Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting

Angela M. Jarrett [1,2], Anum S. Kazerouni [3,14], Chengyue Wu[1], John Virostko[2,4,5], Anna G. Sorace [6,7,8], Julie C. DiCarlo [1,2], David A. Hormuth II [1,2], David A. Ekrut[1], Debra Patt[9], Boone Goodgame[5,10,11], Sarah Avery[12] and Thomas E. Yankeelov [1,2,3,4,5,13] ✉

This protocol describes a complete data acquisition, analysis and computational forecasting pipeline for employing quantitative MRI data to predict the response of locally advanced breast cancer to neoadjuvant therapy in a community-based care setting. The methodology has previously been successfully applied to a heterogeneous patient population. The protocol details how to acquire the necessary images followed by registration, segmentation, quantitative perfusion and diffusion analysis, model calibration, and prediction. The data collection portion of the protocol requires ~25 min of scanning, postprocessing requires 2–3 h, and the model calibration and prediction components require ~10 h per patient depending on tumor size. The response of individual breast cancer patients to neoadjuvant therapy is forecast by application of a biophysical, reaction–diffusion mathematical model to these data. Successful application of the protocol results in coregistered MRI data from at least two scan visits that quantifies an individual tumor's size, cellularity and vascular properties. This enables a spatially resolved prediction of how a particular patient's tumor will respond to therapy. Expertise in image acquisition and analysis, as well as the numerical solution of partial differential equations, is required to carry out this protocol.

Introduction

'Mechanism-based modeling' of cancer implies the incorporation of biological mechanisms into a mathematical model designed to predict the spatial and/or temporal dynamics of tumor characteristics. We (and others) have previously described the various opportunities and barriers for the practical use of imaging data in predictive, mathematical modeling of cancer[1–4]. There is growing evidence that imaging-informed, biophysical mathematical models can accurately predict the development of cancers of the kidney[5], prostate[6,7], brain[8–11], lung[12,13], pancreas[14–17], and breast[18–27]. These studies often aim to evaluate tumor growth or response to therapy on a patient-specific basis without having to first 'train' the model on large population data; that is, the individual patient's data calibrate the model, followed by a model-generated prediction about that individual patient's tumor response. Imaging data are a fundamental enabler of this process as the measurements can be collected in three dimensions (3D) at the time of diagnosis and at multiple time points throughout treatment, allowing for patient-specific calibration and prediction. Such an approach not only has the capacity to potentially forecast response for individual patients, but this strategy may also enable an in silico twin to be established for each patient to test therapeutic regimens and optimize treatment[26,28]. As the majority (85%) of oncology patients receive their care outside of academic research-oriented medical facilities[29], parameterizing models with data accessible in a community setting that is specific to an individual has the potential to dramatically and positively impact patient care.

[1]Oden Institute for Computational Engineering and Sciences, Austin, TX, USA. [2]Livestrong Cancer Institutes, Austin, TX, USA. [3]Departments of Biomedical Engineering, Austin, TX, USA. [4]Departments of Diagnostic Medicine, Austin, TX, USA. [5]Departments of Oncology, Austin, TX, USA. [6]Department of Radiology, University of Alabama at Birmingham, Birmingham, AL, USA. [7]Department of Biomedical Engineering, University of Alabama at Birmingham, Birmingham, AL, USA. [8]O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA. [9]Texas Oncology, Austin, TX, USA. [10]Departments of Internal Medicine, The University of Texas at Austin, Austin, Texas, USA. [11]Seton Hospital, Austin, TX, USA. [12]Austin Radiological Association, Austin, TX, USA. [13]Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [14]Present address: Department of Radiology, University of Washington, Seattle, WA, USA. ✉e-mail: thomas.yankeelov@utexas.edu

This protocol describes how to acquire quantitative MRI data of breast cancer within community-based radiology centers, analyze this data to return spatially resolved maps of tumor physiology, and employ the derived parameters for calibrating a predictive, mechanism-based, mathematical model of tumor growth and treatment response on a patient-specific basis. While the components of this pipeline have been previously described separately[18–27], this the first description of how to apply the methodology as one complete process. There has been a recent increase in the use of medical imaging data to inform patient-specific mathematical models of tumor growth and treatment response (see, e.g., ref. [30],); however, the field has been hampered by a lack of consensus on how the relevant data are collected, processed and modeled. Thus, we hope this protocol will also provide a (practical) learning tool that enables future strategies to be developed that use MRI data in mathematical oncology.

Development of the protocol

Traditionally, high-spatial-resolution imaging data that enable anatomical and morphological assessment have been acquired in the standard-of-care setting. Such data do not typically provide insights into the underlying physiological, cellular and molecular characteristics of cancer, thereby limiting their use for mechanism-based, mathematical modeling. Developing more specific and quantitative measurements related to tumor biology, such as vascular status, perfusion, cellularity, hypoxia, metabolism and proliferation, is a major effort in MRI research[31]. The field has matured over the last decade, with many publications establishing the repeatability and reproducibility of several of these imaging measures[32–35]. The National Cancer Institute's Quantitative Imaging Network was formed to further the clinical validation of such quantitative imaging to enable prediction of tumor response to therapies in clinical trial settings[36].

The present protocol uses two quantitative MRI modalities: dynamic contrast-enhanced MRI (DCE-MRI) and diffusion-weighted MRI (DW-MRI). DCE-MRI acquires images in rapid succession before, during and after the injection of a contrast agent. If the data are acquired at high enough temporal resolution, they can be analyzed with an appropriate pharmacokinetic model to estimate different tissue vascular features on a voxel-specific basis within the imaging volume. DCE-MRI is repeatable and reproducible[37–39], and the output of DCE-MRI analysis has a statistical relationship with the response of breast tumors to neoadjuvant therapy (NAT)[40–42]. In a parallel fashion, DW-MRI acquires data on water mobility in tissues that are linked to the number and quality of cell barriers present, thereby providing a noninvasive readout on tissue cellularity[43]. DW-MRI is also repeatable and reproducible[37,38,44] and can predict the response of breast tumors to NAT[40,45]. These two methods are therefore reasonable candidates for inclusion in mechanism-based, mathematical modeling.

To predict individual prognosis for breast cancer patients, our group has developed mathematical models that use patient-specific MRI data to initialize and constrain model parameters and predictions; i.e., the model parameters are calibrated to the unique characteristics of each patient. The earliest iteration of these modeling efforts employed a simple logistic model that utilized DW-MRI data to estimate tumor cellularity[20]. The logistic model was later expanded to be defined both temporally and spatially in 2D, where the baseline measurements of the tumor, tumor cell movement, and the mechanical properties of the breast tissues were incorporated to constrain the model's predictions of the tumor growth and shape according to each individual patient's anatomy[21]. While this model's predictions were shown to outperform standard measures (such as the Response Evaluation Criteria In Solid Tumors, RECIST) as a prognostic indicator of response to therapy[22], it did not explicitly consider the therapies of each individual patient. Following that iteration, we extended the model to include estimates of drug delivery to each voxel via DCE-MRI, enabling a more accurate assessment of local tumor cell death due to therapy on a patient-specific basis[24]. Most recently, we have used this model to identify theoretical treatment regimens that we hypothesize would outperform the standard-of-care regimen the patient actually received[26].

Details of the mathematical model

The model used here is based on a 3D mathematical model we developed that includes the mechanical coupling of tissue properties to tumor growth and the delivery of systemic therapy[24,26]. This model was designed to be initialized with patient-specific imaging data to predict response of breast cancer patients to NAT[21–23]. The governing equation (a reaction–diffusion type

Table 1 | Description of variables for the model system

Variable	Description
$N_{TC}(\bar{x}, t)$	Number of tumor cells in the voxel at position $\bar{x}$ at time t
D	Diffusion coefficient of tumor cells, where $D = D_0\exp(-\gamma\sigma_{vm}(\bar{x}, t))$ (mm^2/d)
σ_{vm}	Von Mises stress (N/m^2)
$\vec{u}$	Displacement vector due to tumor cell growth (mm)
G	Shear modulus due to breast tissue properties, where $G = E/(2(1-\nu))$ (kPa)
$C_{tissue}^{drug}(\bar{x}, t)$	Concentration of drug in the tissue in the voxel at position $\bar{x}$ at time t (μM)

Parameter	Description	Value
D_0	Diffusion coefficient of tumor cells without stress	Calibrated, mm^2/d
γ	Mechanical coupling coefficient for stress	Assigned at 2.0×10^{-3} (m^2/N)
ν	Poisson's ratio	Assigned, 0.45 (dimensionless)
E	Young's modulus for adipose, fibroglandular and tumor tissues	Assigned, 2 kPa, 4 kPa and 20 kPa, respectively
λ	Coupling constant for displacement of tumor cells	Assigned as 2.5×10^{-3} (N·m/cell)
$k(\bar{x})$	Proliferation rate of tumor cells in the voxel at position $\bar{x}$	Calibrated, 1/d
θ	Carrying capacity of tumor cells per voxel	Defined, cells
α	Efficacy of the drug against tumor cells	Calibrated, 1/(μM·d)

partial differential equation, PDE) for the spatiotemporal evolution of tumor cells $N_{TC}(\bar{x},t)$, (see 'Approximating tumor cellularity' below), with respect to time, t, and per voxel, $\bar{x}$, is:

$$\frac{\partial N_{TC}(\bar{x},t)}{\partial t} = \overbrace{\nabla \cdot (D(\bar{x},t)\nabla N_{TC}(\bar{x},t))}^{\text{Diffusion}} + \overbrace{\underbrace{k(\bar{x})\left(1 - \frac{N_{TC}(\bar{x},t)}{\theta}\right)N_{TC}(\bar{x},t)}_{\text{Proliferation}} - \underbrace{\alpha C_{tissue}^{drug}(\bar{x},t)N_{TC}(\bar{x},t)}_{\text{Death due to treatment}}}^{\text{Reaction}}, \quad (1)$$

where the first term on the right-hand side describes the effects of tumor cell movement (i.e., the diffusion term), and the second term describes the tumor cell proliferation and death in time (i.e., the reaction term). All model parameters and functions are described in Table 1, and the reader is encouraged to refer to it as they move through the description of the mathematical model. This model has been well documented through its evolution[18–27], but we describe the features of the model and the rationale behind the terms in detail here for clarity and completeness.

The function $D(\bar{x}, t)$ represents the random diffusion (movement) of the tumor cells. Previously, this function was simply a constant resulting in isotropic tumor spread[21], but later work found that incorporating individual patient breast anatomy into the diffusion term resulted in statistically significant improvements in the prediction of total tumor cell number when compared with clinical observations[21]. Thus, we mechanically couple the function $D(\bar{x}, t)$ to the breast tissue's material properties via von Mises stress $\sigma_{vm}(\bar{x}, t)$:

$$D(\bar{x}, t) = D_0\exp(-\gamma\sigma_{vm}(\bar{x}, t)), \quad (2)$$

where D_0 is the diffusion coefficient in the absence of external forces, and γ is an empirical coupling constant. The exponential term damps D_0, where the von Mises stress is calculated for the fibroglandular and adipose tissues within the breast, with the fibroglandular tissue assigned a greater stiffness than the adipose tissue[46]. The mechanical coupling is subject to an equilibrium condition dependent upon changes in tumor cell number:

$$\nabla \cdot G\nabla \vec{u} + \nabla \frac{G}{1 - 2\nu}\left(\nabla \cdot \vec{u}\right) - \lambda\nabla N_{TC}(\bar{x},t) = 0, \quad (3)$$

where G is the shear modulus, $G = E/(2(1-\nu))$, with Young's modulus (E) and Poisson's ratio (ν) describing the material properties, $\vec{u}$ is the displacement due to tumor cell growth, and λ is another empirical coupling constant[21–23,30,47–51]. Therefore, the diffusion term encompasses tumor changes such
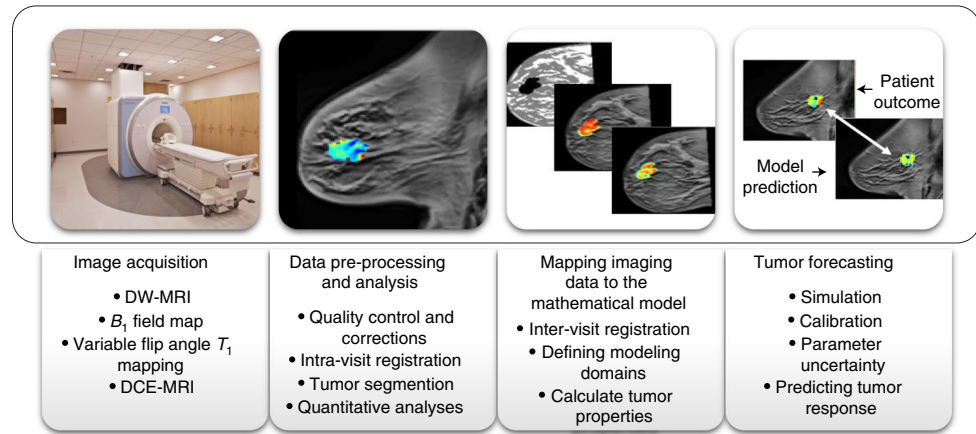
Fig. 1 | Overview of the protocol. Each panel contains summary keywords for each section. The procedure has five major sections: defining the patient population (Step 1, not shown), image acquisition (Steps 2–9), data analysis (Steps 10–25), mapping imaging data to the mathematical model (Steps 26–36), and tumor forecasting (Steps 37–40). Each section of the procedure focuses on specific areas of the protocol, and each section can be adapted for alternative investigations or used independently given specific circumstances in other studies. For example, the image acquisition section can be adapted for MRI studies in other organs. Also, given imaging data that are already acquired and analyzed, the mapping and forecasting sections can be applied. Note that informed consent must be obtained from all subjects. DCE-MRI, dynamic contrast-enhanced MRI; DW-MRI, diffusion-weighted MRI.

as growth or response to therapy that can cause deformations in the surrounding healthy tissues (i.e., fibroglandular and adipose tissues), thereby changing the stress field and the associated expansion of the tumor.

The second term on the right-hand side of Eq. (1) is the reaction term that describes tumor proliferation and therapy response. Due to the nature of the data, logistic growth is defined per voxel. Specifically, for MRI data, measurements are defined per voxel, and for each voxel the volume is known. Therefore, a maximum number of tumor cells can be estimated by using an approximate cell size and packing density (again, see 'Approximating tumor cellularity' below). Using logistic growth, the carrying capacity, θ, is defined per voxel as one value for all voxels, and the proliferation rate k is spatially resolved, $k(\overline{x})$, and derived from the data (see Step 38).

The reaction portion of the model also contains a term for tumor cell death due to therapy. The parameter α is a global parameter that represents the effectiveness of the therapy, and $C_{\text{tissue}}^{\text{drug}}(\overline{x}, t)$ represents the concentration of drug in the tumor tissue at position $\overline{x}$ and time t, as approximated from the Kety–Tofts model and patient-specific parameters (see 'DCE-MRI analysis' below). Importantly, the concentration of contrast agent in the tissue and plasma time courses from the Kety–Tofts model now represent those quantities for the administered drug. This is achieved by replacing concentration of contrast agent in the plasma curve with standard drug concentration curves for individual drugs and calculating the concentration in the tissue with each patient's derived vascular perfusion parameters (again, see 'DCE-MRI analysis' below). Therefore, an approximation of the concentration of drug in the tumor tissue that is spatially non-uniform and temporally varying based on the individual patient's NAT schedule is generated. Thus, the therapy term provides an estimation of the spatiotemporal distribution of drug in the tissue and its effect on the cells of each voxel. This assignment of drug delivery in the model is only a first-order approximation, not unlike other efforts that have sought to characterize the heterogenous delivery of systemic therapy through tissue[52–55]. Note that this approach implicitly assumes that the chemotherapy will extravasate into the tumor tissue in a manner similar to that of the gadolinium-based contrast agent, an assumption that should be relaxed by future model refinements. If a patient's drug concentration in the plasma deviates from population curves, this would affect the calibration of the model parameters; however, α provides a layer of flexibility in the model whereby the population-averaged approximation of the drug in the plasma is modulated by this global parameter.

