Virtual clinical trials of anti-PD-1 and anti-CTLA-4 immunotherapy in advanced hepatocellular carcinoma using a quantitative systems pharmacology model

Richard J Sové [iD],[1] Babita K Verma,[1] Hanwen Wang [iD],[1] Won Jin Ho,[2,3] Mark Yarchoan,[2,3] Aleksander S Popel[1,2]

[1]Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[2]The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[3]Bloomberg-Kimmel Institute for Cancer Immunotherapy, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Correspondence to
Dr Richard J Sové;
rjsove@gmail.com

ABSTRACT

Background Hepatocellular carcinoma (HCC) is the most common form of primary liver cancer and is the third-leading cause of cancer-related death worldwide. Most patients with HCC are diagnosed at an advanced stage, and the median survival for patients with advanced HCC treated with modern systemic therapy is less than 2 years. This leaves the advanced stage patients with limited treatment options. Immune checkpoint inhibitors (ICIs) targeting programmed cell death protein 1 (PD-1) or its ligand, are widely used in the treatment of HCC and are associated with durable responses in a subset of patients. ICIs targeting cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) also have clinical activity in HCC. Combination therapy of nivolumab (anti-PD-1) and ipilimumab (anti-CTLA-4) is the first treatment option for HCC to be approved by Food and Drug Administration that targets more than one immune checkpoints.

Methods In this study, we used the framework of quantitative systems pharmacology (QSP) to perform a virtual clinical trial for nivolumab and ipilimumab in HCC patients. Our model incorporates detailed biological mechanisms of interactions of immune cells and cancer cells leading to antitumor response. To conduct virtual clinical trial, we generate virtual patient from a cohort of 5,000 proposed patients by extending recent algorithms from literature. The model was calibrated using the data of the clinical trial CheckMate 040 (ClinicalTrials.gov number, NCT01658878).

Results Retrospective analyses were performed for different immune checkpoint therapies as performed in CheckMate 040. Using machine learning approach, we predict the importance of potential biomarkers for immune blockade therapies.

Conclusions This is the first QSP model for HCC with ICIs and the predictions are consistent with clinically observed outcomes. This study demonstrates that using a mechanistic understanding of the underlying pathophysiology, QSP models can facilitate patient selection and design clinical trials with improved success.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Advanced hepatocellular carcinoma (HCC) has a poor prognosis with very limited treatment options. However, immune checkpoint inhibitors (ICIs) have shown remarkable advancement in cancer research and the combination of ICIs anti-PD-1 nivolumab and anti-CTLA-4 ipilimumab was approved for HCC. Despite great strides being made in the treatment of HCC, more than 50% of patients do not respond to treatment. Quantitative systems pharmacology (QSP) models are being increasingly used in both academic research and the pharmaceutical industry to develop a mechanistic understanding of the disease-drug interactions and ultimately improve patient outcomes to treatment.

WHAT THIS STUDY ADDS

⇒ This study proposes the first QSP model for HCC with ICIs. A modified virtual patient algorithm is applied that captures interpatient diversity in the clinical trial and predicts potential biomarkers using machine learning approach.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Clinical trials and drug discovery are expensive and time consuming. This framework of QSP model can be used to perform virtual clinical trials and help design clinical trials with improved efficacy by patient selection and best combination therapy.

INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common type of liver cancer, accounting for approximately 80% of primary liver malignancies.[1] HCC is caused by chronic infection of hepatitis B or C virus, excess alcohol consumption, diabetes, obesity, and ingestion of aflatoxin-contaminated food.[2] It is the third-leading cause of cancer-related death worldwide[3] and despite efforts to screen at-risk

individuals, approximately 80% of patients are diagnosed at late stages when curative intervention (ie, surgical resection or transplantation) is no longer possible.[4] Since 2020, the preferred first-line treatment for patients with advanced HCC (aHCC) with preserved liver function is the combination of atezolizumab (anti-PD-L1) and bevacizumab (anti-vascular endothelial growth factor). This combination is associated with improved overall survival, progression-free survival, and quality of life than the historical standard of care, sorafenib.[5 6] However, the prognosis for patients treated with modern systemic therapy continues to lag many other tumor types, with a median progression-free survival of 6.8 months and a medial overall survival of 19 months.[4] Multiple second line therapies were developed for patients receiving first line sorafenib therapy (including regorafenib, cabozantinib, and ramucirumab),[7–9] but the optimal management of patients progressing on first line bevacizumab and atezolizumab remains unclear.

