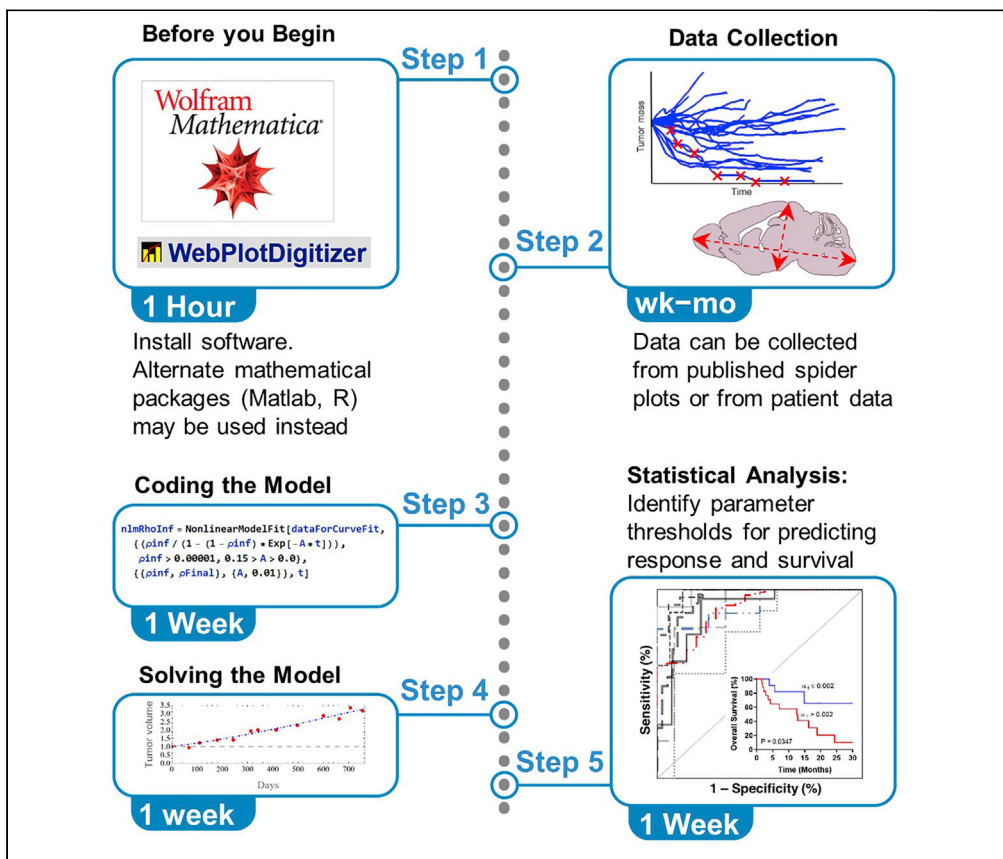


Protocol

Protocol for mathematical prediction of patient response and survival to immune checkpoint inhibitor immunotherapy



Joseph D. Butner,
Maguy Farhat,
Vittorio Cristini,
Caroline Chung,
Zhihui Wang

jbutner@
houstonmethodist.org
(J.D.B.)
cchung3@mdanderson.
org (C.C.)
zwang@
houstonmethodist.org
(Z.W.)

Highlights

Model parameters provide early predictive readout of patient response and survival

Informed using only standard-of-care, non-invasive imaging data

Can be used with as few as two time points after treatment start

Proven reliable across multiple solid tumor and checkpoint inhibitor types

This protocol describes the application of a mechanistic mathematical model of immune checkpoint inhibitor (ICI) immunotherapy to patient tumor imaging data for predicting solid tumor response and patient survival under ICI intervention. We describe steps for data collection and processing, data pipelines, and approaches to increase precision. The protocol is highly predictive as early as the first restaging after treatment start and can be used with standard-of-care imaging measures.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Butner et al., STAR Protocols 3,
101886
December 16, 2022 © 2022
The Author(s).
<https://doi.org/10.1016/j.xpro.2022.101886>



Protocol

Protocol for mathematical prediction of patient response and survival to immune checkpoint inhibitor immunotherapy

Joseph D. Butner,^{1,8,*} Maguy Farhat,² Vittorio Cristini,^{1,3,4,5} Caroline Chung,^{2,*} and Zihui Wang^{1,3,6,7,9,*}

¹Mathematics in Medicine Program, Houston Methodist Research Institute, Houston, TX 77030, USA

²Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

³Neal Cancer Center, Houston Methodist Research Institute, Houston, TX 77030, USA

⁴Department of Imaging Physics, University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA

⁵Physiology, Biophysics, and Systems Biology Program, Graduate School of Medical Sciences, Weill Cornell Medicine, New York, NY 10065, USA

⁶Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA

⁷Department of Medical Education, Texas A&M University School of Medicine, Bryan, TX 77807, USA

⁸Technical contact: jbutner@houstonmethodist.org

⁹Lead contact: zwang@houstonmethodist.org

*Correspondence: jbutner@houstonmethodist.org (J.D.B.), cchung3@mdanderson.org (C.C.), zwang@houstonmethodist.org (Z.W.)
<https://doi.org/10.1016/j.xpro.2022.101886>

SUMMARY

This protocol describes the application of a mechanistic mathematical model of immune checkpoint inhibitor (ICI) immunotherapy to patient tumor imaging data for predicting solid tumor response and patient survival under ICI intervention. We describe steps for data collection and processing, data pipelines, and approaches to increase precision. The protocol is highly predictive as early as the first restaging after treatment start and can be used with standard-of-care imaging measures.

For complete details on the use and execution of this protocol, please refer to Butner et al. (2020)¹ and Butner et al. (2021).²

BEFORE YOU BEGIN

Immune checkpoint inhibitor (ICI)-based immunotherapies have made significant strides in improving cancer therapeutic outcomes. Despite the success of ICIs in several solid tumors (most notably melanoma, lung and colorectal cancers), the majority of patients fail to respond to ICI therapy,³ and reliable methods of identifying patients who will respond early in the course of treatment remain elusive. Mathematical modeling of immunotherapy is yielding valuable insights into dose-response relationships and associated immunotherapy response outcomes,^{4,5} which may help overcome this limitation. However, models that are fully structurally identifiable (e.g., not over-fitted), practically identifiable (e.g., all parameters may be reliably measured),⁶ and immediately usable under current clinical standards have remained elusive. Here, we present a detailed guide to using a mechanistic model of ICI immunotherapy that has been demonstrated to reliably predict patient response and survival under ICI therapy, and that can be used with only already-available, non-invasively collected standard-of-care data.