Overview of the protocol

The procedure is divided into five major components (Fig. 1): identification of patients who would benefit (Step 1, not shown in Fig. 1), image acquisition (Steps 2–9), data preprocessing and analysis

(Steps 10–25), mapping imaging data to the mathematical model (Steps 26–36), and tumor forecasting (Steps 37–40). Each component has been divided into multiple steps for clarity of presentation. In the procedure, we provide detailed descriptions of each component, as well as guidance on avoiding potential pitfalls and suggestions for troubleshooting.

Patient selection

Our protocol was developed for patients recruited from community-based care centers that are eligible for NAT as a component of their care. Such patients are heterogeneous in tumor size, receptor status, age, body mass index and ethnicity. NAT (i.e., any treatment that occurs prior to surgical intervention) is typically indicated for patients with locally advanced breast cancer and consists of one or more regimens given over the course of 4–6 months. For example, in the case of triple-negative breast cancer, the standard of care can include doxorubicin and cyclophosphamide (first regimen), followed by paclitaxel (second regimen). There are, however, many variations in these protocols as determined by treating physicians. (This is, of course, a primary motivator for developing a mathematical forecasting system so that treatments can be optimized on a patient-specific basis[26,28].) The clinical response designations for NAT of pathological complete response (pCR) or residual disease (non-pCR) are determined by surgical pathology. Specifically, pCR is defined and reported as no residual invasive disease in either breast or axillary lymph nodes after NAT.

Image acquisition

This protocol requires the acquisition of quantitative MRI data of a breast cancer patient before and during NAT for calibrating a predictive, mechanism-based mathematical model designed to forecast their individual response. The timing of the imaging time points before and during NAT are of critical importance as they are used to calibrate, simulate and assess predictions of tumor response with the mathematical modeling system. Our protocol recommends for MRI data to be acquired at four time points: (1) prior to the initiation of NAT, (2) after one cycle of NAT, (3) after two to four cycles of NAT, and (4) after one cycle of NAT from scan 3. (Note: by 'cycle' we mean the administration of a single drug or combination of drugs over a designated period of time, typically 2–4 weeks.) These four time points provide data that correspond to the first cycles of each therapeutic regimen for the patients that receive two consecutive regimens (Fig. 2). While three or more imaging time points are encouraged, two imaging time points are all that is required to calibrate the model system and then compare predictions to standard clinical measures (e.g., pathological data from biopsies or surgery) to directly test the modeling predictions.

All image acquisition and patient care (imaging, oncology treatments, etc.) can be performed in community care settings (i.e., not academic, research-oriented medical centers). However, to work within the confines of imaging in standard-of-care settings, certain factors must be considered. For example, in the figures and examples included in this article, two scanners were used: one in an outpatient imaging facility, and the second in a regional hospital that provides both inpatient and outpatient services. While both imaging facilities undertook breast MRI as part of their routine clinical practice (a full diagnostic scan is ~20 min), they are located at different sites and on different service contracts, and have different quality control guidelines. The MRI technologists at such sites are usually responsible for positioning the patients and running the research protocols but might not have prior experience or expertise. Therefore, it is important to establish the repeatability and reproducibility of the required MRI measurements in each new environment[37,38], and implementation of the acquisition protocol requires clear (step-by-step) descriptions to be provided for the MRI technologists performing the scans.

Working with community physicians enables a broader segment of the population to be reached, but scheduling research MRI scans requires close integration and frequent communication regarding the availability of the patients, treating oncologist, referring physicians, nurses, imaging center staff and study staff. In our experience, missing time points or lack of data due to equipment failure, patient health, scheduling issues and/or parties unwilling to provide time is not uncommon. Moreover, community settings do not always employ a research-oriented nurse; therefore, lines of communication need to be clearly defined at the beginning of the study. Despite these challenges, working in community-based radiology settings can be easier for patients, with greater access to different facilities and allowing for more convenient travel to participate in the study.

This protocol requires acquisition of five MRI data types at each scan session: (1) DW-MRI, (2) B_1 field map to correct for radiofrequency (RF) inhomogeneity, (3) variable flip angle T_1-weighted data for generating a precontrast T_1 map, (4) dynamic, high-temporal-resolution, T_1-weighted data before,
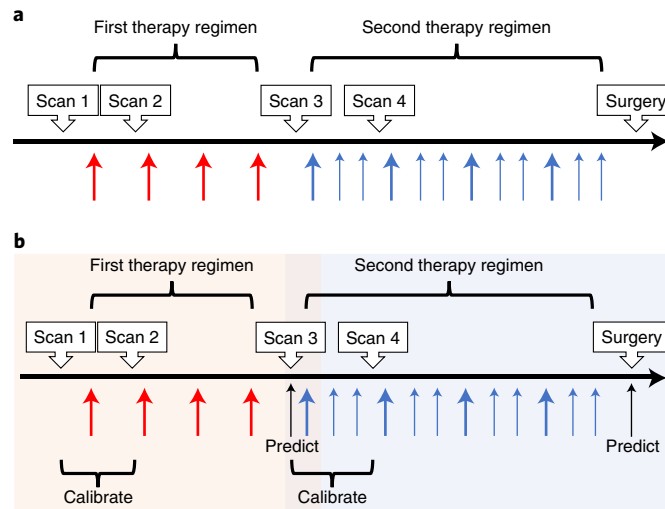
Fig. 2 | Timeline of MRI acquisition with an example standard-of-care NAT regimen for triple-negative breast cancer consisting of two therapeutic regimens. a, Example NAT regimen only. **b**, Example regimen with the protocol's calibration and prediction strategy. For **a** and **b**, wide arrows indicate the first dose and start of each cycle (i.e., the administration of a single drug or combination of drugs over a designated period of time, typically 2–4 weeks), while the narrow arrows indicate any additional doses within each cycle. For the first regimen, red arrows represent combination doxorubicin and cyclophosphamide (typically consisting of four cycles where drugs are administered as single doses separated by 2 weeks). For the second regimen, blue arrows represent paclitaxel (typically consisting of four cycles where therapy is administered every week with each cycle lasting 3 weeks—small arrows represent the additional doses). Some patients are treated with carboplatin in combination with paclitaxel, where carboplatin is administered during the first week of each paclitaxel cycle only (wide blue arrows). After NAT is complete, patients undergo surgery as part of their standard of care to determine pathological response. The protocol has MRI data collected prior to and just after the first cycle of each therapeutic regimen.

during and after the injection of a gadolinium-based contrast agent (DCE-MRI), and (5) high-resolution, pre- and postcontrast, T_1-weighted anatomical scans. These MRI data types were selected to provide reliable and quantitative values for individual breast cancer tumors as they have been well established in the literature. The imaging protocol utilizes standard sequences that are available on all clinical MRI scanners, eliminating the need for work-in-progress sequences or novel sequences that are not universally available.

DW-MRI provides information about the tissue microstructure by quantifying the motion of water molecules. Water molecules freely diffuse at ~3×10^{-3} mm^2/s at 37 °C, but as they encounter various tissue barriers, including large densities of cells, this diffusion rate, known as the apparent diffusion coefficient (ADC), will decrease[31,56,57]. A minimum of two b values (a factor that reflects the strength, duration and timing of the diffusion-encoding gradients in the scan) is required for estimation of ADC (in this protocol, we use three b values of 0, 200 and 800 s/mm^2, which are commonly utilized for breast tissue). DW-MRI acquired with very high b values (>1,000 s/mm^2) may result in low signal-to-noise ratio (SNR) that can adversely affect the ADC estimate, while low b values (<100 s/mm^2) can be affected by tissue perfusion where blood flow in the smallest vessels mimics diffusion, thereby altering the interpretation of the image.

As different tissues exhibit different T_1 relaxation times, T_1 mapping provides a means to differentiate tissue types (e.g., fat, muscle, parenchyma and/or tumor) and provides native T_1 values needed for downstream pharmacokinetic analyses of DCE-MRI data[31,56,57]. We use a standard approach for clinical breast T_1 mapping, involving the collection of multiple T_1-weighted images at variable flip angles. We collect images at ten flip angles ranging from 2° to 20° (in 2° increments) for estimation of typical breast tissue T_1 values (where more flip angles provide more data points for better curve fits, and this range allows for accurate estimations of various tissues in the breast from adipose to tumor). However, this approach is sensitive to inhomogeneities in the RF B_1 magnetic field used to 'tip' the magnetization by various flip angles, potentially leading to inaccurate estimations of native T_1[58]. To address this issue, a B_1 map is acquired to quantify and correct any spatial deviations in the nominal flip angle during acquisition of the T_1-weighted images used in mapping the T_1 parameter. Other T_1 mapping approaches include inversion and saturation recovery sequences (that are not as affected by B_1 inhomogeneities); these methods are the 'gold standard' for the calculation of
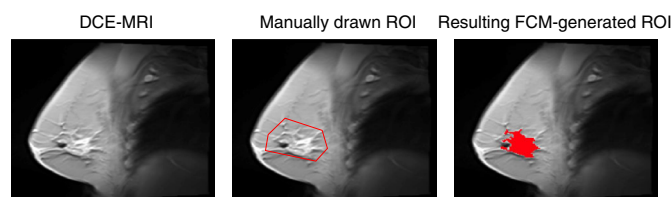
DCE-MRI   Manually drawn ROI   Resulting FCM-generated ROI

Fig. 3 | FCM clustering to generate a tumor ROI. Depicted is the sagittal cross section of a breast for the average DCE-MRI data for one patient (all panels). For the middle panel, a manually drawn ROI is shown, which identifies a conservative bounding polygon for the tumor. The right panel depicts the resulting ROI generated from the FCM algorithm within the manually drawn bounding polygon. Informed consent was obtained from this patient.

T_1, but the time necessary to collect these sequences in multislice or 3D are clinically prohibitive and are not incorporated into our protocol.

In DCE-MRI, a paramagnetic contrast agent is injected into the bloodstream through a peripheral vein. It travels throughout the circulatory system and can extravasate into the tumor, leading to a decreased T_1 relaxation time and corresponding increase in signal intensity in a T_1-weighted image. DCE-MRI data are acquired by collecting T_1-weighted images before, during and after the delivery of contrast agent. DCE-MRI data can then be analyzed to segment different tissues with differing contrast enhancement and also to extract measures characterizing contrast agent pharmacokinetics (details provided below in 'DCE-MRI analysis'). Acquisition parameters for our DCE-MRI measurement were selected to yield a temporal resolution <10 s (7.27 s) for accurate estimation of pharmacokinetic parameters[59,60]. If adjusting this protocol for other tissue types, it is important to bear in mind that an appropriate flip angle that minimizes contrast agent saturation effects must be selected for optimal DCE-MRI results, which may vary depending on the tissue imaged (i.e., breast versus brain).

Data preprocessing and analysis

Image processing starts with quality control, image correction and image registration, before moving to extract tumor-specific characteristics and quantitative descriptions of each tumor's cellular density and vasculature. The methods used include segmentation via clustering techniques as well as analysis of the quantitative MRI data to return maps of quantities reporting on blood flow and water diffusion[40].

Tumor segmentation

To analyze and process data, a tumor region of interest (ROI) is required for each patient and scan session. Using expertly drawn ROIs for the tumor burden is the gold standard. However, if provided with a conservatively drawn 'bounding box' (i.e., a hand-drawn polygon that surrounds the tumor but not its specific contours), thresholding based on enhancement is often used to determine the boundaries of tumors from DCE-MRI data. This threshold is a value chosen for which any voxel with signal intensity above that threshold in the postcontrast image is considered part of the tumor. As thresholding techniques can require manual adjustment for each patient scan and additional information to define patient-specific thresholds[24] (and vary by contrast type and amount), it is best to use as automated an approach as possible. Our protocol employs a fuzzy c-means (FCM) clustering algorithm[61,62]. The FCM algorithm outputs the probability of a voxel being tumor or nontumor, based on DCE-MRI contrast enhancement patterns. As FCM clustering does not partition voxels into clusters, it is more tunable than other 'hard' clustering methods (e.g., k-means clustering). See Fig. 3 for representative images of generating ROIs using FCM.

Registration techniques

Our approach to imaging-based modeling requires that all image sets for each patient be registered to one common spatial coordinate system; i.e., the images must be coregistered. Note that all registration processes do not completely preserve voxel information—even when rigid registration is used—due to multiple interpolations and resampling. To achieve image alignment, we perform two types of registration: intravisit registration (aligning all the data collected within one scan session) and intervisit registration (aligning each of the datasets across all scanning sessions for each patient). Intravisit registration is performed to correct for patient motion during the scanning session and is
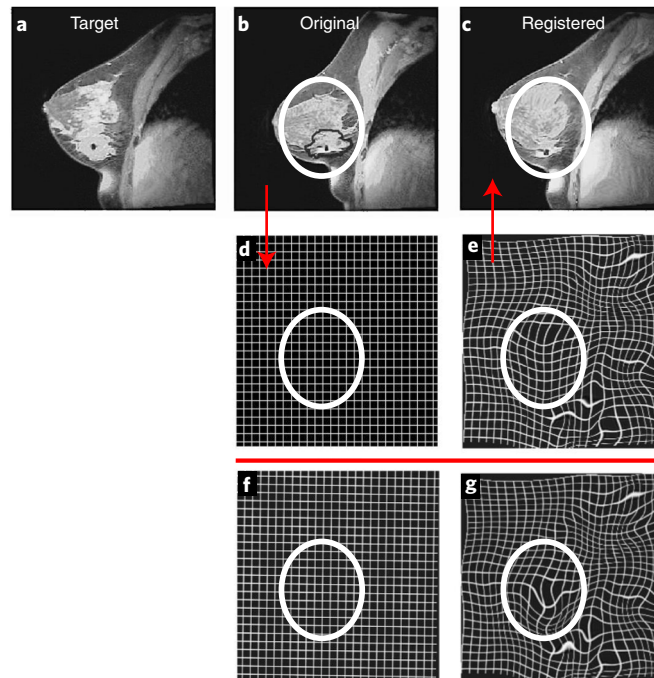
Fig. 4 | Comparison of intervisit registration results with and without tumor ROI penalties incorporated into the registration scheme. a, Target image (scan 2, defined in Fig. 2). **b**, 'Moving' image to be deformed/shifted to align to the target image (scan 3). In **b**, the moving image's tumor ROI is indicated by a black outline. **c**, Result of registering the moving image to the target image using the approach described in Steps 27–31 of the text (i.e., rigid + nonrigid B-spline with a tumor ROI penalty). **d,e**, Representative grid of the original moving image (**d**), and resulting deformed grid after registration (**e**), corresponding to the registered image in **c**. **f**, **g**, Deformations of the representative grid after rigid registration only (translation and rotations only) (**f**) and after the nonrigid B-spline registration without a tumor ROI penalty (**g**). Across **b**–**g**, white circles have been added to aid in comparing the fields for the areas surrounding the tumor. Note that, by including the tumor ROI penalty, there is less deformation of the tumor ROI; i.e., there is less deformation within the white circle in **e** versus **g**. This procedure is applied to the scan 1, 3 and 4 MRI datasets. Informed consent was obtained from this patient.

accomplished through a rigid registration prior to the calculation of quantitative parameters from the data.

