HHS Public Access

Author manuscript

Cancer Res. Author manuscript; available in PMC 2023 March 16.

MRI-based digital models forecast patient-specific treatment responses to neoadjuvant chemotherapy in triple-negative breast cancer

Chengyue Wu[1,*], **Angela M. Jarrett**[1,2], **Zijian Zhou**[3], **Nabil Elshafeey**[4], **Beatriz E. Adrada**[5], **Rosalind P. Candelaria**[5], **Rania M.M. Mohamed**[5], **Medine Boge**[5], **Lei Huo**[6], **Jason B. White**[7], **Debu Tripathy**[7], **Vicente Valero**[7], **Jennifer K. Litton**[7], **Clinton Yam**[7], **Jong Bum Son**[3], **Jingfei Ma**[3], **Gaiane M. Rauch**[4,5], **Thomas E. Yankeelov**[1,2,3,8,9,10]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin. Austin, Texas 78712

[2]Livestrong Cancer Institutes, The University of Texas at Austin. Austin, Texas 78712

[3]Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

[4]Department of Abdominal Imaging, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

[5]Department of Breast Imaging, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

[6]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

[7]Department of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

[8]Department of Biomedical Engineering, The University of Texas at Austin. Austin, Texas 78712

[9]Department of Diagnostic Medicine, The University of Texas at Austin. Austin, Texas 78712

[10]Department of Oncology, The University of Texas at Austin. Austin, Texas 78712

Abstract

Triple-negative breast cancer (TNBC) is persistently refractory to therapy, and methods to improve targeting and evaluation of responses to therapy in this disease are needed. Here, we integrate quantitative magnetic resonance imaging (MRI) data with biologically-based mathematical modeling to accurately predict the response of TNBC to neoadjuvant systemic therapy (NAST) on an individual basis. Specifically, 56 TNBC patients enrolled in the ARTEMIS trial (NCT02276443) underwent standard-of-care doxorubicin/cyclophosphamide (A/C) and then paclitaxel for NAST, where dynamic contrast-enhanced MRI and diffusion-weighted MRI were

[*]**Corresponding author:** Chengyue Wu, Ph.D., Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 E 24th St, Austin, TX 78712, cw35926@utexas.edu.

Conflict of interest disclosure statement: The authors declare no potential conflicts of interest.

acquired before treatment and after two and four cycles of A/C. A biologically-based model was established to characterize tumor cell movement, proliferation, and treatment-induced cell death. Two evaluation frameworks were investigated using: 1) images acquired before and after two cycles of A/C for calibration and predicting tumor status after A/C, and 2) images acquired before, after two cycles, and after four cycles of A/C for calibration and predicting response following NAST. For Framework 1, the concordance correlation coefficients between the predicted and measured patient-specific, post-A/C changes in tumor cellularity and volume were 0.95 and 0.94, respectively. For Framework 2, the biologically-based model achieved an area under the receiver operator characteristic curve of 0.89 (sensitivity/specificity = 0.72/0.95) for differentiating pathological complete response (pCR) from non-pCR, which is statistically superior ($P < 0.05$) to the value of 0.78 (sensitivity/specificity = 0.72/0.79) achieved by tumor volume measured after four cycles of A/C. Overall, this model successfully captured patient-specific, spatiotemporal dynamics of TNBC response to NAST, providing highly accurate predictions of NAST response.

Keywords

Cancer prediction; computational oncology; personalized medicine; data-driven; longitudinal multiparametric imaging

Introduction

Neoadjuvant systemic therapy (NAST) is widely considered the standard-of-care for treatment of stage II-III, locally advanced triple-negative breast cancer (TNBC). NAST increases the success rate for breast conservation surgery by reducing tumor burden and provides the opportunity to treat micrometastases in a naïve state, thereby improving progression-free survival of patients (1, 2). Importantly, TNBC patients who achieve a pathological complete response (pCR) in the neoadjuvant setting have a favorable recurrence-free survival; in contrast, patients who have residual disease after NAST are at increased risk of early recurrence and death (3, 4). Unfortunately, based on recent pooled analysis of 52 studies from 1999 to 2016, only 32.6% of TNBC patients treated with standard taxane/anthracycline-based NAST have a pCR or minimal residual disease at the time of surgical resection (4).

Development of novel neoadjuvant treatment regimens has provided opportunities to tailor treatment for individual patients to improve outcomes for TNBC (5–7). Thus, it is becoming increasingly important to develop techniques which can accurately predict the individual TNBC patient's response to NASTs. If it could be definitively determined that a therapeutic regimen is unlikely to achieve pCR for a patient, then risk-adapted therapy could be adopted, with addition of rationally-based treatments or removal of unnecessary components, potentially improving outcomes and lowering side effects. In particular, the importance of being able to remove patients from ineffective therapies, as early as possible, is difficult to overstate given their significant toxicities (8), including increased likelihood of hospitalization, cardiac damage, leukemia, and even death (9). Moreover, accurately predicting response to NAST would enable identification of exceptional responders who

might benefit from treatment de-escalation (10), including the possibility of non-surgical management of their disease (11).

Numerous efforts have been devoted to investigating approaches that can accurately distinguish pCR from non-pCR patients early in NAST. Imaging biomarkers derived from magnetic resonance imaging (MRI), positron emission tomography, and ultrasound imaging have been shown to be strongly correlated with the response of breast tumors to NAST (12–14). In particular, MRI measurements before and during NAST have been valuable predictors of pCR, especially the functional tumor volume (FTV) derived from dynamic contrast-enhanced (DCE-) MRI and the apparent diffusion coefficient (ADC) derived from diffusion-weighted (DW-) MRI (15–18). More recently, the methods of artificial intelligence have been used to extract features from high-dimensional data to build predictive models for differentiating pCR from non-pCR in breast cancer (19, 20). However, the majority of these approaches for predicting or assessing response have the intrinsic limitation of being *population-based*. Importantly, *population-based* approaches, which rely exclusively on statistical inference from properties of large populations, inevitably obscure conditions specific to the individual patient over time (21), especially for a disease as heterogeneous as cancer (22). Conversely, biologically-based models employing patient-specific data have the potential to shift the paradigm from *population*- to *individual*-based approaches (7). Furthermore, biologically-based mathematical modeling of tumor response can not only predict the changes in global metrics summarizing tumor burden (e.g., total tumor volume and cellularity), but also reveal biologically specific information (e.g., spatially-resolved maps of proliferation, pharmacokinetics, and each patient's sensitivity to the administered therapies). It promises unique opportunities to characterize tumor pathophysiology, to rigorously forecast long-term outcomes, and even optimize treatment plans on a patient-specific basis (23, 24).

In this contribution, we develop a clinical-computational approach to establish patient-specific models to make early predictions on the spatiotemporal development and response of individual TNBC patients to standard-of-care NAST. As this model can represent a physical object (i.e., tumor), predict the behavior of the object given influences (i.e., treatments), and enable decision-making to optimize the future behavior of the object (i.e., improve the treatment outcomes) (25, 26), we posit that our methodology represents a practical manifestation of digital twins in oncology. The approach requires no *population*-based training of models; instead, it integrates an *individual* patient's multiparametric MRI data obtained at multiple time points during their treatment (Fig. 1A) with biologically-based mathematical modeling. In particular, two frameworks were constructed to determine the predictive utility of each patient's digital twin (Fig. 1B). In Framework 1, we seek to employ digital twins to predict the outcome of a single NAST regimen (i.e., A/C); specifically, we evaluate the accuracy of digital twins to predict global metrics related to change in tumor burden, as well as spatiotemporally resolved tumor dynamics at the end of A/C (Fig. 1C). In Framework 2, we seek to employ digital twins to predict the outcome of the entire NAST (i.e., both A/C and paclitaxel); specifically, we evaluate the accuracy of digital twins to predict individual pCR or non-pCR status at the completion of NAST (Fig. 1D).

Materials and Methods

Patients and MRI Data

Treatment-naïve patients with biopsy-confirmed TNBC were enrolled in the prospective, Institutional Review Board–approved clinical trial, "A Robust TNBC Evaluation FraMework to Improve Survival" (ARTEMIS, NCT02276442). Among patients who signed informed consent between June 2018 to January 2020, those who had clinical stage I–III disease, had completed NAST with a complete series of MRI scans, and had known post-surgical pathology were included in this study ($n = 56$).

Each patient underwent multiparametric MRI acquisitions before treatment (baseline/V1), after two cycles (V2), and after four cycles (V3) of the standard-of-care (neoadjuvant) doxorubicin and cyclophosphamide (A/C). (Each "cycle" of A/C is 2 weeks; see Fig. 1A.) Patients with disease progression, or <70% reduction in tumor volume at the end of A/C (27), were offered the opportunity to enroll in a biomarker-guided clinical trial using targeted bio/chemotherapy to complete therapy ($n = 9$). Patients not meeting the criteria for suboptimal response to A/C were recommended to continue standard-of-care paclitaxel weekly for 12 cycles ($n = 43$), or double-dosed every 3 weeks for four cycles ($n = 4$). (The exact paclitaxel regimen for each patient's was determined by their physician.) All patients received surgery after NAST. The post-surgical pathology was used to classify patients as pCR or non-pCR; pCR was defined as the absence of residual invasive and *in situ* cancer on hematoxylin and eosin evaluation of the complete resected breast specimen and all sampled regional lymph nodes following completion of NAST (28).