The modeling effort was built upon our long-standing track record of developing translational biophysical models of drug transport for predicting cancer treatment outcome.^{7–16} In essence, the model describes the immunotherapy treatment system using differential equations based on the



underlying biological and physical mechanisms in the cancer immune response, resulting in greater predictive capabilities than traditional biomarker analysis, and yielding robust results founded upon mechanistic principles. The final model form describes tumor burden (ρ) over time as:

$$d\rho / dt = \rho(\alpha_0 - \mu + \wedge\mu) + \rho^2(-\wedge\mu), \quad (\text{Equation 1})$$

Where α_0 is the intrinsic (baseline) tumor growth rate without treatment intervention, and the key model parameters represent tumor kill rate (μ) by immunotherapy and patient anti-tumor immune state (\wedge), which we have defined as the coupling of immune cell activity and the tumor cell kill (i.e., tumor immunogenicity) scaled by the ratio of tumor cells to intratumoral immune cells at the start of treatment (full derivation may be found in Butner et al.^{1,2}).

Integrating Equation 1 leads to the time-dependent form of the model, which will be used for model solution in this tutorial, such that:

$$\rho(t) = \frac{\rho^\infty}{1 - (1 - \rho^\infty)e^{-A \cdot t}}, \quad (\text{Equation 2})$$

where

$$A = \alpha_0 - \mu + \mu \cdot \wedge \quad (\text{Equation 3})$$

and the long-term solution for tumor burden when $t \rightarrow \infty$ is:

$$\rho^\infty = \frac{\alpha_0 - \mu}{\mu \cdot \wedge} + 1. \quad (\text{Equation 4})$$

In this protocol, we will calculate two growth rates (α) corresponding to growth rates directly before and after start of treatment (t_0 = baseline, when $t = 0$ at start of treatment); these are short-term growth rates. Tumor growth rate before treatment may be calculated as:

$$\alpha_0 = \frac{\ln(\rho_0/\rho_{-1})}{t_0 - t_{-1}}, \quad (\text{Equation 5})$$

where t_{-1} refers to the time of measurement before start of treatment (the specific time will likely vary between patients). Likewise, tumor growth rate directly after treatment may be calculated as:

$$\alpha_1 = \frac{\ln(\rho_1/\rho_0)}{t_1 - t_0}, \quad (\text{Equation 6})$$

where t_1 refers to the time of the first follow-up tumor burden measurement after start of therapy.

Institutional permissions (if applicable)

Approval from the Institutional Review Board at The University of Texas MD Anderson Cancer Center was acquired prior to aggregation and analysis of all patient data.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Renal cell carcinoma measured response (Figure A1 in Motzer et al. ¹⁷)	Motzer et al. ¹⁷	https://doi.org/10.1200/jco.2014.59.0703
Non-small-cell lung cancer measured response (Figure S6 in Borghaei et al. ¹⁸)	Borghaei et al. ¹⁸	https://doi.org/10.1056/nejmoa1507643
Melanoma measured response (Figure 1A in Topalian et al. ¹⁹)	Topalian et al. ¹⁹	https://doi.org/10.1056/nejmoa1200690
Urothelial carcinoma (Figure 2C in Powles et al. ²⁰)	Powles et al. ²⁰	https://doi.org/10.1038/nature13904
Software and algorithms		
Mathematica 13.0.0	Wolfram Research ²¹	https://www.wolfram.com/mathematica/
WebPlotDigitizer	Rohatgi ²²	https://automeris.io/WebPlotDigitizer
RayStation 11B	RaySearch Laboratories ²³	https://www.raysearchlabs.com/rystation/

STEP-BY-STEP METHOD DETAILS

Data collection

⌚ Timing: days to months

The model predicts patient response and/or survival using time-course tumor burden measurements. It may be used with publicly available published data, or with clinically collected patient data for research purposes. Note that although model predictions have been shown to correlate with clinical outcomes, this research remains ongoing, and these predictions should not replace expert clinical opinions. In all cases, the minimum measurements needed are 1) at baseline (start of treatment) and 2) at first restaging (follow-up). In order to obtain unique solutions for all model parameters, 3) one pre-baseline (usually at time of on-study) tumor burden measurement is also necessary for calculating α_0 . We have provided a sample data set from Borghaei et al. and Topalian et al.^{18,19} in [Data S1](#). Sample data.

1. Collecting literature-reported tumor burden data.
 - a. Extract the spider plot of normalized measured tumor burden over time and save locally as an image file.

Note: Sources of spider plots used in published studies, as well as the specific figure we have collected data from in each of these references, are provided in the [key resources table](#). A spider plot graphs the normalized (e.g., percent) change from baseline of the total indexed tumor burden over time for each patient in a study.

- b. Load spider plot into WebPlotDigitizer.
 - i. Once spider plot is loaded, calibrate x-axis and y-axis scales as instructed.
 - ii. In manual extraction mode, place the cursor on each measured point on one curve on the spider plot (including the origin) and place a mark using the 'A' key.
 - iii. Once all points are marked, click 'View Data' to obtain point values. These may be exported as delimited text and imported or copied and pasted directly into spreadsheet software.
 - iv. Repeat for each curve on the spider plot, and then assign each patient a unique identifier number.
2. Collecting in-house clinical patient data.

Note: we have included details specific to our study on melanoma brain metastases treated with Ipilimumab and Nivolumab with or without concurrent stereotactic radiosurgery and measured using magnetic resonance imaging (MRI) from Butner et al.² as an example case study here. Although the details of your study may vary (for example, different cancer types, drugs used, or imaging modalities), the principles used here for patient eligibility and lesion selection serve as useful guidelines for model application to patient data from alternate disease and ICI drug combinations.

- a. Patient selection.
 - i. Retrospectively identify eligible patients who are 18 years of age or older, diagnosed with brain metastases from a histologically confirmed melanoma, and have been treated with a combination of the immune check point inhibitor (ICI) immunotherapies ipilimumab and nivolumab as an initial treatment intervention for brain metastasis, with or without concurrent stereotactic radiosurgery (SRS).