For each patient, all image datasets are registered across time (intervisit) to a common space via a nonrigid registration algorithm with a constraint that preserves the tumor volumes at each time point[63]. If MRI data are obtained at four time points, we choose to register scans 1, 3 and 4 to scan 2 (target), as scan 2 is not an endpoint scan and is acquired early enough so that the patient will (likely) have tumor burden present to help guide alignment. The registration algorithm consists of a rigid registration of the tumor ROIs followed by a deformable b-spline registration with a rigidity penalty on the tumor regions[64,65]. This rigidity penalty is imposed to preserve the tumor volume/size and shape across all scan times. With a fully deformable registration, the tumor ROIs of scans 1, 3 and 4 may be morphed to match the tumor ROI of scan 2. See Fig. 4 for example comparisons of different registration results.

DCE-MRI analysis

The DCE-MRI data are analyzed using models of contrast agent pharmacokinetics to derive quantitative parameters of vascular perfusion and tissue volume fractions[59]. Specifically, the extended Kety–Tofts model is used to perform quantitative analysis of DCE-MRI data. We have found that a temporal resolution of 7.27 s for DCE-MRI data provides sufficient SNR and temporal sampling for the extended Kety–Tofts model to be applied[66]. However, we strongly encourage evaluation of voxel time course fits to determine which model appropriately captures the time course behavior of specific datasets. Pharmacokinetic modeling requires characterization of the time rate of change of the concentration of contrast agent in a feeding artery, i.e., the arterial input function (AIF). We estimate the AIF from the population-averaged signal intensity time course extracted from the axillary artery[67]. Further, we calculate the bolus arrival time (BAT) to shift the population AIF on a

voxel-wise basis to align the enhancement time of the AIF with that of the individual voxel[61], allowing for improved fits and more accurate parameter estimation.

The reference region model[68] is another approach to quantitative DCE-MRI analysis, removing the need for direct measurement of an AIF, where the tumor enhancement curve is compared with that of a reference region, such as pectoral muscle tissue (requiring additional tissue segmentation efforts). Simpler methods to analyze DCE-MRI data, such as the calculation of the signal enhancement ratio[69] and area under the signal intensity time course curve[31], can provide semiquantitative measures of vascular characteristics that have proven informative in distinguishing benign and malignant lesions[70] and predicting disease recurrence[71].

Approximating tumor cellularity

The ADC is calculated from the DW-MRI data, representing the rate at which water molecules diffuse in the tissue. It has been shown to approximate the cellular density of tissue. ADC values are calculated for each voxel via standard methods[72], and there is abundant literature on the inverse relationship between the measured ADC and tumor cellularity[73–77]. However, there is a level of ambiguity in the source of changes in ADC, as many other factors (e.g., cell membrane permeability[78], cell size and tissue tortuosity[79]) in addition to cellularity can also affect the measured ADC. Therefore, the approach of appraising cellularity with ADC is an approximation, and ongoing efforts are designed to eliminate some of the ambiguity in the interpretation of ADC[80–82], which could lead to improvements in the outcomes from applying the protocol.

Mapping imaging data to the mathematical model

After the generation of quantitative maps from each MRI data type, the maps are registered across all scan sessions (i.e., intervisit registration) for each patient so that all imaging data are aligned to a common image space. Once aligned, a breast domain is defined, within which each individual patient's data are used to calculate physical characteristics of each patient's tumor utilized by the mathematical model. The steps for generating tumor characteristics include calculating tumor cellularity, defining masks delineating fibroglandular and adipose tissues, approximating drug distributions, and estimating summary measures for each tumor across all scans.

The mathematical model employs the patient-specific characterization of breast anatomy, cellularity, vascular features and approximate drug distributions. This strategy is motivated by the hypothesis that, if a mathematical model, initialized and constrained by noninvasive imaging data for an individual patient collected early in the course of therapy, can be used to reliably predict tumor response, oncologists may be able to intervene and modify therapy on a patient-specific basis. Additionally, such a model could be used to optimize therapeutic regimens for patients with more robust mathematical methods such as optimal control theory[26,28].

Tumor forecasting

The quantitative maps are then used to initialize and calibrate tumor cell proliferation, drug efficacy and cellular migration within the mathematical model. That is, each patient's own imaging series is used to parameterize or identify growth and response parameters unique to that individual patient. Once parameterized, the model can be run forward for patient-specific predictions of the spatio-temporal evolution of tumor cellularity, allowing for a prediction of treatment response that can be directly compared with the observed outcome for each patient.

Expertise needed to implement the protocol

This protocol is designed to be implemented by a multidisciplinary team with experience and expertise across a wide-range of fields including both the acquisition and analysis of advanced, non-standard-of-care MRI data in the clinical setting, image processing (including segmentation and registration), and the numerical solution of PDEs. If implemented in the community setting, implementation requires coordination with community health providers.

The procedure described below was developed in direct response to the challenges presented for the incorporation of breast cancer MRI data into predictive mathematical models. These methods were therefore developed for data collected in the community setting, across multiple imaging sites, within specific standard-of-care constraints, and for a heterogeneous patient population. We have worked extensively to establish the repeatability and reproducibility of this quantitative MRI protocol at community-based imaging centers[37,38,83]. Furthermore, our protocol utilizes several strategies to

reduce bias and dependency on user interaction, including detailed data acquisition strategies as well as automated or semiautomated computational algorithms. There are, unavoidably, tasks that require expert evaluation (e.g., outlining tumor ROIs by radiologists) and others that rely on the operator's input (e.g., positioning the field of view (FOV) for MRI acquisition), which we have noted throughout the protocol. Throughout our methods, we have worked to include automated quality checks and quantification of the data to reduce overall user/operator influence.

Applications of the protocol

The protocol was designed for predicting response in breast cancer patients receiving NAT as a component of their clinical care. While the details of the protocol presented here are specific to breast cancer, the methods are generally applicable to any solid tumor for which the requisite data are accessible. In fact, this protocol has already been modified for a multimodality breast cancer modeling study[25] and, more recently, for brain cancer[84]. (There have also been many applications of our approach in the preclinical setting[47,48,85–88].)

In addition to the mechanism-based, mathematical modeling for which the protocol was designed, investigators who have previously collected the appropriate imaging data may use portions of the protocol for data processing and analysis to yield (for example) longitudinally aligned quantitative maps of tumor features. The data returned at the completion of Step 24 can then be employed in more conventional statistical studies that assess longitudinal changes in tumor characteristics and separating responders from nonresponders[40,42,45]. Such data can also provide the input data for analysis in the emerging field of radiomics[89,90], in 'habitat imaging'[91,92], and for application in artificial-intelligence-based models.

Limitations

The temporal sampling requirements of the DCE-MRI protocol constrain the achievable spatial resolution of data that can be collected while limiting noise. All images are acquired in the sagittal plane; while the transverse plane is the standard-of-care choice with a bilateral FOV, our imaging protocols use sagittal slices for better resolution over the affected breast within a minimal amount of time. For example, a standard-of-care, T_1-weighted, contrast-enhanced acquisition requires 80–90 s, whereas the same scan in our protocol is <10 s. With this coarser spatial resolution, we are likely unable to measure important anatomical features and may potentially miss small feeding vessels. This, in turn, can limit modeling strategies that aim to incorporate nutrient/oxygen delivery as well as estimates of the distribution of systemic therapies. Of course, ongoing efforts are actively investigating ways to provide faster acquisitions at the spatial resolution and FOV typically acquired in the standard-of-care setting[61,93,94].

Another limitation is that, even if we consider only the United States, each oncology practice or oncologist typically has a set of treatment regimens that are prescribed more often (all still within National Comprehensive Cancer Network guidelines); thus, there may be more treatment options included in the study when working with multiple recruiting physicians and oncology clinics, and especially so if implementing in multiple countries. This can introduce limitations on the type of analyses that can be performed until a large enough cohort of patients has been accrued for statistical power.

With any quantitative measurement, error is expected. Therefore, we have included best practices for encountering and dealing with physiologically implausible results, such as during the interpretation of diffusion-weighted data and pharmacokinetic analysis. This is of particular importance because mathematical models project/propagate errors derived from the data employed for calibration and prediction (which we address in the tumor forecasting section, Steps 37–40). Of course, the MRI data and corresponding quantitative analyses are ultimately approximations of the breast cancer's characteristics at the tissue and cellular scale, and alternative methods and procedures should continue to be explored to reduce error and ambiguity in these quantities.

We emphasize that what we present here is just one of many possible modeling formulations applicable to the dataset produced by this pipeline, and there are various other cancer modeling styles[95]. Also, while the protocol is built to accommodate heterogenous populations of patients, there are certain exclusions applied to make the mathematical modeling as practical as possible. For example, we exclude tumors at stages I and IV due to the fact that stage I tumors may be too small to reliably measure (<2 cm in diameter) with the MRI techniques, and stage IV may require alternative modeling techniques to account for tumor invasion and metastasis. Also, stage I and IV patients do

not typically receive NAT; stage I tumors are treated with surgery and radiation, and stage IV tumors are treated with palliative care. Furthermore, we note that many programming languages and numerical schemes are available for determining solutions for PDEs (such as the finite element method[23]), and for those less interested in the derivation of numerical codes, specific software programs exist to aid in the implementation of PDEs (e.g., FEniCS and MATLAB's PDE solvers).

Materials

Subjects
- Women over the age of 18 years who present with intermediate-to-high-grade invasive primary breast cancer and are considering NAT as a component of their clinical care. Patients with disease from all subtypes and treatment regimens (including immunotherapy, targeted and cytotoxic therapies) may be included if they have disease stage II and III cancers. Patients with a history of kidney disease, with abnormal creatinine or estimated glomerular filtration rate, who are pregnant or who are breast feeding must be excluded. Also exclude any patients who have any non-MR compatible ferromagnetic materials, are acutely ill, and/or for whom an MRI is technically unfeasible (e.g., due to breast volume or obesity) **!CAUTION** Patient data must be collected according to institutional and national regulations, and informed consent must be obtained from patients (for an example, see Supplementary Data 1). In the United States, data must be collected under an institutional review board-approved and American Health Insurance Portability and Accountability Act (HIPAA)-compliant study in which patients give informed consent to participate in a longitudinal MRI (with contrast) study throughout the course of their standard-of-care NAT. For the example data we present here, we obtained health information related to each participant's disease and MRI scans. This information includes laboratory test results, medical imaging reports, and diagnosis and treatment codes.

Reagents
- Gadolinium-based contrast agent (e.g., Multihance (Bracco) or Gadovist (Bayer)) **▲CRITICAL** Required for contrast-enhanced scans.

Equipment
- Power injector for administration of contrast agent (e.g., Medrad)
- MRI scanner
- Personal computer or server **▲CRITICAL** For the data processing and analysis as well as the model simulations, a personal computer may be capable of running the model calibration via the software described below, but we recommend using a server with similar specifications to our server, which has parallelized scripts for computational efficiency: 40 nodes, CPU per node: 2/8 Xeon E5-2680 2.7 GHz (turbo, 3.5) 1/61 Xeon Phi SE10P 1.1 GHz, memory: 32 GB per node.

Equipment setup
MRI scanner
MRI of the breast should be acquired using 3T scanners equipped with an 8- or 16-channel double-breast receive coil. For our studies, we employ Siemens scanners (Siemens Healthcare) with Sentinelle coils (Invivo). Other 3T scanners (Philips, GE) and breast coils may require slight adjustments to replicate the pulse sequences described in Steps 4–9, but as these are standard sequences, the scanners have the capabilities of collecting equivalent imaging protocols. After scanning, Digital Imaging and Communications in Medicine (DICOM) files are stripped of protected health information (PHI), labeled with assigned study identifiers, and uploaded onto a firewalled protected server.

Software
- REDCap (https://www.project-redcap.org/) or similar **▲CRITICAL** All clinical information regarding patient demographics, diagnosis, treatment and surgical outcome must be acquired from the patient clinical record by a nurse at the referring physicians' office. We relay this information to the research team using HIPAA-compliant communication and store in REDCap, a secure, HIPAA-compliant web application for building and managing online databases[96,97]. REDCap is a widely available and utilized service for managing single- and multiinstitutional clinical studies.
- MATLAB (MathWorks, Natick, MA) or similar **▲CRITICAL** We use MATLAB software for the vast majority of our process owing to its portability and convenient functionality across our diverse group of biomedical researchers including experimentalists, computational scientists, engineers, physicists,

Table 2 | Breast MRI acquisition parameters

MRI parameters	DW-MRI	B_1 mapping[a]	High-resolution T_1 (pre- and postcontrast)	T_1 mapping: variable flip angle	DCE-MRI
Scan sequence	Single-shot spin echo, echo planar	TurboFLASH	T_1-weighted, VIBE	3D gradient-echo FLASH	T_1-weighted, VIBE
TR (ms)	3,000	8,680	5.3	7.9[b]	7.02
TE (ms)	52	2	2.3	2.4[b]	4.6
Flip angle (°)	90	8	10	2–20	15
Acquisition matrix	128 × 128	96 × 96	256 × 256	192 × 192	192 × 192
Slice thickness (mm)	5[c]	5[c]	1	5[c]	5[c]
GRAPPA acceleration factor	2	None	2	3	2
Fat suppression	SPAIR	None	SPAIR	None	None

[a]Due to the inclusion of a slice gap in the B_1 mapping protocol, two acquisitions (at 17 s each) were performed to cover the same FOV as the other measurements. [b]These parameters should be set to 'minimum' given the confines of the other pulse sequence settings. [c]If adequate SNR can be maintained, then these values can be reduced. All breast MRI data were acquired in the sagittal plane with a FOV of 256 × 256 mm².

mathematicians and medical professionals. However, alternative software programs are available for the data processing, and with a sufficient programming background, the functions described throughout our protocol can be replicated in those environments.

- Elastix (https://elastix.lumc.nl/) ▲ CRITICAL Elastix is a toolbox for rigid and nonrigid registration of images in 3D[98]. The algorithms used in Elastix for registration are similar to previous approaches used by our group[63,99–101]. However, Elastix is an open-source toolbox, and its functions can be run through the command line via MATLAB allowing it to be integrated into existing data processing scripts.

Procedure

Patient selection ● Timing ~15 min

1   *Ensure patient is eligible for the study.* Patients should be women over the age of 18 years who present with intermediate-to-high-grade invasive primary breast cancer and are considering NAT as a component of their clinical care. Verify with the treating oncologist that each patient has no history of kidney disease and has normal creatinine and estimated glomerular filtration rate within 30 d of imaging studies. Exclude pregnant women and women who are breast feeding. Also exclude any patients who have any non-MR-compatible ferromagnetic materials, are acutely ill and/or for whom an MRI is technically unfeasible (e.g., due to breast volume or obesity).

2   *Obtain informed consent from patient.* Patients should be asked to consent to participate in research. A copy of the consent form used by our ongoing study is supplied in the Supplementary Data 1. **! CAUTION** Informed consent must be obtained from all subjects. Quantitative MRI, PHI and therapeutic regimens are acquired for patients diagnosed with intermediate-to-high-grade invasive breast cancers who are eligible for NAT as a component of their clinical care at community-care oncology sites. The study population consists of women over the age of 18 years who present with primary breast cancer and are considering NAT as a component of their clinical care. Importantly, patients with disease from all subtypes and treatment regimens (including immunotherapy, targeted and cytotoxic therapies) may be included for disease stage II and III cancers.

