scientific reports

OPEN

CombPDX: a unified statistical framework for evaluating drug synergism in patient-derived xenografts

Licai Huang[1,5], Jing Wang[2], Bingliang Fang[3], Funda Meric-Bernstam[4], Jack A. Roth[3] & Min Jin Ha[6]

Anticancer combination therapy has been developed to increase efficacy by enhancing synergy. Patient-derived xenografts (PDXs) have emerged as reliable preclinical models to develop effective treatments in translational cancer research. However, most PDX combination study designs focus on single dose levels, and dose–response surface models are not appropriate for testing synergism. We propose a comprehensive statistical framework to assess joint action of drug combinations from PDX tumor growth curve data. We provide various metrics and robust statistical inference procedures that locally (at a fixed time) and globally (across time) access combination effects under classical drug interaction models. Integrating genomic and pharmacological profiles in non-small-cell lung cancer (NSCLC), we have shown the utilities of combPDX in discovering effective therapeutic combinations and relevant biological mechanisms. We provide an interactive web server, combPDX (https://licaih.shinyapps.io/CombPDX/), to analyze PDX tumor growth curve data and perform power analyses.

Combinations of anticancer drugs have been developed to circumvent mechanisms of resistance to yield clinical benefit and lower toxicity[1,2]. Recently, in vitro high-throughput combinatorial screening data have been enabling the assessment of large number of drug combinations at various dose levels[3–5]. In these experiments, a pair of drugs are plated in a dose–response matrix block and the data at various combinations of dose levels are analyzed to quantify the degree of combination effects. The joint effects are categorized into synergistic, additive, and antagonistic, which imply enhanced, independent, and reduced effect, respectively, when two drugs are present together. The responses obtained from multiple combinations of dose levels are compared against the expected response under null models where no combination effect is present. The classical reference null models include Highest single agent (HSA)[6–8], Loewe additivity[9–12] and Bliss independence (BI)[6,8,13,14]. Recent application developments, such as DrugComb[4], SynergyFinder[15], and Combenefit[16] provide computational tools to analyze drug combination screening data based on these reference models.

Patient-derived xenografts (PDXs) have emerged as reliable preclinical models to develop new treatments and biomarkers in translational cancer research[17]. The PDX models are developed by implanting tumors from patients into mice; this method has been suggested to more accurately reflect clinical outcomes[18,19]. These successes have led to a rapid accumulation and availability of large-scale PDX collections for drug discovery in cancer[18,20,21]. In the PDX experiments to evaluate drug efficacy, tumor volumes of each individual mouse are measured at the initiation of the study and periodically throughout the study. This usually continues until the tumor volume reaches a certain value, resulting in incomplete longitudinal tumor volume data. Due to the high cost of in vivo studies in animals, a common combination experiment for a PDX model includes four treatment groups, control (C), two monotherapies (A and B), and combination therapy (AB) with fixed doses to minimize

[1]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [2]Departments of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [3]Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [4]Department of Investigational Cancer Therapeutics, Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [5]Quantitative Sciences Program, The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences, Houston, TX 77030, USA. [6]Department of Biostatistics, Graduate School of Public Health, Yonsei University, Seoul, South Korea. ✉email: mjha@yuhs.ac
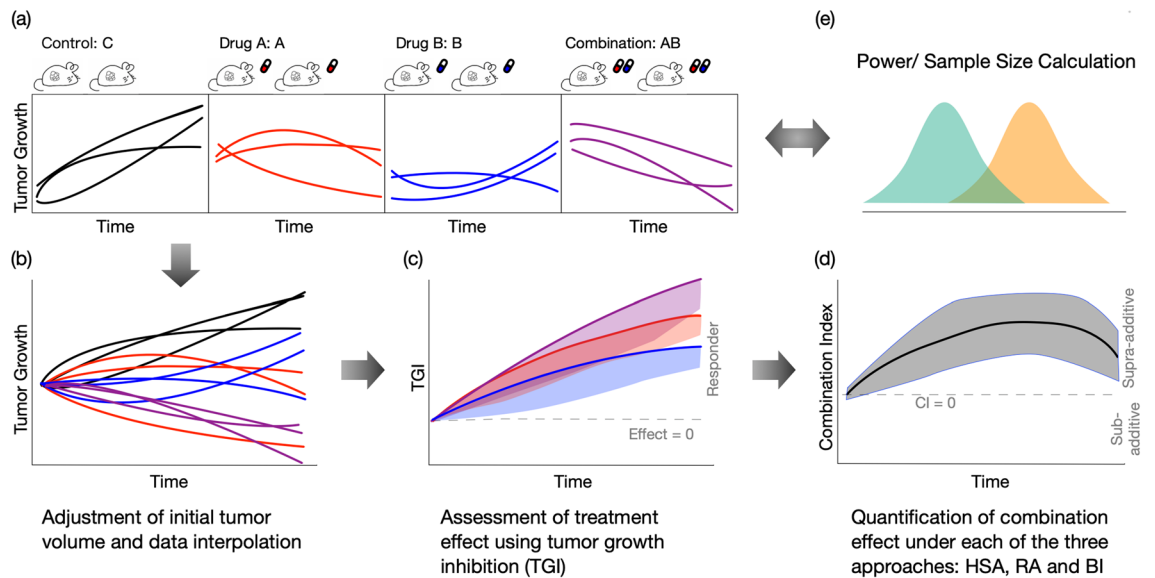
Figure 1. Overview of the analysis pipeline of combPDX. (**a**) Tumor volume measurements from four treatment arms (C, A, B, and AB) are collected from PDX experiments. (**b**) For an individual mouse, the response at each time point is determined by computing relative tumor volume to adjust for heterogeneous initial tumor measurements. (**c**) The drug effect for each treatment group (A, B, and AB) relative to the control group is quantified by tumor growth inhibition (TGI). (**d**) The combination effect under each reference model (HSA, RA, and BI) is assessed using a combination index. In addition, (**e**) sample size calculation and power analysis are provided.

the number of animals required per group. In the fixed dose experiment, the dose–response surface methods are not applicable, and the joint action of drugs should be evaluated at a fixed dose combination.

The statistical framework that assesses the joint action of drug combinations with fixed doses is not well developed. Wu et al.[22,23] proposed interaction indices and the statistical inferential procedure based on surviving fraction of cells and survival endpoints. Demidenko and Miller[24] proposed a log-linear model on the tumor volumes by assuming that they follow exponential growth. However, these are limited to the Bliss independence model. A distinctive set of mathematical definitions might lead to different quantifications of the degree of joint action.

In this article, we propose a comprehensive statistical framework to calculate combination indices (CIs) and the inferential procedures that are robust to potential outliers and errors in PDX experiments. We considered three most well-established reference models: HSA, response additivity (RA), and BI in a unified statistical quantification method. We present a user-friendly web server, combPDX, to compile the joint actions over time from PDX tumor growth data and to provide a power analysis tool to facilitate designs of PDX combination studies. Applying our methods to non-small-cell lung cancer (NSCLC), we show the utilities of our framework in finding underlying mechanisms of combination drug action using gene expression profiles for PDX models.