⚠ **CRITICAL:** Enrolled patients must have at least one baseline pre-treatment (acquired within 4 weeks of starting the treatment of interest) and two follow-up post-treatment but pre-operative 3-dimensional (3D) contrast-enhanced T1-weighted MRI scans.

Note: It is expected that times between scans (both pre-treatment vs. baseline and baseline vs. follow-ups) will vary among scans and among patients. In the study presented in Butner et al.,² times between baseline and first restaging ranged from 5–230 days. Because the method presented herein quantifies tumor growth as a rate instead of a scalar, it is uniquely robust to such variation.

- b. Imaging data collection and transfer.
 - i. Collect all available in-house images, in our case from The University of Texas MD Anderson Cancer Center, as well as outside center acquired images if they are of adequate quality.
 - ii. Export post-contrast enhanced T1-weighted MRI from the picture archiving and communication system (PACS) into a network drive in Digital Imaging and Communications in Medicine (DICOM) format either using manual or automatic batch export.
 - iii. Import the serial DICOM MRI images for each patient into image viewing and annotation software.

Note: We used Raystation 11B in our study, but users will need to choose the one most applicable from those available at their institution. Using a consistent platform along with standard operating procedure will ensure more consistent annotation generation.

- c. Image registration.
 - i. Apply image registration of the serial images for establishing brain location links to ensure accurate identification of the region of interest on the follow-up imaging timepoints.
- d. Lesion segmentation.
 - i. Segment and annotate each visible enhancing brain lesion separately. Use semiautomated or automated approaches, if available.

Note: Semiautomated approaches are commonly based on intensity values (e.g., a region growing tool). If using semiautomated methods, perform manual adjustments using the brush to ensure accurate delineation of each lesion.

△ CRITICAL: Use consistent annotation for each lesion or region of interest for serial imaging. For example, label a melanoma brain metastasis region of interest as “lesion_1” and use consistent labeling as you segment this same lesion on follow-up imaging time points. Another separate lesion can be labeled as “lesion_2”, “lesion_3”, etc.

- e. Volumetric data extraction.
 - i. Extract and record the volume (e.g., in cm³) of each segmented lesion from the serial imaging timepoints into a separate data collection file (e.g., Excel).

Note: This can be done manually or automated by using a script. Although the details of your script will vary based on the imaging viewing and annotation software used, we used Raystation imaging and segmentation/annotation data export tools including Region of Interest (ROI) name, ROI volume to extract lesion volumes from images, and we have found that plentiful open-source resources to export extracted data from Raystation to Excel are available; for example, the Python script export statistics to excel.py for Raystation (our software of choice) may be found on GitHub).

- f. Clinical data collection.
 - i. Collect and record clinical data of interest for each patient, including treatment agents, treatment start and end dates, stereotactic radiosurgery (SRS) details for each lesion, SRS dates, etc.

△ **CRITICAL:** The model requires at least 2 tumor burden measurements at or after start of treatment; however, full statistics on goodness-of-fit can only be obtained if 3 or more measurements are collected. To obtain unique values for all model parameters, one pre-treatment measurement (this can be estimated, see [data processing](#)) is also required.

Note: Often, time of baseline imaging may deviate a few days from actual start of treatment. This data can safely be used as an accurate measure of baseline volume without modification.

Code and directory setup

⌚ **Timing:** hours

3. Obtain and install software.
 - a. Install Mathematica²¹ according to manufacturer guidelines for your operating system; see [key resources table](#).
 - i. Alternatively, other software with differential equation fitting capability may be used (e.g., R, Matlab). However, detailed instructions, guidelines, and hints herein are specific for Mathematica.
 - b. Download and install WebPlotDigitizer²²; link in [key resources table](#).

Optional: The web-based version of WebPlotDigitizer may be used in place of installation on the local host.

- c. A software for storing tumor volume measurements is needed. We recommend Microsoft Excel, as will be presented in this manuscript.

Note: Mathematica is also able to import from alternative formats (i.e., Libre Office Calc), but importing data into Mathematica from these may be less straightforward. Delimited text may also be used.

4. Directory setup.
 - a. We recommend placing the spreadsheet and Mathematica script in the same directory to facilitate data import. Then, instruct Mathematica to look for the data file in its local directory with the following command at the top of the Mathematica notebook:

```
>SetDirectory[NotebookDirectory[]];
```

Data processing

⌚ **Timing:** hours to days

Depending on data source used, post-processing may be necessary before model analysis. If using clinical data, missing data due to lesions not being visible on some scans must be addressed first, and lesions receiving non-ICI ancillary treatment (such as radiation therapy) must be excluded.

5. Post-processing clinical data.
 - a. If lesions are not visible on one or more scans, but are confirmed to be still be present on a subsequent scan, then they can reasonably be assumed to be present at the 'not visible' time. This can be handled by either estimating the missing data via interpolation or excluding the data for all lesions at that time point.

- i. If only a small number of measurement times are available, then the number of unique time measurements should be maximized by estimating lesion value at the missing time; standard interpolation or extrapolation methods may be used.

Note: In order to reduce error introduced in the estimation process, we recommend linear estimation unless your data clearly demonstrates non-linear behavior.

△ **CRITICAL:** In order to reduce estimation error, interpolation should be based on calculated lesion volumes, not by interpolating lesion diameters if using RECIST-based measures. Lesion volume for surrounding times should be calculated as in Coding a script to solve the model: 6.d, missing volume should be interpolated from these, and then the mean diameter for missing lesion can be back-calculated assuming a spherical lesion. If lesion volumes were directly measured via contouring, then interpolation should be done directly between the measured volumes.

- ii. Any lesions that received non-ICI ancillary, targeted treatment, such as radiation therapy or surgical resection, should not be included in this analysis at any time point.

△ **CRITICAL:** Inclusion of any lesions receiving targeted ancillary therapy may introduce dramatic shifts in tumor burden that do not reflect the effects of ICI, and model predictions may not be reliable in this case. Care must be taken to ensure these lesions are excluded.

- b. Lesions that received ancillary treatment, such as targeted radiation therapy or surgery should be excluded, as these do not represent the effects of ICI.

Coding a script to solve the model

⌚ Timing: hours to days

The time-dependent model form presented in Equation 2 is fit to time-course data collected as described above via nonlinear regression fitting methods. In this two-step process, the pre-treatment growth kinetic is first obtained from an exponential fit to pre-treatment data, and then the full model is fit to lesion volume data at baseline and at least 1 post-baseline volume. The protocol described herein is based on model coding in the Mathematica language, and a sample script containing the code described herein is provided in Data S1. Example script; however, other languages (e.g., R, Matlab) may also be used.