Image acquisition ● Timing ~1.5–2 h total per session (patient arrival to departure), 40 min (patient placement and removal from scanner), ~25 min (MRI scanning)

▲ CRITICAL We recommend the scanner be reserved for 40 min to allow sufficient time for patient to be positioned in the breast coil (prone position), be scanned, and exit from the scanner. Preparation for the MRI examination is no different than a typical, standard-of-care examination, including the placement and removal of an intravenous line for administration of the contrast agent.

3   *Insert an intravenous line.* To do this, insert a short peripheral intravenous catheter (20–22 gauge) in the antecubital or forearm area. Check correct positioning of the catheter tip for venous backflow by withdrawing blood and flushing with normal saline.

4   *Establish the FOV.* Obtain several localizer scans at the beginning of the MRI session. First, obtain a localizer scan by acquiring a three-plane (axial, sagittal and coronal) scout image series covering both

breasts. Adjust the FOV to ensure coverage of the tumor in the affected breast without changing scan parameters that would affect resolution or scan time. As all patients have been diagnosed with breast cancer and tumor location is known, it is helpful to identify the approximate location of the lesion (e.g., '7 o'clock, right breast') from the referring physician's office prior to the MRI to aid FOV placement. Often, a metal-induced susceptibility artifact around the biopsy clip/marker can be used to locate the approximate tumor center on the localizer scans. While finding the center of the FOV, adjust the coverage in the anterior–posterior direction to include the axillary artery whenever possible (to allow for characterization of the AIF). Acquire a second localizer scan with a sagittal-only sequence with 30 slices, each 5 mm thick (15 cm of patient right-to-left coverage).

▲ CRITICAL STEP The imaging volume must be centered on the tumor and include all or as much of the tumor as possible. Please see Table 2 for a complete summary of the imaging parameters described in this and the following steps.

? TROUBLESHOOTING

5 *Acquire DW-MRI data.* Acquire DW-MRI over ten slices with 5 mm thickness and no slice gap. We use the following parameters: repetition time/echo time $TR/TE = 3{,}000/52$ ms, flip angle 90°, matrix 128×128 (over a 256×256 mm^2 FOV), and GeneRalized Autocalibrating Partial Parallel Acquisition (GRAPPA) acceleration of 2. We include spectrally selective adiabatic inversion recovery (SPAIR) fat suppression. To allow for approximately equal SNR ratios at all b values, average 6 acquisitions for b values of 0 and 200 s/mm^2, and average 18 acquisitions for a b- value of 800 s/mm^2 for a monopolar, single-shot spin echo, echo planar imaging sequence in a 3D-diagonal diffusion-encoding direction. If using other scanners, use a sequence comparable to GRAPPA, e.g., Autocalibrating Reconstruction for Cartesian imaging (ARC) for GE and image-domain SENSitivity Encoding (SENSE) for Phillips.

? TROUBLESHOOTING

6 *Map the B_1 field.* Use the Siemens TurboFLASH sequence with a preconditioning RF pulse[102] to map the B_1 field with the following acquisition parameters: $TR/TE = 8{,}680/2$ ms, flip angle 8°, matrix 96×96, and slice thickness 5 mm. Due to the inclusion of a slice gap in the B_1 mapping protocol, perform two acquisitions with a 5 mm slice gap, to cover the same FOV as the above measurements without missing coverage in the slice direction. This pulse sequence is only available as a standard product sequence on Siemens scanners at the time of publication. When Philips or GE Healthcare scanners are used, an approximate B_1 map can be computed from assuming a uniform T_1 over adipose tissue and interpolating over the rest of the images as published previously[103,104].

? TROUBLESHOOTING

7 *Acquire a high-resolution T_1-weighted image.* Use a volumetric interpolated breath-hold examination (VIBE; no breath-holding required) sequence with the following acquisition parameters: $TR/TE = 5.3/2.3$ ms, flip angle 10°, acquisition matrix 256×256, slice thickness 1 mm (96 slices), GRAPPA acceleration of 2 in the phase encode direction, and SPAIR fat suppression. Then, acquire images for the precontrast T_1 map without fat suppression with T_1-weighted images from the 3D gradient-echo, FLASH (fast low-angle shot) sequence, also known as an SPGRE (spoiled gradient recalled echo) sequence, at ten flip angles (2°, 4°, 6°, …, 20°) with the following parameters: $TR/TE = 7.9/2.71$ ms and GRAPPA accelerating factor of 3 in the phase encode direction. Use an acquisition matrix of $192 \times 192 \times 10$ over a sagittal square FOV (256 mm^2) with slice thickness of 5 mm.

▲ CRITICAL STEP This precontrast T_1-weighted MRI scan with fat suppression is required for anatomical visualization purposes as well as for radiologists to identify the tumor ROI. The precontrast T_1 map is required for pharmacokinetic modeling of the DCE-MRI data described in the next step.

8 *Acquire a dynamic set of T_1-weighted, VIBE (no breath-holding required) images (this is the DCE-MRI acquisition).* Use the following acquisition parameters: $TR/TE = 7.02/4.6$ ms, flip angle 15°, matrix 192×192, with ten slices of 5 mm thickness each, and GRAPPA acceleration factor of 2 in the phase encode direction yielding a temporal resolution of 7.27 s. (Note that, with further developments in rapid acquisition, slices thinner than 5 mm may be accessible without sacrificing too much SNR.) The comparable sequences on GE and Philips scanners are FAME (Fast Acquisition with Multiphase EFGRE3D) and THRIVE (T1W High Resolution Isotropic Volume Examination), respectively. Start acquisition of this DCE-MRI data. After 1 min, begin to administer a gadolinium-based contrast agent via a power injector at the dosage recommended on the product insert and at the flow rate of 2 mL/s through the IV catheter placed in Step 3, all while continuing to acquire the DCE-MRI data. Follow administration of the contrast agent with a saline flush (20 mL) at a flow rate of 2 mL/s through (again) the IV catheter placed in Step 3.

▲**CRITICAL STEP** Note that different contrast agents or injection rates may lead to different pharmacokinetics. If pharmacokinetic parameters are calculated, used or compared across the patient population, it is imperative to use a consistent contrast agent and injection protocol. It is also important to record the contrast agent dosage used for reference at subsequent visits.

9 *Acquire a postcontrast, high-resolution T_1-weighted image.* Use a VIBE (no breath-holding required) sequence with the following acquisition parameters: *TR/TE* = 5.3/2.3 ms, flip angle 10°, acquisition matrix 256 × 256, slice thickness 1 mm (96 slices), GRAPPA acceleration 2, and SPAIR (Spectral Selection Attenuated Inversion Recovery) fat suppression.
   ▲**CRITICAL STEP** A postcontrast, T_1-weighted MRI scan with fat suppression is required for anatomical visualization purposes as well as for radiologists to identify the tumor ROI.

Image analysis ● Timing ~2–3 h

10 *Upload and save data.* Save as un-anonymized DICOM files onto the Picture Archiving Communication System (PACS) server associated with the radiology practice that is acquiring the scans. Save a separate copy of the DICOM files for which PHI is replaced with a unique subject identifier, and upload these anonymized files onto a firewalled protected server.
   ■**PAUSE POINT** Data analysis can be performed at a later date. For each patient, we advise that the images be processed and analyzed as a set across all visits. See Fig. 5 for a flowchart of all steps in the data analysis pipeline.

11 *Organize data.* Copy DICOM data into individual patient visit directories to ensure original copies of the data taken off the scanner are preserved. Import DICOM data into the MATLAB workspace using built-in MATLAB functions *dicomread* and *dicominfo*. Arrange DICOM slices in ascending order of slice position within the scanner (i.e., not acquisition order but spatial order; slices collected in an interleaved fashion should be reordered by spatial location). Save the final image for each type of MRI scan as a matrix in a MATLAB structure, along with the header information for each DICOM (pulled in using *dicominfo*). Save each scan's structure in floating point precision as .mat files to integrate with the pipeline's image processing and analysis steps. Save the DW-MRI, variable flip angle T_1-weighted and DCE-MRI data as 4D matrices where the fourth dimension represents each *b* value, flip angle and repetition, respectively. Save the final MATLAB structure to each patient/visit directory.
   ■**PAUSE POINT** Data analysis can continue at a later date.

12 *Check the slice positions.* Use the DICOM header information so that each scan is aligned with those of the DCE-MRI data to ensure the same FOV is being analyzed across scans for each patient and visit.
   ▲**CRITICAL STEP** At the scanner, several types of errors can occur, including a shifted FOV (compared with previous scans) or slices offset by a value different than the 5 mm slice thickness.
   ■**PAUSE POINT** Data analysis can continue at a later date.
   ? TROUBLESHOOTING

13 *Up-sample the DW-MRI and B_1 map data to match the DCE-MRI spatial resolution.* Use a nearest-neighbor interpolation for 2D gridded data (*interp2*, MATLAB). Save the resulting interpolated data to the .mat file in the MATLAB structure for that patient and visit.
   ▲**CRITICAL STEP** In this protocol, the B_1 map and DW-MRI data are acquired at a lower spatial resolution than that of the DCE-MRI data (due to time constraints and SNR considerations). These scans must match the resolution of the target image for successful intravisit registration.
   ■**PAUSE POINT** Data analysis can continue at a later date.

14 *Align DW-MRI data to DCE-MRI data.* Use the $b = 0$ image (as the SNR is higher for the $b = 0$ s/mm^2 data compared with the $b = 200$ or 800 s/mm^2 data). Register each slice of the $b = 0$ s/mm^2 diffusion-weighted scan to the corresponding slice of the first repetition of DCE-MRI data using the MATLAB rigid registration algorithm with function *imregtform*. Apply the transformation obtained from the function's output to the $b = 200$ and 800 s/mm^2 DW-MRI data using the MATLAB *imwarp* function. Save the resulting 4D matrix of DW-MRI data registered to the DCE-MRI data as a separate .mat file.
   ■**PAUSE POINT** Data analysis can continue at a later date.

15 *Align the variable flip angle T_1-weighted MRI to the DCE-MRI data.* Repeat Step 14, but use the variable flip angle T_1-weighted MRI data in place of the DW-MRI data. Register each slice of each T_1-weighted image to the first repetition of DCE-MRI data (*imregtform, imwarp*, MATLAB). Save the variable flip angle data as a 4D matrix, and then export as a .mat file.
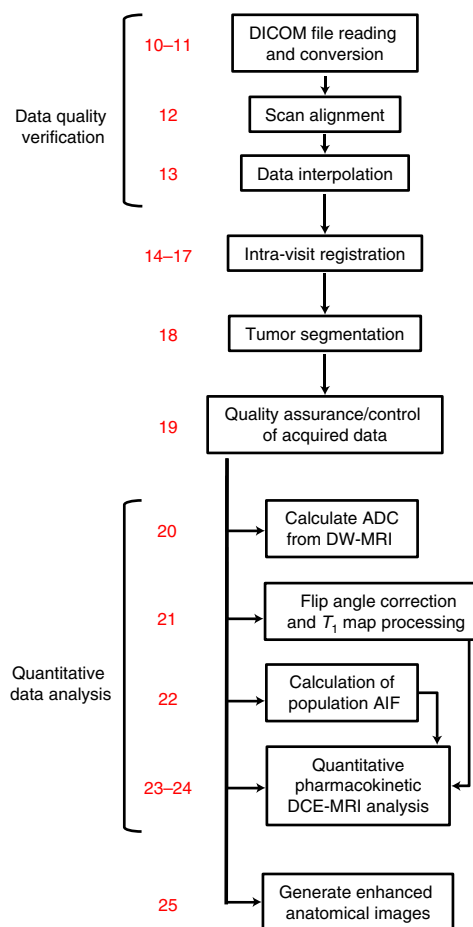   ■**PAUSE POINT** Data analysis can continue at a later date.

Fig. 5 | Flowchart of the data analysis steps of the protocol. Step numbers are listed in red, and step names are listed in black. While the early data analysis steps (10–19) rely on the completion of the previous steps, several of the latter steps (20–24) can be performed in parallel.

16 *Align B₁ maps.* The B_1 mapping sequence utilized for this study outputs a proton-density-weighted image and a calculated map of the estimated flip angle each voxel in the image experienced. To align the B_1 maps to the DCE-MRI data, register each slice of the proton-density-weighted image to the first repetition of DCE-MRI data (*imregtform*, *imwarp*, MATLAB—see Step 14) owing to the higher SNR of the proton-density-weighted image compared with the flip angle map. Apply the resulting geometric transformation to the flip angle map. Save the resulting data, and export as a .mat file.
    ■ **PAUSE POINT** Data analysis can continue at a later date.

17 *Correct for motion within the DCE-MRI scan.* Align each slice of each repetition of the DCE-MRI sequence to the corresponding slice of the first repetition (*imregtform*, *imwarp*, MATLAB—see Step 14). Save the resulting data as a 4D matrix, and export as a .mat file.
    ▲ **CRITICAL STEP** Breast motion due to respiration or patient movement can occur during the DCE-MRI acquisition.
    ■ **PAUSE POINT** Data analysis can continue at a later date.

18 *Generate tumor ROIs.* Manually draw a bulk ROI (this must be done by a board-certified radiologist) to conservatively outline the lesion under investigation. Then, apply FCM clustering to the voxels within the drawn ROI with the class number set to 2 (one each for the lesion and nonlesion tissue types) using MATLAB's *fcm* function. After the binary mask is identified, postprocess to fill holes (i.e., zeros in the mask that are surrounded by ones) via the MATLAB function *imfill*. Additionally, eliminate regions <8.45 mm^3 (i.e., $1 \times 1 \times 1$ voxels) via the MATLAB function *bwareaopen*. Save the resulting ROI to the .mat file in the MATLAB structure for that patient and visit.
    ▲ **CRITICAL STEP** The following lesion segmentation process was adapted from ref. [62].
    ■ **PAUSE POINT** Data analysis can continue at a later date.
    ? **TROUBLESHOOTING**

19 *Ensure registration is successful, and identify any artifacts (due to silicone implants, cardiac motion, etc.).* Manually visualize each set of pre- (from Steps 11–13) and postregistration (from Steps 14–17) scans. Remove patient datasets with substantial artifacts (deforming the images) from the study. To evaluate the results of tumor segmentation, visualize ROIs by overlaying them on background-subtracted DCE-MRI data as well as high-resolution postcontrast T_1-weighted images.
   ■ **PAUSE POINT** Data analysis can continue at a later date.