Results

Overview. The pipeline of combPDX includes three steps to assess combination effects as well as the power analysis procedure (Fig. 1). Longitudinal raw tumor volume measurements are collected from four treatment groups (C, A, B and AB) and the tumor growth curves are displayed (Fig. 1a). For each individual mouse, the response at each time point is determined by computing the relative tumor volume to adjust for heterogeneous initial tumor measurements across animals, and the missing relative tumor volumes at time t are interpolated using the neighboring measurements (Fig. 1b). In some PDX experiments, missing values may occur due to multiple inconsistent measurements or missing assessments at a given time point. At each time point, we determine the treatment effects of A, B, and AB compared to the control group C (Fig. 1c). Based on the treatment effects, we provide the CIs under HSA, RA, and BI, and the corresponding 95% confidence intervals (Fig. 1d). We finally implement a power analyses tool under the three reference models (Fig. 1e).

Simulation studies. We conducted a series of simulations to examine the performance of the proposed approaches. The expected tumor volumes under singe-agent treatment or control group were generated by the Gompertz tumor growth model over time[25,26]. Then the expected tumor volumes for combination drug were generated under each of the reference model (HSA, RA, and BI). The expected tumor growth curves for these simulation settings are shown in Fig. S1. For each setting, we generated 1000 replicate datasets with sample sizes of 5 or 10 per group. The detailed data generating process is described in Section S1. We evaluated the CIs with the confidence intervals at day 21 obtained from all the three reference models under each of the simulation scenarios. We compared the inferential methods with/without the bootstrap procedures based on 1,000 replications. The coverage probabilities are summarized in Table S1. When sample size is 5 in each group, the confidence intervals without the bootstrap procedure were slightly narrower than nominal level, resulting in inflated
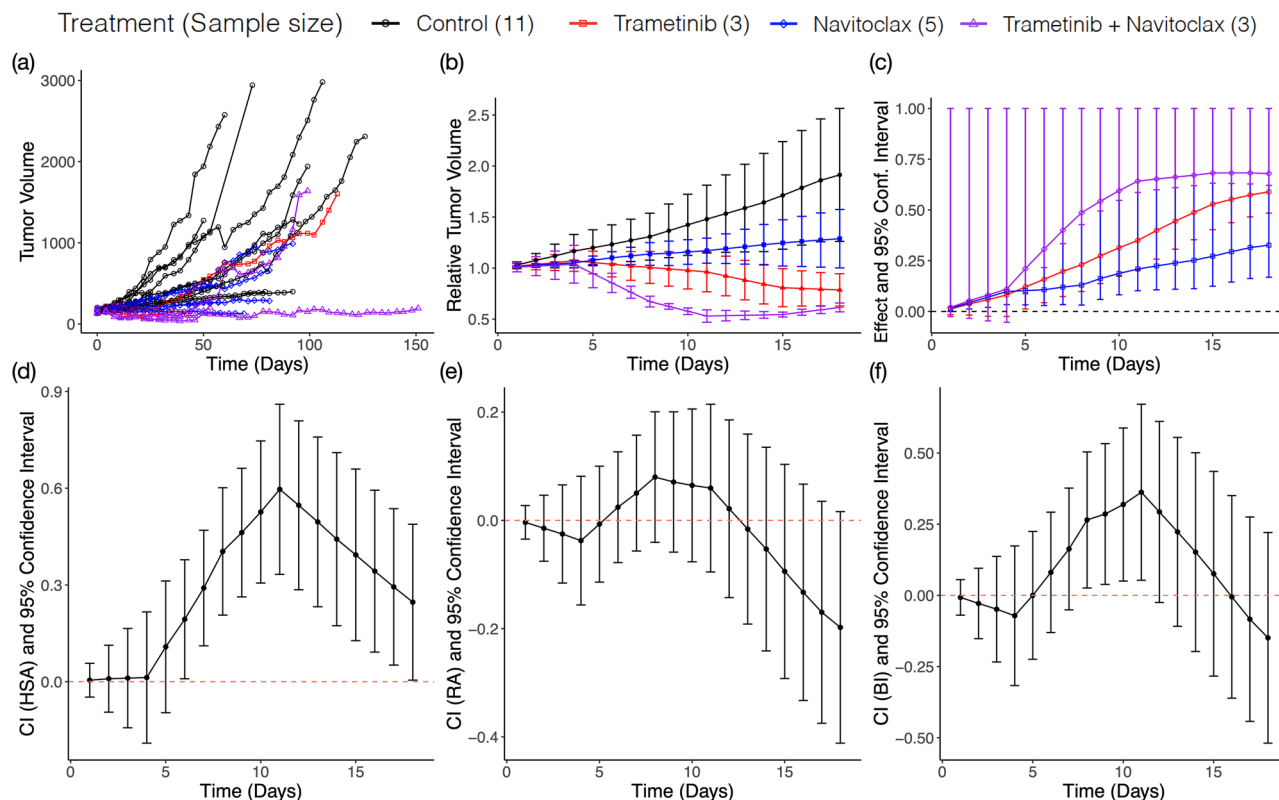
Figure 2. Effect and Combination Indices for trametinib plus navitoclax. (**a**) Profile plot of tumor volume. (**b**) Profile plot of relative tumor volume. Y-axis shows the mean ± standard error of relative tumor volume within each treatment group. (**c**) The drug effect for each treatment group relative to control group is quantified by tumor growth inhibition (TGI). The vertical line indicates the one-sided 95% confidence interval. (**d–f**) The joint action of combination drug under each reference model (HSA, RA, and BI) is assessed using a combination index. The vertical line indicates the two-sided 95% confidence interval.

type I error under the null hypothesis. When we have more sample size (n = 10) in each group, confidence intervals without bootstrap become close to the nominal coverage probability, and the confidence intervals with and without bootstrap tend to agree on each other. Overall, the bootstrap procedure helps in constructing more accurate confidence intervals when small sample size while both procedures provide valid statistical inferential performance when sample size becomes larger.

Evaluation of drug combinations of KRT232, navitoclax, and trametinib in NSCLC. We performed the analysis of a real study of the antitumor activity of the combination of KRT232, navitoclax, and trametinib using the PDX models for NSCLC, where a total of 28 PDX models were tested[27–29]. For each PDX tumor model, mice were randomized to the four treatment arms, and the tumor volume for each mouse was recorded every 2–3 days. 20 combination therapy experiments having sufficient samples sizes were selected (Section S2). All analyses were performed in R v4.0.2[30]. The treatment information for the selected experiments is summarized in Table S2. Table S3 summarizes the resulting CIs at day 10 and gCIs under the three reference models for each PDX model. trametinib and navitoclax combination shows the synergism and Fig. 2 displays the output from our combPDX analysis. Raw tumor volumes for all mice across the four groups are displayed in panel (a). After the preprocessing step, the average relative tumor volumes with 1-s.d. bars are displayed in panel (b). Then the effect sizes of each treatment groups A, B, AB relative to the control group are determined with 95% confidence intervals in panel (c), implying that all treatments trametinib, navitoclax, and the combination treatments have the significant antitumor activity compared to the control in the PDX experiment. Trametinib and navitoclax combination has the synergistic effect under HSA and BI from day 6 to day 18 and from day 8 to day 11, respectively (Fig. 2d–f and Table S3). This combination is currently under clinical investigation for the treatment of NSCLC[31].