- 6. Importing and processing data for model analysis.
 - a. First, specify the worksheet and tab where data will be imported from. If the workbook is in the same directory as the Mathematica notebook and working directory has been set as above, this is done as follows:

```
>data=Table[Import["ExcelFile.xlsx", {"Sheets", "Sheet1"}]]
```

If data is in a different directory, then the file path must also be specified.

- b. Once data is imported, select the data for only one patient, based on the assigned unique patient ID value (see data collection) in the appropriate column in the spreadsheet (A=1, B=2, etc.). This is easily done with the Cases[] function.

```
>singlePatientData=
Cases[data, x_ /; (x[[IDNumberColumn]]==patientID)];
```

Note: Mathematica identifies columns numerically instead of alphabetically, such that A=1, B=2, etc. All column numbers in this section are converted accordingly.

- c. Identify unique lesion numbers and time points.

```
>lesionNumbers=DeleteDuplicates[
  singlePatientData[[All,lesionColumnNumber]];
>uniqueTimes=DeleteDuplicates[
  singlePatientData[[All,timeColumnNumber]]];
```

- d. If lesion axes were measured (i.e., RECIST; see [data collection](#)), then these must be converted into an estimated volume, one for each indexed lesion. If lesion volume was measured by contouring, skip to the next step.
 - i. Long axis (RECIST) and approximately perpendicular short axis should be averaged, and lesion volume may be estimated as a 3D sphere.

$$\text{Volume} = \frac{4}{3}\pi \left(\frac{1}{2} \left((\text{long axis} + \text{short axis})/2 \right) \right)^3$$

- e. Sum lesion volumes at each time point to get a total tumor burden at each time point, and append each volume to a list. In this case, we might name the list `allPatientVolumes`.

```
>AppendTo[allPatientVolumes,totaLesionVolume];
```

- f. Data must be normalized by the volume at time $t = 0$, so that the volume at $t = 0$ is equal to 1. If your data has one pre-treatment measurement, then this would be at index `[[2]]` on the summed volume list.

```
>normalizedVolumesSummedLesions=
  allPatientVolumes/(allPatientVolumes[[2]]);
```

Note: Baseline (start of treatment) has been defined at $t = 0$. This means that if pre-treatment lesion measurements are available, they will be at negative times relative to baseline.

Note: Mathematica starts indexing at `[1]`; if you are using a language that starts indexing at `[0]` please adjust the above accordingly.

- g. MapThread the normalized volumes together with the time coordinates for model analysis in order to obtain a single list of linked `{x,y}` coordinates.

```
>normalizedDataPoints=MapThread[{#1,#2}&,
  {uniqueTimes,normalizedVolumesSummedLesions}}];
```

Note: this step may not be necessary if using another mathematical language (e.g., this data format is generally not used for regression analysis in R or Matlab).

7. Solving the model.

- a. If pre-treatment volume data is available, solve for the pre-treatment growth kinetic (α_0). Alternatively, if pre-treatment tumor burden is unknown, the pre-treatment growth kinetic (α_0) should be estimated by calculating the growth kinetic for the fastest growing tumor burden(s) from a subset of patients within the population (we used the top 10% fastest growing in Butner et al.²) shortly after start of treatment as described in 7.b, using Equation 6.

Note: Details on this approach can be found in Butner et al.^{2,24}

- i. The model assumes exponential pre-treatment growth between lesion measurement directly preceding baseline and at baseline (see Equation 5 for details). After extracting these two points, exact solution for the exponential growth rate is accomplished by taking the log of the y-coordinate and performing linear regression.

```
>dataToComputeAlpha=
  normalizedDataPoints[{{1,2},{1,2}}];
>logAlphaData=
  {#[[1]],Log[#[[2]]]}&/@dataToComputeAlpha;
>alphaMinusFit=
  FindFit[logAlphaData,cx+d,{c,d},x];
```

Note: This approach only uses one pre-treatment volume measurement, even if others are available. In this case, earlier times should be removed from the data. We recommend at least 2 weeks' time between pre-treatment measurement and baseline measurement.

Note: This same method is used to calculate the growth kinetic after start of treatment (α_1).

- ii. Extract the numeric value of α_0 from the exact solution.

```
>alphaZero=FromDigits[{c}/.alphaExpFit];
```

- b. Extracting the first restaging growth kinetic (α_1).
 - i. Repeat the steps described in step 7.a., but using the baseline ($t = 0$) and first restaging/follow-up (after start of treatment: $t > 0$) points (see Equation 6).

Note: In the case of unavailable pre-treatment measurements, α_0 should be estimated from the α_1 of the fastest growing tumor burdens within the patient population (these are assumed to have the least response to therapy, and therefore to give a reasonable estimate of untreated growth rate). We used the average of the top 10% fastest growing in Butner et al.²

- c. Fitting the model to data.
 - i. In order to obtain an exact model solution, the time-dependent model form (see Equation 2) is fit numerically to all tumor burden time course data with only $t \geq 0$ (dataPositive) using the function NonlinearModelFit.

```
>nlm=NonlinearModelFit[dataPositive,
  {(pinf/(1-(1-pinf)*Exp[-A*t]))},{pinf,A},t]
```

Note: This equation may be expanded such that $n_{lm} = f(\alpha_0, A, \mu)$. This should be avoided, as doing so will introduce a region where the model is undefined (due to dividing by zero) and may cause the algorithm to fail. ρ_{inf} and A should also be constrained to be > 0 .

Note: Once obtained, the solution should be plotted against the data to ensure goodness of fit. If the solution is not satisfactory, it may be improved by further refining the numerical method. Mathematica offers many ways to do this, we suggest a dual phase evolution-based method we've found to reliably produce satisfactory results here.

```
Method->{NMinimize,Method->
  {"DifferentialEvolution","ScalingFactor"->0.7,
  "CrossProbability"->0.3,
  "PostProcess"->{FindMinimum,AccuracyGoal->0,
  PrecisionGoal->20,
  Method->"LevenbergMarquardt"}}}
```

Note: This method seeks to avoid local minima to find the global best fit. Suggestions to further optimize this approach are provided in [troubleshooting](#).