There has been an increased interest in immunotherapy due to the success of a programmed cell death protein 1 (PD-1) blockade trial for the treatment of advanced melanoma.[10] This led to many successful immunotherapy trials for the treatment of various solid tumors, with the main targets for immune checkpoint blockade being PD-1, programmed cell death ligand 1 (PD-L1) and cytotoxic T lymphocyte-associated antigen 4 (CTLA-4).[11 12] This promotes antitumor response in T cells and prevents T cell exhaustion.[13] For HCC, the US Food and Drug Administration (FDA) has recently granted accelerated approval of immune checkpoint inhibitors pembrolizumab (anti-PD-1), nivolumab (anti-PD-1) and the combination of nivolumab and ipilimumab (anti-CTLA-4) as second line treatment based on the results of KEYNOTE-224[14] and CheckMate 040[15 16]

Nivolumab treatment for aHCC was first evaluated in CheckMate 040, a phase I/II study that included 48 patients in the dose-escalation phase and 214 patients in the dose-expansion phase; this study showed an objective response rate (ORR) of 20%.[15] The median overall survival times were 28.6 months for first-line therapy and 15 months for patients who were previously treated with sorafenib,[17] which led to the FDA approval for nivolumab as a second-line therapy for aHCC following the failure of sorafenib treatment. Nivolumab is currently being evaluated for first-line therapy in two phase III trials (CheckMate 459, NCT02576509 and CheckMate 9DX, NCT03383458). The combination of nivolumab and ipilimumab is currently being evaluated in HCC as second-line therapy (CheckMate 040, NCT01658878) and has been granted accelerated approval by the FDA for patients previously treated with sorafenib. This combination has shown promising results in other cancers starting with its success in melanoma.[18] Despite the success of this treatment, response rate is moderate; 33% at a median follow-up time of 28 months in the ongoing CheckMate 040 trial (NCT01658878).[19]

There is clearly a need for biomarker identification to increase the success rate of the treatment of aHCC.

Additionally, the complex biological mechanisms explaining the success of different combinations of drugs remain unknown.[20] Having a qualitative understanding of the interplay between the immune cells and cancer cells is not sufficient to fully optimize the treatment. This is where quantitative mechanistic mathematical and computational models can become useful.

Mathematical and computational modeling has been an essential tool in drug development for decades. In conventional pharmacokinetic/pharmacodynamic (PK/PD) modeling, the uptake, distribution and target interactions are modeled to create a link between the pharmacology and clinical outcomes. These models, however, do not take full advantage of existing mechanistic knowledge. Quantitative systems pharmacology (QSP) modeling integrates mechanistic knowledge of molecular and cellular interactions and PK/PD modeling to mechanistically link drug pharmacology to clinical outcomes. QSP is becoming an increasingly popular tool in both academia and industry for studying disease response to drugs. Pharmaceutical companies use QSP models to make decisions at all stages in the development cycle, and they may become standard practice and used in the approval by regulatory agencies. QSP models have shown to be useful in both prospective and retrospective analyses of immunotherapies in multiple cancer types.[21–28] Rule-based QSP model was developed by Abrams et al, to predict dose-response relationship for trispecific antibodies.[29] For a detailed review of QSP modeling in immuno-oncology (IO), see the recent review article by Chelliah et al.[30] While there are multiple computational models that exists for hepatic diseases[31–34] there is only one QSP model for liver, non-alcoholic fatty liver disease (NAFLDsym),[35 36] derived from DILIsym[37] that incorporates key components of steatosis and lipotoxicity in NAFLD patients. There is an unmet need for a detailed large-scale QSP model for HCC to search for novel therapies and design of clinical trials.

In this study, we develop a QSP model for aHCC using our platform, QSP-IO,[38] to simulate therapy with immune checkpoint inhibitors nivolumab and ipilimumab using virtual patients generated from the CheckMate 040 trial results of nivolumab monotherapy in aHCC.[15] We then use the model in a detailed analysis of the CheckMate 040 clinical trial for different dose regimens of the combination therapy with the objective of determining predictive biomarkers that could facilitate patient selection and ultimately improve the success of this treatment.

