Results

IV. Discussion

V. Conclusion

Glossary

Symbol	Definition
y_{ij}	The observed log-normalized tumor volume for subject *i* at time *j*.
ψ_i	The vector of all individual parameters for subject *i*.
$\emptyset_i$	The vector of the log-transformed parameters for subject *i*.
μ	The vector of fixed effects (population means of the log-parameters).
η_i	The vector of random effects for subject *i*.
Ω	The variance-covariance matrix of the random effects.
σ^2	The variance of the residual error term, ϵ_{ij}.
a	Exponential growth rate parameter.
$\alpha_1,\ \alpha_2$	Gompertz growth parameters (initial growth rate and deceleration rate).
K, r	Logistic growth parameters (carrying capacity and growth rate).
$\beta_g,\ k_g$	Drug effect parameters (effect magnitude and decay rate) for drug *g*.
θ	A vector of model parameters. Used to denote either the set of all population parameters, or a subset of parameters for a specific model component
D_i	The distance-based statistic for subject *i*.
$L(t, \hat{\theta}_0)$	The predicted mean tumor growth trajectory under the null hypothesis.
δ_{int}	Interaction parameter in the synthetic data generation

Bibliography

Comets E, Lavenu A, Lavielle M (2017) Parameter estimation in nonlinear mixed effect models using saemix, an R implementation of the SAEM algorithm Journal of Statistical Software 80:1–41

Drikvandi R (2017) Nonlinear mixed-effects models for pharmacokinetic data analysis: assessment of the random-effects distribution Journal of pharmacokinetics and pharmacodynamics 44:223–232

Gao H et al. (2015) High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response Nature medicine 21:1318–1325

Huang L, Wang J, Fang B, Meric-Bernstam F, Roth JA, Ha MJ (2022) CombPDX: a unified statistical framework for evaluating drug synergism in patient-derived xenografts Scientific reports 12:12984

Mer AS et al. (2019) Integrative pharmacogenomics analysis of patient-derived xenografts Cancer research 79:4539–4550

Mokhtari RB, Homayouni TS, Baluch N, Morgatskaya E, Kumar S, Das B, Yeger H (2017) Combination therapy in combating cancer Oncotarget 8:38022

Mould DR, Upton RN (2013) Basic concepts in population modeling, simulation, and model-based drug development—part 2: introduction to pharmacokinetic modeling methods CPT: pharmacometrics & systems pharmacology 2:1–14

Oberg AL et al. (2021) Statistical analysis of comparative tumor growth repeated measures experiments in the ovarian cancer patient derived xenograft (PDX) setting Scientific reports 11:8076

Ranjan Mukhopadhyay A (2008) Multivariate attribute control chart using Mahalanobis D^2 statistic Journal of Applied Statistics 35:421–429

Ribba B et al. (2014) A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis CPT: pharmacometrics & systems pharmacology 3:1–10

Williams JD, Woodall WH, Birch JB, Sullivan JH (2006) Distribution of Hotelling's T^2 statistic based on the successive differences estimator Journal of Quality Technology 38:217–229