iScience

CellPress
OPEN ACCESS

Perspective

Designing clinical trials for patients who are not average

Thomas E. Yankeelov,[1,2,3,4,5,9,*] David A. Hormuth II,[4,5] Ernesto A.B.F. Lima,[4,6] Guillermo Lorenzo,[4,10] Chengyue Wu,[4] Lois C. Okereke,[4] Gaiane M. Rauch,[7] Aradhana M. Venkatesan,[7] and Caroline Chung[8]

SUMMARY

The heterogeneity inherent in cancer means that even a successful clinical trial merely results in a therapeutic regimen that achieves, on average, a positive result only in a subset of patients. The only way to optimize an intervention for an individual patient is to reframe their treatment as their own, personalized trial. Toward this goal, we formulate a computational framework for performing personalized trials that rely on four mathematical techniques. First, mathematical models that can be calibrated with patient-specific data to make accurate predictions of response. Second, digital twins built on these models capable of simulating the effects of interventions. Third, optimal control theory applied to the digital twins to optimize outcomes. Fourth, data assimilation to continually update and refine predictions in response to therapeutic interventions. In this perspective, we describe each of these techniques, quantify their "state of readiness", and identify use cases for personalized clinical trials.

THE FUNDAMENTAL PROBLEM WITH CLINICAL TRIALS

Clinical trials typically consist of four phases in which a new intervention is investigated in human subjects to determine its safety and efficacy. This system is notoriously inefficient; the average cost per patient is $59,500 and takes more than 10 years[1] with only 10% of drugs in phase I studies eventually approved.[2] Even if this incredibly resource-intensive process is successful, resulting in an improved intervention, there may be little consideration given to how the new treatment protocol can be optimally delivered to each patient. This is because clinical trials are designed to experimentally test a limited set of interventions to find the one that works best—on average—for a particular population of patients. Furthermore, clinical trials specifically establish a treatment delivery protocol with any adjustments considered a protocol deviation; thus, personalized adjustments are explicitly discouraged. For a disease as heterogeneous as cancer,[3] this is a fundamental limitation; unfortunately, a fixed treatment protocol yielding the best average outcome for the population will frequently not yield the best outcome for an individual patient.[4,5] The design and timing of therapeutic interventions must be based on the unique characteristics of each patient—not the average characteristics of the population.

PERSONALIZING THE APPROACH TO CLINICAL TRIALS

It is instructive to consider the computational framework outlined by the I-SPY consortium.[6] In particular, I-SPY 2 is an "adaptive" clinical trial designed to reduce the time required to determine which candidate regimens are most effective for particular breast cancer subtypes. Its adaptive design varies the ratio of patients in the experimental arm to the control arm by assigning a higher proportion to the arm that yields higher rates of response—and does so as soon as possible in the life of the trial. The computational framework outlined below is designed to develop a rigorous method to identify optimal treatment strategies on a per patient basis, rather than the population. That is, we seek to predict, early in the course of an intervention, which individuals will benefit from a particular regimen—just as the I-SPY 2 trial seeks to assess, early in the course of the trial, which populations would benefit from a particular regimen. Importantly, I-SPY 2 has been successful in their effort;[7] we posit that a patient-specific, prediction-based approach will hasten similar advances for other malignancies.

We[5,8–10] and others[4,11–13] have developed biology-based mathematical models[14,15] designed to be calibrated with patient-specific data to make patient-specific predictions of the spatiotemporal dynamics of tumor response. Such models can realize patient-specific digital twins

[1]Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, USA
[2]Department of Diagnostic Medicine, The University of Texas at Austin, Austin, TX 78712, USA
[3]Department of Oncology, The University of Texas at Austin, Austin, TX 78712, USA
[4]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA
[5]Livestrong Cancer Institutes, The University of Texas at Austin, Austin, TX 78712, USA
[6]Texas Advanced Computer Center, The University of Texas at Austin, Austin, TX 78712, USA
[7]Department of Abdominal Imaging, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
[8]Department of Radiation Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
[9]Division of Diagnostic Imaging, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
[10]Department of Civil Engineering and Architecture, University of Pavia, 27100 Pavia, Italy
*Correspondence: thomas.yankeelov@utexas.edu
https://doi.org/10.1016/j.isci.2023.108589

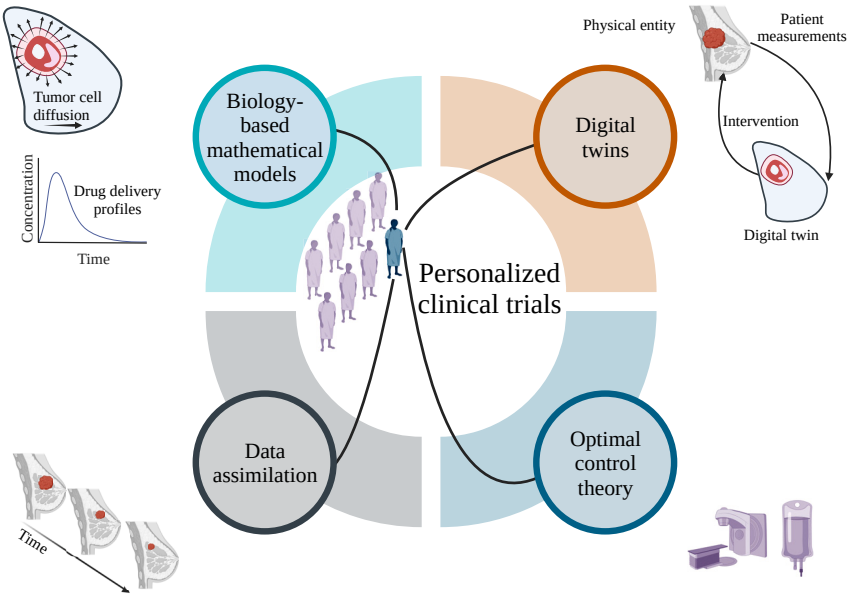

Figure 1. Four tools for the next generation of clinical trials

We propose four techniques to provide the computational foundation to practically achieve an $N = 1$ clinical trial. The first technique is biology-based mathematical modeling, which explicitly describes how tumor cells proliferate, move, and respond to therapy in the language of mathematics. The second technique is the construction of digital twins, which are a dynamic, bidirectional interaction between the physical tumor and the digital representation of the tumor. The third technique is optimal control theory, which is a technique that can be used to identify the optimal dose and schedule for an individual patient while balancing treatment efficacy with adverse effects. The final technique is data assimilation, which provides guidance on how to integrate newly acquired data (e.g., imaging, clinical, molecular, or genetic) throughout clinical care within the mathematical model at the core of the patient's digital twin to improve model predictions.