20 *Calculate ADC maps.* Fitting the DW-MRI data to Eq. (4):

$$S(b) = S_0 \cdot \exp(-\mathrm{ADC} \cdot b), \tag{4}$$

where $S(b)$ is the signal intensity in the presence of diffusion gradients at strength b, S_0 is the signal intensity in the absence of diffusion gradients, and b is the strength of the diffusion gradients. Fit Eq. (4) to the signal intensities from the $b = 200$ and $b = 800$ s/mm^2 DW-MRI data on a voxel-by-voxel basis. In particular, for datasets that have only two b values, the ADC values can be calculated directly (i.e., no fitting). Alternatively, use linear regression with $ln(S(b))$ and b to determine the *ADC* map via MATLAB function *regression*. While the $b = 0$ s/mm^2 data are acquired, they are not used in the ADC calculation for this protocol—due to low b values being affected by tissue perfusion—but are used in the registration of the diffusion-weighted data (details in Step 14). Save the resulting ADC map to the .mat file in the MATLAB structure for that patient and visit.
   ■ **PAUSE POINT** Data analysis can continue at a later date.
   ? TROUBLESHOOTING

21 *Calculate the B_1-corrected T_1 values for each voxel.* Fit (*lsqcurvefit*, MATLAB) the measured multi-flip angle, signal intensity data to the gradient-recalled echo signal, S, equation:

$$S = S_0 \cdot \left[\frac{\sin(f \cdot a) \cdot (1 - \exp(-\mathrm{TR}/T_1))}{1 - (\exp(-\mathrm{TR}/T_1) \cdot \cos(f \cdot a))} \right] \exp\left(-TE/T_2^*\right), \tag{5}$$

where S_0 is a constant related to scanner gain and proton density, α is the prescribed set of flip angles, and f is the flip angle correction factor for a given voxel that accounts for inhomogeneity in B_1 ($TE << T_2^*$ is assumed, so $\exp(-TE/T_2^*)$ can be set to 1). The flip angle correction factor (f) is calculated by dividing the B_1 map by the flip angle (a). The fitting process yields the flip-angle-corrected T_1 map and S_0. Save the resulting corrected T_1 map to the .mat file in the MATLAB structure for that patient and visit.
   ■ **PAUSE POINT** Data analysis can continue at a later date.

22 *Calculate the population AIF.* For each patient's dataset, generate a subtraction image from the contrast-enhanced data where the average precontrast dynamics are subtracted from the average postcontrast dynamics. For each scan, identify the axillary artery in the subtraction images. Manually select a single voxel within the axillary artery as a 'seed' for a 3 × 3 kernel centered on the seed. Save the location of the seed in a vector. For an adjacent slice, select a 5 × 5 ROI centered on the voxel corresponding to the seed voxel location. For each voxel within the 5 × 5 ROI, define a 3 × 3 window centered on each voxel. Compare the average signal intensity time courses of each window (25 windows) to that of the seed voxel's kernel using the correlation coefficient (MATLAB's function *corr*). For the voxel window with the greatest correlation to the seed, save the location of the voxel in the vector and set the voxel as the new seed. Repeat this process through all of the slices. For the voxels saved in the vector, remove any voxels with the following criteria:
   • the maximum signal intensity does not occur within the first ~45 s of the postcontrast dynamics
   • the maximum signal intensity is approximately <20 times greater than the standard deviation of the first three precontrast dynamics
   • the average signal intensity of the last ~120 s is approximately >40% of the maximum
   For the voxels that remain in the vector after the three bulleted steps, average them at each time point to yield one single, average time course. Save the resulting time course as the individual AIF to the .mat file in the MATLAB structure for that patient and visit. To generate the population AIF, average the resulting individual AIFs across all patient sets available.
   ▲ **CRITICAL STEP** This calculation of the population AIF makes use of the semiautomatic algorithm developed by Li et al.[63]. Calculation of the population AIF can be updated with each acquisition of new patient data[67].
   ■ **PAUSE POINT** Data analysis can continue at a later date.
   ? TROUBLESHOOTING

23 *Determine the BAT.* Fit the signal intensity time course from each tumor voxel (using MATLAB's *lsqcurvefit*) to the half-logistic function

$$S_{hl}(t) = \begin{cases} 0, & 0 < t \leq \tau_{AT} \\ \frac{2A}{1 + \exp(-B(t - \tau_{BAT}))} - A, & t > \tau_{AT} \end{cases}, \tag{6}$$

where A and B control the amplitude and slope of the uptake portion of the $S_{hl}(t)$ time course, respectively, and τ_{BAT} is the BAT at a specific voxel location. Temporally align the population AIF curve for each voxel with:

$$C_p(t) = \begin{cases} 0, & 0 < t \leq delay_{pop} \\ AIF(t - delay_{pop}), & t > delay_{pop} \end{cases}, \tag{7}$$

where $delay_{pop} = \tau_{BAT} - \tau_{pop}$ and τ_{pop} is the BAT of the population AIF[61].

24 *Convert the measured signal intensity for the DCE-MRI experiment to the concentration of contrast agent to enable the pharmacokinetic modeling described below.* The signal intensity measured in the DCE-MRI experiment is described by Eq. (5), where we note that $T_1 \equiv T_1(t)$ is how the measured T_1 value at time t changes due to the concentration of contrast agent according to:

$$1/T_1(t) = r_1 C_t(t) + 1/T_{10}, \tag{8}$$

where r_1 is the relaxivity constant specific to the contrast agent, and T_{10} is the native T_1 (from the precontrast T_1 map obtained from Step 21). Finally, $C_t(t)$ is described using the BAT-shifted AIF (i.e., $C_p(t)$ from Eq. (7)) and fitting each voxel's concentration time course to the extended Kety–Tofts model,

$$C_t(\overline{x}, t) = K^{trans}(\overline{x}) \int_0^t C_p(u) \exp(-(K^{trans}(\overline{x})/v_e(\overline{x})(t - u))) du + v_p(\overline{x}) C_p(t), \tag{9}$$

where $C_t(\overline{x},t)$ and $C_p(t)$ are the contrast agent concentrations in the tissue and plasma, respectively, at position x and time t; $K^{trans}(\overline{x})$ is the volume transfer constant of contrast agent from the plasma to the tissue space at position x, $v_e(\overline{x})$ is the extravascular extracellular volume fraction at position $\overline{x}$, and $v_p(\overline{x})$ is the plasma volume fraction at position $\overline{x}$. The measured signal intensity time course for each voxel within the tumor ROI can now be fit by this nested set of three equations (i.e., Eq. (9) substituted into Eq. (8), which is then substituted into Eq. (5)) and $C_p(t)$ from Eq. (7), using *lsqcurvefit* (MATLAB). This fitting routine provides a value of K^{trans}, v_e and v_p at each voxel position, $\overline{x}$, within the tumor ROI. Calculate the efflux constant, $k_{ep}(\overline{x})$ as $K^{trans}(\overline{x})/v_e(\overline{x})$. Save the resulting four pharmacokinetic parameter maps to the .mat file in the MATLAB structure for that patient and visit.
    ■ **PAUSE POINT** Data analysis can continue at a later date.
    ? TROUBLESHOOTING

25 *Generate enhanced anatomical images.* Enhance contrast by applying a local-statistics-based transfer function to each voxel of a subtraction anatomical image (average precontrast images minus average postcontrast images)[105]. Specifically, use the transfer function

$$I_{enh} = \begin{cases} I_c - \sqrt{I_c^2 - I_{orig}^2}, & 0 < X \leq X_c \\ I_c - \sqrt{(I_{max} - I_c)^2 - (I_{max} - I_{orig})^2}, & X_c < X \leq X_{max} \end{cases} \tag{10}$$

where I_{org} is the original image, I_{max} the maximum intensity in the original image, and I_{enh} the resulting enhanced image. This transfer function is S-shaped, so I_c is the inflection point defined as the 95th percentile of the intensity. This strategy for histogram normalization ensures that foreground enhancement as well as background suppression can be achieved simultaneously. Save the resulting enhanced anatomical images to the .mat file in the MATLAB structure for that patient and visit.
    ■ **PAUSE POINT** Data analysis can continue at a later date.

Mapping the imaging data to the mathematical model ● Timing ~1 d (Steps 28–30 are the rate-limiting steps as they require the use of all patient data in a study)
▲ **CRITICAL** The following steps (i.e., Steps 26–36) describe the methods required to convert the processed MRI data into a single modeling domain and derive relevant, physical quantities for model simulations starting with intervisit registration.

26 *Manually compare the slices across the different visits evaluating each patient's anatomy to determine a rough slice alignment.* Note that this is the first-time data from different visits are compared; i.e.,

one cannot complete this step or continue to the next steps until data are acquired from at least two visits. Using enhanced images, utilize unique characteristics in the patient's anatomy (visible structures in the tissue and vessels) and the tumor-bearing slices to determine which of the ten slices for each visit correspond to the slices across all visits. Performing this initial alignment improves the ability of the registration algorithm to converge to a solution. Save the corresponding slices to the .mat file in the MATLAB structure for that patient and visit.

▲ CRITICAL STEP This is required prior to applying the registration algorithm.

■ PAUSE POINT Analysis can continue at a later date.

? TROUBLESHOOTING

27 *Convert files to .mhd files.* This can be accomplished by using MATLAB's Medical Imaging Toolbox function *write_mhd*. To use this function, save the tumor ROIs and enhanced anatomical images for the slices determined in Step 26 for registration. Additionally, define parameter files for each patient and scan to be registered (further instructions for formatting of these files can be found at http://elastix.bigr.nl/wiki/index.php/Parameter_file_database). Once these files are generated, perform registration by calling the *elastix* function, which first rigidly aligns the tumor ROIs and corresponding anatomical images followed by the b-spline registration with rigidity penalty over the tumor ROI. Once the deformation field is generated, apply the deformation field to all remaining maps and images for each patient set by calling the *transformix* function.

▲ CRITICAL STEP This is required prior to running Elastix.

28 *Explore a range of rigidity penalty weights.* To select an appropriate weight for the rigidity penalty, apply the above registration functions (Step 27) for a range of values for subset of patients, where scan 2 is the 'target' image and scan 1 is the image to be registered.

29 *Evaluate rigidity penalty weights.* Evaluate the results of the different penalty weights (from Step 28) using the metrics detailed below. Select a weight that maximizes metrics 1 and 3 while minimizing metric 2. For metric 3, divide histograms (probability density functions) of the original target and registered maps into the same 100 segments (i.e., converting the histograms into vectors), and then normalize to calculate their inner product as similarity.

- Metric 1: similarity or normalized mutual information between reference image and registered target image
- Metric 2: sum of the Jacobian determinant of the final deformation field within the tumor region
- Metric 3: consistency between the distributions of quantitative parameters before and after registration

30 *Select a rigidity penalty weight.* Compare the resulting registered images (Step 28) to identify a 'tolerant range' of the studied configuration parameter (i.e., weight of the rigidity penalty term). Within this tolerant range, the performance of registration is reasonable and similar; outside the range, the registration either fails or substantially changes values within the tumor ROI (per the metrics described in Step 29). Note that the tolerant range varies between different patients, as well as between different visits for individual patients. Therefore, choose a value that falls in the tolerant range for all the patients used in this parameter investigation.

■ PAUSE POINT Analysis can continue at a later date.

31 *Register the images.* Using the details of Step 27 and the penalty weight derived from Steps 28–30, register the enhanced images of scans 1, 3 and 4 to scan 2 (target). Apply the resulting deformation fields to all corresponding patient parameter maps. Save the resulting registered data to new .mat files (separate from previous data analysis files) in the MATLAB structure for that patient and visit.

■ PAUSE POINT Analysis can continue at a later date. The following steps describe defining and calculating modeling quantities.

? TROUBLESHOOTING

32 *Define modeling domains.* Manually draw breast ROIs to define the domain for mathematical modeling. Using the enhanced images, for each slice, carefully outline the breast region outside of the chest wall and within the skin of the organ using the MATLAB function *roipoly*. Save the outlined shapes as binary masks, with zeros outside the breast ROI. If the tumor is not near the chest wall, segmentation can simply eliminate the chest region. Save the resulting breast masks to the .mat file in the MATLAB structure for that patient and visit.

■ PAUSE POINT Analysis can continue at a later date.

? TROUBLESHOOTING

33 *Calculate tumor cells per voxel.* Using the MATLAB function *imfilter*, smooth the registered ADC maps for each slice using a Gaussian filter with size 3 × 3 voxels. Convert the ADC value for each

voxel within the tumor (as segmented using the above methods) to an estimate of the number of tumor cells per voxel at each position $\overline{x}$ and time t, $N_{TC}(\overline{x},t)$, via established methods:[20–22,24,27,72]

$$N_{\text{TC}}(\overline{x}, t) = \theta\left(\frac{\text{ADC}_w - \text{ADC}(\overline{x}, t)}{\text{ADC}_w - \text{ADC}_{\min}}\right), \tag{11}$$

where ADC_w is the ADC of free water (3×10^{-3} mm^2/s)[106], $\text{ADC}(\overline{x}, t)$ is the ADC value for the voxel at position $\overline{x}$ and time t, and $\text{ADC}_{\min}$ is the minimum (positive) ADC value over all tumor voxels for the patient across all scans[27,74]. The parameter θ is the carrying capacity describing the maximum number of tumor cells that can physically fit within a voxel; its numerical value can be determined by assuming a spherical packing density of 0.7405[107], a nominal tumor cell radius of 10 μm[108] and a voxel volume of 8.45 mm^3 (using the DCE-MRI resolution). Save the resulting tumor cell map ($N_{\text{TC}}(\overline{x}, t)$) to the .mat file in the MATLAB structure for that patient and visit.

■ **PAUSE POINT** Analysis can continue at a later date.

34 *Segment the fibroglandular and adipose tissues.* Use the enhanced imaging data and a two-class *k*-means clustering (MATLAB function *kmeans*) to generate initial masks for fibroglandular and adipose tissues[37]. To suppress noise and eliminate voxels containing both tissues, apply *k*-means clustering for a second time on the adipose region segmented by the first clustering to erode the edges. Finally, for the fibroglandular tissue mask, eliminate small 'islands' (≤10 connected voxels) using MATLAB function *bwareaopen*. Save the resulting segmentation masks to the .mat file in the MATLAB structure for that patient and visit.

■ **PAUSE POINT** Analysis can continue at a later date.

35 *Approximate drug delivery.* To approximate concentrations of drug delivered throughout the tumor tissue, utilize the derived physiological parameters from the perfusion/diffusion analysis and each patient's individual therapeutic regimen. Using Eq. (9) (the Kety–Tofts model) whereby the DCE-MRI-derived parameters (K^{trans}, v_e and v_p) are available for each voxel (from the analysis described in Step 24), replace the $C_p(t)$ term in Eq. (9) with the concentration of drug in the plasma from measured population curves for each drug the patient received (for example, see refs. [109–112]) with repeated doses according to each patient's specific therapeutic regimen. Save the resulting drug distributions as a 4D matrix (time being the fourth dimension) to the .mat file in the MATLAB structure for that patient and visit.

■ **PAUSE POINT** Analysis can continue at a later date.

? **TROUBLESHOOTING**

36 *Calculate tumor summary measures.* Calculate the total tumor cellularity by summing the number of tumor cells across all the voxels in the tumor ROI. Approximate the tumor volume as the product of the total number of voxels within the segmented tumor ROI and the voxel volume (8.45 mm^3 using the DCE-MRI spatial resolution described above in section 7). To calculate the longest axis of each tumor, evaluate the 3D tumor ROI using the function *regionprops3* within MATLAB, which approximates the longest possible axis within a 3D object. These measures of cellularity, volume and longest axis are to be applied to all of the model's predictions to enable direct comparison with the clinically measured data. Note that implementing this process of image segmentation and longest diameter calculation makes the evaluation as rigorous as possible.

■ **PAUSE POINT** Mathematical modeling can continue at a later date. The following steps provide the methods for implementing a mathematical model to utilize the above patient-specific, MRI-derived quantities to generate individual patient response predictions.