- d. Obtaining exact model parameter solutions.
 - i. Extract numerical values for model parameters from the NonlinearModelFit solution, i.e.,

```
>rhoInf=pinf/.nlmRhoInf["BestFitParameters"];
```

- ii. If pre-treatment tumor burden was known, use these values and the value for α_0 obtained in 7.a to solve the system of equations in [Equations 3](#) and [4](#); this is easily done with the `Solve[]` function.

EXPECTED OUTCOMES

The model may be informed using non-invasively obtained tumor burden imaging data, either from in-house generated patient data or from publicly-available literature or database sources, and the model may be solved through numerical fitting to measured data. Once solutions are obtained, statistical analysis may be performed to find binary classifier threshold values for each model parameter to predict patient response, and patient survival if survival data is available.

Model parameters A and μ have been shown to be associated with patient response to ICI therapy, and parameter α_1 is also shown to be predictive of patient survival ([Figure 1](#)). The exact numerical cutoff for optimal response and survival prediction is likely disease and therapy specific; these may be identified for applications of interest as described in [quantification and statistical analysis](#).

QUANTIFICATION AND STATISTICAL ANALYSIS

⌚ Timing: hours to days

Model users will likely want to identify parameter thresholds for binary classification of patient response (responder: tumor burden is reduced vs. non-responder: tumor burden increases, at time of last available follow-up) or survival after ICI therapy. While determination of this threshold for response is always possible with this approach (if it exists within the population being studied), the ability to predict patient survival is inherently dependent on the availability of survival data (e.g.,

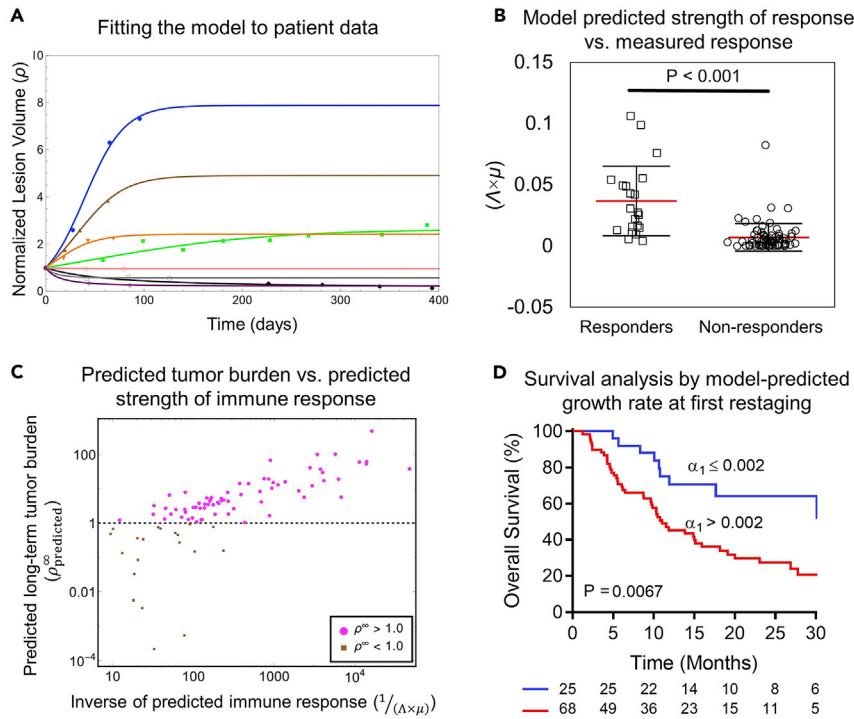


Figure 1. Examples of model analysis and results

(A) Model fits to normalized time-course tumor burden data.

(B) Model-estimated strength of immune response ($\lambda \times \mu$) is expected to be significantly different between patients by measured tumor response at last follow-up (responders: tumor burden reduced vs. baseline, non-responders: tumor burden increased vs baseline). In the example shown, $P < 0.001$ by Wilcoxon rank-sum test, and error bars show mean and standard deviation (red and black, respectively).

(C) Model-estimated strength of immune response is also expected to correlate with model-predicted long-term tumor burden ($\rho^{\infty}_{\text{predicted}}$, see Equation 4).

(D) Model parameter α_1 (model-derived growth rate at first restaging; Equation 6) reliably predicts patient survival no later than first restaging. P values (Mantel-Cox) for separation shown in inset. Reproduced with permission from Butner et al.^{1,2}

time to death or censor), which can usually be obtained when using in-house data but cannot be gleaned from published spider-plots.

In order to make model-based response predictions, specific model parameter ranges for the drug-disease combination being studied should be established. Reliable ranges for multiple drug-disease combinations have already been reported^{1,2,24}; however, values may vary slightly depending on disease and ICI,² and this should still be performed for further validation. Parameter ranges that predict for responders vs. non-responders may be separated by a binary classifier, which represents a specific value for each parameter with highest likelihood for accurate separation of response categories.

△ CRITICAL: Care must be taken to select a binary classifier that separates the data in a reasonable fashion. Values on extreme ends of the data should be avoided in most cases. This is usually a function of the desired predictive focus. For example, maximizing sensitivity may lead to unacceptably low specificity, and vice versa. We recommend that users start by maximizing the accuracy statistic to establish a reasonable estimation of where the binary classifier should be found.

1. Identifying binary classifier values for patient response prediction.

- a. Determine if there is a significant difference between model parameter values for different response categories.
 - i. Define desired response categories. Examples could be tumor volume increases or decreases relative to baseline (at time of first ICI treatment) or sorting based on standard RECIST criteria (e.g., stable or progressive disease vs. partial or complete response).
 - ii. Sort model parameter values based on the chosen response criteria, where response category is determined based on measured tumor burden at last restaging.
 - iii. Perform appropriate statistical test for significant difference between values in two categories.

Note: In our observations, data is usually nonparametric (not normally distributed). Standard tests for normalcy include skewness and kurtosis or Shapiro-Wilk tests; these may be accomplished using your statistical analysis software of choice. Based on the results, significant difference may be determined based on parametric tests (e.g., Student's *t* test) or non-parametric (e.g., Wilcoxon rank-sum, Mann-Whitney, etc.). If a significant difference is found between categories (i.e., $p < 0.05$ or other desired significance level) proceed to step b next; otherwise, this parameter is not supported as a binary classifier in your data set.