to which optimal control theory can be applied to identify personalized therapeutic regimens (i.e., personalized clinical trials) designed to dramatically outperform standard-of-care protocols. Furthermore, using the methods of data assimilation the predictions produced by the digital twins can be repeatedly updated as more patient-specific data becomes available, thereby providing new opportunities to guide subsequent interventions. To the best of our knowledge, there are no examples in the literature outlining how these four concepts could be leveraged to execute clinical trials at the patient-specific level. We hope that providing a practical, prescriptive methodology for connecting these four concepts can hasten the arrival of clinical trials for patients that are not average.

TOOLS FOR THE NEXT GENERATION OF CLINICAL TRIALS

In the following sections we systematically describe each of the four techniques in our framework (i.e., biology-based mathematical models, digital twins, optimal control theory, and data assimilation) with one figure, one example application, and a summary statement on their "state-of-readiness". Each of these four mathematical techniques were developed and refined in the mathematical and physical sciences as well as in engineering over decades. Thus, there is a strong and sophisticated foundational framework for their implementation. Furthermore, the computational power that is now readily accessible has made it possible to implement such a paradigm in practice. Figure 1 provides a summary of these four techniques and their "state-of-readiness". Figures 2, 3, 4, and 5 then show how they can be applied to the particular use case of predicting the response of locally advanced breast cancer to neoadjuvant therapy. We note that the decision to focus on breast cancer was merely made on the basis of clarity and continuity of presentation as these methods are (of course) agnostic to the disease site to which they are applied; all that is required is the requisite domain knowledge and access to the appropriate data.

Biology-based mathematical models

We refer to mathematical models that explicitly account for the biological mechanisms underlying the growth and treatment response of cancer as "biology-based" models.[14,15] This approach is different from modeling strategies that rely exclusively on statistical relationships amongst the data. For the present discussion, the most relevant feature of a biology-based mathematical model is that it can be "calibrated" with patient-specific data and need not require averages obtained from a population of patients. Biology-based mathematical models are usually governed by a collection of variables and parameters that can change over time and/or the spatial geometry of the tumor-harboring tissue; these entities can describe (for example) how the tumor cells proliferate, move given the mechanical properties of the tissue, and die in response to treatment.[4,5,8–13] Indeed, the latter also constitutes a central subject of the mature field of pharmacodynamics. Such models can

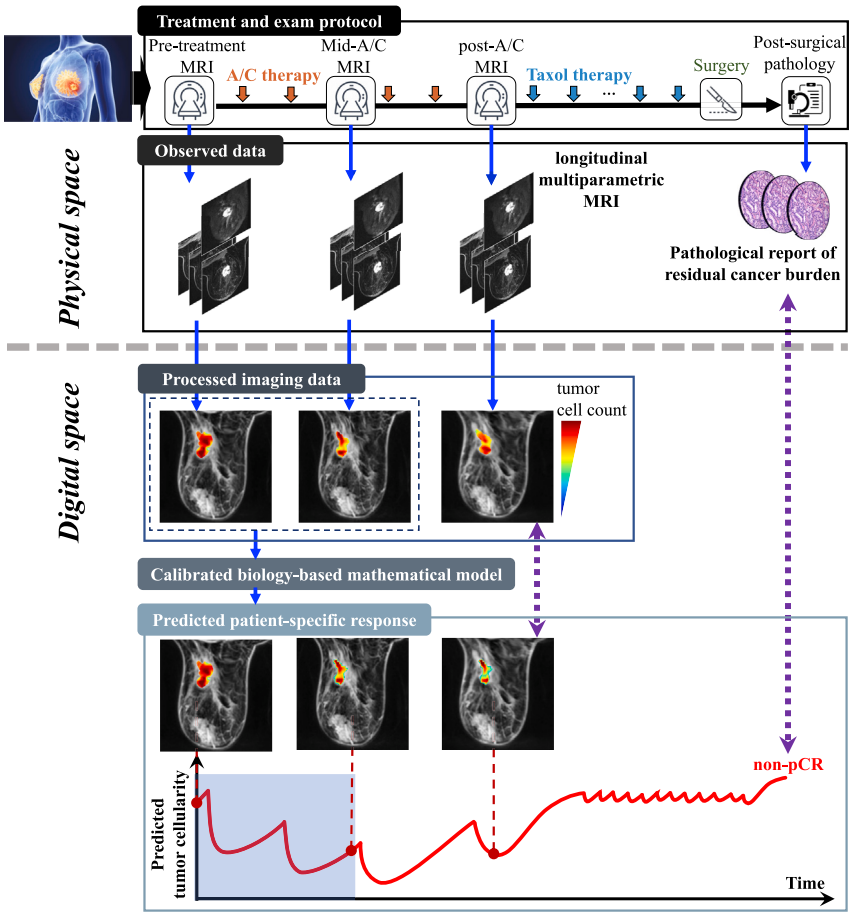

Figure 2. A mathematical model for predicting response

The figure illustrates how to employ longitudinal MRI data to calibrate a biology-based mathematical model to predict the response of a locally advanced breast cancer to neoadjuvant therapy. Multiparametric MRI data are processed through an automated pipeline. The processed data from the early phase of the treatment, as well as the specific treatment regimen for the patient under investigation, are used to calibrate the biology-based mathematical model which captures how the tumor cells proliferate, move, and respond to therapy, thereby establishing a personalized digital twin. The digital twin then provides patient-specific predictions for the spatiotemporally-resolved response of the tumor to given therapies, along with the post-therapy pathological status. The predictions can be validated by comparing to the measurements obtained before and after neoadjuvant therapy.

further describe these mechanisms for other cell types interacting with tumor cells, such as the endothelial cells in tumor-supporting vasculature[15] responsible for delivering systemic treatments (a central subject of the mature field of pharmacokinetics).