MRI was performed on a GE Discovery MR750 or MR750w whole-body scanner (GE Healthcare) with an eight-channel bilateral breast coil. In particular, the DCE-MRI data was acquired using a 3D DISCO sequence with bipolar readouts (29) (Fig. 2A) and the following scan parameters: field-of-view = 30×30 cm^2, matrix size = 320×320, slice thickness/spacing = 3.2/−1.6 mm, number of slices = 140, flip angle = 12°, repetition/echo time = 8/2 ms. After one pre-contrast phase was obtained, a single bolus of contrast agent (Gadovist, Bayer HealthCare) was injected (0.1 mL/kg at ~ 2 mL/second followed by saline flush) at the start of the post-contrast acquisition. The temporal resolution of the DISCO series ranged from 8 to 15.5 s (median = 11 s), depending on the slice coverage, which resulted in the number of post-contrast phases varying from 32 to 64 (18). The DW-MRI was acquired using a 2D spin-echo sequence of the diseased breast with the following scan parameters: field-of-view = 16×16 cm^2, matrix size = 80×80, slice thickness/spacing = 4/0 mm, number of slices = 16, flip angle = 90°, repetition/echo time = 4000/70 ms. The *b*-values used were 100 and 800 s/mm^2. The apparent diffusion coefficient (ADC) map was calculated using a GE AW server (v3.2, GE Healthcare, Milwaukee, WI). The tumors were manually segmented by two board-certified breast radiologists with 4–12 years of experience (authors MB, RMM). Tumor segmentations were reviewed by two breast fellowship trained radiologists with 19 and 20 years of experience (authors GMR, BEA). The entire tumor volume was segmented on two early phases (60 and 150 s) of the DCE-MRI using an in-home software package developed in MATLAB (R2021b, Mathworks, Natick, MA). All segmentations were further refined using the thresholding tool

of the package to exclude non-tumor voxels that were determined by radiologists. Necrotic regions and artifacts from the biopsy clip were manually segmented and excluded by two radiologists (authors MB, RMM).

Image processing

All MRI data from each patient were processed through a pipeline that consists of three components: 1) pre-processing, 2) inter-visit registration, and 3) post-processing (Fig. 2B–D). This highly automated pipeline allows efficient processing of multi-visit, multi-parametric MRI with minimal user input.

First, the multiparametric images are co-localized to the same imaging grid to align slices and voxel locations. Specifically, DCE-MRI collected bilaterally were trimmed to the DW-MRI field-of-view covering the diseased breast, and the slices of DW-MRI and ADC maps were linearly interpolated to match the slice locations of DCE-MRI. A rigid registration was applied on the DCE-MRI to align all phases in one scan to the pre-contrast phase (MATLAB function, *imregtform*). A rigid registration was applied between the co-localized DCE-MRI and DW-MRI data to remove the small mismatches between the image volumes.

Second, an inter-visit image registration was performed to account for the change of breast tissue shape and patient position across MRI visits. Specifically, the registration was performed to align the images from V1 and V3 to the images from V2. The algorithm consisted of a rigid registration of the tumor ROIs for initial alignment, followed by a deformable b-spline, non-rigid registration on the whole breast with a rigidity penalty on the tumor regions (30–32). This rigidity penalty was imposed to preserve the tumor volume and shape across all visits. This registration was developed based on MATLAB and an open-source, command-line software, Elastix (33).

Third, the post-processing was performed as preparation for the subsequent predictive modeling. Specifically, a semi-automatic segmentation of the breast contour was performed on the pre-contrast frame of DCE-MRI based on a manually chosen intensity threshold, followed by a smoothing of the segmented mask edge (MATLAB function, *imgaussfilt*). A two-class *k*-means clustering (MATLAB function, *kmeans*) was used to segment fibroglandular and adipose tissues in each pre-contrast DCE-MRI (34). The enhancement of each DCE-MRI was calculated by subtracting the pre-contrast phase from the average of the post-contrast phases. A tumor cellularity map, $N(\boldsymbol{x},t)$, was estimated based on the measured ADC map of each MRI visit (35):

$$N(\boldsymbol{x},t)=\theta\left(\frac{ADC_w-ADC(\boldsymbol{x},t)}{ADC_w-ADC_{min}}\right), \quad [1]$$

where ADC_w is the ADC of free water (3×10^{-3} mm^2/s), $ADC(\boldsymbol{x},t)$ is the ADC value for the voxel at position $\boldsymbol{x}$ and time t, and ADC_{min} is the minimal (positive) ADC value in the tumor for the patient across all visits. The carrying capacity, θ, describes the maximum number of tumor cells that can physically fit within a voxel, which was determined by assuming a spherical packing density of 0.7405 and a nominal tumor cell radius of 10 μm (30).

Image-guided biologically-based modeling

We have previously developed a biologically-based mathematical model to represent the spatiotemporally-resolved dynamics of tumor growth and response to NAST (30, 36). Specifically, a reaction-diffusion partial differential equation is used to describe the evolution of tumor cells, $N(\boldsymbol{x},t)$, in response to the therapies, as shown in Eq. [2]:

$$\frac{\partial N(\boldsymbol{x},t)}{\partial t} = \overbrace{\underbrace{\nabla\cdot\left(D(\boldsymbol{x},t)\nabla N(\boldsymbol{x},t)\right)}_{\text{Movement}}}^{\text{Tumor cell diffusion}} + \overbrace{\underbrace{k\left(x\right)\left(1-\frac{N(\boldsymbol{x},t)}{\theta}\right)N\left(\boldsymbol{x},t\right)}_{\text{Proliferation}} - \underbrace{N\left(\boldsymbol{x},t\right)\sum_i \lambda_i\left(\boldsymbol{x},t\right)}_{\text{treatment-induced death}}}^{\text{Reaction}}, \quad [2]$$

where a detailed list of variables, parameters, their definitions and assignments are given in Supplemental Table S1. The first term on the right-hand side of Eq. [2] describes the movement of tumor cells, as well as the compression of surrounding tissue, by a diffusion process coupled to tissue mechanical properties (37) *via* Eq. [3]:

$$D(\boldsymbol{x},t) = D_0 e^{-\gamma\sigma(\boldsymbol{x},t)}, \quad [3]$$

where D_0 is the tumor cell diffusion coefficient in the absence of external forces. The exponential term reduces tumor cells' mobility due to the surrounding tissue stiffness *via* the von Mises stress, $\sigma(\boldsymbol{x},t)$, and an empirical coupling constant, γ. The von Mises stress was calculated for the fibroglandular and adipose tissues within the breast, with the fibroglandular tissue assigned a greater stiffness than the adipose tissue. Technical details on the mechanical-coupled diffusion can be found in (37). The proliferation of the tumor cells (second term on right-hand side of Eq. [2]) is described by logistic growth with a spatially varying proliferating rate, $k(\boldsymbol{x})$, and a global carrying capacity, θ. The effects of administered therapies (third term on the right-hand side of Eq. [3]) is modeled as the treatment-induced death rates of tumor cells, $\lambda_i(\boldsymbol{x},t)$. The death rates are determined by the concentration of the administered drugs and the exponential decay of their efficacies:

$$\lambda_i(\boldsymbol{x},t) = \alpha_i \sum_{j=1}^{J_i} e^{-\beta_i\left(t-\tau_{i,j}\right)} C\left(\boldsymbol{x},\tau_{i,j}\right), i = 1, 2, \text{and } 3, \quad [4]$$

where α_i is the efficacy of the i^{th} administered drug, and i = 1, 2, and 3 refers to doxorubicin, cyclophosphamide, and paclitaxel, respectively. As each drug was administered multiple times during the therapeutic regimen, the total effectiveness of drugs at time t was an accumulation of all administered cycles and their decays; that is, the j^{th} administration of the i^{th} drug was conducted at time $\tau_{i,j}$ at a total of J_i times. The decay of drug efficacy from each administration is represented by the decay rate of β_i. The spatial distribution of i^{th} drugs caused by the j^{th} administration of this therapy, $C(\boldsymbol{x}, \tau_{i,j})$, is determined by the enhancement of DCE-MRI (30). Specifically, from the last DCE-MRI data collected before the injection at $\tau_{i,j}$, the area under the DCE-MRI time course is calculated at each voxel and normalized by the maximum value within the tumor. The voxel-wise normalized area

under the curve represents the map of drug concentration induced by this injection (see Supplemental Section 1.1 for details).

Eqs. [2–4] are personalized by the imaging and clinical data of each patient. Specifically, the geometry of the computational domain is determined by the segmented breast contour, tumor, and fibroglandular/adipose tissues. The tumor cellularity map at each imaging time point is determined from the ADC map obtained at the corresponding time point *via* Eq. [1]. The sequential cellularity maps are then used for patient-specific calibration of the model parameters (i.e., $k(\boldsymbol{x})$, D_0, α_i, and β_i). Additionally, the DCE-MRI acquired during NAST were used for updating the spatial distribution of the administered drugs and the mechanical properties of breast tissues. The model constrained by patient-specific data was implemented in MATLAB and solved with the finite difference method. Details of the numerical implementation can be found in (30, 36).