Tumor Forecasting ● Timing ~10 h per patient, can be up to several days depending on tumor size

37 *Implement the model (i.e., Eqs. (1–3)) in 3D.* Use a fully explicit finite difference scheme with central difference in space and forward difference in time. With voxel dimensions defined by the size of the DCE-MRI voxel grid, a maximum diffusion coefficient of 0.001 mm^2/d, a time step of $\Delta t = 0.25$ d will ensure numerical stability. Set the size of the computational domain by a square whose dimensions are determined by the size of the breast domain for each patient. Assign the tissue stiffnesses in Table 1 to the tumor, fibroglandular and adipose tissue ROIs for the mechanical coupling. Apply a no-flux boundary condition for the tumor cells and zero tissue displacement (i.e., $\overrightarrow{u} = 0$ at the *x*-boundary) on the breast domain boundary. For computational efficiency, solve the mechanical coupling (Eq. (3)) to the diffusion (Eq. (2)) every 20 timesteps. For additional details on

implementing these numerical methods, see ref. [50]. Please also see the troubleshooting table for additional details to ensure simulation accuracy.

■ **PAUSE POINT** Model calibration can continue at a later date.

? TROUBLESHOOTING

38 *Calibrate model parameters for each patient.* Utilize two of the MRI datasets for each patient to calibrate the mathematical model and then simulate the model to a later scan or the time of surgery to make a prediction of tumor response. For example, the datasets from visits 1 and 2 are used for calibration to the first therapeutic regimen, thereby enabling simulating the model and predicting the measured response of the tumors at the time of visit 3. Similarly, use datasets from visits 3 and 4 for calibration to the second therapeutic regimen to simulate the model and predict the response of tumors at the time of surgery (as determined by pathology). See Fig. 2 for an illustration of this calibration and prediction strategy.

Use the cellularity maps of the two scans for each patient (derived from the ADC of the DW-MRI data—see Steps 20 and 33) to calibrate model parameters, where the tumor from the earlier imaging visit initializes the calibration (with the corresponding tissue maps for mechanical coupling and distributions of drugs) and the later tumor ROI is the target for calibration. Calibrate the D_0 and α parameters globally, and the k parameter map spatially, where the remaining parameters are assigned to the literature values in Table 1 and θ is directly calculated (Step 33). Utilize a Levenberg–Marquardt least-squares, nonlinear optimization for calibration[113], where the sum of squared errors between the simulated tumor cell numbers from the model and the calculated number of tumor cells from the imaging data is minimized with the following parameters: maximum number of iterations 200, initial lambda 10^{-20}, and lambda increment factors 9 and 11, and perform the Jacobian calculation every 25 iterations. Additionally, bound the global parameters D_0 and α to be greater than zero, and bound $D_0 < 0.001$ mm^2/d. Stopping criteria are the sum of the squared errors < 0.001, concordance correlation coefficient of 1 between the model simulation and target distribution of tumor cells, and/or after the maximum number of iterations has been realized. See ref. [50] for additional details on the implementation of the Levenberg–Marquardt method for model calibration.

■ **PAUSE POINT** Tumor forecasting can continue at a later date.

? TROUBLESHOOTING

39 *Assess uncertainty.* For the following predictions and evaluations, uncertainty related to the calibrated parameters and data measurements must be considered. Choose three representative datasets from the cohort. Using the calibrated parameters, simulate the model from scan 1 to scan 2 (target scan for calibration) for each tumor. Add an appropriate range of noise (determined by, for example, repeatability studies for DW-MRI data[37]) to voxels in each tumor cell map of the model generated scan 2 results using a normal distribution (MATLAB function *randn*). Calibrate the model to the noisy scan 2 tumor cell map, and save the resulting parameter values. Repeat for a total of 100 sets for each tumor in the three representative datasets ($N = 300$ total). Calculate the percent difference for each parameter between the corresponding original parameter values and each of the results from fitting the noisy data. Calculate the 95% confidence interval for each parameter using all samples. Determine whether a uniform or normal distribution is representative of the resulting parameter difference distributions.

For each patient in the cohort, generate 50 random parameter sets by sampling distributions of the parameters defined by the calculated 95% confidence intervals centered about each patient's calibrated parameter set (use MATLAB's *rand* for uniform distributions or *randn* for normally distributed values). Simulate the model from scan 2 to scan 3 for each of the 50 randomized parameter sets for each patient. Calculate the 95% confidence interval across all 50 resulting tumor predictions for each patient for total cellularity, volume and longest axis. Therefore, simulation results will include confidence intervals based on the uncertainty in the parameter estimates.

40 *Forecast tumor response.* Using the corresponding calibrated parameters and maps (tissues, tumor cellularity, drug distributions) from scan 2, simulate the model from the time of scan 2 to scan 3. Evaluate the resulting 3D prediction with measured tumor response using the summary measures (total tumor cellularity, total tumor volume and longest axis, Step 36). Additionally, compare the predicted tumor and measured scan 3 tumor with the Dice coefficient (measuring overlap of the predicted ROI and measured ROI; a Dice of zero indicates no overlap, whereas a Dice of 1 indicates perfect overlap), and/or the concordance correlation coefficient to directly compare the prediction to the measurement for each patient. The predicted percent change from baseline to scan 3 can be compared with actual response defined by RECIST.

Using the corresponding calibrated parameters and maps (tissue, cellularity, drug distribution) from scan 4, simulate the model from the time of scan 4 to scan surgery. Compare the resulting summary measures and their corresponding predicted percent changes with surgically defined responses for the cohort (e.g., pCR versus non-pCR groups) using appropriate statistical comparison test for the cohort size (i.e., t tests, Wilcoxon rank sum test, Kendall and Pearson correlation coefficients, etc.).

Troubleshooting

Troubleshooting advice can be found in Table 3.

Table 3 | Troubleshooting table

Step	Problem	Possible reason	Solution
4	Unable to identify the tumor	Breast tissue anatomy and tumor sizes/shapes vary widely between patients, and it may be very difficult to identify the tumor	Biopsy clips and radiological notes are frequently helpful for identifying the tumor location. Biopsy clips create a dark signal void surrounded by brighter tissue signal and are easily spotted
5	Low SNR	The resolution of the image may be too fine	Boost SNR by choosing a lower resolution or regridding the image post acquisition
6	When a product sequence for measuring B_1 is not available	Scanner is not equipped with a B_1-mapping sequence	After an initial uncorrected T_1 map is made using the variable-flip angle sequence approach, use thresholding to select only those voxels that most likely belong to adipose fat. Based on the assumption that T_1 values should be a consistent value for these voxels, compute an estimate of the flip angle error by assuming the acquired map includes B_1 error. Once the B_1 map is filled in for adipose fat voxels, use spatial interpolation to estimate B_1 at other (non-fat) voxels. This correction, while suboptimal to performing a true B_1 map, should serve to reduce the magnitude of flip angle error when B_1 mapping sequences are not available on the study scanner[103,104]
12	A certain scan acquisition contains more slices than included for the DCE-MRI data	This often happens during scanning for variable flip angle T_1 mapping data, the high-resolution T_1-weighted data	Only utilize the aligned slices or do a 3D interpolation to realign the images for downstream analysis, and save the data in a new structure
18	Too much or too little tumor results from segmentation	FCM algorithms have a threshold that can be tuned for belonging to different partitions	Adjust the membership threshold in the FCM algorithm
20	Unable to derive expected values for ADC	Using low b values (<100 s/mm^2) can be affected by tissue perfusion, thus altering the interpretation of the image	Including more than two b values can improve the accuracy of parameter estimation
		Poor data fitting	May need to use interpolation methods to populate ADC values for voxels with bad fits
21	Signal drop in the AIF	Standard-of-care examinations may administer high concentrations of contrast agent, which may induce T_2^* effects in the signal intensity curves measured from the large vessels (i.e., those typically employed for AIF estimation)	May need to assume an analytical form of the concentration of the contrast agent in the plasma and fit to the average signal intensity time course[67]
22	Unphysiological values from the pharmacokinetic model fits K^{trans}, v_e, v_p, $k_{ep} < 0$, v_e, $v_p > 1$, and/or K^{trans}, $k_{ep} > 5$	Could be in regions of necrosis where the Kety–Tofts model does not account for such phenomena[114]	Bound calibration method or filtered-out values beyond the physiological range from downstream analysis (may need to use interpolation methods to populate filtered-out voxels)
26	Alignment of tumors/anatomy is off by several slices	Can occur despite best efforts to center the FOV on the center of the tumor during scanning—due to the tumor dramatically changing in size and shape during the course of therapy	Additional blank slices may be needed to 'pad' the ends of any particular scan set to achieve alignment and maintain equal dimensions across datasets

Table continued

Table 3 (continued)

Step	Problem	Possible reason	Solution
31	Intervisit registration failure	Despite careful initial alignment of slices and selection of the tumor ROI weight factor, the registration can fail (especially for tumors that have greatly reduced in size by the time of scan 4)	Additional manipulations may be required, including trimming the moving images to remove large amounts of image outside the breast and/or remove large amounts of the image that includes the chest cavity. Even with these additional corrections, registration can fail or simply be unable to align tumor ROIs. In these cases, simple rigid registration or manual alignment may be necessary. The MATLAB functions *imregtform* and *imwarp* can be used for rigid registration
32	Unusual breast domain	Some breasts contain implants near the tumor ROI that would not be a part of the modeling domain	For breasts containing implants, two domains can be drawn: the first containing the entire breast region and the second where the breast implant is segmented out. The first breast mask can be used for image/display purposes, but the mathematical model simulates on the second domain as we assume the tumor cannot grow into the breast implant
35	DCE-MRI data do not have the requisite temporal resolution for pharmacokinetic analysis	We note that certain MRI protocols do not allow for the requisite temporal resolution for pharmacokinetic analysis	In these cases, to approximate a normalized drug distribution in each voxel of tissue, a map of the blood volume can be calculated by computing the area under the dynamic curve (AUC) of the baseline-subtracted time course for each voxel and normalized by the AIF AUC value (as an approximation of blood volume) from contrast-enhanced data[25]. The signal enhancement ratio[71] may also be used in this way to estimate drug distribution
37	Numerically unstable results	This can be due to several numerical scheme errors	Check that the diffusion coefficient, grid size and time step selected maintain numerical stability (see ref. [50] for stability equations specific to this scheme). Verify boundary conditions where tumor cells are conserved and not lost when proliferation is zero and diffusion is large
38	Calibration is computationally expensive	Calibration can take >1 d per patient when using a spatially defined proliferation map for large tumors and domains	To reduce computation time for all calibrations, the voxel matrix within the designated domain can be down-sampled. Parallelize codes to further reduce computational cost by utilizing multiple available processors concurrently on a server. Focus calibration on a smaller domain containing the tumor ROI and not over the whole breast. Broyden's method can also be used to approximate the update of the Jacobian for each iteration
	Calibration will not converge	Using maps from the previous scan may limit the calibration scheme to hit the target (e.g., using scan 1 maps to calibrate to scan 2). Different tumors may be more difficult for the calibration to converge	Can use the starting number of tumor cells and shape from scan 1 (scan 3) with the tissue and drug distribution maps of scan 2 (scan 4) for calibration. The lambda increment factors increased or decreased are often problem-specific and need to be empirically determined to improve convergence. Larger tumors may require more iterations

Timing

All times are approximate.

Step 1, patient accrual: highly variable depending on local method

Step 2, patient consent: 15 min

Step 3, placement of IV line: 10 min

Step 4, FOV determination for MRI examination: 2 min

Step 5, DW-MRI data acquisition: 1 min 39 s

Step 6, B_1 map data acquisition: 34 s

Step 7, pre-contrast high-resolution T_1-weighted scan, and variable flip angle T_1-weighted images: 3 min 13 s and 1 min 39 s, respectively

Step 8, DCE-MRI data acquisition: 8 min

Step 9, precontrast high-resolution T_1-weighted scan: 3 min 13 s

Step 10, upload to PACS: 5–10 min

Step 11, DICOM reading and conversion: 6 min

Step 12, checking of slice positions using the DICOM header information: 1 min

Step 13, upsampling of the DW-MRI and B_1 map data to the resolution of the DCE-MRI data: 3–4 min

Step 14, alignment of the DW-MRI data to the DCE-MRI data: 1 min

Step 15, alignment of the variable flip angle T_1-weighted images to the DCE-MRI data: 7 min

Step 16, alignment of the B_1 data to the DCE-MRI data: 1 min

Step 17, motion correction within the DCE-MRI data: 9 min

Step 18, tumor segmentation: 20 min

Step 19, quality control of the acquired data: 5 min

Step 20, calculation of the *ADC* maps: 4 min

Step 21, computation of the B_1-corrected T_1 maps: 22 min

Step 22, construction of a patient's AIF: 5 min

Step 23, BAT determination for each voxel in the DCE-MRI dataset: 5 min

Step 24, pharmacokinetic analysis for each voxel in the DCE-MRI data: 25 min

Step 25, enhancement of the anatomical images: 2 min

Step 26, comparison of slices across all visits for each patient: 10 min

Step 27, conversion of file types to .mhd: 1 min

Steps 28–30, identification of longitudinal registration weighting for all patient data from all patient visits: 1 d

Step 31, alignment of parameter maps from all visits for each patient: 15 min

Step 32, selection of domain for mathematical modeling for each patient set (i.e., all scan sessions from one patient): 5–10 min

Step 33, conversion of ADC maps to estimates of cell number: 1 min

Step 34, segmentation of breast tissue into fibroglandular and adipose tissues: 1 min

Step 35, estimation of spatiotemporal distribution of drug concentration: 1 min

Step 36, calculatation of the estimate of tumor cells, total tumor volume and longest axis: 1 min

Step 37, simulation of the model forward in space and time: 5 min

Step 38, calibration of the model to the patient data: several hours to 2 d, depending on the domain size

Step 39, assessment of uncertainty in model predictions: 1 d

Step 40, evaluation of predictive ability: 5 min

Anticipated results

In the following sections, we describe and present illustrative results associated with processing one dataset through the entire protocol. The example dataset used is that of a breast cancer patient with triple negative (estrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 negative) invasive ductal carcinoma in the left breast. At the time of the first imaging session, the patient was 59 years of age with a body mass index of 28.3 kg/m^2. All four scans were acquired over a period of 6 months during which the patient was treated with NAT. Surgery determined that the patient had residual disease (i.e., a non-pCR outcome) at the conclusion of NAT.

Image acquisition

See Fig. 6 for examples of the resulting MRI data collected for the patient's first scan from the data acquisition steps. Figure 6a shows the lower *b*-value diffusion-weighted image (see Step 5), while Fig. 6b presents the B_1 map quantifying the difference between the prescribed flip angle and that actually experienced by each voxel (Step 6). Figure 6c shows a single 10° flip angle image from the multi-flip angle data acquired to estimate the T_1 map (Step 7), and Fig. 6d displays the average of all images acquired during the DCE-MRI sequence (Step 8). Note the differences between the various MRI modalities; in particular, observe how the DCE-MRI data (Fig. 6d) allow for visualization of the tumor and other tissue structures.