Patient-specific data is then used to either directly assign values to the variables and parameters, or to infer the parameter values through the process of model calibration.[14,18,19] Once the model is calibrated (and henceforth personalized), it can be run forward in time to make a prediction of a quantity of interest (e.g., tumor cell density maps, tumor size or cellularity). These forecasts can then be directly compared to observations from each patient for validation or model update. Figure 2 illustrates this process using quantitative magnetic resonance imaging (MRI) data to calibrate a biology-based mathematical model to predict the response of a locally advanced breast cancer to neoadjuvant therapy.

The literature now contains many successful applications of this approach in the clinical setting,[4,5,8–10,12–15] and it is our view that the "state-of-readiness" of this technology is 'high'. That is, the mathematical and computational formalism has matured to the point where those skilled in the art can readily apply it to appropriate clinical data. Given a mathematical model that can accurately recapitulate the spatiotemporal development of a tumor, the next step is to use that model to construct digital twins[20] to (for example) simulate a patient's outcome over a range of therapeutic options.

Digital twins

While a myriad of definitions of digital twins exist,[20–22] that provided by Boulos and Zhang is particularly instructive to the present discussion: "A digital twin is a virtual model of a physical entity, with dynamic, bidirectional links between the physical entity and its corresponding twin in the digital domain.[23]" The mathematical system described in Figure 2 is the virtual model of the tumor (i.e., the "physical entity").

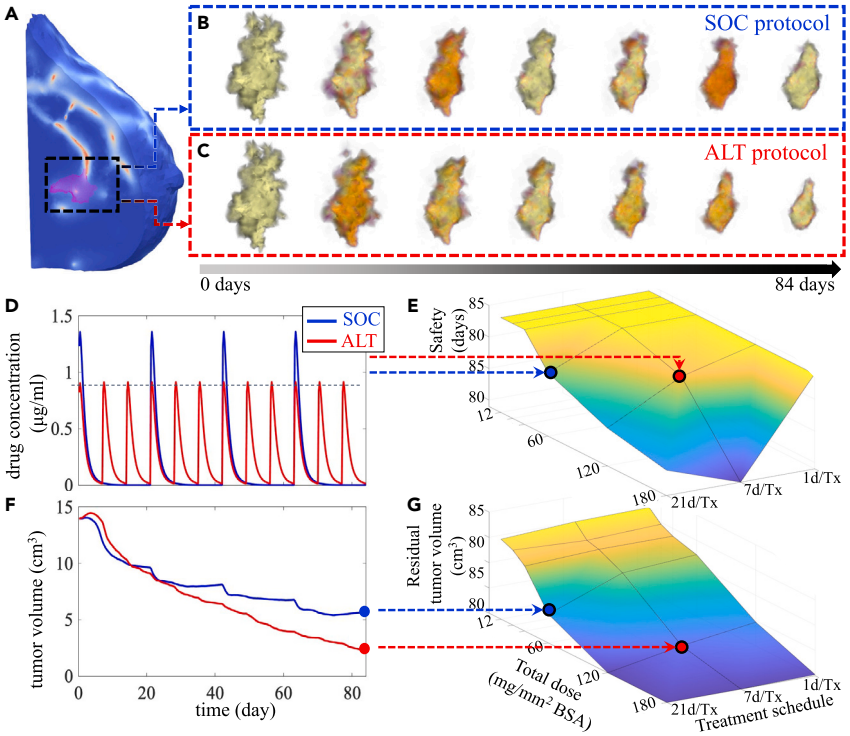

Figure 3. Systematically investigating therapeutic interventions for a breast cancer patient

MRI data quantitatively characterize tumor physiology (e.g., the blood perfusion, interstitial tissue density, and tumor cellularity) and can be used to personalize parameters in a biology-based mathematical model capturing how tumor cells proliferate, move, and respond to therapy (panel A). Then, the personalized model can predict the response of this particular tumor to various therapeutic regimens. For example, panels B and C visualize the 12-week dynamics of tumor response to a standard-of-care (SOC) treatment protocol for breast cancer (i.e., 60 mg/mm^2 body surface area of doxorubicin per 3 weeks) and an alternative (ALT) treatment protocol (i.e., 20 mg/mm^2 body surface area per week), respectively. Comparing the predicted median drug concentration in the healthy breast tissue under the SOC and ALT protocols (panel D), the SOC and ALT protocols result in a "Safety" metric (i.e., duration of drug concentration lower than an empirical toxic concentration[16]) of 79 and 81 days, respectively. Comparing the predicted tumor volume (panel F), the SOC and ALT protocols result in residual tumor volumes of 5.69 and 2.68 cm^3, respectively. This demonstrates that the ALT protocol outperforms the SOC protocol in both response and safety. This prediction can be performed for a wide range of combinations of drug administration doses and schedules, leading to surfaces for quantities of interests as showing in panels E and G. Additionally, treatment optimization can be adapted to focus the particular needs for each patient within the framework of optimal control theory.[17]

"Bidirectional dynamics" are realized by taking measurements from the individual patient to calibrate the model, using the model to make a prediction of how the physical tumor will evolve in space and time, and then using that prediction to intervene on the patient.[20,24] Once the digital twin exists, various interventions can be simulated to identify an optimal solution. It is critical to note that this is simply not possible to investigate experimentally; that is, one cannot give multiple interventions to a patient simultaneously to determine what is the best path forward—it simply has to be done through digital twins that are powered by biologically-based, patient-specific models. Figure 3 illustrates the systematic investigation (through simulation) of a wide range of therapeutic interventions for an individual breast cancer patient. Notice how the indicates alternative combination of dose and schedule is quite remote from that determined by the standard-of-care.