Digital twin frameworks

As shown in Fig. 1C, Framework 1 focuses on employing digital twins to predict the outcome of a single NAST regimen (i.e., A/C). Specifically, for each patient, the processed images from V1 and V2 are imported with the treatment regimen and calibrated to Eqs. [2–4]. The calibrated model is the digital twin as it represents patient-specific pathophysiological properties of tumor growth and response, including pre-treatment tumor shape and cellularity, the proliferation rate and mobility of tumor cells, and the efficacy and decay of administered drugs (A/C). As the efficacy and decay rates were strongly coupled and challenging to simultaneously calibrate with only two-time-points, the decay rates of A/C were randomly sampled five times from ranges found in the published literature (38): $\beta_1 \in [0.01, 0.6]$ day^{-1}, $\beta_2 \in [1.0, 5.4]$ day^{-1}. With each set of sampled β_1 and β_2, the efficacy of A/C was calibrated, resulting in five parameter sets (see Supplemental Section 2.1). By applying the remainder of the A/C (i.e., all A/C after V2; Fig. 1C) to the digital twin, the patient-specific spatiotemporally-resolved tumor cell distributions are predicted up to the end of A/C; one prediction is given by each parameter set, resulting in a median and a range of predictions. Furthermore, for each patient, global metrics—total tumor cellularity (TTC) and total tumor volume (TTV)—are derived from the spatiotemporally-resolved predictions.

The predictive accuracy of Framework 1 is evaluated both temporally and spatially. The temporal accuracy is assessed by the agreement between the predicted and measured global metrics (i.e., TTC and TTV). In particular, the concordance correlation coefficient (CCC; see Supplemental Section 1.2) is computed between the predicted and measured changes of TTC at the end of A/C; similarly, the CCC is computed between the predicted and measured changes of TTV. The spatial accuracy is assessed by the difference between the predicted and measured spatially-resolved tumor cell distributions. In particular, for each patient, the percent change of tumor cell counts from baseline (V1) to the end of A/C (V3) can be calculated at each location, $\boldsymbol{x}$. We calculate the change from predicted and measured tumor cell distributions (i.e., $\Delta\text{TCD}_\text{p}(\boldsymbol{x})$ and $\Delta\text{TCD}_\text{m}(\boldsymbol{x})$, respectively), then the spatially-resolved difference is given by $\Delta\text{TCD}_\text{p}(\boldsymbol{x}) - \Delta\text{TCD}_\text{m}(\boldsymbol{x})$. In each patient, this difference is reported as the median and interquartile range within the original tumor region; across the cohort,

the differences are evaluated by the mean and 95% confident interval (CI) of the median difference from each patient. (Considering five predictions are given for each patient, all the evaluations above are based on the median of the five predictions.)

As show in Fig. 1D, Framework 2 focuses on employing digital twins to predict the outcome of the entire NAST (i.e., both A/C and paclitaxel). Specifically, for each patient, the processed images from V1, V2, and V3 are imported into the mechanism-based model for initialization and calibration. The calibrated model is the digital twin. (Note that the three-time-point data enables calibration of the efficacy and decay rates simultaneously within the same ranges assumed in Framework 1, so the sampling scheme in Framework 1 is not needed for Framework 2.) As no imaging data were available during the paclitaxel regimen, the efficacy of paclitaxel was set to literature value, $\alpha_3 = 0.3\ \text{day}^{-1}$ (43), and the decay rate of paclitaxel was assumed as the average of calibrated A/C decay rate; $\beta_3 = (\beta_1 + \beta_2)/2$. (See Supplemental Section 2.1 for details.) By applying paclitaxel to the digital twin, it predicts the patient-specific spatiotemporally-resolved tumor cell distributions, TTC, and TTV at the end of NAST.

The output of Framework 2 is evaluated by the ability of the digital twins to differentiate pCR and non-pCR. Specifically, receiver operating characteristic (ROC) analysis is performed on the predicted TTC and TTV. We report the area under ROC curve (AUC), sensitivity (i.e., the ability to correctly identify residual tumor at final surgical pathology), and specificity (i.e., the ability to correctly identify pCR at final surgical pathology) based on the optimal cut-off. Additionally, the AUC/sensitivity/specificity from the predicted TTC and TTV were compared to those obtained by the measured TTC and TTV. The 95% CIs of the AUCs were computed and compared *via* DeLong's method (39), with $P < 0.05$ considered statistically significant.

Data availability statement

Raw data for this study were generated at the University of Texas MD Anderson Cancer Center. Raw data are not publicly available due to IRB restrictions of data containing information that could compromise research participant privacy and/or consent. The derived data that support the findings of this study are available from the corresponding author upon reasonable request.

Results

Framework 1: Patient-specific predictions of the spatiotemporal response to A/C

A cohort of 50 patients was used in Framework 1. Six patients were excluded from the entire patient cohort ($n = 56$) due to image acquisition error or artifacts ($n = 1$), unsuccessful inter-visit registration caused by a large change of breast shape between visits ($n = 1$), and a complete response of the tumor at V2 which led to no data for model calibration ($n = 4$).

Each patient's digital twin provided a range of estimated treatment efficacies of A/C (e.g., Fig. 3A–B), simulating a range of treatment outcomes (e.g., Fig. 3C–D). Specifically, Fig. 3C presents a patient who had a suboptimal response to A/C (i.e., V3 imaging showed a <70% reduction in tumor volume). The digital twin predicted a mean (range) TTC and

TTV of 3.17×10^8 (3.05×10^8 – 3.29×10^8) cells and 3.19×10^3 (3.15×10^3 – 3.23×10^3) mm^3 at V3, respectively. This corresponds to predicted decreases of 49.55% (47.60% – 51.50%) and 36.14% (35.44% – 36.84%) for TTC and TTV at V3, respectively. In comparison, the measured decreases are 38.99% and 32.58% for TTC and TTV at V3, respectively. In contrast, Fig. 3D presents a patient who had a good response to A/C (i.e., V3 imaging showed > 70% reduction in tumor volume). The digital twin predicted a TTC and TTV of 3.82×10^6 (1.69×10^6 – 5.94×10^6) cells and 12.63 (0.00 – 25.27) mm^3 at V3, respectively. This corresponds to predicted decreases of 98.00% (96.88% – 99.11%) and 99.24% (98.48% – 100.00%) for TTC and TTV at V3, respectively. In comparison, the measured decrease is 100% for both TTC and TTV at V3. Across the cohort, the CCC between the predicted and measured changes in TTC at V3 was 0.95 (Fig. 3E); the CCC between predicted and measured changes in TTV was 0.94 (Fig 3F). These results indicate a high predictive accuracy and precision (i.e., uncertainty in the model's prediction; see Supplemental Section 2.2 for interpretation) of the temporal dynamics of the response of TNBC to neoadjuvant A/C (Table 1).

Importantly, the personalized digital twins provide not only the global metrics summarized in the previous paragraph, but also the spatiotemporal evolvement of each patient's tumor. Fig. 4A and 4B show the measured and predicted tumor cell distributions, respectively, from the central slice of the same two example patients; and Fig. 4C and 4D present the 3D rendering of measured and predicted tumor volumes, respectively. In both cases, the digital twins successfully capture the lack (Fig. 4A) or presence (Fig. 4B) of response. Quantitatively, for the first patient, the difference between the predicted and measured change of tumor cell distributions at V3 has a median (interquartile range) of −3.30% (−22.07% – 0.00%); for the second patient, the difference has a median (interquartile range) of 0.00% (0.00% – 0.00%). The median differences across all patients had a mean (95% CI) of 0.20% (−20.35% – 20.75%) at V3 (Fig. 4E). These results indicate high predictive accuracy and precision of spatially-resolved predictions of tumor cell distributions in TNBC patients in response to neoadjuvant A/C (Table 1).

Framework 2: Patient-specific prediction of final pathological response

A cohort of 37 patients (18 pCR, 19 non-pCR) was used in Framework 2. After excluding the six patients from the entire cohort ($n = 56$) as conducted in Framework 1, thirteen more patients were excluded due to presumed non-response to further standard-of-care chemotherapy and enrollment into clinical trials ($n = 9$; part of ARTEMIS schema), missing the schedule of paclitaxel ($n = 2$), and errors in image acquisition ($n = 2$; ADC maps did not cover the entire tumor at V3).

For each patient, the personalized digital twin estimated treatment efficacies based on the V1 – V3 images (Fig. 5A), and represented tumor dynamics in response to A/C. These results were then used to predict the response to paclitaxel and, therefore, final treatment outcome after all NAST (Fig. 5C). For example, Fig. 5C shows the digital twin was able to be calibrated during the A/C regimen and used to predict no regrowth during paclitaxel for a patient that did, in fact, achieve pCR after NAST. In contrast, Fig. 5D shows the digital twin captured an initial response and then subsequent regrowth during A/C, and predicted

strong regrowth before and during paclitaxel, resulting in a predicted TTC and TTV values of 9.03×10^8 and TTV of 5.21×10^3 mm^3, respectively, after NAST.