METHODS
Computational model overview
The proposed QSP model was developed using QSP-IO, our modular MATLAB (MathWorks, Natick, MA) SimBiology-based platform for developing QSP models in IO.[38] The resulting model structure is similar to that of models introduced in our previous studies.[21–26 39] The model is a four-compartment model describing relevant
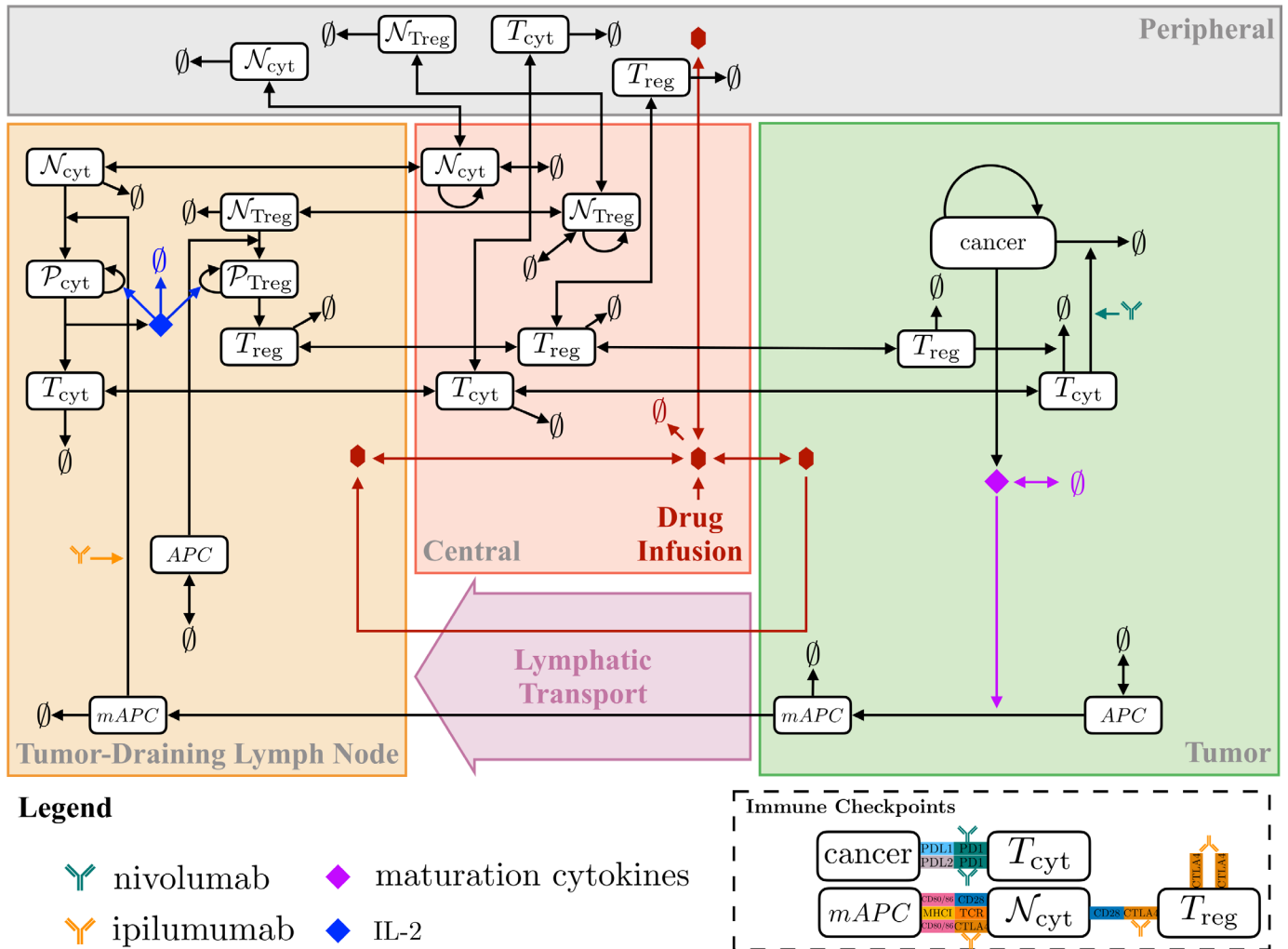
Figure 1 Diagram of the immuno-oncology QSP model interactions. APC, antigen presenting cells; mAPC, mature antigen presenting cells; MHC, major histocompatibility complex; N, naïve T cells; P, proliferating T cells; QSP, quantitative systems pharmacology.

immune interactions in tumor, tumor-draining lymph node (LN), central (blood), and peripheral (all other organs) compartments, thus representing the entire patient.

Our model considers logistic tumor growth dynamics with a tumor cell death-dependent release of antigens which are taken up by antigen-presenting cells (APCs), for example, dendritic cells. Antigens within dendritic cell endosomes are broken down and can bind major histocompatibility complex molecules, which translocate to the surface to be presented to naïve T cells. Naïve T cells are modeled by a zeroth order source term in the central compartment and can migrate to the other compartments. Naïve T cells in the LN compartment can interact with APCs to become activated where they undergo proliferation and become mature cytotoxic T cells that can migrate between compartments; the model also considers regulatory T cell (Treg) dynamics that is similar to that of the cytotoxic T cells. The model also incorporates immune checkpoints PD-1, PD-L1, PD-L2 and CTLA-4. Figure 1 summarizes the interactions considered in the model. For completeness and reproducibility,

all equations and parameter values are included in online supplemental information, and the code used to produce the results can be found at www.github.com/popellab/qspio/.

The base model for HCC consists of 69 ordinary differential equations, and 116 parameters plus 12 parameters related to each of the immune checkpoint inhibitor for a total of 140 parameters. Physiologically realistic parameter ranges were estimated from the literature; the details for each parameter are given in online supplemental table S6.