While digital twins have been deployed in multiple industrial applications,[21,22] the oncology literature contains only a modest number of entries showing successful applications of digital twins in the clinical setting.[20,24,25] Thus, it is our view that the "state-of-readiness" of digital twins for cancer is 'low'. Indeed, a National Academy of Sciences effort is currently underway to identify foundational research gaps and future directions for digital twins in biomedical science.[26] While the technology certainly exists to generate meaningful digital twins, more comprehensive "field-testing" within oncology must be performed before a higher level of readiness can be assigned. However, once more successful examples become available, the next step to select an optimal therapy for each individual patient is within grasp, as the digital twin can be used to generate a range of treatment plans from which the most promising can be selected via the established methods of optimal control theory.

Optimal control theory

Optimal control theory[14,17,27] seeks to determine the specific factors controlling a system (e.g., how to treat a tumor) so that specific outcomes can be achieved (e.g., maximizing tumor control). We can state our optimal control problem as follows: given a model that is capable of

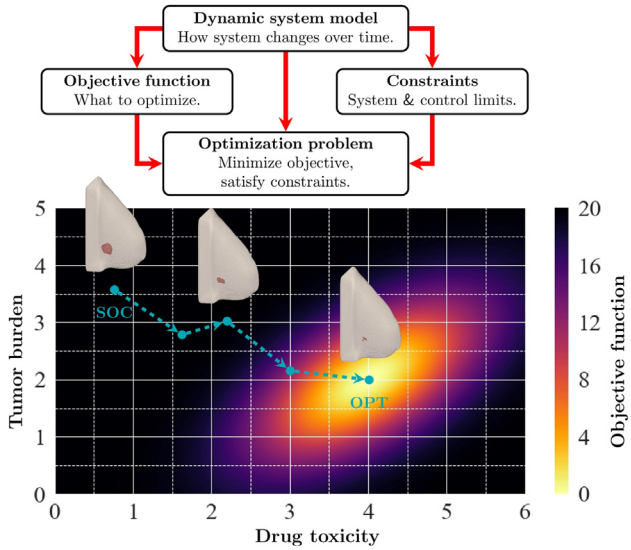

Figure 4. Optimization of a breast cancer treatment protocol

This figure illustrates the path (highlighted in blue) of an optimization problem with two objectives: minimizing tumor burden and minimizing drug toxicity. The standard-of-care (SOC) protocol is represented at the beginning of the path (top left). An intermediate protocol is shown in the middle of the path, which balances the optimization objectives better than the SOC. At the end of the path, we present the optimal (OPT) protocol obtained after solving the optimization problem, which minimizes the objective function while satisfying the imposed constraints. Thus, the optimization process requires a biology-based mathematical model of the dynamic system of breast cancer growth and drug delivery (see Figure 2 and the text), defining the objective function (e.g., minimize the tumor burden while reducing drug toxicity) and constraints (e.g., minimize drug dose), and solving the resulting optimization problem. The objective function weights the trade-off between tumor size and drug toxicity, which must be balanced to achieve the optimal protocol for each individual patient. Hence, the optimal protocol provides personalized treatment recommendations for the breast cancer patient that are hypothesized to outperform the one-size-fits-all approach of the SOC approach, thereby leading to better clinical outcomes.

accurately predicting the response of a tumor to a set of interventions, find the order, timing, and dosing of the treatment regimen that minimizes the tumor cell number at the conclusion of therapy. A second problem we can formulate is: given a model that is capable of accurately predicting the response of a tumor to a specific set of interventions, maximally reduce the amount of intervention (e.g., drug or radiation dose) while maintaining the tumor control achieved by the full dose. This latter problem is particularly important in certain pediatric settings where toxicity-induced side effects can adversely affect quality of life in later years.[28] Note that solving either of these two critically important problems requires a biologically based mathematical model capable of predicting tumor growth and therapeutic response, which can be further leveraged to construct a patient-specific digital twin. It is extraordinarily difficult to conceive of this framework from an AI/Big Data perspective because there are simply not enough patients or resources to experimentally test all the dosing and treatment options at the population level for a given cancer type. (See below for more discussion on this important point.) Fortunately, each patient's digital twin allows us to computationally identify their optimal therapeutic intervention. Figure 4 presents the mathematical formulation of this problem for the particular case of identifying alternative therapeutic neoadjuvant regimens for treating breast cancer.

Multiple studies have been investigating the mathematical formulation and properties as well as the simulated outcomes of the application of optimal control theory to cancer treatments.[14,17,29,30] There are, though, few examples where optimal control theory has been applied to a digital twin to identify alternative treatment regimens that are hypothesized to produce greater tumor control than what was achieved clinically[4,5,12,16,31] or in experiments,[32,33] but there are no clinical studies validating these predictions for individual patients. One possible way forward is to integrate the notion of a mathematically defined "optimal" treatment protocol with clinically accepted regimens. To do so, one could first use optimal control theory to identify a theoretically optimal treatment protocol. Of course, this protocol may not be clinically feasible (and/or tolerable), so the next step is to merely select the protocol from the set of currently accepted therapeutic options for the patient that provides the outcome closest to the one identified by optimal control theory, thereby yielding the "optimal practical" therapeutic regimen.[29] This is a particular exciting—and practical—prospect which could have enormous impact on the field.

As the literature is only beginning to explore the practical role of optimal control theory in pre-clinical and clinical oncology and that only established biology-based models are a requisite for its application, it is thus our view that the "state-of-readiness" of this technology is 'intermediate'.