Across the cohort, as shown in Table 2, the predicted TTC and TTV (based on the model calibrated with the V1-V3 data) from the digital twins both achieved an AUC of 0.89 (0.78 – 0.99) for distinguishing pCR from non-pCR (with post-surgical pathology as the ground truth). In comparison, the measured TTV or TTC (based on the V3 data) achieved an AUC of 0.78 (0.62 – 0.94) for distinguishing pCR from non-pCR. Furthermore, using the predicted TTC and TTV, the specificity was 0.95 and 0.89, respectively. In comparison, if using the measured TTV or TTC, the specificity was only 0.79. These results demonstrate that compared to the directly measured data, the digital twins improved the prediction of final response. Specifically, the AUC was improved by 14.28% for TTC ($P = 0.04$; significant) and 13.83% for TTV ($P = 0.07$). The specificity was improved by 20.25% and 12.66% for TTC and TTV, respectively; the sensitivity was unchanged.

Discussion

We have developed a digital twin approach to achieve early, patient-specific, spatiotemporally-resolved predictions of the response of TNBC patients to neoadjuvant doxorubicin, cyclophosphamide, and paclitaxel. This approach was based on a biologically-based mathematical model calibrated with multi-visit, multiparametric MRI acquired for the individual patient. Thus, the methodology represents a significant step away from *population-based* predictions, and towards *individual-based* predictions.

Framework 1 shows how the digital twin employs the imaging data from individual patients acquired early in the course of neoadjuvant A/C to make very accurate predictions of tumor status at the conclusion of A/C. The CCC between the predicted and measured values of total tumor burden and total tumor volume were 0.95 and 0.94, respectively. This strongly indicates that early changes during the A/C regimen contain sufficient information to calibrate the digital twin and confidently predict tumor response at the end of A/C. This observation aligns with previous reports demonstrating that metrics from early-treatment MRI are strong predictors of NAST response in breast cancer (15–18, 40). This provides strong support that our approach could be used to adjust treatment regimens on a patient-specific basis. For example, our approach could be applied after the first two cycles of A/C to predict if further dosing with A/C should be continued, or if an alternative intervention should be considered.

Framework 2 shows how the digital twin employs the imaging data from individual patients acquired during the A/C portion of NAST to make accurate predictions of their final pathological status (i.e., pCR or non-pCR) at the completion of all NAST, with an AUC = 0.89. Importantly, using the tumor volume measured at the end of A/C yielded only an AUC = 0.78; thus, the digital twin provided a substantial improvement on the AUC (14%). Comparing to previous MRI-based predictions of breast cancer response to NAST (17, 18, 40, 41), the digital twin also shows an improved accuracy. Both our previous study (18) and an I-SPY study (41) reported the best discrimination between pCR/non-pCR in TNBC using the optimized mid-treatment FTV, with AUC = 0.85. Moreover, the addition of post-NAST

ADC to FTV showed an improvement for predicting response in TNBC, increasing AUC from 0.71 to 0.81 (17). Combining a pharmacokinetic parameter (i.e., k_{ep}) with ADC measured post one-cycle NAST yielded an AUC = 0.88 in breast cancer (40). Moreover, the predictive accuracy of the digital twin is comparable to state-of-the-art, machine learning (ML)-based predictions. For example, Ravichandran *et al.* applied a convolutional neural network (CNN) to predict pCR from pre-treatment DCE-MRI and achieved an AUC of 0.77 in a total of 166 breast cancer patients (42). A more recent CNN-based study using both pre- and post-treatment DCE-MRI achieved an AUC of 0.91 in a cohort of 42 breast cancer patients (43).

Importantly, our digital twin approach has several inherent advantages comparing to ML algorithms. First, ML methods rely on access to a large patient population to train the algorithm and this training dataset must include all relevant pathophysiological features of the disease under investigation, have high-quality annotations, and be generalizable from one population to the next. In contrast, our approach calibrates a biologically-based model using patient-specific data to make patient-specific predictions, which does not require population-based training or annotation labels. Second, ML can be difficult to interpret biologically due to complex modeling features. In contrast, the digital twins provide an accurate prediction not only of pCR status at the conclusion of NAST, but also of the mechanistic interpretation of the tumor development during NAST. For example, our modeling framework can capture the initial and subsequent responses to A/C as depicted in Fig. 3A–D for patients with very different response dynamics. Third, because the digital twin parameters quantify and elucidate the observed tumor response dynamics, it provides another potential application: predicting response to multiple candidate therapeutic regimens. Thus, digital twins built on quantitative imaging data can provide—early in the course of NAST—a practical way to optimize individual treatment plans and hasten truly personalized cancer care (44).

There are, however, a few areas of our digital twin that require further investigation. First, instead of utilizing DCE-MRI to inform the spatial distribution of delivered drug, coupling the tumor-response model (i.e., Eq. [2]) with drug delivery and/or tumor angiogenesis models (45, 46) could estimate drug distribution more accurately. Second, instead of assuming cell death is merely proportional to drug concentration (i.e., the last term on the right-hand side of Eq. [2]), a model accounting for detailed therapeutic mechanisms and pharmacodynamics could improve predictive accuracy (47, 48). Third, our methodology is not currently capable of predicting invasion to the lymph nodes or axilla, which limits the accuracy for predicting pathological response or residual cancer burden (see Supplemental Section 2.3 for additional analysis). Of course, building more comprehensive models necessitates either more assumptions or more measurements, so a careful balance of model complexity and predictive accuracy must be sought.

While the dataset employed in this study is much larger and more homogeneous than our previous preliminary study (36), there are still a few points about the cohort that could affect the analysis. Framework 1 excluded four patients with no visible tumor at V2. This decision was made so we did not overestimate the accuracy of our predictions as such patients also show no visible tumor at V3,thereby resulting in a 100%-accurate prediction without really testing the modeling. Additionally, the absence of pathological evaluation at

the intermediate time point led to the lack of "ground-truth" for Framework 1. Framework 2 excluded nine patients due to enrollment in other trials, causing the cohort to be enriched with pCR patients. This could lead to an overestimation of accuracy for differentiating pCR/non-pCR (see Supplemental Section 2.4 for additional analysis). Future efforts will seek to apply our digital twin formalism to larger patient population, with a more typical partition of pCR/non-pCR, for further validation.

While we (and others) have worked hard to characterize errors in the measurements themselves (34, 49, 50), the processing (e.g., segmentation, registration) steps are also potential sources of error that can be propagated through the modeling pipeline and lead to bias in the prediction. A detailed investigation suggested that an un-anticipated changes in ADC values and distribution, as well as the appearance of necrotic regions, are potential sources of error (see Supplemental Section 2.5). Additionally, Framework 1 involved sampling the drug decay rates, which introduces an uncertainty in the predicted tumor response and limits the accuracy for predicting final pathology (see Supplemental Section 2.6). However, compared to previous attempts of simultaneously calibrating the drug efficacy and decay (36), this procedure not only ensures a more robust model calibration, but also allows the uncertainty in tumor dynamics to be quantified and interpreted. Another source of uncertainty is the setting of drug efficacy and decay rate of paclitaxel in Framework 2, due to the lack of imaging during the paclitaxel portion of NAST. One solution is to incorporate more measurements during NAST, especially after alternating the therapies, so that the digital twin can be updated to preserve an accurate prediction. Of course, it is important to note the goal of the digital twin is not to provide a perfect reproduction of the patient's situation. Rather, a realistic goal is to provide an accurate and practical formalism that provides clinically actionable insights. In the present contribution, we have achieved this goal in the context of predicting the response of early-stage triple-negative breast cancer to neoadjuvant systemic chemotherapy.

This study also supports the value of longitudinal MRI in cancer care. Currently, only pre-treatment MRI is standard for assessing breast cancer patients. While increasing to multiple MRIs will increase the cost for imaging, the benefit of early detection of (for example) chemo-resistance would help to avoid unnecessary toxicity and costs. Both this and prior studies showed that follow-up imaging after the first 1 – 3 cycles (15–18, 40, 41), of NAST is helpful for the early prediction of response. Including even one follow-up MRI (enough to enable model calibration) would be of great benefit.

Conclusion

We have developed patient-specific digital twins *via* integrating longitudinal, multiparametric MRI data with biologically-based mathematical modeling. This technique accurately captures the spatiotemporal response of TNBC to NAST and achieves high accuracy and specificity for predicting the final pathological status for each individual patient. The success of this approach demonstrates the potential of digital twins to shift the paradigm from *assessing* response to *predicting* and, eventually, *optimizing* response.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the National Institutes of Health for funding through NCI U01CA142565, U01CA174706, and U24CA226110. We thank the Cancer Prevention and Research Institute of Texas for support through CPRIT R120005. T.E.Y. is a CPRIT Scholar in Cancer Research. We thank the Oncological Data and Computational Sciences collaboration between the Oden Institute for Computational Engineering and Sciences at The University of Texas at Austin, the MD Anderson Cancer Center, and the Texas Advanced Computing Center for providing seed funding on this project. We thank the generous philanthropic contributions of the University of Texas MD Anderson Cancer Center Moon Shots Program™. J.K.L discloses the following financial relationships: Grant or research support from Novartis, Medivation/Pfizer, Genentech, GSK, EMD-Serono, Astra-Zeneca, Medimmune, Zenith, Merch; Speaker's Bureau for MedLearning, Physician's Education Resource, Prime Oncology, Medscape, Clinical Care Options, Medpage; Royalty from UpToDate.