Parameter selection and virtual patient generation

We generated virtual patients following the methods by Allen *et al*[40] with minor modifications to their original approach described below. In their method, Allen *et al* randomly generated a parameter set such that the values of all parameters were physiologically realistic. Next, they used simulated annealing (SA), a Markov Chain Monte Carlo optimization algorithm, to ensure that the simulation outputs from the parameter set were also within physiologically realistic range. This was accomplished using an

objective function that evaluated the squared Euclidean distance from the physiological output range; such that the contribution of the simulation outputs would be zero to the objective function if the outcome is within the physiological range. This reduces the computational cost of minimization since the optimization is not to specific points but within the plausible range. This process was repeated to get a large number of plausible patients; the final virtual patient population was then chosen such that the distribution of selected outputs matched that of the data. Selection was done using a Bayesian-based rejection sampling procedure[40] where the probability of inclusion of a plausible patient as virtual patient is given by

$$P(r) = \beta \frac{f(r)}{\rho_5(r)} \quad 12 \backslash^* \text{ MERGEFORMAT ()}$$

where $f(r)$ is the empirical probability density function of the clinical data, $\rho_5(r) = \frac{5}{V_5(r)}$ approximates the density of plausible patients (as number of patients per unit volume) at the output value, r, using $V_5(r)$ which is the volume of a hypersphere whose radius is the distance to the 5th nearest neighbor. The proportionality factor, β, was fit using SA with the Kolmogorov-Smirnov (KS) statistic from a two-sample univariate KS test as the objective function.

In contrast, we used Latin Hypercube Sampling (LHS) to sample the parameter space using different estimated distributions for each of the parameter with descriptive statistics adjusted within the physiological ranges. The details of the parameters used for virtual patient generation are shown in online supplemental table S3. Out of 140 parameters, 89 parameters were selected to be sampled based on their relevance to HCC and to account for the biological mechanism of interaction of the drugs nivolumab (anti-PD-1) and ipilimumab (anti-CTLA-4) in the tumor. The detailed description of model parameters and their relevance has been highlighted in online supplemental tables S3 and S6. There are a total of 17 parameters that are cancer-specific out of which 16 were included in the sampling for virtual patient population to capture variability across HCC patients. Simulations that did not reach the initial tumor diameter during the initialization procedure (7%), failed to converge due to numerical instabilities (3%), or whose T cells densities were outside of a normal range for patients with HCC (18%) were rejected from the plausible patient population. To estimate the plausible patient cohort from proposed patient, we accept or reject the plausible patient as virtual patient based on physiological T cell densities in aHCC. Activated CD8+T cell density in the tumor was in the range of 0–63,000 cell/μL[41] and CD4+FoxP3+ T cell density in the blood was 0–50 cell/μL.[42] For our analysis we assumed two-fold CD4+ cell density in the blood rather than the range reported in the literature.[42] This is due to the small sample size of the patients analyzed and to account for patient variability. In the study by Kalathil et al,[42] blood samples of only 19 patients with aHCC were analyzed. T cell density in the tumor was calculated from area densities in immunohistochemistry images following the procedure in the study by Mi et al..[43]

To determine the virtual patient population, we employed the tumor size data of target lesion at individual patient level from the waterfall plot of CheckMate 040 in Yau et al.[16] The best percent change in tumor diameter of the plausible patient population was compared with data obtained from the CheckMate 040 clinical trial for patients with aHCC treated with nivolumab and ipilimumab.[16] Rejection sampling was used to determine the final virtual patient population using equation 1 with given a priori distribution of plausible population to select from, as done by Allen et al. Since any given value of β can result in different virtual populations, the objective function was the mean KS statistic out of five virtual populations generated from β. A graphical representation of the algorithm is presented in figure 2.

Virtual clinical trial

The virtual clinical trial consisted of six treatment groups: no drugs, 3 mg/kg of nivolumab given every 2 weeks starting at $t=0$,[15] 1 mg/kg of ipilimumab given every 3 weeks starting at $t=0$ and three groups for the combination regimes of nivolumab and ipilimumab corresponding to the three arms (A, B, C) in the clinical trial, CheckMate 040. In arm A, 1 mg/kg of nivolumab and 3 mg/kg of ipilimumab were administered every 3 weeks (for 4 doses), followed by 240 mg of nivolumab every 2 weeks (flat dose). In arm B, 3 mg/kg of nivolumab and 1 mg/kg of ipilimumab were administered every 3 weeks (for four doses), followed by 240 mg of nivolumab every 2 weeks (flat dose). In arm C, 3 mg/kg of nivolumab was administered every 2 weeks with 1 mg/kg of ipilimumab administered every 6 weeks.[16]

Response Evaluation Criteria in Solid Tumors (RECIST)

The analyses in this study were based on the RECIST V.1.1 classification,[44] presented here for convenience. RECIST provides a means for classifying solid tumors based on their response to treatment; the classification is determined by the percent change in tumor diameter compared with baseline/smallest diameter recorded. Tumor diameter is defined as the sum of the longest diameters for each lesion of interest. In this study, we consider only one lesion and calculate its diameter based on the tumor volume assuming the tumor is spherical. The baseline tumor diameter is above 10 mm for the lesions to be measurable.