Data assimilation

One intervention is seldom all that is required to control a patient's cancer. Thus, a methodology for updating predictions as more data becomes available is required within the digital twin. Data assimilation is a set of techniques in which newly acquired data are integrated with the

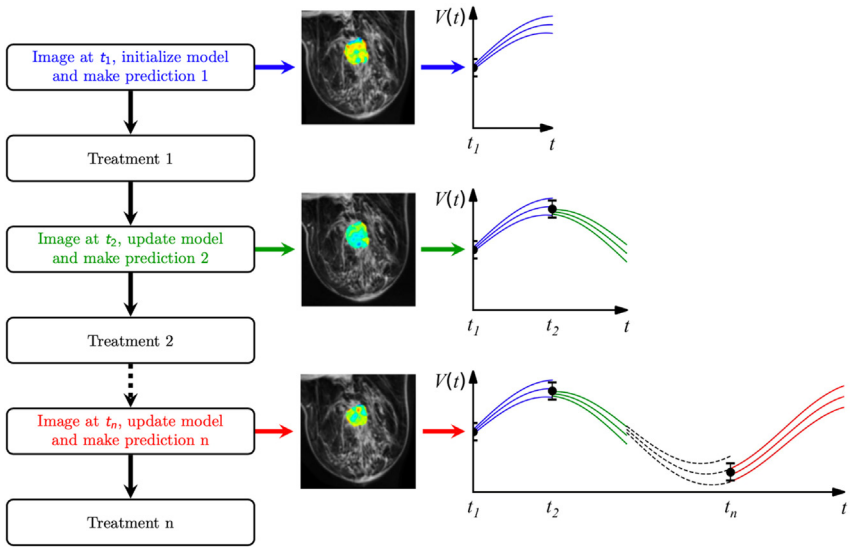

Figure 5. Updating model prediction through data assimilation

Beginning in the upper left-hand corner of the figure, a patient presents at time t_1 with locally advanced breast cancer and their measured initial conditions (e.g., imaging, clinical, molecular, or genetic) are used to assign model parameters in a biology-based mathematical model. This patient-specific model is then used to make the first of several patient-specific predictions of the spatiotemporal response of the tumor (in this example, the colored pixels indicate tumor cell density). The patient then proceeds to Treatment 1 and, after a period of time, a new set of data is collected which allows the model to be re-calibrated to make a refined prediction 2 at time t_2. The patient then proceeds to Treatment 2. This procedure can be repeated n times during the entire course of therapy to continually update model predictions to provide the most accurate and up-to-date guidance on tumor status. (We note that while Figures 2, 3, 4, and 5 are all focused on breast cancer, this was done for clarity as the methodology is applicable to any solid tumor for which the requisite data are available.).

predictions of a mathematical model to update the characterization of the (spatio-) temporal dynamics of a system.[34] While this type of analysis is widely used in (for example) weather forecasting,[35] it is not commonly employed in cancer—though it does provide a straightforward way to constantly update and refine patient-specific predictions. Figure 5 illustrates one approach to data assimilation for updating patient-specific predictions for a breast cancer patient receiving neoadjuvant therapy. Recall that to make a patient-specific prediction, values for all model parameters are required (see Figure 2). Since these parameters are obtained by calibrating sequential data from an individual, it is difficult to make an accurate prediction from just the initial measurement. This can be overcome by calibrating a relevant biology-based model to each patient in a training set and then using (for example) a neural network to assign model parameters from the training set to a new patient in a testing set given their initial conditions. Then, when further data becomes available for the patient in the test set, the model can be calibrated to their sequential data and the resulting parameters can then be weighted with the population-based data. The process is repeated each time new data becomes available for the individual and, as more confidence is gained in the individual's data, the contribution from the population data approaches zero. This process of model updating is, of course, critical for patients receiving treatment as the initial (pre-treatment) predictions will rapidly lose relevance and the changes in tumor characteristics must be followed throughout treatment.

The notion of updating the patient-specific parameters and ensuing predictions of a biology-based mathematical model as new patient data is collected during the oncological management of their disease is not uncommon in the literature.[8,9,36–39] An early (and instructive) example, from Kostelich et al.[39] applied the Kalman filter to reaction-diffusion type models to simulate the growth of glioblastoma. When accounting for both model and measurement (i.e., MRI) error, they were able to accurately predict tumor growth (*in silico*) for one year with 60-day forecast/update cycles. More recently, Zahid et al.[37] developed a method to combine historical population-based responses with weekly measurements of patient-specific volumes to accurately predict outcomes for an individual test patient. In spite of such examples, there is still a dearth of studies investigating the application of robust data assimilation techniques to recapitulate, predict, and optimize cancer growth and treatment outcomes.[24,40] Thus, our view is that the "state-of-readiness" of this technology is 'intermediate' as data assimilation techniques only require a validated biology-based model for their used and they have been applied in a number of biomedical applications,[24,40–43] but an accepted formalism has not been applied in the clinical setting. Thus, we do not see this methodology as a rate-limiting step for realizing personalized clinical trials.

OPPORTUNITIES FOR COMBINING BIOLOGY-BASED AND DATA-DRIVEN MODELING

Methods of machine learning and "big data"—on their own—are not capable of achieving practical personalized clinical trials for optimizing interventions on a patient-specific basis. By its very nature, statistical inference relies on properties of large populations that can obscure conditions specific to the individual, especially for a disease as heterogeneous as cancer. More specifically, machine learning algorithms rely on

inductive reasoning to identify patterns observed in past data, whereas biology-based models rely on deductive reasoning to extrapolate patterns that were not present in past data.[44] The practical manifestation of these philosophical musings is that the "big data only" approach cannot predict sequelae that are not in the training set—and as interventions get progressively more targeted, and cancer itself is further divided in subtypes, the training sets required for applying an AI/Big Data approach to an individual patient's particular cancer subtype simply do not exist. This does not mean, of course, that AI and big data are without utility in realizing personalized clinical trials. Perhaps the most immediate use of leveraging both historical and individual data already has a developing literature—clinical trials based on Bayesian, rather than frequentist, statistics. In Bayesian-based clinical trials, patient heterogeneity is explicitly accounted for in the trial design leading to improvements in efficacy and reductions in toxicity by making treatment decisions based specific features of a subgroup rather features averaged over a population.[45]