Molecular biomarkers associated with CIs. Although KRT232 showed no synergistic signal in combination with navitoclax and trametinib (Table S3), we systematically investigated pathway-level signatures of the combination action in a framework of pharmacogenomic analysis utilizing multiple PDX experiments performed for the combinations. We conducted gene set variation analysis (GSVA)[32] based on C2 collection of curated biological pathways as provided by the Broad Institute's collection[33]. The pathway enrichment score (ES) by GSVA provides single-number summaries of pathway activity for each sample and each pathway. The PDX
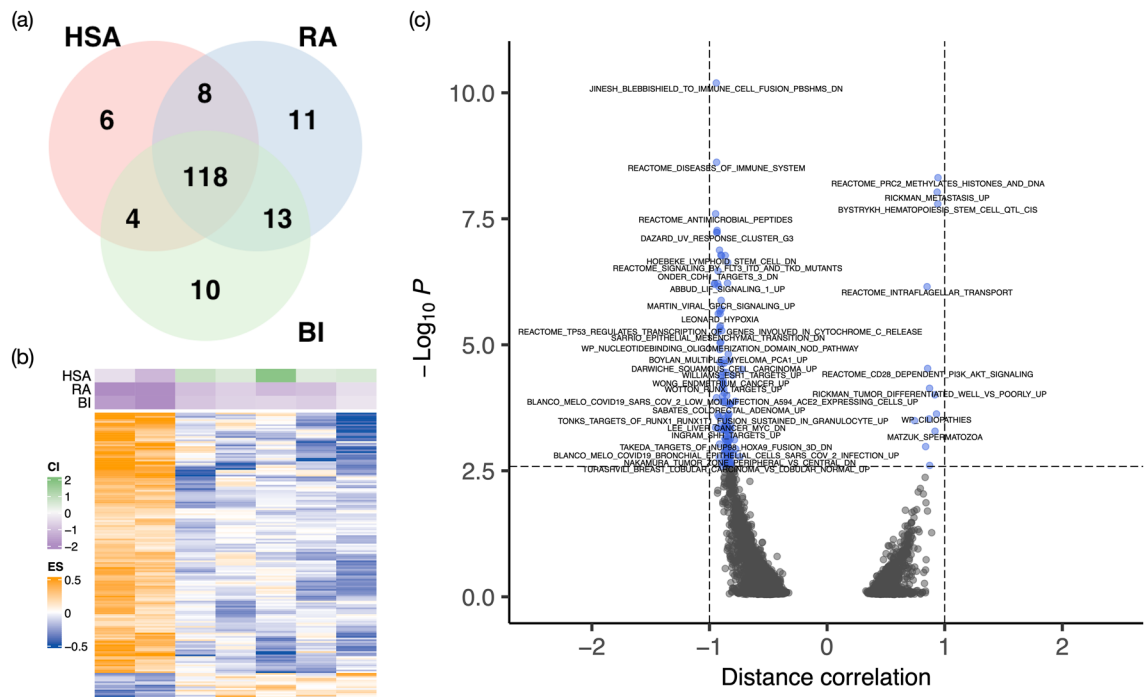
Figure 3. Differentially activated pathways at FDR of 0.1 for KRT232 plus navitoclax. (**a**) Venn diagram showing distribution of significantly expressed pathways. The figure illustrates the number of statistically significantly expressed pathways associated with each gCI (**b**) Heatmap of baseline pathway enrichment score from GSVA with the row annotation to be gCI. Each column represents a PDX model, and each row represents a gene whose enrichment score was significantly correlated with one or more gCI using the distance correlation test. (**c**) Volcano plot showing that p-value versus distance correlation for BI (the direction of the correlation is obtained by Spearman correlation). Blue dots represent significant differential pathways and black dots represent insignificant pathways. Panel (**a**) was generated using the R package VennDiagram[40] (version 1.6.0 https://cran.r-project.org/web/packages/VennDiagram/index.html), Panel (**b**) was generated using the R package ComplexHeatmap[41] (version 2.6.2 https://github.com/jokergoo/ComplexHeatmap), and Panels (**c**) was generated using the R package EnhancedVolcano[42] (version 1.8.0 https://github.com/kevinblighe/EnhancedVolcano).

models were profiled for their expression of 15,732 genes using RNA-seq after filtering out genes whose 75% percentile was less than 20. Pathways with less than 10 genes were excluded, which resulted in 5,164 pathways in total. We selected PDX models that had RNA-seq data available and fulfilled the Bliss assumption ($0 \le \delta_g \le 1$) across all time points. We came up with seven and nine PDXs in KRT232 plus navitoclax and KRT232 plus trametinib (Section S2).

We performed distance correlation tests[34] for detecting linear/nonlinear associations between pathways and the combination indices (Section S2). In combination treatment KRT232 plus navitoclax in NSCLC, controlling FDR at 0.1, resulting in 136, 150, and 145 pathways were significantly associated with HSA, RA, and BI, respectively, with 118 intersecting pathways across the three reference models (Fig. 3a). Heatmap of those pathways associated with at least one of the CIs shows clear pattern of two clusters of PDX models (Fig. 3b). The top significant pathways include those related to the therapeutic target or the prognosis in NSCLC (Fig. 3c and Data S1). For example, the p53 pathway interplays with MDM2 inhibitor KRT-232 to suppress tumor cell growth[27,35,36]. Moreover, IRAK and HIF-1 pathways are associated with the development of tumor in NSCLC[37–39].

Similar analysis was performed in combination treatment KRT232 plus trametinib. Controlling FDR at 0.1, resulting in 27, 12, and 13 pathways were significantly associated with HSA, RA, and BI, respectively, with seven intersecting pathways across the three reference models (Fig. S2a and Data S1). Heatmap of those 32 pathways also shows clear pattern of two clusters of PDXs (Fig. S2b). The seven intersecting pathways include those related to the therapeutic target in NSCLC. For example, several members in NFAT gene family differentially expressed in tumor vs. normal cells[43], and FGFR3 is a potential therapeutic target in NSCLC[44,45].