To perform the virtual patient evaluation by RECIST V.1.1 criteria, we assessed tumor diameter every 6 weeks with the first measurement at week 6 and baseline at time 0. The frequency of tumor measurement (ie, 6 weeks) corresponds patient evaluation in the CheckMate 040 trial. We assume tumor diameter less than or equal to 2 mm as complete disappearance of tumor. The criteria are summarized as follows. Complete response (CR) is characterized by a complete disappearance of the tumor with confirmation at consecutive assessment of 6 weeks. Partial
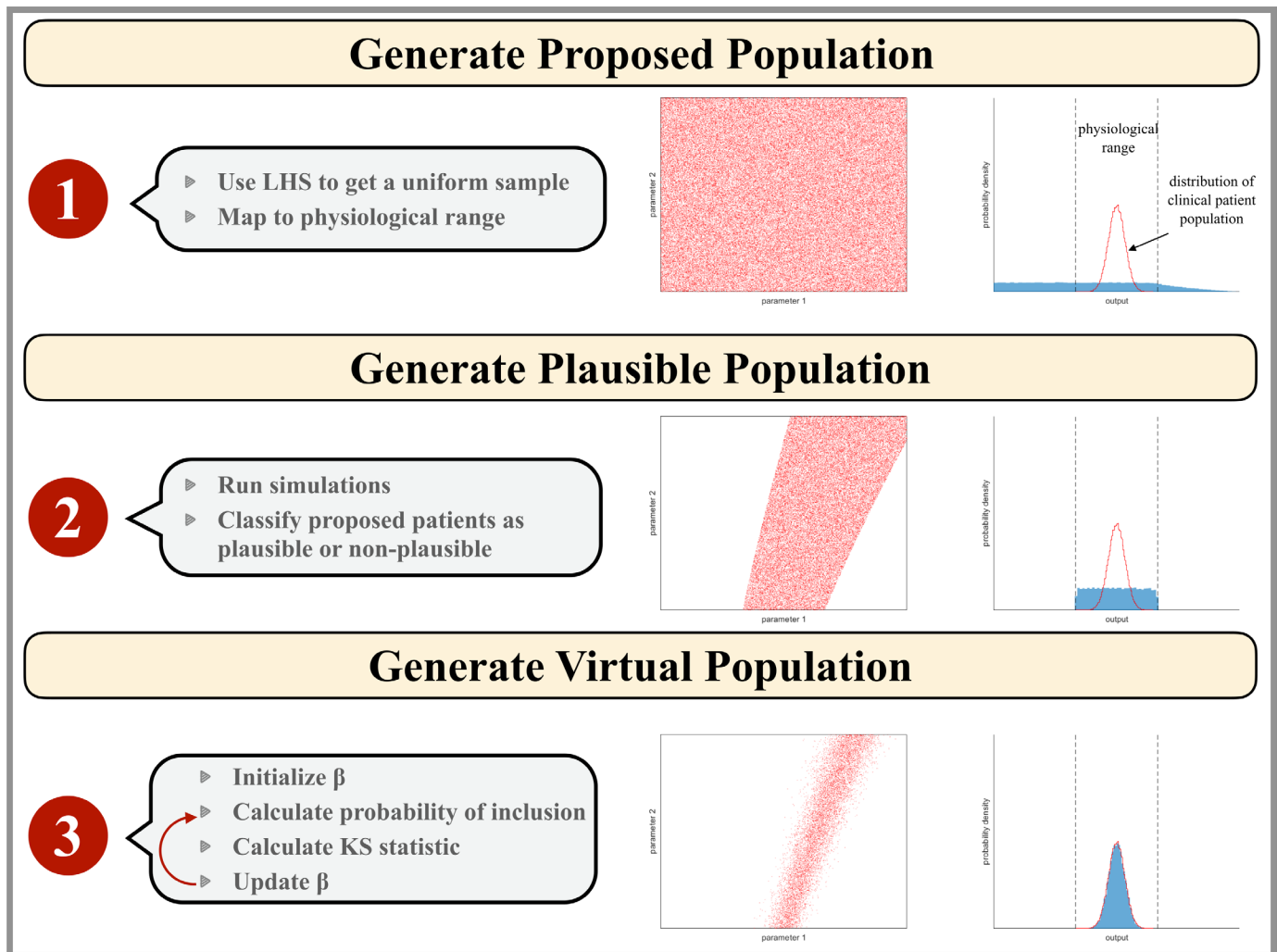
Figure 2 Illustration of the virtual patient generation algorithm. Input parameter values are represented as red points in a two-dimensional parameter space which would correspond to a hyperspace for our model analysis. The blue histogram represents the model outputs that correspond to the clinical patient data available. The physiological range of output values is outlined by black dashed lines and the empirical distribution of the clinical population data is shown as a red line. Step 1: the proposed patient population is generated by mapping a Latin Hypercube Sampling (LHS) on the interval (0–1) to parameter space assuming a physiological parameter range for each parameter. Step 2: the plausible patient population is the subset of the proposed patient population whose simulation outputs are physiologically realistic. Step 3: the virtual patient population is a subset of the plausible patient population such that the distribution of virtual patient outputs reflects that of the clinical patient population.