The developing field of scientific machine learning (SML) also provides guidance on integrating biology-based and data-driven modeling.[46] SML connotes the synthesis of mechanism-based modeling (frequently couched in the language of differential equations) with data-driven models (frequently couched in the language of neural networks) to provide interpretable, domain-aware, approaches to assist decision making and/or reduce the computational cost of mechanism-based modeling. The goal is to exploit the ability to explain and extrapolate provided by the former, with the ability to describe datasets when the underling mechanisms are not known provided by the latter. A natural opportunity for this approach is in the digital twin setting which may require many thousands of model evaluations for optimal control theory and uncertainty quantification in model predictions. The computational cost of these approaches could be lessened, for example, through data-driven surrogate modeling for dynamical systems and inverse problems.[47,48] Other significant applications of SML in cancer research and clinical oncology include data-driven mechanistic model discovery[49,50] as well as the integration of multimodal, multi-scale data and mechanistic models to obtain personalized forecasts of tumor growth.[46,51] (These issues are all thoroughly explored in.[20]). One particularly promising application for SML is to use molecular and genomic data within mechanism-based models. This approach has been successfully employed to integrate mathematical models of cellular dynamics within a machine-learning framework. The resulting network models capture key features of known interactions, increases interpretability, and may be able to predict cell response to treatments not included in the training set.[52]

CONCLUSION

The care of a cancer patient always involves treating, followed by observation to learn how the treatment affects the patients which, in turn, guides subsequent interventions. Thus, we have attempted to offer an outline of how this approach to patient care could be formalized with mathematical methods to render superior results. Specifically, we propose that the personalization of patient care can be built on four key mathematical and computational techniques: biology-base mathematical modeling, digital twins, optimal control theory, and data assimilation. We posit that this framework will help move clinical trials from identifying treatment regimens designed to treat the population average, to *predicting* and *optimizing* treatment response for each patient (i.e., personalized clinical trials). This transformation from *population*-to *patient*-based therapy is inevitable and the methods of computational science will hasten its arrival.

ACKNOWLEDGMENTS

We thank the National Cancer Institute for funding through 1U24 CA226110, 1R01 CA235800, 1U01CA253540, 1R01CA260003, and 1R01CA276540. We thank the Cancer Prevention and Research Institute of Texas for support via CPRIT RR160005 and RP220225. T.E.Y. is a CPRIT Scholar of Cancer Research. G.L. acknowledges funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 838786.

AUTHOR CONTRIBUTIONS

Conceptualization, T.E.Y.; Writing – Original Draft, T.E.Y., D.A.H., E.A.B.F., G.L., C.W., and L.O.; Writing – Review and Editing, all authors.; Funding Acquisition, T.E.Y. and D.A.H.; Visualization, D.A.H., E.A.B.F., G.L., and C.W.; Supervision, T.E.Y.

DECLARATION OF INTERESTS

Effective July 2023, A.M.V. will receive an Honorarium from Elsevier Ltd. T.E.Y., D.A.H., and C.W. have a pending patent (US Provisional Patent Application No. 63/247233 filed September 22, 2021, pending). T.E.Y. and G.L. have a pending patent (US Provisional Patent Application No. 18/135580 filed August 8, 2023, pending). No other author declares a potential conflict of interest.

REFERENCES

1. Practice, B.T.P. (2015). Biopharmaceutical industry-sponsored clinical trials: Impact on state economies. Technical report.
2. Dowden, H., and Munro, J. (2019). Trends in clinical success rates and therapeutic focus. Nat. Rev. Drug Discov. *18*, 495–496.
3. Marusyk, A., Janiszewska, M., and Polyak, K. (2020). Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell *37*, 471–484.
4. Corwin, D., Holdsworth, C., Rockne, R.C., Trister, A.D., Mrugala, M.M., Rockhill, J.K., Stewart, R.D., Phillips, M., and Swanson, K.R. (2013). Toward Patient-Specific, Biologically Optimized Radiation Therapy Plans for the Treatment of Glioblastoma. PLoS One *8*, e79115.
5. Jarrett, A.M., Hormuth, D.A., Wu, C., Kazerouni, A.S., Ekrut, D.A., Virostko, J., Sorace, A.G., DiCarlo, J.C., Kowalski, J., Patt, D., et al. (2020). Evaluating patient-specific neoadjuvant regimens for breast cancer via a mathematical model constrained by quantitative magnetic resonance imaging data. Neoplasia *22*, 820–830.