Discussion

In this article, we have proposed a comprehensive statistical framework to quantify the joint action of two drugs in standard in vivo combination experiments with fixed doses, where the dose–response models such as the Chou-Talalay method[46,47] and Isobologram[14] are not applicable. Our framework is generally applicable to tumor xenograft designs, including PDX experiments that is a newly designed novel model system for drug development and individualized treatment. The usual practice to decide combination effect in in vivo designs is based on p-values obtained from two-group comparisons of the combination group versus the control and

monotherapy groups. However, the p-value subthresholding approach does not directly quantify the magnitude of the combination effect that is useful for further statistical modeling in pharmacogenomic setting where the driving molecular mechanisms of variable drug responses are systematically studied.

The combPDX web-application provides the visualization and analysis pipeline of the longitudinal tumor volume data at fixed dose levels, as well as power analysis tool to design in vivo combination experiments. Our framework is inspired by effect-based approaches that compare the effect from the combination of two drugs AB to the effects from its individual mono-drugs A and B, following the determination of efficacy of each treatment compared to the control. Various metrics have been suggested to summarize each individual growth curve to a value, e.g., the adjusted area under the tumor growth curve (aAUC)[48], which can be employed in our CI calibrations based on the fact that the aAUC is interpreted as the mean tumor volume across time.

There is no global consensus on defining drug synergism/antagonism in the field. Different reference models may lead to inconsistent statistical significance of combination effects based on different underlying assumptions. We have extensively studied the differences of these three models in the mathematical formulations. With fixed group means of C, A, B, and AB, the HSA model always provides the highest CI value, while RA has the lowest CI in that it provides the most conservative procedure to declare a synergistic effect (see Section S3.6). The CI derived under HSA is formulated similarly to t-tests that compare tumor volume data of combination therapy group vs. the better monotherapy group although the bootstrap procedure of HSA is more robust than the t-test. Both methods have the limitation of not utilizing tumor volume measurements from control group. Thus, the synergism declared from these methods should be considered as the minimal evidence and used for drugs which mono-therapeutic effects have been proved sufficiently in the field. In a counter example of drug antagonism (Fig. S3) where the two drugs present a clear antagonistic pattern, HSA shows additivity because the tumor volume of the combination drug is close to that of trametinib, however, RA and BI, declared antagonism by adjusting the tumor volumes in control group (Table S3).

Using tumor volume data from NSCLC PDX experiments that evaluate two-drug combinations of KRT232, navitoclax, and trametinib, we have shown the utilities of calibrating CIs in finding underlying mechanisms of combination drug response based on gene expression data. We found that trametinib plus navitoclax had the synergistic effects in HSA and BI models, which is along with the currently undergoing clinical trials for the treatment of NSCLC[31]. Moreover, the integrative analysis of KRT-232 plus navitoclax pharmacological data with gene expression data provided highly concordant pathway signatures across the three reference interaction models. These pathways included major driving mechanisms of the combination therapy in NSCLC.

In summary, combPDX represents an important step towards combination effect calibration for PDX models. It provides comprehensive data analysis and result visualization for in vivo combination drug testing. Coupled with molecular profile, combPDX facilitates the discovery of new biomarkers for combination therapy. Going forward, the knowledge of the biological mechanisms will add promise to the identification of the optimal personalized treatment.

Method

Animal experiments. Fresh lung cancer samples were collected in 2012 and 2017 from surgically resected specimens under approved research protocols with informed consent from the patients[27,28]. The protocols for the use of clinical specimens and data in the current study were approved by the institutional review board at The University of Texas MD Anderson Cancer Center. All animal studies were carried out in accordance with the Guidelines for the Care and Use of Laboratory Animals (National Institutes of Health Publication 85-23) and the institutional guidelines of MD Anderson.

Tumor volume data processing. The combPDX requires a long data matrix as an input where each of the tumor volume measures are stacked by rows with four columns: mice ID, treatment, day, and tumor volume. Table S4 presents the description of these metrics and Table S5 shows an example of input data. Due to the variation of initial tumor volumes across mice, the response for an individual animal at each time point is defined by the relative tumor volume, which is the raw tumor volume divided by the initial tumor volume of the mouse. We denote v_t as the relative tumor volume for a mouse at time t (Section S3.1). Unfortunately, missing values may arise from data entry errors. For example, discrepant records may be entered for the same subjects or missing assessment for a subject at a given time point. For subjects with no tumor volume measurement at time t, but with flanking volume measurements at time t_0 and t_1, we use linear interpolation to impute the relative tumor volume at time t:

$$v_t = v_{t_0} + \frac{v_{t_1} - v_{t_0}}{t_1 - t_0}(t - t_0)$$

An alternative and likely more accurate interpolation approach is to model the tumor growth curves while accounting for the innate tumor environment. This is often infeasible because quantifying the innate tumor environment is often unavailable. If such information is available, the linear interpolation can be replaced, and then the treatment effect and combination index can be inferred based on the desired imputed values.

Missing value interpolations are used to maintain sample size by leveraging information from existing data points. In this manner, additional assumptions are imposed on the data analysis. For example, the innate tumor environment of a mouse is similar to that of its nearby time points, or the log tumor growth is at the same rate as its neighborhood. Therefore, care should be taken if an inference is carried out with multiple interpolations of missing values at a given time point. We recommend that no more than half of the values be interpolated data points within a subgroup at a given time for data analysis.

Denote the relative tumor volume for an individual mouse in group $g = C, A, B, AB$ as v_g, which follows an independent identical distribution with mean μ_g and variance σ_g^2. The mean and variance parameters μ_g and σ_g^2 can be estimated by the sample mean $\bar{v}_g$ and sample variance $\hat{\sigma}_g^2$. Our analytic inference procedure in this paper is based on the central limit theorem, $\bar{v}_g$ follows normal distribution with mean μ_g and variance σ_g^2/n_g, where n_g is the sample size in treatment group g (Section S3.1). Combined with Bootstrap procedures, we re-calibrate the null distribution of test statistics to propose a robust statistical framework to any deviations from the theoretical distribution.

Determination of treatment effect. To access the combination effect of two drugs, we take effect-based approaches that directly compare the effect of the combination to the effects of its individual components. The effect of a treatment group (A, B or AB) is defined by the antitumor activity compared with the control group (C). At a given time point, treatment effect for a group g is quantified by the mean reduction in the relative tumor volumes between treatment and control groups divided by the control mean

$$\delta_g = \frac{\mu_C - \mu_g}{\mu_C} \text{ for } g \in A, B, AB \tag{1}$$

A large δ_g value indicates a strong treatment effect. For combination experiments, we expect that the mean relative tumor volumes of the treatment groups μ_A, μ_B and μ_{AB} are less than the control tumor volume μ_C, which results in δ_A, δ_B and δ_{AB} located between 0 and 1. While this has been widely used as the tumor growth inhibition (TGI) with predetermined cutoffs of declaring an antitumor activity[49–51], we incorporate statistical inferential procedures by constructing 95% confidence intervals. The lower bound of a one-sided $100(1-\alpha)\%$ confidence interval for a combination index can be calculated using the Delta method,

$$\hat{\delta}_g - z_{1-\alpha}\hat{se}\left(\hat{\delta}_g\right)$$

where $\hat{\delta}_g = (\bar{v}_C - \bar{v}_g)/\bar{v}_C$, $\hat{se}\left(\hat{\delta}_g\right) = \left(\bar{v}_g^2\hat{\sigma}_C^2/\left(\bar{v}_C^4 n_C\right) + \hat{\sigma}_g^2/\left(\bar{v}_C^2 n_g\right)\right)^{1/2}$, and $z_{1-\alpha}$ is the $(1-\alpha)$-th quantile of standard normal distribution (Section S3.2).