response (PR) is defined as percent change between −30% and −100%, with reference as the baseline tumor diameter, PR is confirmed at the subsequent assessment. Progressive disease (PD) is characterized by an increase of at least 20% in the tumor diameter with reference to the smallest measured diameter; in addition, there must be an absolute increase of 5 mm in the tumor diameter. Stable disease (SD) is characterized by neither sufficient decrease nor increase in the tumor diameter to be classified as PR or PD, with smallest tumor diameter as the reference. We assume a minimum duration of 6 weeks for stable disease as followed in the clinical study. The overall response rate is calculated as the proportions of virtual patients with complete and PR. In the virtual clinical trial, we followed the above criteria, to closely mimic the clinical situation.

Random forest

In this study, we used random forest, a supervised machine learning technique, to assess the importance of biomarkers in determining whether a patient will respond to treatment. Briefly, random forests are an ensemble of decision trees generated using a subset of the predictors and a bootstrapped sample of the data; the decision of the random forest is the majority decision of the decision trees. The TreeBagger function in MATLAB was used to train the random forests. The number of trees was determined by using the out-of-bag (OOB) error, an estimate of the prediction error. OOB error converges as the number of trees are increased; online supplemental figure S6 shows the convergence of the OOB error for 1000 trees.

The random forest was trained using the simulation results from the QSP model of HCC for all the virtual patients treated with the combination of nivolumab and ipilimumab. The accuracy of the random forest was assessed using the simulation results from 50 additional randomly generated virtual patients; only the results of successful simulations were considered. The importance of the predictors was estimated by randomly permuting all the observations of the predictor and calculating the influence of the predictor on the model's ability to predict patient response. This is performed by finding the change in OOB error which would be affected on permutating the values of the predictor. Predictors with the largest changes in OOB error were considered to be important predictors, that is, biomarker in our application.

RESULTS

Five thousand parameterizations of the model were generated by varying 89 of the 140 parameters by using a LHS that was mapped to a physiological range. Details of the parameterizations are given in online supplemental table S1. Each model parameterization was simulated six times: once without treatment, once with nivolumab treatment, once with ipilimumab treatment and three times with the combination of nivolumab and ipilimumab for the three arms with different dose regimen (see methods) in CheckMate 040 clinical trial. All our simulations were performed in MATLAB SimBiology Toolbox using the ode15s solver.

Unlike actual clinical trials, in virtual clinical trials the same patient can be treated with different treatments. To predict the outcome of all the treatments, we generate a cohort of virtual population from the initial 5000 proposed patients using the clinical information of the patients with HCC and data from the CheckMate 040 trial, for the combination therapy of nivolumab and ipilimumab. Of the 5000 parameter sets, 3622 sets were successful in all six treatment groups, giving us a large set to perform virtual trials. The remaining proposed patients either did not reach the initial tumor size or failed to converge or their T cell count was outside the normal range of that in patients with HCC. These parameter set are the common proposed patients that are used as input for determination of the plausible patients based on the physiologically reasonable range of T cell counts in HCC patients. Out of these proposed patients, 1747 were identified as plausible patients common to all the treatment groups. The final virtual cohort was generated using the information of the best percent change in tumor diameter for the combination treatments of all the three arms of CheckMate 040.[16] We compare the distribution of the best percent change in tumor diameter from the simulations of the combination treatment arm A to combination treatments of CheckMate 040. We chose arm A since among the three arms in CheckMate 040 with different dose regimens, only arm A was approved by the FDA.[45] The distribution of the plausible patient population which were used as
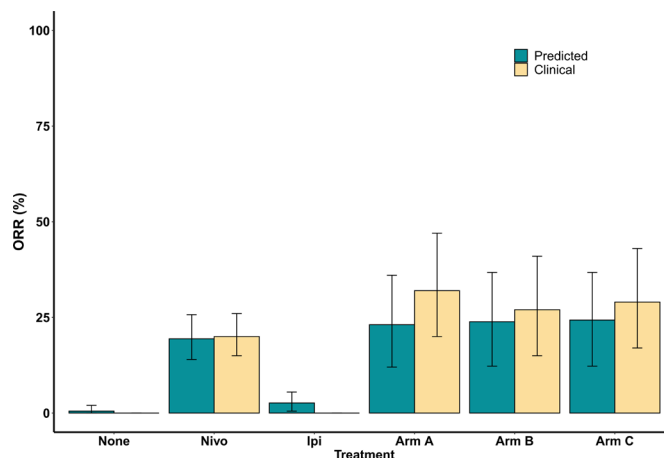

Figure 3 Overall objective response rate (ORR) comparison between model prediction and clinical trial CheckMate 040 results for all the treatment groups.

initial virtual population was significantly different from that of CheckMate 040 (p<0.05; KS=0.28). After fitting, the distribution of final virtual patient population was the same as that of CheckMate 040 as assessed by the KS goodness-of-fit test (p<0.05; KS=0.004) with beta=1080.5.