Data processing and analysis

See Fig. 7 for how the example images in Fig. 6 are processed using the protocol (Steps 10–25) to identify various properties of the patient's tumor. In particular, the ADC map is derived from the DW-MRI data acquisitions (Step 20), the tumor ROI is defined using the DCE-MRI data and FCM algorithm (Step 18), a corrected T_1 map is generated from the variable flip angle T_1 images and the B_1 map (Step 21), and the resulting Kety–Tofts parameters characterizing the vasculature within the tumor are derived utilizing the DCE-MRI data, tumor ROI and corrected T_1 map (Step 23).
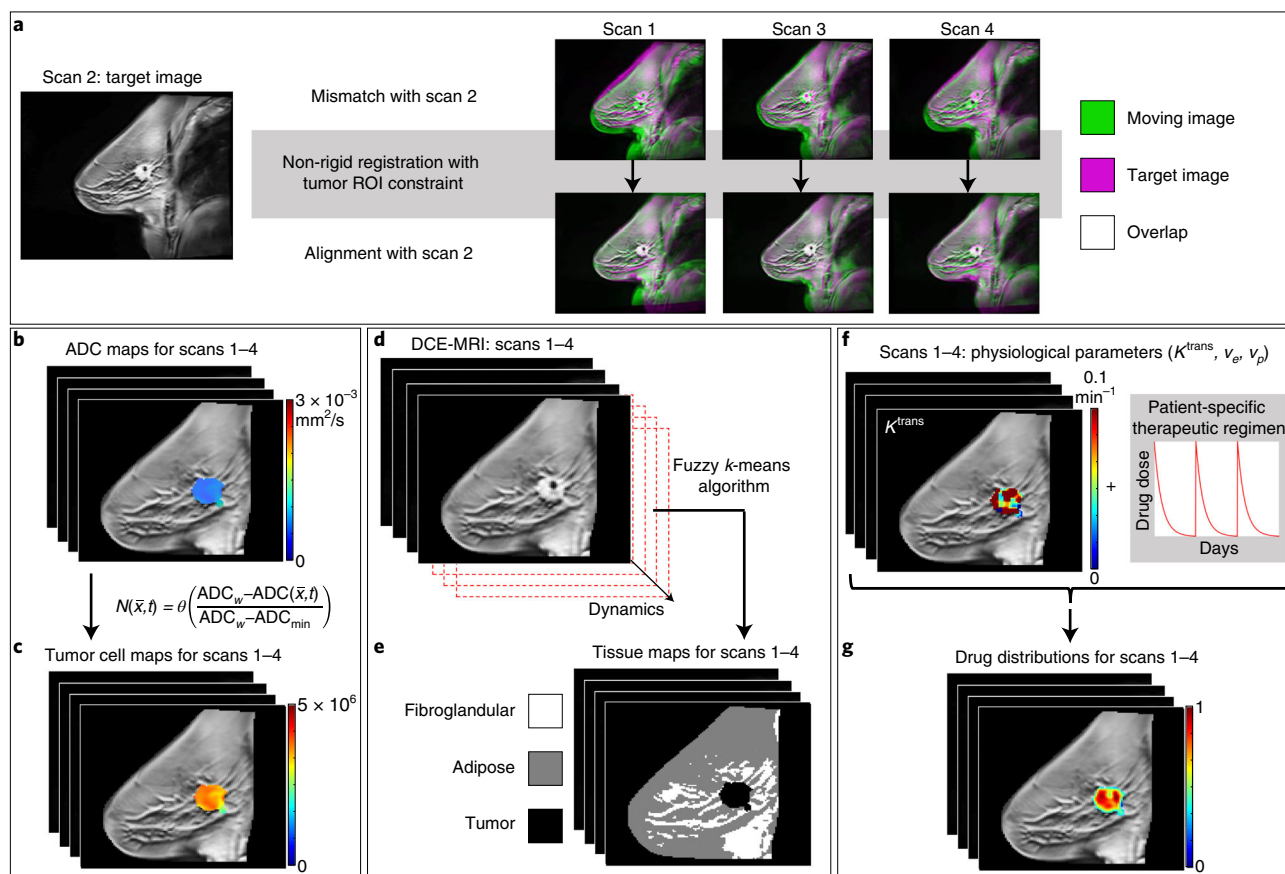
Fig. 6 | Example image acquisition results. a–d, A central slice for an illustrative patient depicting the 200 s/mm^2 b value from DW-MRI (**a**), the flip angle (FA) ratio from the B_1 map (**b**), the 10° T_1-weighted acquisitions (**c**) and the average signal intensity for the DCE-MRI data across all dynamics (**d**). The tumor burden is indicated with the red box. Informed consent was obtained from this patient.

Fig. 7 | Example results from the data analysis. a–h, Each of the MRI data is aligned, interpolated to the same resolution, and registered across visits using a rigid registration algorithm. The DW-MRI data are used to calculate the *ADC* map (**a**, **e**). The DCE-MRI data are used to identify the tumor ROIs (**b**, **f**). The multi-flip angle (MFA) T_1 scans and the B_1 map correction are used to calculate a T_1 map (**c**, **d** and **h**, respectively). The DCE-MRI data (**b**) along with the T_1 map (**h**) are used to calculate the Kety–Tofts model parameters (**g**) within the tumor ROI (from **f**). Informed consent was obtained from this patient.

Mapping imaging data to the mathematical model

See Fig. 8 for example intermediary and subsequent images for the steps that convert the quantitative data maps to quantities utilized within the mathematical modeling system. Here, for example, images for all four scans can be visualized in the intervisit registration Fig. 8a (Steps 26–31). After intervisit registration, a modeling domain is identified over the breast (Step 32), ADC values within the tumor ROI are converted to tumor cellularity (Step 33), masks for the fibroglandular and adipose tissues are generated using a *k*-means clustering algorithm (Step 34), and drug distribution in the patient's tissue is approximated using the Kety–Tofts model and plasma concentration curves of the patient's therapeutic regimen (Step 35). For this patient, the resulting summary measures described in Step 36 for visits 1–4 are: total cellularity (cells) of 1.53×10^9, 1.42×10^9, 1.13×10^9 and 1.20×10^9; volumes (mm^3) of 14,462, 11,831, 5,813 and 4,551; and longest axes (mm) of 36, 34, 32 and 29, respectively.

Fig. 8 | Converting imaging data to physical quantities for the mathematical model. a–g, Prior to deriving modeling quantities, intervisit registration is required to align the images across all visits (**a**; details provided in Fig. 4). Once aligned, the ADC maps (**b**) are used to calculate the tumor cellularity (**c**). DCE-MRI data (**d**) are used to identify fibroglandular and adipose tissues (**e**) using a fuzzy *k*-means algorithm. The Kety–Tofts model parameters, specific to each patient, are used along with each patient's individual therapeutic regimen (**f**) to derive approximate drug distributions in the tumor tissue (**g**). Informed consent was obtained from this patient.

Tumor forecasting

See Fig. 9 for an example comparison of the prediction from the mathematical model and the experimentally measured data for three central slices at the time of scan 3 for the same patient data presented in Figs. 6–8. Specifically, the mathematical model was calibrated using the patient's data from her first two scans, and then, with the resulting patient-specific parameters, the model was simulated forward in time from scan 2 to the time of scan 3 (as described in Steps 37–40). Note that the model is able to predict the patient's tumor response with errors <17% for the three summary tumor measures (total cellularity, volume and longest axis) and has strong statistical correlation with the shape and cellular densities of the tumor (see figure caption for more details).

Data availability

We have made available one patient dataset that has been fully preprocessed (i.e., Steps 1–36) and ready for the calibration and prediction components (i.e., Steps 37–40). This will enable the interested investigator to verify that the calibration and prediction code is working on the individual investigator's platform. The dataset is available at https://github.com/ChengyueWu/Quantitative-MRI-of-breast-cancer-patients-to-forecast-response-to-therapy.

Code availability

We have made the code for the calibration and prediction components (i.e., Steps 37–40) available without charge to anyone for academic, research, experimental or personal use. This code and license
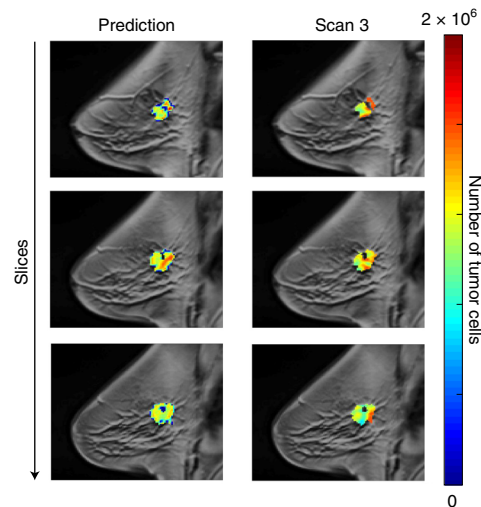
Fig. 9 | Results of the 3D model predictions (over three central slices, left column) compared with the observed results at the third scan time (right column) for one example patient. The number of tumor cells is indicated by the color overlay on each anatomical image. Notice that the model captures the correct shape of the tumor (Dice coefficient = 0.79), and while areas of higher and lower cellularity may not directly match, there are similar scales of the cellular densities between the model's predictions and the patient's actual tumor, resulting in Pearson and concordance correlation coefficients of 0.80 and 0.78, respectively. Comparing the summary measures of the predicted and measured tumor at scan 3 yields 13%, 17% and 4% difference in total cellularity, volume and longest axis, respectively. Informed consent was obtained from this patient.

may be found at https://github.com/ChengyueWu/Quantitative-MRI-of-breast-cancer-patients-to-forecast-response-to-therapy. To distribute or make other use of the software, including commercial use, a license must be obtained from The University of Texas at Austin (by contacting licensing@otc.utexas.edu).

References

1. Yankeelov, T. E. et al. Toward a science of tumor forecasting for clinical oncology. *Cancer Res.* **75**, 918–923 (2015).
2. Yankeelov, T. E. et al. Clinically relevant modeling of tumor growth and treatment response. *Sci. Transl. Med.* **5**, 187ps9 (2013).
3. Rockne, R. C. et al. The 2019 mathematical oncology roadmap. *Phys. Biol.* **16**, 041005 (2019).
4. Alfonso, J. C. L. et al. The biology and mathematical modelling of glioma invasion: a review. *J. R. Soc. Interface* **14**, 20170490 (2017).
5. Chen, X., Summers, R. M. & Yao, J. Kidney tumor growth prediction by coupling reaction-diffusion and biomechanical model. *IEEE Trans. Biomed. Eng.* **60**, 169–173 (2013).
6. Lorenzo, G. et al. Tissue-scale, personalized modeling and simulation of prostate cancer growth. *Proc. Natl Acad. Sci. USA* **113**, E7663–E7671 (2016).
7. Lorenzo, G. et al. Computer simulations suggest that prostate enlargement due to benign prostatic hyperplasia mechanically impedes prostate cancer growth. *Proc. Natl Acad. Sci. USA* **116**, 1152–1161 (2019).
8. Yuan, J. & Liu, L. Brain glioma growth model using reaction-diffusion equation with viscous stress tensor on brain MR images. *Magn. Reson. Imaging* **34**, 114–119 (2016).
9. Rockne, R. et al. Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: a mathematical modeling approach. *Phys. Med. Biol.* **55**, 3271–3285 (2010).
10. Clatz, O. et al. Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation. *IEEE Trans. Med. Imaging* **24**, 1334–1346 (2005).
11. Baldock, A. L. et al. From patient-specific mathematical neuro-oncology to precision medicine. *Front. Oncol.* **3**, 62 (2013).
12. Mi, H. et al. Prediction of lung tumor evolution during radiotherapy in individual patients with PET. *IEEE Trans. Med. Imaging* **33**, 995–1003 (2014).
13. Mi, H. et al. Joint tumor growth prediction and tumor segmentation on therapeutic follow-up PET images. *Med. Image Anal.* **23**, 84–91 (2015).
14. Liu, Y. X. et al. Patient specific tumor growth prediction using multimodal images. *Med. Image Anal.* **18**, 555–566 (2014).

15. Wong, K. C. et al. Pancreatic tumor growth prediction with elastic-growth decomposition, image-derived motion, and FDM-FEM coupling. *IEEE Trans. Med. Imaging* **36**, 111–123 (2017).

16. Liu, Y. et al. Multimodal image driven patient specific tumor growth modeling. *Med. Image Comput. Comput. Assist. Interv.* **16**, 283–290 (2013).

17. Wong, K. C. et al. Tumor growth prediction with reaction-diffusion and hyperelastic biomechanical model by physiological data fusion. *Med. Image Anal.* **25**, 72–85 (2015).

18. Atuegwu, N. C. et al. Incorporation of diffusion-weighted magnetic resonance imaging data into a simple mathematical model of tumor growth. *Phys. Med. Biol.* **57**, 225–240 (2012).

19. Atuegwu, N. C., Gore, J. C. & Yankeelov, T. E. The integration of quantitative multi-modality imaging data into mathematical models of tumors. *Phys. Med. Biol.* **55**, 2429–2449 (2010).

20. Atuegwu, N. C. et al. Parameterizing the logistic model of tumor growth by DW-MRI and DCE-MRI data to predict treatment response and changes in breast cancer cellularity during neoadjuvant chemotherapy. *Transl. Oncol.* **6**, 256–264 (2013).

21. Weis, J. A. et al. A mechanically coupled reaction-diffusion model for predicting the response of breast tumors to neoadjuvant chemotherapy. *Phys. Med. Biol.* **58**, 5851–5866 (2013).

22. Weis, J. A. et al. Predicting the response of breast cancer to neoadjuvant therapy using a mechanically coupled reaction-diffusion model. *Cancer Res.* **75**, 4697–4707 (2015).

23. Weis, J. A., Miga, M. I. & Yankeelov, T. E. Three-dimensional image-based mechanical modeling for predicting the response of breast cancer to neoadjuvant therapy. *Comput. Methods Appl. Mech. Eng.* **314**, 494–512 (2017).

24. Jarrett, A. M. et al. Incorporating drug delivery into an imaging-driven, mechanics-coupled reaction diffusion model for predicting the response of breast cancer to neoadjuvant chemotherapy: theory and preliminary clinical results. *Phys. Med. Biol.* **63**, 105015 (2018).

25. Jarrett, A. M. et al. Towards integration of ^{64}Cu-DOTA-Trasztusumab PET-CT and MRI with mathematical modeling to predict response to neoadjuvant therapy in HER2+ breast cancer. *Sci. Rep.* **10**, 20518 (2020).

26. Jarrett, A. M. et al. Evaluating patient-specific neoadjuvant regimens for breast cancer via a mathematical model constrained by quantitative magnetic resonance imaging data. *Neoplasia* **22**, 820–830 (2020).

27. Atuegwu, N. C. et al. Integration of diffusion-weighted MRI data and a simple mathematical model to predict breast tumor cellularity during neoadjuvant chemotherapy. *Magn. Reson. Med.* **66**, 1689–1696 (2011).

28. Jarrett, A. M. et al. Optimal control theory for personalized therapeutic regimens in oncology: background, history, challenges, and opportunities. *J. Clin. Med.* **9**, 1314 (2020).

29. Copur, M. S. et al. Impact of the National Cancer Institute Community Cancer Centers Program on clinical trial and related activities at a community cancer center in rural Nebraska. *J Oncol. Pract.* **12**, 67–68 (2016). e44-51.

30. Hormuth, D. A. et al. Mechanism-based modeling of tumor growth and treatment response constrained by multiparametric imaging data. *JCO Clin. Cancer Inform.* **3**, 1–10 (2019).

31. Yankeelov, T. E., Pickens, D. R. & Price, R. R. *Quantitative MRI in Cancer* (CRC Press, 2012).

32. Huang, W. et al. Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. *Transl. Oncol.* **7**, 153–166 (2014).

33. Bane, O. et al. Accuracy, repeatability, and interplatform reproducibility of T. *Magn. Reson. Med.* **79**, 2564–2575 (2018).

34. Newitt, D. C. et al. Multisite concordance of apparent diffusion coefficient measurements across the NCI Quantitative Imaging Network. *J. Med. Imaging (Bellingham)* **5**, 011003 (2018).