Combination index (CI). Based on the treatment effects δ_g evaluated for all three treatment groups A, B and AB, we aim to access the superiority of using drug combination AB to individual drugs A and B. Although there is no consensus on defining the synergistic action of two drugs[52], we derive combination indices under the three popular reference models: (1) HSA (CI_{HSA}), (2) RA (CI_{RA}), and (3) BI (CI_{BI}). All the CIs under the three models are calibrated to have the same implication: $CI < 0, = 0, > 0$ represent antagonistic, independent, and synergistic effects, respectively.

Highest Single Agent (HSA)[6–8,53] shows that the synergistic combination effect occurs when the combined effect is greater than the more effective individual component: $\max(\delta_A, \delta_B)/\delta_{AB} < 1$. Under the HSA, we derive CI with respect to group means:

$$CI_{HSA} = \log\left(\mu_g\right) - \log(\mu_{AB}),$$

where g is chosen from A or B that has larger effect δ. If the two single agent effects are equal, $\delta_A = \delta_B$, we choose the one that has the narrower confidence interval evaluated. The CI_{HSA} is independent of the control experiment.

Response Additive (RA)[53,54] assumes that the fixed-dose two-drug combination has a linear additive effect under independence. A combination drug is considered synergistic if it shows a more enhanced effect than the sum of the two monotherapies' effects: $(\delta_A + \delta_B)/\delta_{AB} < 1$. The corresponding CI can be derived with respect to group means as:

$$CI_{RA} = \log(\mu_A + \mu_B) - \log(\mu_{AB} + \mu_C).$$

The Bliss Independence (BI) approach[6,8,13,14,53] is based the definition of independence on its probabilistic interpretation. Two drugs are independent if one drug's presence does not affect the probability of another drug's effect on tumor growth decay. Wu et al.[22] proposed an interaction index under such a definition using relative tumor volume. Assume that the treatment effects δ_g are the outcomes of a probabilistic process such that $0 \leq \delta_g \leq 1$. Note that δ_g in Eq. (1) takes a value between 0 and 1 if $\mu_C > \mu_g$ for $g = A, B, AB$. Under BI, if the combination therapy is considered synergistic, we have $(\delta_A + \delta_B - \delta_A\delta_B)/\delta_{AB} < 1$. The combination index is given by.

$$CI_{BI} = \log(\mu_A) + \log(\mu_B) - \log(\mu_C) - \log(\mu_{AB}).$$

Calibration of confidence intervals for CIs. We develop statistical inferential procedures by deriving confidence intervals for the CIs under the asymptotic normality in addition to the bootstrap method. Using the Delta method, the standard errors of the indices are approximated by:

$$\hat{se}\left(\widehat{CI}_{HSA}\right) \simeq \left(\frac{1}{\bar{v}_g^2}\frac{\hat{\sigma}_g^2}{n_g} + \frac{1}{\bar{v}_{AB}^2}\frac{\hat{\sigma}_{AB}^2}{n_{AB}}\right)^{1/2},$$

$$\widehat{se}\left(\widehat{CI}_{RA}\right) \simeq \left(\frac{1}{(\bar{v}_A + \bar{v}_B)^2}\frac{\hat{\sigma}_A^2}{n_A} + \frac{1}{(\bar{v}_A + \bar{v}_B)^2}\frac{\hat{\sigma}_B^2}{n_B} + \frac{1}{(\bar{v}_C + \bar{v}_{AB})^2}\frac{\hat{\sigma}_C^2}{n_C} + \frac{1}{(\bar{v}_C + \bar{v}_{AB})^2}\frac{\hat{\sigma}_{AB}^2}{n_{AB}}\right)^{1/2},$$

$$\widehat{se}\left(\widehat{CI}_{BI}\right) \simeq \left(\frac{1}{\bar{v}_A^2}\frac{\hat{\sigma}_A^2}{n_A} + \frac{1}{\bar{v}_B^2}\frac{\hat{\sigma}_B^2}{n_B} + \frac{1}{\bar{v}_C^2}\frac{\hat{\sigma}_C^2}{n_C} + \frac{1}{\bar{v}_{AB}^2}\frac{\hat{\sigma}_{AB}^2}{n_{AB}}\right)^{1/2}.$$

A two-sided $100(1-\alpha)\%$ confidence interval is constructed as $\left(\widehat{CI} - z_{1-\alpha/2}\widehat{se}\left(\widehat{CI}\right), \widehat{CI} + z_{1-\alpha/2}\widehat{se}\left(\widehat{CI}\right)\right)$ (see Derivations in Sections S3.3–3.5).

The standard intervals based on asymptotic approximation to normal distributions can be inaccurate in practice due to skewness and heavier tails of tumor volume data. A robust bootstrap procedure is provided to construct confidence intervals[55]. Bootstrap is a resampling method that samples the original data with replacements iteratively to estimate test statistics or null distributions. To calculate bootstrap-t interval at a given time point, we repeat the following steps B times. For a given reference model, HSA, RA or BI, at the iteration b, we

(1) Sample a size of n_g animals with replacement within each treatment arm g, and the corresponding tumor volume measurement for an animal is denoted by v_g^{*b}.
(2) Calculate combination index $\widehat{CI}^{*b}$ based on the bootstrap samples from the step 1, v_g^{*b} for g = C, A, B, AB.
(3) Compute $Z^{*b} = (\widehat{CI}^{*b} - \widehat{CI})/\widehat{se}^*\left(\widehat{CI}^{*b}\right)$ where $\widehat{se}^*\left(\widehat{CI}^{*b}\right)$ is calculated by the theoretical standard error calibrated.

The α-th percentile of Z^{*b} is estimated by the value $\hat{t}^{(\alpha)}$ such that $\#\left\{Z^{*b} \leq \hat{t}^{(\alpha)}\right\}/B = \alpha$. Finally, the bootstrap-t confidence interval is constructed by $\left(\widehat{CI} - \hat{t}^{(1-\alpha/2)}\widehat{se}\left(\widehat{CI}\right), \widehat{CI} - \hat{t}^{(\alpha/2)}\widehat{se}\left(\widehat{CI}\right)\right)$. This bootstrap procedure adjusts for derivations from the asymptotic distribution of CI by recalibrating the percentiles.