References

1. Liu SV, Melstrom L, Yao K, Russell CA, Sener SF. Neoadjuvant therapy for breast cancer. J Surg Oncol 2010;101:283–291. [PubMed: 20187061]
2. Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 2008;26:1275–1281. [PubMed: 18250347]
3. Wu K, Yang Q, Liu Y, Wu A, Yang Z. Meta-analysis on the association between pathologic complete response and triple-negative breast cancer after neoadjuvant chemotherapy. World J Surg Oncol 2014;12:95. [PubMed: 24731479]
4. Spring LM, Fell G, Arfe A, Sharma C, Greenup RA, Reynolds KL, et al. Pathological complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: a comprehensive meta-analysis. Clin Cancer Res 2020;26:1–11 [PubMed: 31900310]
5. Schmid P, Adams S, Rugo HS, Schneeweiss A, Barrios CH, Iwata H, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. N Engl J Med 2018;379(22):2108–2121. [PubMed: 30345906]
6. Schmid P, Salgado R, Park YH, Muñoz-Couselo E, Kim SB, Sohn J, et al. Pembrolizumab plus chemotherapy as neoadjuvant treatment of high-risk, early-stage triple-negative breast cancer: results from the phase 1b open-label, multicohort KEYNOTE-173 study. Ann Oncol 2020;31(5):569–581. [PubMed: 32278621]
7. Hormuth DA, Jarrett AM, Lorenzo G, Lima EA, Wu C, Chung C, et al. Math, magnets, and medicine: Enabling personalized oncology. Expert Rev Precis Med Drug Dev 2021;6(2):79–81. [PubMed: 34027102]
8. Nyrop KA, Deal AM, Shachar SS, Basch E, Reeve BB, Choi SK, et al. Patient-Reported Toxicities During Chemotherapy Regimens in Current Clinical Practice for Early Breast Cancer. The Oncologist 2019;24:762–771. [PubMed: 30552158]
9. MCandrew N, DeMichele A. Neoadjuvant chemotherapy considerations in triple-negative breast cancer. J Target Ther Cancer 2018;7:52–69. [PubMed: 29577076]
10. Carey LA. Research priorities in prediction of response in early breast cancer. The Breast 2019;48:S31–S33. [PubMed: 31839156]
11. van la Parra RF, Tadros AB, Checka CM, Rauch GM, Lucci A Jr, Smith BD, et al. Baseline factors predicting a response to neoadjuvant chemotherapy with implications for non-surgical management of triple-negative breast cancer. Br J Surg 2018;105(5):535–543. [PubMed: 29465744]
12. Marinovich ML, Sardanelli F, Ciatto S, Mamounas E, Brennan M, Macaskill P, et al. Early prediction of pathologic response to neoadjuvant therapy in breast cancer: systematic review of the accuracy of MRI. Breast 2012;21:669–677. [PubMed: 22863284]

13. Khairalseed M, Javed K, Jashkaran G, Kim JW, Parker KJ, Hoyt K. Monitoring Early Breast Cancer Response to Neoadjuvant Therapy Using H-Scan Ultrasound Imaging: Preliminary Preclinical Results. J Ultrasound Med 2019;38:1259–1268. [PubMed: 30280391]
14. Wang Y, Zhang C, Liu J, Huang G. Is 18F-FDG PET accurate to predict neoadjuvant therapy response in breast cancer? A meta-analysis. Breast Cancer Res Treat 2012;131(2):357–369. [PubMed: 21960111]
15. Virostko J, Hainline A, Kang H, Arlinghaus LR, Abramson RG, Barnes SL, et al. Dynamic contrast-enhanced magnetic resonance imaging and diffusion-weighted magnetic resonance imaging for predicting the response of locally advanced breast cancer to neoadjuvant therapy: a meta-analysis. J Med Imaging 2018;5(1):011011.
16. Hylton NM, Blume JD, Bernreuter WK, Pisano ED, Rosen MA, Morris EA, Weatherall PT, Lehman CD, Newstead GM, Polin S, Marques HS. Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy—results from ACRIN 6657/I-SPY TRIAL. Radiology 2012;263(3):663–672. [PubMed: 22623692]
17. Li W, Newitt DC, Wilmes LJ, Jones EF, Arasu V, Gibbs J, et al. Additive value of diffusion-weighted MRI in the I-SPY 2 TRIAL. J Magn Reson Imaging 2019;50(6):1742–1753. [PubMed: 31026118]
18. Musall BC, Abdelhafez AH, Adrada BE, Candelaria RP, Mohamed RM, Boge M, et al. Functional Tumor Volume by Fast Dynamic Contrast-Enhanced MRI for Predicting Neoadjuvant Systemic Therapy Response in Triple-Negative Breast Cancer. J Magn Reson Imaging 2021;54(1):251–260. [PubMed: 33586845]
19. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. Invest Radiol 2019;54(2):110. [PubMed: 30358693]
20. Sammut SJ, Crispin-Ortuzar M, Chin SF, Provenzano E, Bardwell HA, Ma W, et al. Multi-omic machine learning predictor of breast cancer therapy response. Nature 2022;601:623–629. [PubMed: 34875674]
21. Clare SE, Shaw PL. “Big Data” for breast cancer: where to look and what you will find. NPJ breast cancer 2016;2(1):1–5.
22. Coveney PV, Dougherty ER, Highfield RR. Big data need big theory too. Philos Trans A Math Phys Eng Sci 2016;374(2080):20160153. [PubMed: 27698035]
23. Yankeelov TE, Atuegwu N, Hormuth D, Weis JA, Barnes SL, Miga MI, et al. Clinically relevant modeling of tumor growth and treatment response. Sci Transl Med 2013;5(187):187ps9.
24. Yankeelov TE, Quaranta V, Evans KJ, Rericha EC. Toward a science of tumor forecasting for clinical oncology. Cancer Res 2015;75:918–923. [PubMed: 25592148]
25. Kapteyn MG, Pretorius JV, Willcox KE. A probabilistic graphical model foundation for enabling predictive digital twins at scale. Nat Comput Sci 2021;1(5):337–347.
26. Wu C, Lorenzo G, Hormuth DA, Lima EA, Slavkova KP, DiCarlo JC, Virostko J, Phillips CM, Patt D, Chung C, Yankeelov TE. Integrating mechanism-based modeling with biomedical imaging to build practical digital twins for clinical oncology. Biophysics Reviews 2022;3(2):021304. [PubMed: 35602761]
27. Yam C, Yen EY, Chang JT, Bassett RL, Alatrash G, Garber H, et al. Immune phenotype and response to neoadjuvant therapy in triple-negative breast cancer. Clin Cancer Res 2021;27(19):5365–5375. [PubMed: 34253579]
28. U.S. Department of Health and Human Services, Food and Drug Administration, Oncology Center of Excellence, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Pathological Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval Guidance for Industry. Silver Spring, MD: United States Food and Drug Administration, Oncology Center of Excellence; 2020. p.5.
29. Saranathan M, Rettmann DW, Hargreaves BA, Clarke SE, Vasanawala SS. Differential subsampling with Cartesian ordering (DISCO): A high spatio-temporal resolution Dixon imaging

sequence for multiphasic contrast enhanced abdominal imaging. J Magn Reson Imaging 2012;35(6):1484–1492. [PubMed: 22334505]