- b. If significant difference is found in step 1.a.iii, the binary classifier threshold for response prediction may be optimized by generating a Receiver Operator Classifier (ROC) curve, where response is sorted based on a proposed binary classifier threshold vs. measured response.

Note: The optimal binary classifier threshold is identified by performing a parameter sweep of the binary classifier value in order to identify the value where a desired ROC statistic (e.g., sensitivity, specificity, accuracy, F-score, Youden's *J* statistic, etc.) is maximized; the statistic used should be chosen based on your desired outcome or prediction need.

- i. Sensitivity quantifies the percentage of patients predicted to be non-responders by the first restaging growth kinetic (α_1).
- ii. Specificity quantifies the percentage of patients predicted to be non-responders by model parameters μ and Δ .
- iii. If a balanced optimization between accuracy in predicting responders and non-responders is desired, accuracy or Youden's *J* statistic may be used.

Note: We recommend users start with maximizing the accuracy statistic first; see **Critical**, above.

2. Identifying binary classifier threshold for patient survival prediction.

- a. We recommend predicting survival to median survival time in your patient cohort; that is, this survival time will determine 'good' vs. 'bad' survival outcomes. Survival prediction will be based on model-derived growth kinetic after start of treatment (α_1), and accuracy statistic should be optimized as described.
- b. Calculate the α_1 parameter for each patient as specified [coding a script to solve the model](#): 7.a. However, in this case, the growth rate is estimated between times at baseline (start of treatment) and time of first restaging after start of treatment. Assuming data is structured as in the example so far, in Mathematica this is done as:

```
>dataToComputeAlphaOne=normalizedDataPoints[{{2,3}}];
>logAlphaOneData=
  {#[[1]],Log[#[[2]]]}&/@dataToComputeAlphaOne;
>alphaOneFit=FindFit[logAlphaOneData,cx+d,{c,d},x];
```

- c. Determine median survival time in your population. This may be done by generating a single Kaplan-Meier curve for the full data set in your statistics package of choice and extracting the time when 50% of the population is surviving or censored.

Note: We recommend generating a Kaplan-Meier curve instead of using methods like simply calculating the median (e.g., the MEDIAN() function in Microsoft Excel) because the Kaplan-Meier method will handle censored data correctly to determine true 50% survival time.

- d. Categorize data based on if the patient survived to the calculated median survival time. Sort patients known to have survived to times later than the 50% survival time into the positive category and patients who died or were censored before this time into the negative category.
- e. Generate an ROC curve based on positive vs. negative categories identified.
- f. Perform another parameter sweep to determine optimal binary classifier threshold for predicting survival with α_1 using the same procedure detailed in step 1.b.
- g. To assess the performance of α_1 for predicting patient survival, divide patients based on the identified binary classifier threshold and generate Kaplan-Meier curves for both in your statistical analysis software of choice and assess statistics for separation between groups. We recommend using the standard threshold of p -value < 0.05 .

LIMITATIONS

As previously mentioned, this procedure is dependent on data availability. If pre-treatment tumor burden measurements are unavailable, then unique solutions for all model parameters cannot be obtained; however, patient response prediction may still be possible through parameter combinations or estimation of pre-treatment growth rate. In the case of insufficient numbers of tumor measurement times (restagings) or errors in lesion volume measurement, model accuracy will understandably be reduced; however, we have shown that Spearman coefficient between model fits with full vs. truncated data are strong for all values of model parameter μ , but less so for cases of extreme truncation (baseline + 1 more measurement after start of treatment) of Δ , while variations of $\pm 10\%$ of model parameter values were found to correspond to $\pm 9.2\%$ in tumor volumes, demonstrating good stability when tumor volume measurement fluctuates.²⁴ Additionally, the numeric values of binary classifier thresholds identified with this procedure may be disease-drug specific, although we have shown some crossover in binary classifier thresholds among different disease-ICI combinations.¹ Finally, we note that response prediction has been shown to increase in accuracy with additional, longer-time follow-up data.²⁴ Although this is to be reasonably expected, early-time prediction is still shown to be highly reliable, demonstrating accuracy as high as 88% at time of first restaging.¹

TROUBLESHOOTING

Problem 1

Variable treatment response measurement.

Despite efforts to standardize treatment response measurements (even for clinical trials), the widely accepted standards (such as RECIST or RANO) are vulnerable to interpretation and categorization of tumors based on threshold-based changes in 2-dimensional metrics. Tumor segmentation for volumetric measurement is subject to inter- and intra-individual variability.

Potential solution

Ensuring a clear standard, consistent operating procedure (SOP) for the tumor measurements across all observers that also including cross-review of measurements for agreement to address interobserver variability will improve consistency in tumor response measurement. Development,

validation, and implementation of automated tools can enhance the consistency and performance of tumor measurements with human-in-the-loop operational procedures. Treatment response rubrics should also be codified and standardized across all users when calibrating model parameters for response prediction in order to maximize accuracy of identified binary classifier thresholds. Users may define response categories based on existing standards, or using simplified response categories (e.g., responders vs. non-responders, as shown in [Figure 1](#)).

Problem 2

Variable imaging data acquisition.

If data is being acquired from patients who are receiving standard of care therapy, and even in clinical trials, there may be time points when some patients did not have the imaging required to make tumor measurements for the model. Additionally, even if imaging was acquired, the image acquisition protocol and scanner may not be consistent across all time points, leading to variability in tumor detection and measurement. For example, variability in slice thickness can dramatically impact tumor detection and volume measurement, particularly for small brain metastatic lesions.²⁵

Potential solution

Patients without the required imaging time points or imaging sequences that follow pre-specified imaging protocols to minimize inconsistency in tumors measurements may be excluded. Missed or missing measurement times do not prevent model-based prediction, as long as these are not the critical times of one pre-treatment scan, baseline (at start of treatment), and at least one follow-up scan after treatment start. In the case of inconsistent imaging protocols across the patient cohort, there are also ongoing efforts to harmonize imaging data acquired with differing imaging protocols.^{26,27}

Problem 3

Imaging-based tumor measurements are missing or not visible at one or more time points.

Commonly, a lesion may not be visible at one or more imaging time (restaging); this can cause an artificially low tumor burden and associated overestimation of tumor kill.