Global assessment of combination effect. The CI values assess combination effects at a given time point. We extend the procedure to a global assessment for any given study intervals of interest. Given reference model, the global CI is defined as the average of the CIs within the study interval of interest for time points $1, \ldots, T$:

$$gCI = \frac{1}{T}\sum_{t=1}^{T} CI_t,$$

where CI_t is the combination index at time t. Due to correlations of tumor volume measurements within individual mouse, we conduct nested bootstrap procedure to construct a confidence interval for gCI without analytically specifying the variance $var(gCI)$[55]. The algorithm to construct confidence interval for gCI is similar to that in previous section, with an additional nested layer to estimate $\widehat{se}^*\left(\widehat{gCI}^{*b}\right)$. At the iteration b, we

(1) Sample a size of n_g animals with replacement within each treatment arm g, the corresponding growth curve data for an animal are denoted by $\mathbf{v_g^{*b}} = \left(v_{g,1}^{*b}, v_{g,2}^{*b}, \ldots, v_{g,T}^{*b}\right)$.
(2) Calculate combination index $\widehat{gCI}^{*b}$ based on the $\mathbf{v_g^{*b}}$, g = C, A, B, AB. We repeat the following nested bootstrap procedure L = 25 times to estimate $\widehat{se}^*\left(\widehat{gCI}^{*b}\right)$,

　(a) Sample a size of n_g growth curves with replacement from $\mathbf{v_g^{*b}}$ within each group and denote the corresponding growth curve data as $\mathbf{v_g^{*b,l}} = \left(v_{g,1}^{*b,l}, v_{g,2}^{*b,l}, \ldots, v_{g,T}^{*b,l}\right)$ for each animal.
　(b) Calculate combination index $\widehat{gCI}^{*b,l}$ based on $\mathbf{v_g^{*b,l}}$, g = C, A, B, AB
　　The standard error of each resampled data set can be estimated by $\widehat{se}^*\left(\widehat{gCI}^{*b}\right) = \sqrt{\sum_l\left(\widehat{gCI}^{*b,l} - \widehat{gCI}^{*b,(\cdot)}\right)^2/(L-1)}$, where $\widehat{gCI}^{*b,(\cdot)} = \sum_l\widehat{gCI}^{*b,l}/L$.

(3) Compute $Z^{*b} = (\widehat{gCI}^{*b} - \widehat{gCI})/\widehat{se}^*\left(\widehat{gCI}^{*b}\right)$.

The α-th percentile of Z^{*b} is estimated by the value $\hat{t}^{(\alpha)}$ such that $\#\{Z^{*b} \leq \hat{t}^{(\alpha)}\}/B = \alpha$. Finally, the bootstrap-t confidence interval is $\left(\widehat{gCI} - \hat{t}^{(1-\alpha/2)}\widehat{se}\left(\widehat{gCI}\right), \widehat{gCI} - \hat{t}^{(\alpha/2)}\widehat{se}\left(\widehat{gCI}\right)\right)$.

Power analysis. Based on the CIs under the HSA, BI and RA reference models, we provide power analysis tool to design PDX combination experiments. Under the null hypothesis of independent action of a drug combination, $H_0 : CI = 0$, where the CI follows asymptotic normal distribution $N(0, var(CI))$. Assume that under
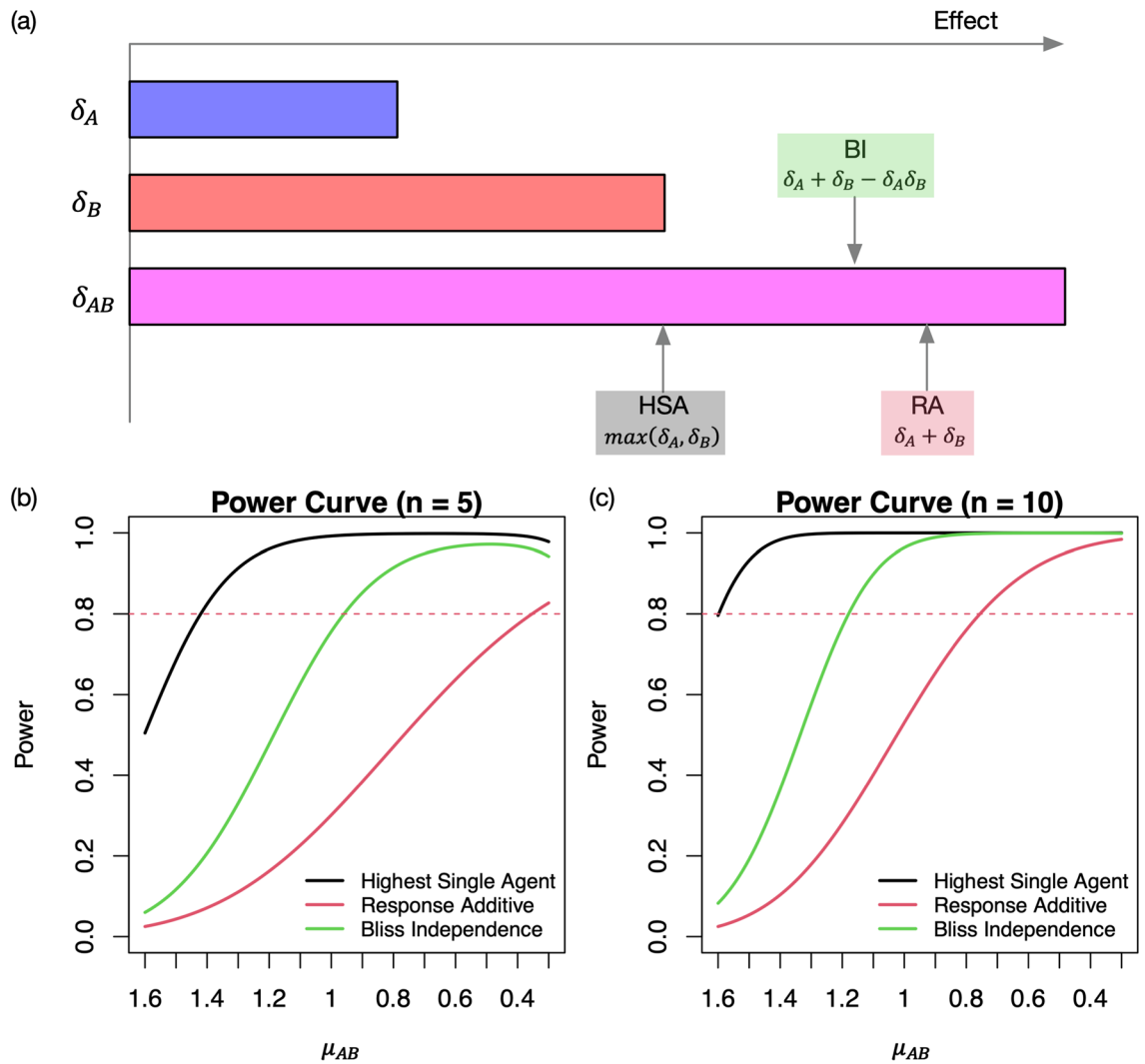
Figure 4. Comparison of HSA, BI and RA models under the hypothetical scenario where $\mu_C = 2.8, \mu_A = 2.4$ and $\mu_B = 2$. (**a**) Given δ_A and δ_B, the arrows indicate the minimum δ_{AB} values that have synergistic effects (CI > 0). (**b**, **c**) Statistical power varies by the mean tumor volume for the combination AB when the sample size is 5 and 10 at the 0.05 significance level. The standard errors of relative tumor volumes are set to 0.7 for the control group C and 0.3 for the treatment groups A, B and AB.