30. Jarrett AM, Kazerouni AS, Wu C, Virostko J, Sorace AG, DiCarlo JC, et al. Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. Nat Protoc 2021;16(11):5309–5338. [PubMed: 34552262]
31. Li X, Dawant BM, Welch EB, Chakravarthy AB, Freehardt D, Mayer I, et al. A nonrigid registration algorithm for longitudinal breast MR images and the analysis of breast tumor response. Magn Reson Imaging 2009;27(9):1258–1270. [PubMed: 19525078]
32. Li X, Dawant BM, Welch EB, Chakravarthy AB, Xu L, Mayer I, et al. Validation of an algorithm for the nonrigid registration of longitudinal breast MR images using realistic phantoms. Med Phys 2010;37(6):2541–2552. [PubMed: 20632566]
33. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. Elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2009;29(1):196–205. [PubMed: 19923044]
34. Sorace AG, Wu C, Barnes SL, Jarrett AM, Avery S, Patt D, et al. Repeatability, reproducibility, and accuracy of quantitative mri of the breast in the community radiology setting. J Magn Reson Imaging 2018;48(3):695–707.
35. Atuegwu NC, Arlinghaus LR, Li X, Welch EB, Chakravarthy BA, Gore JC, et al. Integration of diffusion-weighted MRI data and a simple mathematical model to predict breast tumor cellularity during neoadjuvant chemotherapy. Magn Reson Med 2011;66(6):1689–1696. [PubMed: 21956404]
36. Jarrett AM, Hormuth II DA, Wu C, Kazerouni AS, Ekrut DA, Virostko J, et al. Evaluating patient-specific neoadjuvant regimens for breast cancer via a mathematical model constrained by quantitative magnetic resonance imaging data. Neoplasia 2020;22(12):820–830. [PubMed: 33197744]
37. Hormuth DA II SL, Weis JA, Miga MI, Yankeelov TE. Mechanically coupled reaction-diffusion model to predict glioma growth: methodological details. Methods in molecular biology (Clifton, NJ) 2018;1711:225.
38. Lorenzo G, Jarrett AM, Meyer CT, Tyson DR, Quaranta V, Yankeelov TE. Identifying relevant parameters that characterize the early response to NAT in breast cancer patients using a novel personalized mechanistic model integrating in vitro and in vivo imaging data. Cancer Res 2021;81(4 Supplement):PS13–44.
39. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845. [PubMed: 3203132]
40. Li X, Abramson RG, Arlinghaus LR, Kang H, Chakravarthy AB, Abramson VG, Farley J, Mayer IA, Kelley MC, Meszoely IM, Means-Powell J. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. Investigative radiology. 2015;50(4):195–204. [PubMed: 25360603]
41. Li W, Arasu V, Newitt DC, Jones EF, Wilmes L, Gibbs J, Kornak J, Joe BN, Esserman LJ, Hylton NM. Effect of MR imaging contrast thresholds on prediction of neoadjuvant chemotherapy response in breast cancer subtypes: a subgroup analysis of the ACRIN 6657/I-SPY 1 TRIAL. Tomography 2016;2(4):378–387. [PubMed: 28066808]
42. Ravichandran K, Braman N, Janowczyk A, Madabhushi A. A deep learning classifier for prediction of pathological complete response to neoadjuvant chemotherapy from baseline breast DCE-MRI. In Medical imaging 2018: computer-aided diagnosis 2018;10575:105750C.
43. El Adoui M, Drisis S, Benjelloun M. Multi-input deep learning architecture for predicting breast tumor response to chemotherapy using quantitative MR images. Int J Comput Assist Radiol Surg 2020;15(9):1491–1500. [PubMed: 32556920]
44. Hernandez-Boussard T, Macklin P, Greenspan EJ, Gryshuk AL, Stahlberg E, Syeda-Mahmood T, et al. Digital twins for predictive oncology will be a paradigm shift for precision cancer care. Nat Med 2021;27(12):2065–2066. [PubMed: 34824458]
45. Wu C, Hormuth DA, Pineda F, Karczmar GS, Yankeelov TE. Towards patient-specific optimization of neoadjuvant treatment protocols for breast cancer based on image-guided fluid dynamics. IEEE Trans Biomed Eng 2022. doi:10.1109/TBME.2022.3168402.

46. Hormuth DA, Phillips CM, Wu C, Lima EA, Lorenzo G, Jha PK, et al. Biologically-Based Mathematical Modeling of Tumor Vasculature and Angiogenesis via Time-Resolved Imaging Data. Cancers 2021;13(12):3008. [PubMed: 34208448]

47. Jarrett AM, Bloom MJ, Godfrey W, Syed AK, Ekrut DA, Ehrlich LI, et al. Mathematical modelling of trastuzumab-induced immune response in an in vivo murine model of HER2+ breast cancer. Math Med Biol 2019;36(3):381–410. [PubMed: 30239754]

48. West J, Schenck RO, Gatenbee C, Robertson-Tessi M, Anderson AR. Normal tissue architecture determines the evolutionary course of cancer. Nat Commun 2021;12(1):1–9. [PubMed: 33397941]

49. Shukla-Dave A, Obuchowski NA, Chenevert TL, Jambawalikar S, Schwartz LH, Malyarenko D, 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 2019;49(7):e101–e121. [PubMed: 30451345]

50. Newitt DC, Zhang Z, Gibbs JE, Partridge SC, Chenevert TL, Rosen MA, et al. Test–retest repeatability and reproducibility of ADC measures by breast DWI: Results from the ACRIN 6698 trial. Journal of Magnetic Resonance Imaging. 2019;49(6):1617–1628. [PubMed: 30350329]

Statement of Significance

Integrating MRI data with biologically-based mathematical modeling successfully predicts breast cancer response to chemotherapy, suggesting digital twins could facilitate a paradigm shift from simply assessing response to predicting and optimizing therapeutic efficacy.

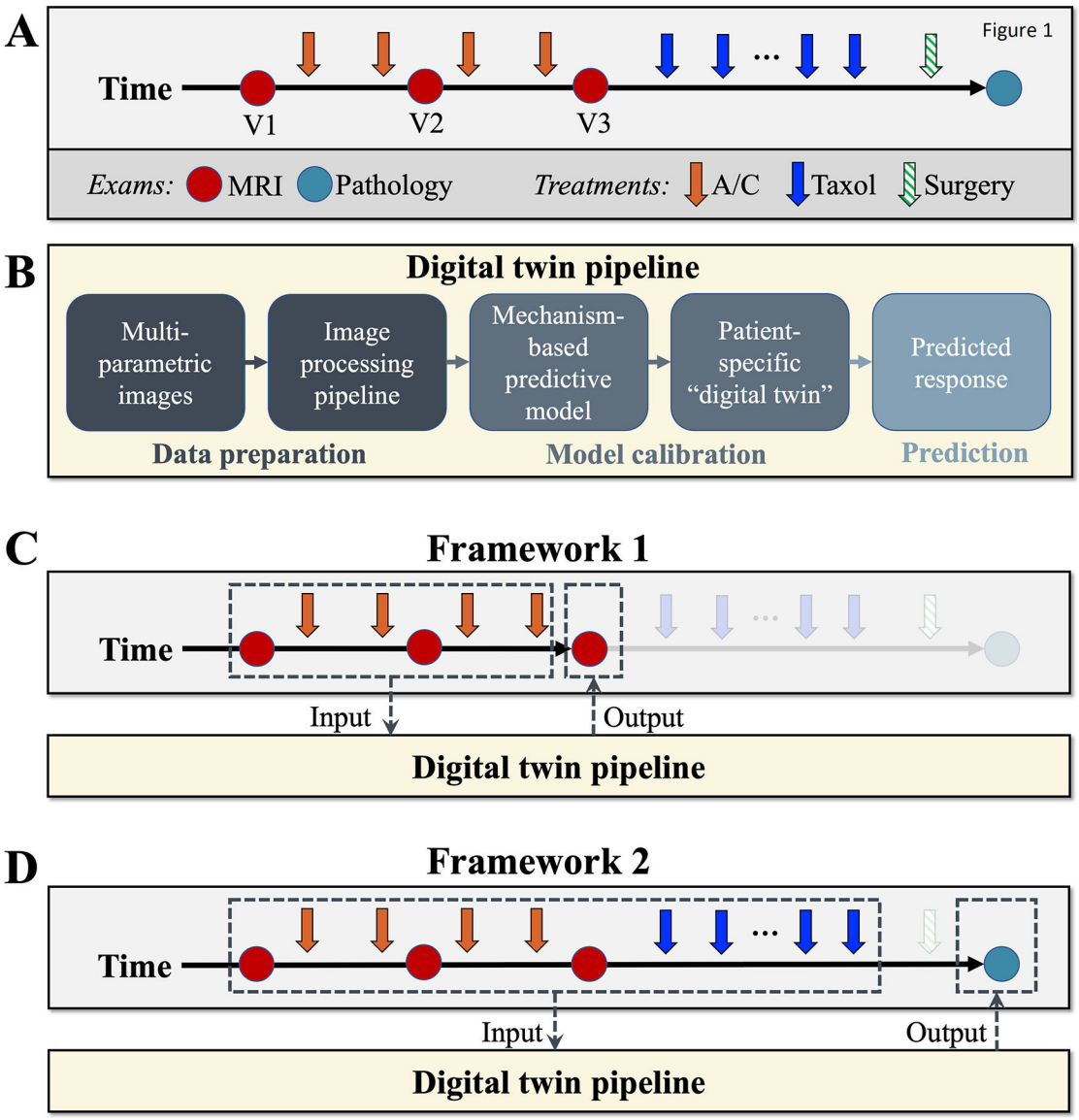

Figure 1. The two Frameworks to predict patient-specific response to NAST.

Panel **A** shows the timeline of treatment administration and data acquisition for each patient. Panel **B** illustrates the processing-modeling pipeline to generate patient-specific digital twins. Two frameworks are established to evaluate the predictive ability of the digital twins. Framework 1 (Panel **C**) employs digital twins to predict the outcome of the doxorubicin and cyclophosphamide (A/C) regimen. Patient-specific images from visits 1 (V1) and 2 (V2) along with the schedule of A/C provide the input to which the digital twin is calibrated. Once calibrated, the digital twin outputs a prediction of the spatiotemporal development of the tumor in response to A/C. The prediction is then directly compared to the V3 images. Framework 2 (Panel **D**) employs digital twins to predict the outcome of the entire NAST. Images from V1, V2, and V3, along with the schedule of both A/C and paclitaxel, are given as input, and the digital twin outputs a prediction of whether the patient will achieve a pCR. The prediction is then directly compared to the post-surgical pathological response.