35. Bell, L. C. et al. Evaluating multisite rCBV consistency from DSC-MRI imaging protocols and post-processing software across the NCI quantitative imaging network sites using a digital reference object (DRO). *Tomography* **5**, 110–117 (2019).

36. Yankeelov, T. E. et al. Quantitative imaging in cancer clinical trials. *Clin. Cancer Res.* **22**, 284–290 (2016).

37. Sorace, A. G. et al. Repeatability, reproducibility, and accuracy of quantitative MRI of the breast in the community radiology setting. *J. Magn. Reson. Imaging* https://doi.org/10.1002/jmri.26011 (2018).

38. Virostko, J. et al. Magnetization transfer MRI of breast cancer in the community setting: reproducibility and preliminary results in neoadjuvant therapy. *Tomography* **5**, 44–52 (2019).

39. Jones, K. M., Pagel, M. D. & Cárdenas-Rodríguez, J. Linearization improves the repeatability of quantitative dynamic contrast-enhanced MRI. *Magn. Reson. Imaging* **47**, 16–24 (2018).

40. Li, X. et al. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. *Invest. Radiol.* **50**, 195–204 (2015).

41. Kim, Y. et al. Early prediction of response to neoadjuvant chemotherapy using dynamic contrast-enhanced MRI and ultrasound in breast cancer. *Korean J. Radiol.* **19**, 682–691 (2018).

42. Ah-See, M. L. et al. Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer. *Clin. Cancer Res.* **14**, 6580–6589 (2008).

43. Padhani, A. R. et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia* **11**, 102–125 (2009).

44. Galbán, C. J. et al. Multi-site clinical evaluation of DW-MRI as a treatment response metric for breast cancer patients undergoing neoadjuvant chemotherapy. *PLoS One* **10**, e0122151 (2015).

45. Partridge, S. C. et al. Diffusion-weighted MRI findings predict pathologic response in neoadjuvant treatment of breast cancer: the ACRIN 6698 multicenter trial. *Radiology* **289**, 618–627 (2018).

46. McKnight, A. L. et al. MR elastography of breast cancer: preliminary results. *AJR Am. J. Roentgenol.* **178**, 1411–1417 (2002).

47. Hormuth, D. A. et al. A mechanically coupled reaction-diffusion model that incorporates intra-tumoural heterogeneity to predict in vivo glioma growth. *J. R. Soc. Interface* **14**, 20161010 (2017).

48. Hormuth, D. A. II et al. Biophysical modeling of in vivo glioma response following whole brain radio-therapy in a murine model of brain cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **100**, 1270–1279 (2018).

49. Hormuth, D. A. et al. Predicting in vivo glioma growth with the reaction diffusion equation constrained by quantitative magnetic resonance imaging data. *Phys. Biol.* **12**, 046006 (2015).

50. Hormuth, D. A. et al. Mechanically coupled reaction-diffusion model to predict glioma growth: metho-dological details. *Methods Mol. Biol.* **1711**, 225–241 (2018).

51. Hormuth, D. II et al. Predicting in vivo tumor growth and angiogenesis with an MRI calibrated biophysical model. *Neuro-Oncol.* **19**, vi23 (2017).

52. Swanson, K. R., Alvord, E. C. & Murray, J. D. Quantifying efficacy of chemotherapy of brain tumors with homogeneous and heterogeneous drug delivery. *Acta Biotheor.* **50**, 223–237 (2002).

53. Kim, M., Gillies, R. J. & Rejniak, K. A. Current advances in mathematical modeling of anti-cancer drug penetration into tumor tissues. *Front. Oncol.* **3**, 278 (2013).

54. Owen, M. R. et al. Mathematical modeling predicts synergistic antitumor effects of combining a macro-phage-based, hypoxia-targeted gene therapy with chemotherapy. *Cancer Res.* **71**, 2826–2837 (2011).

55. Shah, A. B., Rejniak, K. A. & Gevertz, J. L. Limiting the development of anti-cancer drug resistance in a spatial model of micrometastases. *Math. Biosci. Eng.* **13**, 1185–1206 (2016).

56. Bernstein, M. A., King, K. F. & Zhou, X. J. *Handbook of MRI Pulse Sequences* (Elsevier, 2004).

57. Nishimura, D. G. *Principles of Magnetic Resonance Imaging* (Stanford University, 1996).

58. Whisenant, J. G. et al. Bloch-Siegert B1-mapping improves accuracy and precision of longitudinal relaxation measurements in the breast at 3 T. *Tomography* **2**, 250–259 (2016).

59. Yankeelov, T. E. & Gore, J. C. Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples. *Curr. Med. Imaging Rev.* **3**, 91–107 (2009).

60. Shukla-Dave, A. et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J. Magn. Reson. Imaging* **49**, e101–e121 (2019).

61. Wu, C. et al. Quantitative analysis of vascular properties derived from ultrafast DCE-MRI to discriminate malignant and benign breast tumors. *Magn. Reson. Med.* **81**, 2147–2160 (2019).

62. Chen, W., Giger, M. L. & Bick, U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Acad. Radiol.* **13**, 63–72 (2006).

63. Li, X. et al. Validation of an algorithm for the nonrigid registration of longitudinal breast MR images using realistic phantoms. *Med. Phys.* **37**, 2541–2552 (2010).

64. Gubern-Mérida, A. et al. Automated localization of breast cancer in DCE-MRI. *Med. Image Anal.* **20**, 265–274 (2015).

65. Staring, M., Klein, S. & Pluim, J. P. A rigidity penalty term for nonrigid registration. *Med. Phys.* **34**, 4098–4108 (2007).

66. Li, X. et al. Statistical comparison of dynamic contrast-enhanced MRI pharmacokinetic models in human breast cancer. *Magn. Reson. Med.* **68**, 261–271 (2012).

67. Li, X. et al. A novel AIF tracking method and comparison of DCE-MRI parameters using individual and population-based AIFs in human breast cancer. *Phys. Med. Biol.* **56**, 5753–5769 (2011).

68. Yankeelov, T. E. et al. Quantitative pharmacokinetic analysis of DCE-MRI data without an arterial input function: a reference region model. *Magn. Reson. Imaging* **23**, 519–529 (2005).

69. Li, K. L. et al. Kinetic assessment of breast tumors using high spatial resolution signal enhancement ratio (SER) imaging. *Magn. Reson. Med.* **58**, 572–581 (2007).

70. Sorace, A. G. et al. Distinguishing benign and malignant breast tumors: preliminary comparison of kinetic modeling approaches using multi-institutional dynamic contrast-enhanced MRI data from the International Breast MR Consortium 6883 trial. *J. Med. Imaging (Bellingham)* **5**, 011019 (2018).

71. Li, K. L. et al. Invasive breast cancer: predicting disease recurrence by using high-spatial-resolution signal enhancement ratio imaging. *Radiology* **248**, 79–87 (2008).

72. Whisenant, J. G. et al. Assessing reproducibility of diffusion-weighted magnetic resonance imaging studies in a murine model of HER2+ breast cancer. *Magn. Reson. Imaging* **32**, 245–249 (2014).

73. Sugahara, T. et al. Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. *J. Magn. Reson. Imaging* **9**, 53–60 (1999).

74. Anderson, A. W. et al. Effects of cell volume fraction changes on apparent diffusion in human cells. *Magn. Reson. Imaging* **18**, 689–695 (2000).

75. Guo, Y. et al. Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging. *J. Magn. Reson. Imaging* **16**, 172–178 (2002).

76. Humphries, P. D. et al. Tumors in pediatric patients at diffusion-weighted MR imaging: apparent diffusion coefficient and tumor cellularity. *Radiology* **245**, 848–854 (2007).

77. Barnes, S. L. et al. Correlation of tumor characteristics derived from DCE-MRI and DW-MRI with histology in murine models of breast cancer. *NMR Biomed.* **28**, 1345–1356 (2015).

78. Latour, L. L. et al. Time-dependent diffusion of water in a biological model system. *Proc. Natl Acad. Sci. USA* **91**, 1229–1233 (1994).

79. van der Toorn, A. et al. Dynamic changes in water ADC, energy metabolism, extracellular space volume, and tortuosity in neonatal rat brain during global ischemia. *Magn. Reson. Med.* **36**, 52–60 (1996).

80. Xu, J. et al. Magnetic resonance imaging of mean cell size in human breast tumors. *Magn. Reson. Med.* **83**, 2002–2014 (2020).

81. Reynaud, O. et al. Surface-to-volume ratio mapping of tumor microstructure using oscillating gradient diffusion weighted imaging. *Magn. Reson. Med.* **76**, 237–247 (2016).

82. Bedair, R. et al. Assessment of early treatment response to neoadjuvant chemotherapy in breast cancer using non-mono-exponential diffusion models: a feasibility study comparing the baseline and mid-treatment MRI examinations. *Eur. Radiol.* **27**, 2726–2736 (2017).

83. Virostko, J. et al. The rate of breast fibroglandular enhancement during dynamic contrast-enhanced MRI reflects response to neoadjuvant therapy. *Eur. J. Radiol.* **136**, 109534 (2021).

84. Hormuth, D. A. et al. Abstract 5487: a biologically-motivated mathematical model for forecasting chemoradiationresponse in high grade gliomas: initial results. *Cancer Res.* **80**, 5487 (2020).

85. Barnes, S.L. et al. DCE- and DW-MRI as early imaging biomarkers of treatment response in a preclinical model of triple negative breast cancer. *NMR Biomed.* https://doi.org/10.1002/nbm.3799 (2017).

86. Whisenant, J. G. et al. Evaluating treatment response using DW-MRI and DCE-MRI in trastuzumab responsive and resistant HER2-overexpressing human breast cancer xenografts. *Transl. Oncol.* **7**, 768–779 (2014).

87. Hormuth, D. A. et al. Calibrating a predictive model of tumor growth and angiogenesis with quantitative MRI. *Ann. Biomed. Eng.* **47**, 1539–1551 (2019).

88. Hormuth, D. A., Jarrett, A. M. & Yankeelov, T. E. Forecasting tumor and vasculature response dynamics to radiation therapy via image based mathematical modeling. *Radiat. Oncol.* **15**, 4 (2020).

89. Valdora, F. et al. Rapid review: radiomics and breast cancer. *Breast Cancer Res. Treat.* **169**, 217–229 (2018).

90. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: images are more than pictures, they are data. *Radiology* **278**, 563–577 (2016).

91. Gillies, R. J. & Balagurunathan, Y. Perfusion MR imaging of breast cancer: insights using "habitat imaging". *Radiology* **288**, 36–37 (2018).

92. Syed, A. K. et al. Multiparametric analysis of longitudinal quantitative MRI data to identify distinct tumor habitats in preclinical models of breast cancer. *Cancers (Basel)* **12**, 1682 (2020).

93. Onishi, N. et al. Ultrafast dynamic contrast-enhanced mri of the breast using compressed sensing: breast cancer diagnosis based on separate visualization of breast arteries and veins. *J. Magn. Reson. Imaging* **47**, 97–104 (2018).

94. Onishi, N. et al. Ultrafast dynamic contrast-enhanced breast MRI may generate prognostic imaging markers of breast cancer. *Breast Cancer Res.* **22**, 58 (2020).

95. Jarrett, A. M. et al. Mathematical models of tumor cell proliferation: a review of the literature. *Expert Rev. Anticancer Ther.* **18**, 1271–1286 (2018).

96. Harris, P. A. et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).

97. Harris, P. A. et al. The REDCap consortium: building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).

98. Klein, S. et al. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* **29**, 196–205 (2010).

99. Li, X. et al. Analyzing spatial heterogeneity in DCE- and DW-MRI parametric maps to optimize prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Transl. Oncol.* **7**, 14–22 (2014).

100. Li, X. et al. A nonrigid registration algorithm for longitudinal breast MR images and the analysis of breast tumor response. *Magn. Reson. Imaging* **27**, 1258–1270 (2009).

101. Atuegwu, N. C. et al. Longitudinal, intermodality registration of quantitative breast PET and MRI data acquired before and during neoadjuvant chemotherapy: preliminary results. *Med. Phys.* **41**, 052302 (2014).

102. Chung, S. et al. Rapid B1+ mapping using a preconditioning RF pulse with TurboFLASH readout. *Magn. Reson. Med.* **64**, 439–446 (2010).

103. Sung, K. et al. Simultaneous T(1) and B(1) (+) mapping using reference region variable flip angle imaging. *Magn. Reson. Med.* **70**, 954–961 (2013).

104. Pineda, F. D. et al. B1 and T1 mapping of the breast with a reference tissue method. *Magn. Reson. Med.* **75**, 1565–1573 (2016).

105. Yu, Z. and C. Bajaj. A fast and adaptive method for image contrast enhancement. in *International Conference on Image Processing* (IEEE) **2**, 1001–1004 (2004).

106. Hagmann, P. et al. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics* **26**, S205–S223 (2006).

107. Martin, I. et al. Computer-based technique for cell aggregation analysis and cell aggregation in in vitro chondrogenesis. *Cytometry* **28**, 141–146 (1997).

108. Das, C. M. et al. Dielectrophoretic segregation of different human cell types on microscope slides. *Anal. Chem.* **77**, 2708–2719 (2005).

109. Ohtsu, T. et al. Clinical pharmacokinetics and pharmacodynamics of paclitaxel: a 3-hour infusion versus a 24-hour infusion. *Clin. Cancer Res.* **1**, 599–606 (1995).

110. Juma, F. D., Rogers, H. J. & Trounce, J. R. Pharmacokinetics of cyclophosphamide and alkylating activity in man after intravenous and oral administration. *Br. J. Clin. Pharmacol.* **8**, 209–217 (1979).
111. Green, R. F. et al. Plasma pharmacokinetics of adriamycin and adriamycinol: implications for the design of in vitro experiments and treatment protocols. *Cancer Res.* **43**, 3417–3421 (1983).
112. Oguri, S. et al. Clinical pharmacokinetics of carboplatin. *J. Clin. Pharmacol.* **28**, 208–215 (1988).
113. Marquardt, D. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1**, 431–441 (1963).
114. Woodall, R. T. et al. The effects of intravoxel contrast agent diffusion on the analysis of DCE-MRI data in realistic tissue domains. *Magn. Reson. Med.* **80**, 330–340 (2018).

Author contributions

T.E.Y. conceived the project and obtained funding. J.C.D., J.V., A.G.S. and T.E.Y developed the procedures for the MRI acquisition with input from D.P., B.G. and S.A. A.S.K., C.W., D.A.H. and D.A.E. developed procedures for the data processing with input from A.M.J., J. V., A.G.S. and T.E.Y. A.M.J., C.W. and D.A.H. developed the procedures for generating modeling quantities with input from A.S.K. and T.E.Y. A.M.J., D.A.H. and T.E.Y. developed the procedures for tumor forecasting. A.M.J., J.C.D., J.V. and A.G.S. collected the data. A.M.J., A.S.K. and T.E.Y. organized the manuscript. All authors reviewed and edited the manuscript and approved the final draft.

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/s41596-021-00617-y.

Correspondence and requests for materials should be addressed to Thomas E. Yankeelov.

Peer review information *Nature Protocols* thanks Lihua Li and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

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.

Related links
Key references using this protocol
Weis, J. et al. A. *Cancer Res.* **74**, 4697–707 (2015): https://doi.org/10.1158/0008-5472.CAN-14-2945
Jarrett, A. et al. *Neoplasia.* **22**, 820–830 (2020): https://doi.org/10.1016/j.neo.2020.10.011