the alternative hypothesis $H_A : CI = \gamma$, where CI follows normal distribution $N(\gamma, \text{var(CI)})$. Therefore, with prespecified values μ_g, σ_g, n_g for each group g, the power is calculated as

$$1 - \beta = \Pr\left(\left|\frac{\gamma}{\sqrt{\text{var(CI)}}}\right| \geq z_{\frac{\alpha}{2}}\Big|H_1\right) = \Pr\left(Z \geq z_{\frac{\alpha}{2}} - \frac{\gamma}{\sqrt{\text{var(CI)}}}\Big|H_1\right) + \Pr\left(Z \geq -z_{\frac{\alpha}{2}} - \frac{\gamma}{\sqrt{\text{var(CI)}}}\Big|H_1\right),$$

where α, β are the desired Type I and Type II error rates.

Given mean tumor volumes, Fig. 4a (Fig. S4) illustrates the minimum δ_{AB} (maximum tumor volume μ_{AB}) having synergistic effect under each reference model given δ_A and δ_B, and we also mathematically compared CI_{HSA}, CI_{RA} and CI_{BI} in Section S3.6. Given mean tumor volumes, we have $0 \leq CI_{RA} \leq CI_{BI} \leq CI_{HSA}$, which implies that HSA model provides the most relaxed procedure while RA is the most conservative. Combined with the statistical inference procedure using the asymptotic normality, the power curves, Fig. 4b,c for sample sizes of 5 and 10, respectively indicate that RA (HSA) require the largest (smallest) effect sizes of the combination to achieve the same statistical power when the effect sizes of monotherapies and standard deviations are fixed.

Software. We provide an interactive web server, combPDX (https://licaih.shinyapps.io/CombPDX/), to analyze PDX tumor growth curve data and perform power analyses. The combPDX currently includes five tabs: Upload Dataset, Visualize Results, Power calculation, Download Results and Batch Analysis (Figs. S5–S7). These allow the users to test treatment effect as well as combination effect, view output figures and tables, download result and run batch analysis. The tutorial is presented in Section S4.

Data availability

References

1. Dawson, J. C. & Carragher, N. O. Quantitative phenotypic and pathway profiling guides rational drug combination strategies. *Front. Pharmacol.* **5**, 118. https://doi.org/10.3389/fphar.2014.00118 (2014).
2. Wright, G. D. Antibiotic adjuvants: Rescuing antibiotics from resistance. *Trends Microbiol.* **24**, 862–871. https://doi.org/10.1016/j.tim.2016.06.009 (2016).
3. Mathews Griner, L. A. *et al.* High-throughput combinatorial screening identifies drugs that cooperate with ibrutinib to kill activated B-cell-like diffuse large B-cell lymphoma cells. *Proc. Natl. Acad. Sci. USA* **111**, 2349–2354. https://doi.org/10.1073/pnas.1311846111 (2014).
4. Zagidullin, B. *et al.* DrugComb: An integrative cancer drug combination data portal. *Nucleic Acids Res.* **47**, W43–W51. https://doi.org/10.1093/nar/gkz337 (2019).
5. Narayan, R. S. *et al.* A cancer drug atlas enables synergistic targeting of independent drug vulnerabilities. *Nat. Commun.* **11**, 2935. https://doi.org/10.1038/s41467-020-16735-2 (2020).
6. Berenbaum, M. C. What is synergy?. *Pharmacol. Rev.* **41**, 93–141 (1989).
7. Lehár, J. *et al.* Chemical combination effects predict connectivity in biological systems. *Mol. Syst. Biol.* **3**, 80. https://doi.org/10.1038/msb4100116 (2007).
8. Geary, N. Understanding synergy. *Am. J. Physiol.-Endocrinol.* **304**, E237-253. https://doi.org/10.1152/ajpendo.00308.2012 (2013).
9. Frei, W. Versuche über Kombination von Desinfektionsmitteln. *Z. Hyg. Infektionskr.* **75**, 433–496. https://doi.org/10.1007/BF02207053 (1913).
10. Loewe, S. Die quantitativen probleme der pharmakologie. *Ergebnisse der Physiologie* **27**, 47–187 (1928).
11. Loewe, S. T. & Muischnek, H. Über kombinationswirkungen. *Naunyn Schmiedebergs Arch. Exp. Pathol. Pharmakol.* **114**, 313–326 (1926).
12. Loewe, S. The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung* **3**, 285–290 (1953).
13. Bliss, C. I. The toxicity of poisons applied jointly. *Ann. Appl. Biol.* **26**, 585–615 (1939).
14. Greco, W. R., Bravo, G. & Parsons, J. C. The search for synergy: A critical review from a response surface perspective. *Pharmacol. Rev.* **47**, 331–385 (1995).
15. Ianevski, A., He, L., Aittokallio, T. & Tang, J. SynergyFinder: A web application for analyzing drug combination dose-response matrix data. *Bioinformatics* **33**, 2413–2415. https://doi.org/10.1093/bioinformatics/btx162 (2017).
16. Di Veroli, G. Y. *et al.* Combenefit: An interactive platform for the analysis and visualization of drug combinations. *Bioinformatics* **32**, 2866–2868. https://doi.org/10.1093/bioinformatics/btw230 (2016).
17. Jung, J., Seol, H. S. & Chang, S. The generation and application of patient-derived xenograft model for cancer research. *Cancer Res. Treatment* **50**, 1 (2018).
18. Gao, H. *et al.* High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat. Med.* **21**, 1318–1325. https://doi.org/10.1038/nm.3954 (2015).
19. Izumchenko, E. *et al.* Patient-derived xenografts effectively capture responses to oncology therapy in a heterogeneous cohort of patients with solid tumors. *Ann. Oncol.* **28**, 2595–2605. https://doi.org/10.1093/annonc/mdx416 (2017).
20. Doroshow, J. H. Abstract IA12: NCI's patient-derived cancer models repository. *Clin. Cancer Res.* **22**, IA12. https://doi.org/10.1158/1557-3265.Pdx16-ia12 (2016).
21. Conte, N. *et al.* PDX Finder: A portal for patient-derived tumor xenograft model discovery. *Nucleic Acids Res.* **47**, D1073-d1079. https://doi.org/10.1093/nar/gky984 (2019).
22. Wu, J., Tracey, L. & Davidoff, A. M. Assessing interactions for fixed-dose drug combinations in tumor xenograft studies. *J. Biopharm. Stat.* **22**, 535–543. https://doi.org/10.1080/10543406.2011.556285 (2012).
23. Wu, J. Assessing interactions for fixed-dose drug combinations in subcutaneous tumor xenograft studies. *Pharm. Stat.* **12**, 115–119. https://doi.org/10.1002/pst.1559 (2013).
24. Demidenko, E. & Miller, T. W. Statistical determination of synergy based on Bliss definition of drugs independence. *PLoS ONE* https://doi.org/10.1371/journal.pone.0224137 (2019).
25. Winsor, C. P. The Gompertz curve as a growth curve. *Proc. Natl. Acad. Sci. USA* **18**, 1–8. https://doi.org/10.1073/pnas.18.1.1 (1932).
26. Ribba, B. *et al.* A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT Pharmacometrics Syst. Pharmacol.* **3**, 113. https://doi.org/10.1038/psp.2014.12 (2014).
27. Zhang, X. *et al.* KRT-232 and navitoclax enhance trametinib's anti-Cancer activity in non-small cell lung cancer patient-derived xenografts with KRAS mutations. *Am. J. Cancer Res.* **10**, 4464–4475 (2020).
28. Chen, Y. *et al.* Tumor characteristics associated with engraftment of patient-derived non–small cell lung cancer xenografts in immunocompromised mice. *Cancer* **125**, 3738–3748 (2019).
29. Hao, C. *et al.* Gene mutations in primary tumors and corresponding patient-derived xenografts derived from non-small cell lung cancer. *Cancer Lett.* **357**, 179–185 (2015).
30. Team, R. C. R: A language and environment for statistical computing. https://www.R-project.org/. Accessed 3 Jan 2022.
31. Kim, C. & Giaccone, G. MEK inhibitors under development for treatment of non-small-cell lung cancer. *Expert Opin. Investig. Drugs* **27**, 17–30. https://doi.org/10.1080/13543784.2018.1415324 (2018).
32. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinform.* **14**, 7. https://doi.org/10.1186/1471-2105-14-7 (2013).
33. Liberzon, A. *et al.* The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
34. Székely, G. J. & Rizzo, M. L. The distance correlation t-test of independence in high dimension. *J. Multivar. Anal.* **117**, 193–213. https://doi.org/10.1016/j.jmva.2013.02.012 (2013).
35. Rew, Y. & Sun, D. Discovery of a small molecule MDM2 inhibitor (AMG 232) for treating cancer. *J. Med. Chem.* **57**, 6332–6341 (2014).
36. Sun, D. *et al.* Discovery of AMG 232, a potent, selective, and orally bioavailable MDM2–p53 inhibitor in clinical development. *J. Med. Chem.* **57**, 1454–1472 (2014).
37. Shimoda, L. A. & Semenza, G. L. HIF and the lung: Role of hypoxia-inducible factors in pulmonary development and disease. *Am. J. Respir. Crit. Care Med.* **183**, 152–156. https://doi.org/10.1164/rccm.201009-1393PP (2011).
38. Kurtipek, E. *et al.* (European Respiratory Society, 2016).
39. Zhang, X., Dang, Y., Li, P., Rong, M. & Chen, G. Expression of IRAK1 in lung cancer tissues and its clinicopathological significance: A microarray study. *Int. J. Clin. Exp. Pathol.* **7**, 8096–8104 (2014).