For each treatment group, we calculated the response rate and classified the plausible and virtual patients into four categories according to RECIST V.1.1: CR, PR, stable disease (SD), and PD. The overall response rate for monotherapy of nivolumab, ipilimumab and combination therapy of nivolumab plus ipilimumab with dose regimen of arm A for plausible patient cohort was 16.05% (95% CI 11% to 22%), 3.29% (95% CI 1% to 6%), and 19.62% (95% CI 10% to 32%), respectively. The corresponding response rate for virtual patient cohort was 19.41% (95% CI 14% to 26%), 2.64% (95% CI 1% to 6%), and 23.08% (95% CI 12% to 36%), respectively. The overall response rate reported in CheckMate 040[15] for monotherapy of nivolumab (dose expansion phase, 3 mg/kg) and combination of nivolumab and ipilimumab with arm A regimen was 20% (95% CI 15% to 26%) and 32% (95% CI 20% to 47%), respectively. This shows an agreement between the model prediction and the clinical results of CheckMate 040 trial. The response of virtual cohort is improved as compared with that of the plausible cohort implying that the virtual cohort is a close representation of the patients in the clinical trial. The comparison of the response rate predicted by the model with the clinical results of all the treatment groups is shown in figure 3. Notably, there has not been a clinical trial for monotherapy of ipilimumab in HCC. Here, the 95% CIs were calculated using bootstrap sampling of the virtual patients. The bootstrap sampling size was set to the number of patients enrolled in the respective clinical study with the assumption of 200 patients in both no treatment and monotherapy of ipilimumab. The corresponding two-sided CIs were estimated using the Clopper-Pearson method in the clinical trial.

In the above analysis, we followed the RECIST V.1.1 criteria of confirming the complete and partial responders

Table 1 Summary of the endpoint results of the virtual patients in the simulations and CheckMate 040 for monotherapy of nivolumab (dose expansion phase, 3 mg/kg) and combination treatment of nivolumab and ipilimumab (arm A)

	CheckMate 040	Simulations	CheckMate 040	Simulations
	Nivolumab		Combination: arm A	
No of patients	214	1365	50	1365
Complete response, no (%)	3 (1)	23 (1.7)	4 (8)	16 (1.2)
Partial response, no (%)	39 (18)	242 (17.7)	12 (24)	299 (21.9)
Stable disease, no (%)	96 (45)	618 (45.3)	9 (18)	673 (49.3)
Progressive disease, no (%)	68 (32)	482 (35.3)	20 (40)	377 (27.6)
ORR % (95% CI)	20 (15 to 26)	19.4 (14.0 to 25.7)	30.6 (21.8 to 48.9)	23.1 (12 to 36)

ORR, objective response rate.

at subsequent tumor measurement To compare we also assessed the response status without confirmation of complete and partial responder at subsequent tumor measurement. We found that the overall response for virtual cohort for monotherapy of nivolumab and combination treatment (arm A) was 23.1 and 27.8% (online supplemental table S2) which is higher than the abovementioned results. We suspect this is because some patients may relapse after initial response to the treatment and so at subsequent assessment they are not considered as responders.

The simulation outcomes of all the virtual patients were classified into different response status based on RECIST V.1.1. Our model can predict both partial and complete responders. The virtual patient cohort is common in all the treatment groups while the number of plausible patients varied. There were 1915 plausible patients in no treatment group out of which 1,365 were determined as virtual patients (620 PD; 738 SD; 7 PR; 0 CR); 2,143 plausible patients with nivolumab treatment with 1,365 virtual patients (482 PD; 618 SD; 242 PR; 23 CR); 1,916 plausible with ipilimumab treatment with 1365 virtual patients (485 PD; 844 SD; 36 PR; 0 CR); 2,171 plausible with arm A combination treatment with 1365 virtual patients (377 PD; 673 SD; 299 PR; 16 CR); 2,185 plausible patients for arm B combination treatment with 1,365 virtual patients (373 PD; 666 PD; 309 PR; 17 CR); and 2,193 with arm C combination treatment with 1,365 virtual patients (362 PD; 671 PD; 315 PR; 17 CR). Table 1 shows the comparison of this classification for the simulation results of nivolumab monotherapy and the combination of nivolumab and ipilimumab (arm A) with that of the clinical trial of CheckMate 040.[16] Online supplemental table S2 shows the results of virtual patients for the remaining four treatment groups.

The predictions from the simulations for the per cent changes in tumor diameters for the course of the treatment in each class (CR, PR, SD, and PD) are shown in figure 4. The solid line is the mean for individual class and the shaded region is the 95% prediction interval. Figure 5 shows the waterfall and spider plots for patients from the CheckMate 040 clinical trial, the virtual clinical trial of nivolumab only, the virtual clinical trial of ipilimumab only, and the virtual clinical trial of nivolumab in combination with ipilimumab for the three arms of CheckMate 040. Outcomes of the virtual patients from the simulations are consistent with those reported in the clinical trial.