Potential solution

The investigator must assess if they believe the tumor is still there but not visible or if the tumor has shrunk below minimum resolution size; this can be inferred by confirming the presence or absence of the tumor at subsequent follow-ups. If the tumor is measured to be similar sizes before and after the not visible time point, then it is likely present. In this case, the tumor volume can be estimated by linear interpolation between surrounding measured volumes. Conversely, if the tumor is not seen at later time points, or demonstrates a long delay before reemergence, then it is likely that the tumor was reduced below the visible threshold, and its volume may be assumed to be zero at these times (unless this causes the total tumor volume to be zero).

Problem 4

Variability in treatment course.

Many patients do not get the full immunotherapy treatment course for 4 cycles often due to side effects. Many patients receive SRS concurrently with immunotherapy which shows direct shrinkage and possible effect on other distant lesions.

Potential solution

The investigator can account for variability in treatment durations within the model or set eligibility criteria that sets a threshold for minimum treatment received. Developing a model that predicts

treatment response at a lesion level rather than patient level would account for metastases treated with local therapies such as SRS or surgical resection.

Problem 5

Model curve does not fit data well.

We have observed that local minima may cause problems with regression analysis of our model, which is also usually the cause of poor fits.

Potential solution

There are many other methods to overcome local minima in regression analysis, which often function by introducing random noise into the algorithm to increase likelihood of finding the global minimum. Well-suited algorithms include stochastic optimization approaches such as simulated annealing or stochastic gradient methods.^{28,29} We have provided a dual phase evolution method³⁰ to do this in Mathematica in [coding a script to solve the model](#): 7.c. If another programming language is used, then investigators will need to select an available method for the platform of choice.

Problem 6

Total measured tumor burden goes to zero.

Total tumor burden may be reported to be zero, too small to be seen, or not visible at one or more time points. Attempting the numerical fit to this data will result in a 'divide by zero' error.

Potential solution

Remove all data points where total volume = zero and rerun the fitting script. Note that the model does not predict recurrence that may happen after total tumor burden is reduced below observable levels or eliminated.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zhihui Wang (zwang@houstonmethodist.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

A detailed guideline regarding how to parameterize the model and then use it to make predictions on immunotherapy outcome has been provided, so users should be able to implement the model in their own computational language of choice. We have also provided a sample Mathematica script ([Data S1](#). Example script) and literature-obtained data for ICI treatment of non-small cell lung cancer (NSCLC) and Melanoma as representative cases ([Data S1](#). Sample data), so interested users may test the model out of the box.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xpro.2022.101886>.

ACKNOWLEDGMENTS

This research has been supported in part by the National Science Foundation Grant DMS-1930583 (V.C. and Z.W.), the National Institutes of Health (NIH) Grants 1R01CA253865 (V.C. and Z.W.), 1R01CA226537 (V.C. and Z.W.), 1R01CA222007 (V.C. and Z.W.), 1R01AI165372 (Z.W.),