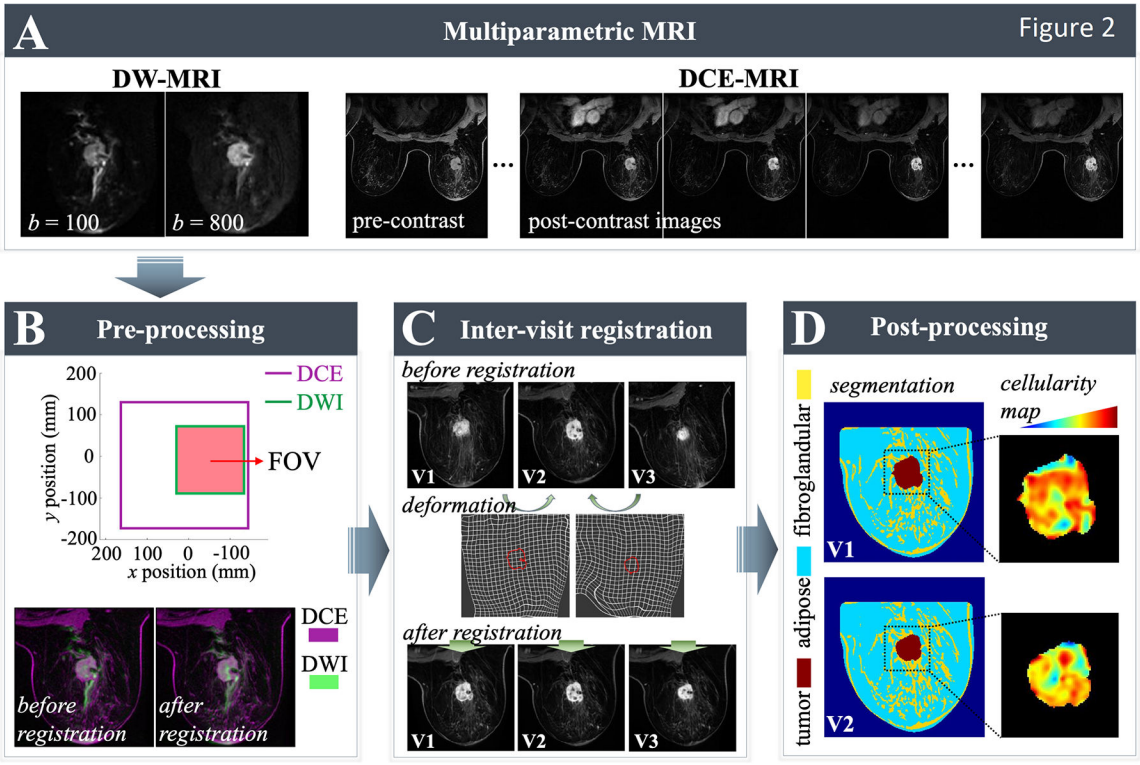

Figure 2. Flow charts of MRI data processing.

Panel **A** shows an example set of DW-MRI and DCE-MRI data acquired at one visit of a patient. Panels **B-D** illustrates the three steps of the processing pipeline, respectively. In panel **B**, multiparametric images are trimmed to the same field-of-view (FOV) and registered. In panel **C**, images from V1 and V3 are registered to those from V2. In panel **D**, tissue segmentation and calculation of tumor cellularity (from the DW-MRI data) are performed. These steps prepare the data for calibration with the biologically-based mathematical model and establishing each patient's digital twin.

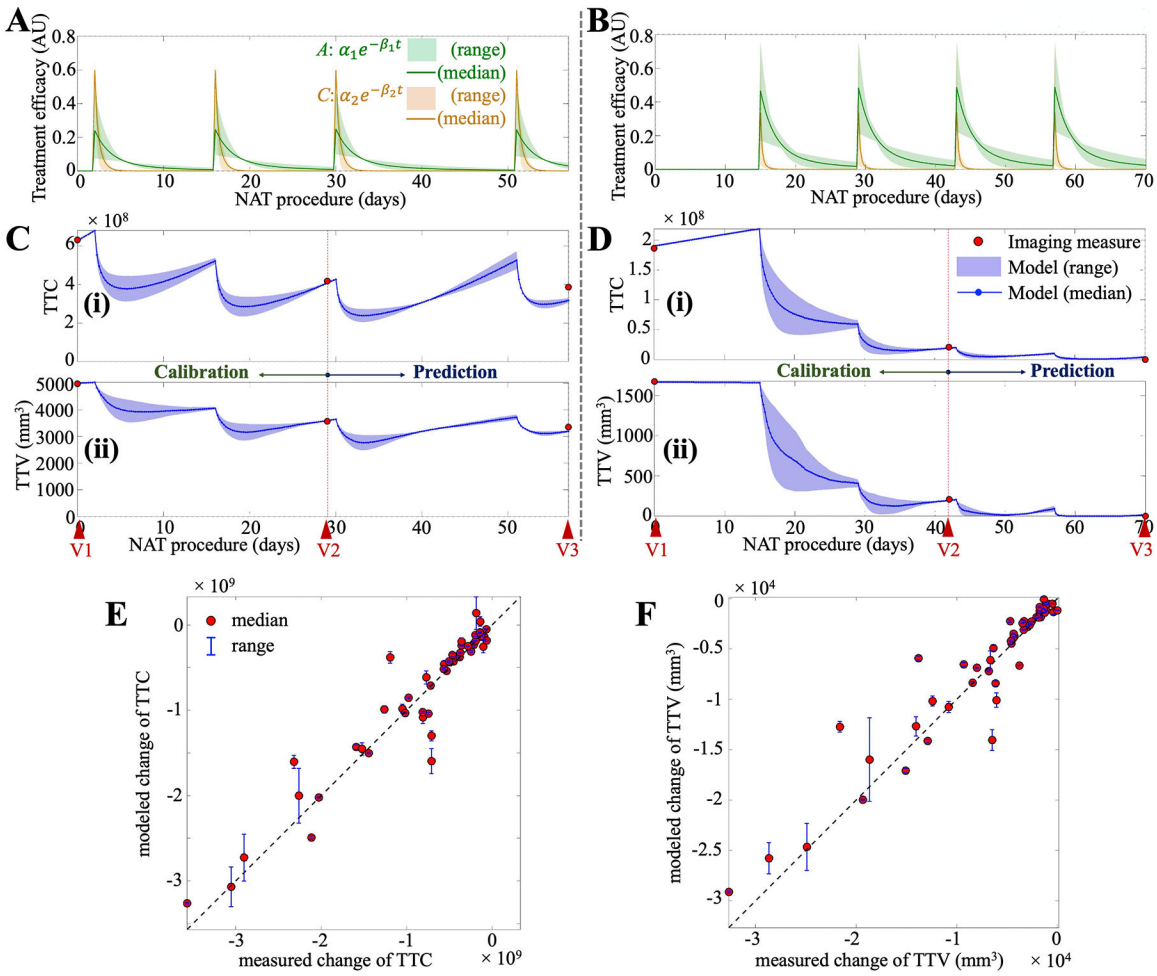

Figure 3. Temporal accuracy of patient-specific predictions of the response of TNBC to A/C.

Panels **A** and **B** show the time courses of calibrated therapeutic efficacies over the A/C regimen in two representative patients, respectively. Panels **C** and **D** present the temporal dynamics predicted by the digital twins of the same two patients, in which subpanels **(i)** and **(ii)** represent the change of tumor cellularity (TTC) and tumor volume (TTV), respectively. In each panel, red circles present the measured TTC or TTV at certain time points, while blue curves and shadows present the predicted median and range of dynamics, respectively. Very small differences are observed between the measured and predicted changes of TTC and TTV over time in the example patients. Panels **E** and **F** plot the correlation between the measured and predicted changes of TTC and TTV (CCC = 0.95 and 0.94), respectively, in the cohort. These results indicate high precision and accuracy for predicting patient-specific temporal dynamics of TNBC in response to A/C.