40. Chen, H. & Boutros, P. C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* **12**, 1–7 (2011).
41. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**. R package version 2.6.2. https://github.com/jokergoo/ComplexHeatmap (2016). Accessed 3 Jan 2022.
42. Blighe, K. *et al.* EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.8.0. https://github.com/kevinblighe/EnhancedVolcano (2020). Accessed 3 Jan 2022.
43. Chen, Z.-L. *et al.* Expression and unique functions of four nuclear factor of activated T cells isoforms in non-small cell lung cancer. *Chin. J. Cancer* **30**, 62–68. https://doi.org/10.5732/cjc.010.10156 (2011).
44. Chandrani, P. *et al.* Drug-sensitive FGFR3 mutations in lung adenocarcinoma. *Ann. Oncol.* **28**, 597–603. https://doi.org/10.1093/annonc/mdw636 (2017).
45. Shinmura, K. *et al.* A novel somatic FGFR3 mutation in primary lung cancer. *Oncol. Rep.* **31**, 1219–1224 (2014).
46. Chou, T.-C. & Talalay, P. Analysis of combined drug effects: A new look at a very old problem. *Trends Pharmacol. Sci.* **4**, 450–454 (1983).
47. Chou, T.-C. & Talalay, P. Quantitative analysis of dose-effect relationships: The combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzyme Regul.* **22**, 27–55 (1984).
48. Evrard, Y. A. *et al.* Systematic establishment of robustness and standards in patient-derived xenograft experiments and analysis. *Can. Res.* **80**, 2286–2297. https://doi.org/10.1158/0008-5472.Can-19-3101 (2020).
49. Houghton, P. J. *et al.* The pediatric preclinical testing program: Description of models and early testing results. *Pediatr. Blood Cancer* **49**, 928–940. https://doi.org/10.1002/pbc.21078 (2007).
50. Mer, A. S. *et al.* Integrative pharmacogenomics analysis of patient-derived xenografts. *Can. Res.* **79**, 4539–4550. https://doi.org/10.1158/0008-5472.Can-19-0349 (2019).
51. Ortmann, J. *et al.* KuLGaP: A selective measure for assessing therapy response in patient-derived xenografts. *bioRxiv.* https://doi.org/10.1101/2020.09.08.287573 (2020).
52. Roell, K. R., Reif, D. M. & Motsinger-Reif, A. A. An introduction to terminology and methodology of chemical synergy-perspectives from across disciplines. *Front. Pharmacol.* **8**, 158. https://doi.org/10.3389/fphar.2017.00158 (2017).
53. Foucquier, J. & Guedj, M. Analysis of drug combinations: Current methodological landscape. *Pharmacol. Res. Perspect.* **3**, e00149. https://doi.org/10.1002/prp2.149 (2015).
54. Slinker, B. K. The statistics of synergism. *J. Mol. Cell. Cardiol.* **30**, 723–731. https://doi.org/10.1006/jmcc.1998.0655 (1998).
55. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (CRC Press, 1994).

Acknowledgements

Author contributions

M.J.H. conceived the project. M.J.H. and L.H. developed the statistical framework and L.H. developed software and performed bioinformatics analysis. J.W., B.F., F.M.-B. and J.R. provided the expertise and edited the manuscript. M.J.H. and L.H. wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-16933-6.

Correspondence and requests for materials should be addressed to M.J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.