1R01DK132104 (Z.W.), and 1R01DK133610 (Z.W.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, Z.W.; Methodology, J.D.B., V.C., C.C., Z.W.; Software, J.D.B.; Investigation, J.D.B., M.F., V.C., C.C., Z.W.; Visualization and data interpretation, J.D.B., M.F., C.C., Z.W.; Writing, J.D.B., M.F., C.C., Z.W.; Supervision, C.C., Z.W.; Funding acquisition, V.C., Z.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Butner, J.D., Elganainy, D., Wang, C.X., Wang, Z., Chen, S.H., Esnaola, N.F., Pasqualini, R., Arap, W., Hong, D.S., Welsh, J., et al. (2020). Mathematical prediction of clinical outcomes in advanced cancer patients treated with checkpoint inhibitor immunotherapy. *Sci. Adv.* 6, eaay6298. <https://doi.org/10.1126/sciadv.aay6298>.
- Butner, J.D., Wang, Z., Elganainy, D., Al Feghali, K.A., Plodinec, M., Calin, G.A., Dogra, P., Nizzero, S., Ruiz-Ramírez, J., Martin, G.V., et al. (2021). A mathematical model for the quantification of a patient's sensitivity to checkpoint inhibitors and long-term tumour burden. *Nat. Biomed. Eng.* 5, 297–308. <https://doi.org/10.1038/s41551-020-00662-0>.
- Haslam, A., and Prasad, V. (2019). Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Netw. Open* 2, e192535. <https://doi.org/10.1001/jamanetworkopen.2019.2535>.
- Altrock, P.M., Liu, L.L., and Michor, F. (2015). The mathematics of cancer: integrating quantitative models. *Nat. Rev. Cancer* 15, 730–745. <https://doi.org/10.1038/nrc4029>.
- Konstorum, A., Vella, A.T., Adler, A.J., and Laubenbacher, R.C. (2017). Addressing current challenges in cancer immunotherapy with mathematical and computational modelling. *J. R. Soc. Interface* 14, 20170150. <https://doi.org/10.1098/rsif.2017.0150>.
- Wieland, F.-G., Hauber, A.L., Rosenblatt, M., Tönsing, C., and Timmer, J. (2021). On structural and practical identifiability. *Curr. Opin. Syst. Biol.* 25, 60–69. <https://doi.org/10.1016/j.coisb.2021.03.005>.
- Cristini, V., Koay, E., and Wang, Z. (2017). *An Introduction to Physical Oncology: How Mechanistic Mathematical Modeling Can Improve Cancer Therapy Outcomes* (CRC Press).
- Pascal, J., Bearer, E.L., Wang, Z., Koay, E.J., Curley, S.A., and Cristini, V. (2013). Mechanistic patient-specific predictive correlation of tumor drug response with microenvironment and perfusion measurements. *Proc. Natl. Acad. Sci. USA* 110, 14266–14271. <https://doi.org/10.1073/pnas.1300619110>.
- Pascal, J., Ashley, C.E., Wang, Z., Brocato, T.A., Butner, J.D., Carnes, E.C., Koay, E.J., Brinker, C.J., and Cristini, V. (2013). Mechanistic modeling identifies drug-uptake history as predictor of tumor drug resistance and nano-carrier-mediated response. *ACS Nano* 7, 11174–11182. <https://doi.org/10.1021/nn4048974>.
- Wang, Z., Kerketta, R., Chuang, Y.L., Dogra, P., Butner, J.D., Brocato, T.A., Day, A., Xu, R., Shen, H., Simbawa, E., et al. (2016). Theory and experimental validation of a spatio-temporal model of chemotherapy transport to enhance tumor cell kill. *PLoS Comput. Biol.* 12, e1004969. <https://doi.org/10.1371/journal.pcbi.1004969>.
- Brocato, T.A., Coker, E.N., Durfee, P.N., Lin, Y.S., Townson, J., Wyckoff, E.F., Cristini, V., Brinker, C.J., and Wang, Z. (2018). Understanding the connection between nanoparticle uptake and cancer treatment efficacy using mathematical modeling. *Sci. Rep.* 8, 7538. <https://doi.org/10.1038/s41598-018-25878-8>.
- Brocato, T.A., Brown-Glaberman, U., Wang, Z., Selwyn, R.G., Wilson, C.M., Wyckoff, E.F., Lomo, L.C., Saline, J.L., Hooda-Nehra, A., Pasqualini, R., et al. (2019). Predicting breast cancer response to neoadjuvant chemotherapy based on tumor vascular features in needle biopsies. *JCI insight* 5, 126518. <https://doi.org/10.1172/jci.insight.126518>.
- Dogra, P., Butner, J.D., Nizzero, S., Ruiz Ramírez, J., Noureddine, A., Peláez, M.J., Elganainy, D., Yang, Z., Le, A.D., Goel, S., et al. (2020). Image-guided mathematical modeling for pharmacological evaluation of nanomedicines and monoclonal antibodies. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 12, e1628. <https://doi.org/10.1002/wnan.1628>.
- Goel, S., Ferreira, C.A., Dogra, P., Yu, B., Kuttyreff, C.J., Siamof, C.M., Engle, J.W., Barnhart, T.E., Cristini, V., Wang, Z., and Cai, W. (2019). Size-optimized ultrasmall porous silica nanoparticles depict vasculature-based differential targeting in triple negative breast cancer. *Small* 15, e1903747. <https://doi.org/10.1002/sml.201903747>.
- Dogra, P., Butner, J.D., Ruiz Ramírez, J., Chuang, Y.L., Noureddine, A., Jeffrey Brinker, C., Cristini, V., and Wang, Z. (2020). A mathematical model to predict nanomedicine pharmacokinetics and tumor delivery. *Comput. Struct. Biotechnol. J.* 18, 518–531. <https://doi.org/10.1016/j.csbj.2020.02.014>.
- Dogra, P., Ramírez, J.R., Butner, J.D., Peláez, M.J., Chung, C., Hooda-Nehra, A., Pasqualini, R., Arap, W., Cristini, V., Calin, G.A., et al. (2022). Translational modeling identifies synergy between nanoparticle-delivered miRNA-22 and standard-of-care drugs in triple-negative breast cancer. *Pharm. Res.* 39, 511–528. <https://doi.org/10.1007/s11095-022-03176-3>.
- Motzer, R.J., Rini, B.I., McDermott, D.F., Redman, B.G., Kuzel, T.M., Harrison, M.R., Vaishampayan, U.N., Drabkin, H.A., George, S., Logan, T.F., et al. (2015). Nivolumab for metastatic renal cell carcinoma: results of a randomized phase II trial. *J. Clin. Oncol.* 33, 1430–1437. <https://doi.org/10.1200/jco.2014.59.0703>.
- Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D.R., Steins, M., Ready, N.E., Chow, L.Q., Vokes, E.E., Felip, E., Holgado, E., et al. (2015). Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* 373, 1627–1639. <https://doi.org/10.1056/NEJMoa1507643>.
- Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 366, 2443–2454. <https://doi.org/10.1056/NEJMoa1200690>.
- Powles, T., Eder, J.P., Fine, G.D., Braiteh, F.S., Lortot, Y., Cruz, C., Bellmunt, J., Burris, H.A., Petrylak, D.P., Teng, S.L., et al. (2014). MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* 515, 558–562. <https://doi.org/10.1038/nature13904>.
- Wolfram Research (2021). *Mathematica 13.0.0* (Wolfram Research, Inc.).
- Rohatgi, A. (2021). WebPlotDigitizer 4.5. <https://automeris.io/WebPlotDigitizer>.
- RaySearch Laboratories (2021). *RayStation 11B* (RaySearch Laboratories).
- Butner, J.D., Martin, G.V., Wang, Z., Corradetti, B., Ferrari, M., Esnaola, N., Chung, C., Hong, D.S., Welsh, J.W., Hasegawa, N., et al. (2021). Early prediction of clinical response to checkpoint inhibitor therapy in human solid

- tumors through mathematical modeling. *Elife* 10, e70130. <https://doi.org/10.7554/eLife.70130>.
25. Thrower, S.L., Al Feghali, K.A., Luo, D., Paddick, I., Hou, P., Briere, T., Li, J., McAleer, M.F., McGovern, S.L., Woodhouse, K.D., et al. (2021). The effect of slice thickness on contours of brain metastases for stereotactic radiosurgery. *Adv. Radiat. Oncol.* 6, 100708. <https://doi.org/10.1016/j.adro.2021.100708>.
26. Garcia-Dias, R., Scarpazza, C., Baecker, L., Vieira, S., Pinaya, W.H.L., Corvin, A., Redolfi, A., Nelson, B., Crespo-Facorro, B., McDonald, C., et al. (2020). Neuroharmony: a new tool for harmonizing volumetric MRI data from unseen scanners. *Neuroimage* 220, 117127. <https://doi.org/10.1016/j.neuroimage.2020.117127>.
27. Dewey, B.E., Zhao, C., Reinhold, J.C., Carass, A., Fitzgerald, K.C., Sotirchos, E.S., Saidha, S., Oh, J., Pham, D.L., Calabresi, P.A., et al. (2019). DeepHarmony: a deep learning approach to contrast harmonization across scanner changes. *Magn. Reson. Imaging* 64, 160–170. <https://doi.org/10.1016/j.mri.2019.05.041>.
28. Pennisi, M., Catanuto, R., Pappalardo, F., and Motta, S. (2008). Optimal vaccination schedules using simulated annealing. *Bioinformatics* 24, 1740–1742. <https://doi.org/10.1093/bioinformatics/btn260>.
29. Moles, C.G., Mendes, P., and Banga, J.R. (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res.* 13, 2467–2474. <https://doi.org/10.1101/gr.1262503>.
30. Paperin, G., Green, D.G., and Sadedin, S. (2011). Dual-phase evolution in complex adaptive systems. *J. R. Soc. Interface* 8, 609–629. <https://doi.org/10.1098/rsif.2010.0719>.