6. The I-SPY Trials. https://www.ispytrials.org/.
7. Nanda, R., Liu, M.C., Yau, C., Shatsky, R., Pusztai, L., Wallace, A., Chien, A.J., Forero-Torres, A., Ellis, E., Han, H., et al. (2020). Effect of Pembrolizumab Plus Neoadjuvant Chemotherapy on Pathologic Complete Response in Women With Early-Stage Breast Cancer: An Analysis of the Ongoing Phase 2 Adaptively Randomized I-SPY2 Trial. JAMA Oncol. *6*, 676–684.
8. Wu, C., Jarrett, A.M., Zhou, Z., Elshafeey, N., Adrada, B.E., Candelaria, R.P., Mohamed, R.M.M., Boge, M., Huo, L., White, J.B., et al. (2022). MRI-Based Digital Models Forecast Patient-Specific Treatment Responses to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. Cancer Res. *82*, 3394–3404.
9. Hormuth, D.A., II, Feghali, K.A.A., Elliott, A.M., Yankeelov, T., and Chung, C. (2021). Image-based personalization of computational models for predicting response of high-grade glioma to chemoradiation. Sci. Rep. *11*, 1–14.
10. Lorenzo, G., Heiselman, J.S., Liss, M.A., Miga, M.I., Gomez, H., Yankeelov, T.E., Hughes, T.J., and Reali, A. (2022). Abstract 5064: Patient-specific forecasting of prostate cancer growth during active surveillance using an imaging-informed mechanistic model. Cancer Res. *82*, 5064.
11. Wong, K.C.L., Summers, R.M., Kebebew, E., and Yao, J. (2017). Pancreatic Tumor Growth Prediction With Elastic-Growth Decomposition, Image-Derived Motion, and FDM-FEM Coupling. IEEE Trans. Med. Imaging *36*, 111–123.
12. Lipková, J., Angelikopoulos, P., Wu, S., Alberts, E., Wiestler, B., Diehl, C., Preibisch, C., Pyka, T., Combs, S.E., Hadjidoukas, P., et al. (2019). Personalized Radiotherapy Design for Glioblastoma: Integrating Mathematical Tumor Models, Multimodal Scans, and Bayesian Inference. IEEE Trans. Med. Imaging *38*, 1875–1884.
13. Scheufele, K., Subramanian, S., and Biros, G. (2021). Fully Automatic Calibration of Tumor-Growth Models Using a Single mpMRI Scan. IEEE Trans. Med. Imaging *40*, 193–204.
14. Lorenzo, G., Hormuth, D.A., Jarrett, A.M., Lima, E.A.B.F., Subramanian, S., Biros, G., Oden, J.T., Hughes, T.J.R., and Yankeelov, T.E. Quantitative in vivo imaging to enable tumor forecasting and treatment optimization. In Cancer, Complexity, Computation, A. A. I. Balaz, ed. (Springer)
15. Hormuth, D.A., Phillips, C.M., Wu, C., Lima, E.A.B.F., Lorenzo, G., Jha, P.K., Jarrett, A.M., Oden, J.T., and Yankeelov, T.E. (2021). Biologically-Based Mathematical Modeling of Tumor Vasculature and Angiogenesis via Time-Resolved Imaging Data. Cancers *13*, 3008.
16. Wu, C., Hormuth, D., Lorenzo, G., Jarrett, A., Pineda, F., Howard, F.M., Karczmar, G., and Yankeelov, T.E. (2022). Towards patient-specific optimization of neoadjuvant treatment protocols for breast cancer based on image-guided fluid dynamics. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. *1*.
17. Jarrett, A.M., Faghihi, D., Ii, D.A.H., Lima, E.A.B.F., Virostko, J., Biros, G., Patt, D., and Yankeelov, T.E. (2020). Optimal Control Theory for Personalized Therapeutic Regimens in Oncology: Background, History, Challenges, and Opportunities. J. Clin. Med. *9*, 1314.
18. Hormuth, D., II, Eldridge, S.B., Weis, J., Miga, M.I., and Yankeelov, T.E. (2018). Mechanically Coupled Reaction-Diffusion Model to Predict Glioma Growth: Methodological Details. In Springer Methods and Protocols: Cancer Systems Biology, L. von Stechow, ed. (Springer New York), pp. 225–241.
19. Jarrett, A.M., Kazerouni, A.S., Wu, C., Virostko, J., Sorace, A.G., DiCarlo, J.C., Hormuth, D.A., Ekrut, D.A., Patt, D., Goodgame, B., et al. (2021). Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. Nat. Protoc. *16*, 5309–5338.
20. Wu, C., Lorenzo, G., Hormuth, D.A., Lima, E.A.B.F., Slavkova, K.P., DiCarlo, J.C., Virostko, J., Phillips, C.M., Patt, D., Chung, C., and Yankeelov, T.E. (2022). Integrating mechanism-based modeling with biomedical imaging to build practical digital twins for clinical oncology. Biophys. Rev. *3*, 021304.
21. Niederer, S.A., Sacks, M.S., Girolami, M., and Willcox, K. (2021). Scaling digital twins from the artisanal to the industrial. Nat. Comput. Sci. *1*, 313–320.
22. Tao, F., Zhang, H., Liu, A., and Nee, A.Y.C. (2019). Digital Twin in Industry: State-of-the-Art. IEEE Trans. Industr. Inform. *15*, 2405–2415.
23. Kamel Boulos, M.N., and Zhang, P. (2021). Digital Twins: From Personalised Medicine to Precision Public Health. J. Pers. Med. *11*, 745.
24. Chaudhuri, A., Pash, G., Hormuth, D.A., II, Lorenzo, G., Kapteyn, M., Wu, C., Lima, E.A.B.F., Yankeelov, T.E., and Willcox, K. (2023). Predictive Digital Twin for Optimizing Patient-Specific Radiotherapy Regimens under Uncertainty in High-Grade Gliomas. Front. Artif. Intell. *6*, 1222612.
25. Wembacher-Schroeder, E., Kerstein, N., Bander, E.D., Pandit-Taskar, N., Thomson, R., and Souweidane, M.M. (2021). Evaluation of a patient-specific algorithm for predicting distribution for convection-enhanced drug delivery into the brainstem of patients with diffuse intrinsic pontine glioma. J. Neurosurg. Pediatr. *28*, 34–42.
26. National Academies of Sciences, Engineering, and Medicine (2023). In Opportunities and Challenges for Digital Twins in Biomedical Research: Proceedings of a Workshop–in Brief, L. Casola, ed. (The National Academies Press).
27. Lenhart, S., and Workman, J.T. (2007). Optimal Control Applied to Biological Models.
28. Geenen, M.M., Cardous-Ubbink, M.C., Kremer, L.C.M., van den Bos, C., van der Pal, H.J.H., Heinen, R.C., Jaspers, M.W.M., Koning, C.C.E., Oldenburger, F., Langeveld, N.E., et al. (2007). Medical assessment of adverse health outcomes in long-term survivors of childhood cancer. JAMA *297*, 2705–2715.
29. Colli, P., Gomez, H., Lorenzo, G., Marinoschi, G., Reali, A., and Rocca, E. (2021). Optimal control of cytotoxic and antiangiogenic therapies on prostate cancer growth. Math. Models Methods Appl. Sci. *31*, 1419–1468.
30. Cunningham, J.J., Brown, J.S., Gatenby, R.A., and Staňková, K. (2018). Optimal control to develop therapeutic strategies for metastatic castrate resistant prostate cancer. J. Theor. Biol. *459*, 67–78.
31. Angaroni, F., Graudenzi, A., Rossignolo, M., Maspero, D., Calarco, T., Piazza, R., Montangero, S., and Antoniotti, M. (2020). An Optimal Control Framework for the Automated Design of Personalized Cancer Treatments. Front. Bioeng. Biotechnol. *8*, 523.
32. Lima, E.A.B.F., Wyde, R.A.F., Sorace, A.G., and Yankeelov, T.E. (2022). Optimizing combination therapy in a murine model of HER2+ breast cancer. Comput. Methods Appl. Mech. Eng. *402*, 115484.
33. Randles, A., Wirsching, H.G., Dean, J.A., Cheng, Y.K., Emerson, S., Pattwell, S.S., Holland, E.C., and Michor, F. (2021). Computational modelling of perivascular-niche dynamics for the optimization of treatment schedules for glioblastoma. Nat. Biomed. Eng. *5*, 346–359.
34. Guzzi, R. (2016). Data Assimilation: Mathematical Concepts and Instructive Examples (Springer International Publishing).
35. Lahoz, W.A., and Schneider, P. (2014). Data assimilation: making sense of Earth Observation. Front. Environ. Sci. *2*.
36. Lorenzo, G., di Muzio, N., Deantoni, C.L., Cozzarini, C., Fodor, A., Briganti, A., Montorsi, F., Pérez-García, V.M., Gomez, H., and Reali, A. (2022). Patient-specific forecasting of postradiotherapy prostate-specific antigen kinetics enables early prediction of biochemical relapse. iScience *25*, 105430.
37. Zahid, M.U., Mohsin, N., Mohamed, A.S.R., Caudell, J.J., Harrison, L.B., Fuller, C.D., Moros, E.G., and Enderling, H. (2021). Forecasting Individual Patient Response to Radiotherapy in Head and Neck Cancer with a Dynamic Carrying Capacity Model. Int. J. Radiat. Oncol. Biol. Phys. *111*, 693–704.
38. Brady-Nicholls, R., Nagy, J.D., Gerke, T.A., Zhang, T., Wang, A.Z., Zhang, J., Gatenby, R.A., and Enderling, H. (2020). Prostate-specific antigen dynamics predict individual responses to intermittent androgen deprivation. Nat. Commun. *11*, 1750.
39. Kostelich, E.J., Kuang, Y., McDaniel, J.M., Moore, N.Z., Martirosyan, N.L., and Preul, M.C. (2011). Accurate state estimation from uncertain data and models: an application of data assimilation to mathematical models of human brain tumors. Biol. Direct *6*, 64.
40. Liu, J., Ii, D.A.H., Yang, J., Yankeelov, T.E., Liu, J., Ii, D.A.H., Yang, J., and Yankeelov, T.E. (2023). A data assimilation framework to predict the response of glioma cells to radiation. MBE *20*, 318–336.
41. Albers, D.J., Levine, M., Gluckman, B., Ginsberg, H., Hripcsak, G., and Mamykina, L. (2017). Personalized glucose forecasting for type 2 diabetes using data assimilation. PLoS Comput. Biol. *13*, e1005232.
42. Arthurs, C.J., Xiao, N., Moireau, P., Schaeffter, T., and Figueroa, C.A. (2020). A flexible framework for sequential estimation of model parameters in computational hemodynamics. Adv. Model. Simul. Eng. Sci. *7*, 48.
43. D'Elia, M., Mirabella, L., Passerini, T., Perego, M., Piccinelli, M., Vergara, C., and Veneziani, A. (2012). Applications of variational data assimilation in computational hemodynamics. In Modeling of Physiological Flows MS&A — Modeling, Simulation and Applications, D. Ambrosi, A. Quarteroni, and G. Rozza, eds. (Springer Milan), pp. 363–394.
44. Baker, R.E., Peña, J.M., Jayamohan, J., and Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol. Lett. *14*, 20170660.
45. Thall, P.F. (2019). Bayesian Utility-Based Designs for Subgroup-Specific Treatment Comparison and Early-Phase Dose