Figure 4

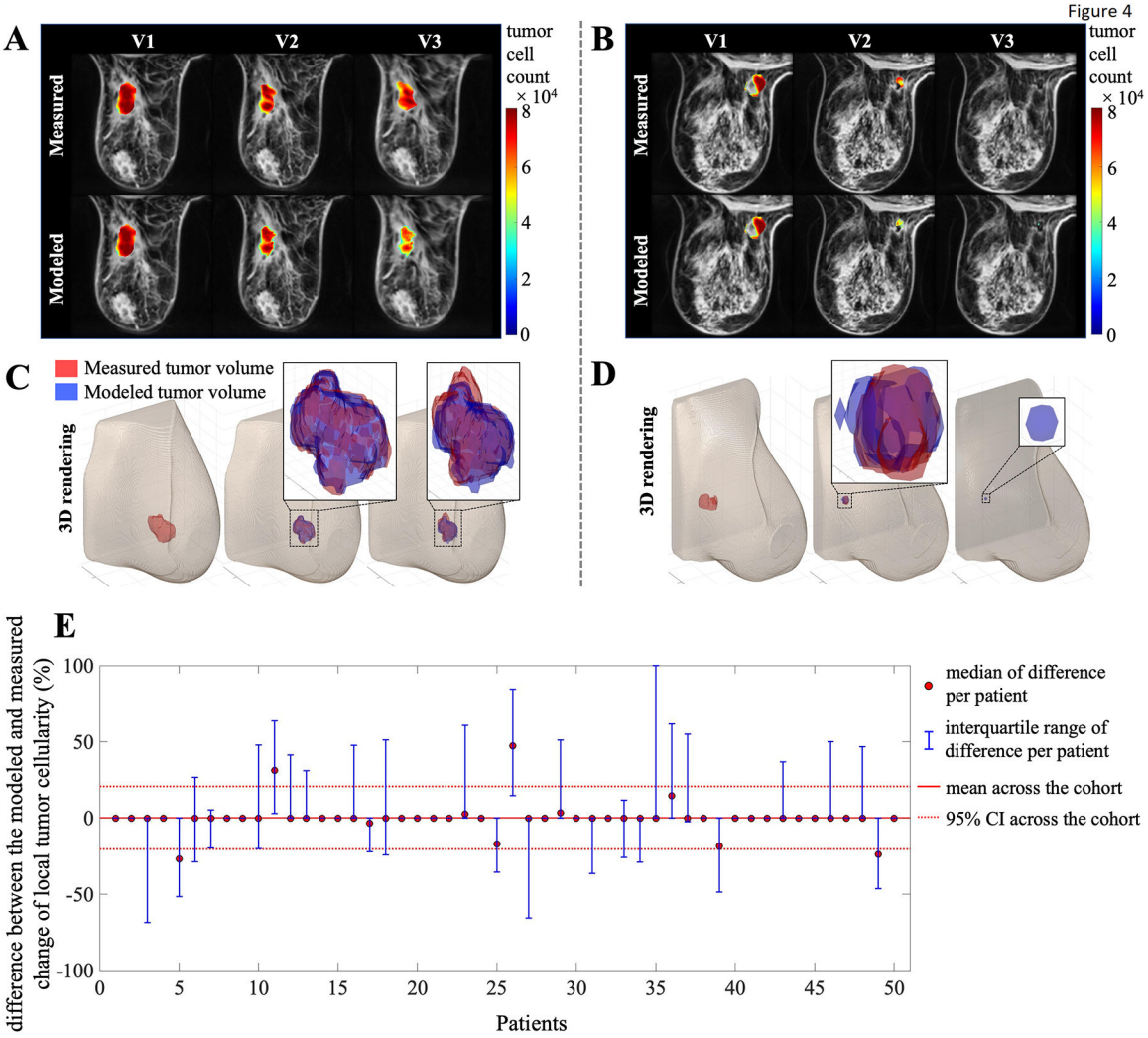

Figure 4. Spatial accuracy of patient-specific predictions of the response of TNBC to A/C.

Panels **A** and **B** show the measured and predicted tumor cell distributions on the central tumor slice for two patients. Panels **C** and **D** show the 3D renderings of the measured and predicted change of those two tumor shapes. Very small differences are observed between the measured and predicted tumor cell distributions or tumor shapes in the patients. Panel **E** presents the difference between the measured and predicted change of tumor cell distributions in the cohort. The median (red circle) and interquartile range (blue bar) of difference within each patient's tumor region are presented. The difference across all patients has a mean (95% CI) of 0.20% (−20.35% – 20.75%). These results indicate a high accuracy of the digital twins for predicting patient-specific spatial dynamics of TNBC in response to A/C.

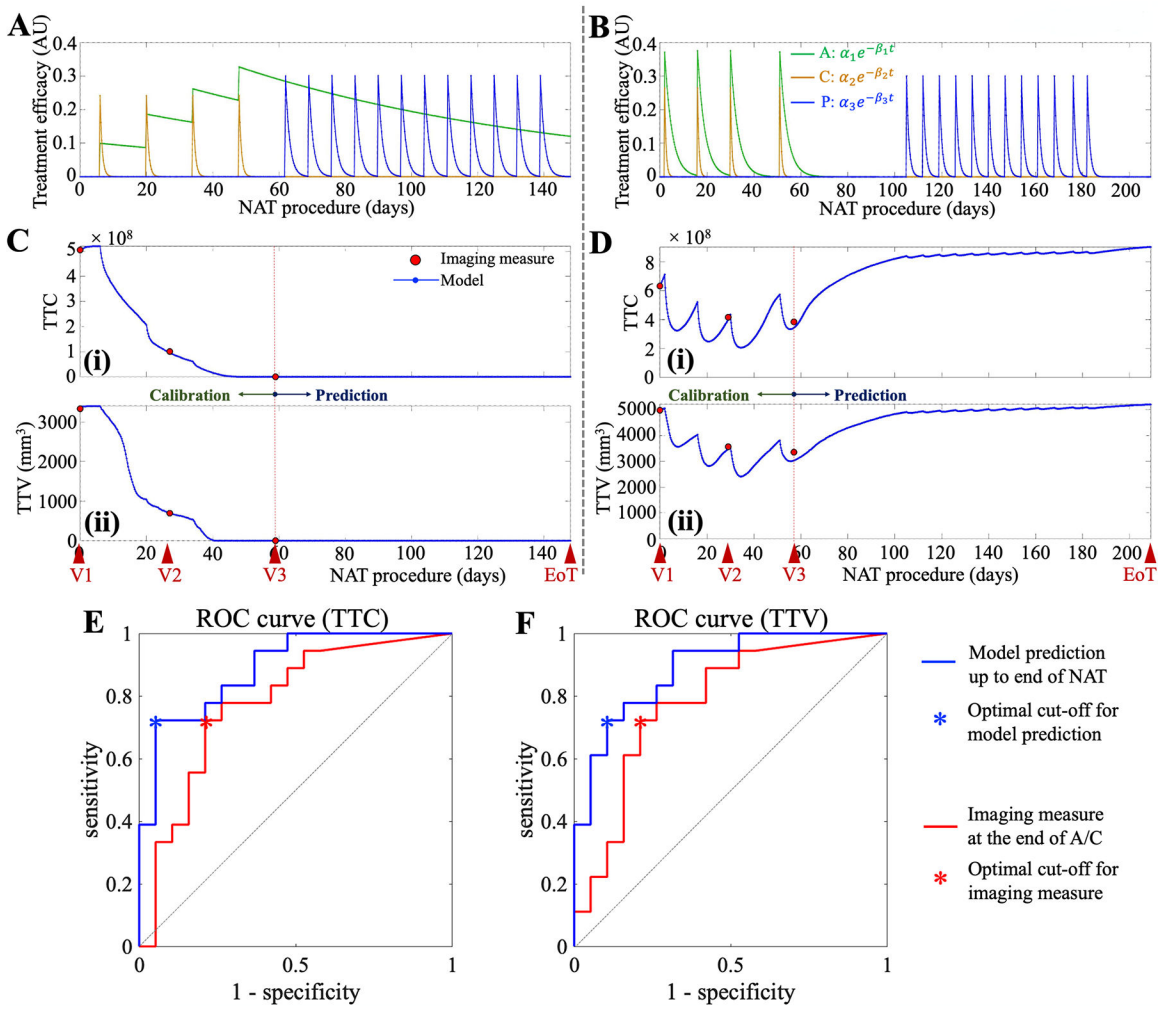

Figure 5. Accuracy of patient-specific prediction of final pathological response.

Panels **A** and **B** show the time courses of calibrated therapeutic efficacies during NAST in two example patients, respectively. Panels **C** and **D** present the temporal dynamics predicted by the digital twins for the same two example patients, in which subpanels **(i)** and **(ii)** represent the change of tumor cellularity (TTC) and tumor volume (TTV), respectively. Panel **E** presents the ROC analysis of differentiating pCR from non-pCR based on predicted (blue) and measured (red) TTC; Similarly, panel **F** presents the ROC analysis based on predicted and measured TTV. The larger AUCs of blue curves compared to red curves in both panels **E** and **F** indicates superior accuracy for predicting final pathological response.

Table 1.

Statistical evaluation of patient-specific predictions of the TNBC response to A/C

time point \ metric	temporal accuracy		spatial accuracy
	CCC between predicted and measured ΔTTC[a]	**CCC between predicted and measured ΔTTV[b]**	**mean (95% CI) of the median differences between predicted and measured ΔTCD[c]**
V2	1.00	0.99	0.22% (−3.34% – 3.78%)
V3	0.95	0.94	0.20% (−20.35% – 20.75%)

Note:

[a] ΔTTC = change of total tumor cellularity from pre-treatment to time point V2 or V3

[b] ΔTTV = change of total tumor volume from pre-treatment to time point V2 or V3

[c] ΔTCD = change of spatially-resolved tumor cell distribution within original tumor region from pre-treatment to time point V2 or V3

Table 2.

Statistical evaluation of patient-specific prediction of final pathological response

metric \ statistic		AUC (95% CI)	Sensitivity	Specificity
TTC	Measured	0.78 (0.62 – 0.94)	0.72	0.79
	Predicted	0.89 (0.78 – 0.99)	0.72	0.95
TTV	Measured	0.78 (0.62 – 0.94)	0.72	0.79
	Predicted	0.89 (0.78 – 0.99)	0.72	0.89