Optimization in Oncology Clinical Trials. JCO Precis. Oncol. *1–7*, 1–7.

46. Alber, M., Buganza Tepole, A., Cannon, W.R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W.W., Perdikaris, P., Petzold, L., and Kuhl, E. (2019). Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. NPJ Digit. Med. *2*, 115.
47. Viguerie, A., Grave, M., Barros, G.F., Lorenzo, G., Reali, A., and Coutinho, A.L.G.A. (2022). Data-Driven Simulation of Fisher-Kolmogorov Tumor Growth Models Using Dynamic Mode Decomposition. J. Biomech. Eng. *144*, 121001.
48. Ezhov, I., Mot, T., Shit, S., Lipkova, J., Paetzold, J.C., Kofler, F., Pellegrini, C., Kollovieh, M., Navarro, F., Li, H., et al. (2022). Geometry-Aware Neural Solver for Fast Bayesian Calibration of Brain Tumor Models. IEEE Trans. Med. Imaging *41*, 1269–1278.
49. Brummer, A.B., Xella, A., Woodall, R., Adhikarla, V., Cho, H., Gutova, M., Brown, C.E., and Rockne, R.C. (2023). Data driven model discovery and interpretation for CAR T-cell killing using sparse identification and latent variables. Front. Immunol. *14*, 1115536.
50. Nardini, J.T., Lagergren, J.H., Hawkins-Daarud, A., Curtin, L., Morris, B., Rutter, E.M., Swanson, K.R., and Flores, K.B. (2020). Learning Equations from Biological Data with Limited Time Samples. Bull. Math. Biol. *82*, 119.
51. Gaw, N., Hawkins-Daarud, A., Hu, L.S., Yoon, H., Wang, L., Xu, Y., Jackson, P.R., Singleton, K.W., Baxter, L.C., Eschbacher, J., et al. (2019). Integration of machine learning and mechanistic models accurately predicts variation in cell density of glioblastoma using multiparametric MRI. Sci. Rep. *9*, 10063.
52. Yuan, B., Shen, C., Luna, A., Korkut, A., Marks, D.S., Ingraham, J., and Sander, C. (2021). CellBox: Interpretable Machine Learning for Perturbation Biology with Application to the Design of Cancer Combination Therapy. Cell Syst. *12*, 128–140.e4.