To determine potential biomarkers that can predict if a patient will respond to treatment, we performed subgroup analysis using 11 measurable pretreatment biomarkers. Virtual patients were divided into subgroups based on the median value for each biomarker, and the ORR was calculated for each subgroup (figure 6). Based on the analysis of the combination therapy, the biomarkers responsible for the largest changes in ORR were T cell density, Treg density and antigen binding affinity. The analyses for nivolumab and ipilimumab as monotherapy and as combination therapy for dose regimen of arm B and C are given in the online supplemental information and figures S1–S4, respectively.

A random forest consisting of 1,000 trees using the 11 biomarkers as predictors was trained using the simulation results from the combination therapy arm A (1,365 virtual patients). The prediction error was 22.9%, estimated by the OOB error (online supplemental figure S6). To get a more accurate assessment of the prediction error, a test set of 50 new virtual patients was generated. The random forest was able to correctly classify 95.7% of the patients (4.3% false positive rate and 0% false negative rate) resulting in a prediction error of 4.3%. The importance of each biomarker was assessed by calculating the mean increase in OOB error caused by the random permutation of all the observation of each predictor; the results are shown in figure 7. A similar assessment was performed for all the treatment groups to identify the important biomarkers (online supplemental figure S5).

A second random forest (1,000 trees; OOB error 22.7%) was created using the six most important biomarkers in all the treatment groups: cytotoxic T cell density, Treg density, T cell to Treg ratio, number of tumor-specific T cell clones, PD-L1 expression, and exhausted T cell density. This forest was able to predict the classification of 91.3% of the patients (6.5% false positive rate and
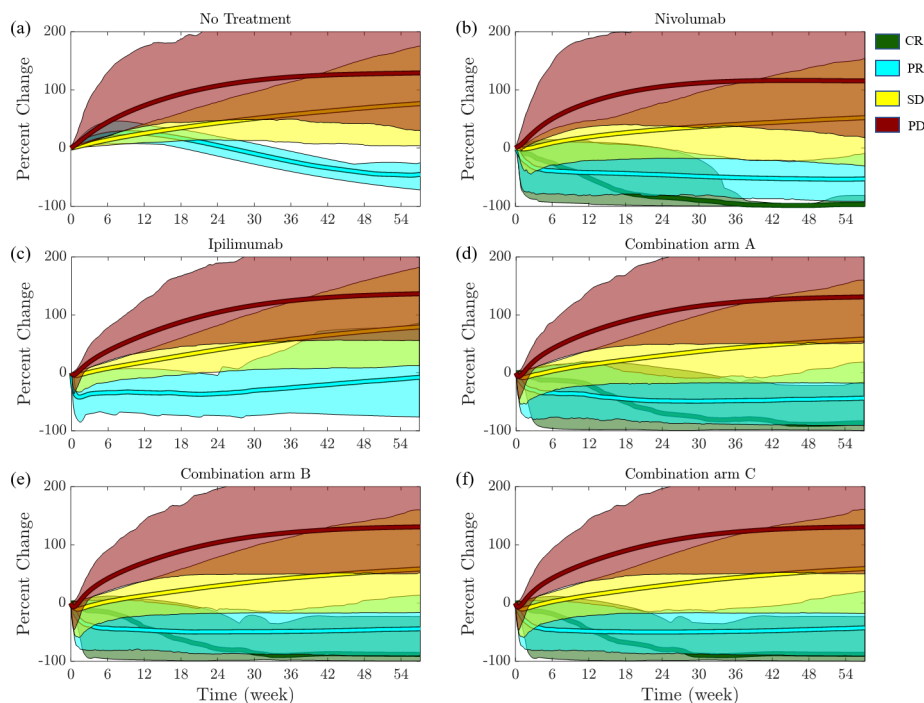
Figure 4 Mean per cent change in tumor diameter as a function of the time after the beginning of treatment for virtual patients. Simulations were grouped according to RECIST criteria; progressive disease (PD; $x>20\%$; red), stable disease (SD; $-30\% \leq x \leq 20\%$; yellow), partial response (PR; $x < -30\%$; cyan), and complete response (CR; $x = -100\%$; green). Shaded regions correspond to the 95% CI of the simulations in the corresponding RECIST category. RECIST, Response Evaluation Criteria in Solid Tumors.

2.2% false negative rate) resulting in a prediction error of 8.7%. The convergence of OOB error with the number of trees used to generate the random forest for all the treatment groups is shown in online supplemental figure S6. The pretreatment and post-treatment distribution of the top six predictive biomarkers estimated by the random forest for combination therapy of nivolumab and ipilimumab (arm A) is shown in figure 8. This shows that virtual patients with high T cell count and high T cell to Treg ratio respond better to the treatment. These can potentially be important predictive biomarkers in aHCC patients for anti-PD1 and anti-CTLA4 immunotherapy.

DISCUSSION

In this study, we presented an analysis of the treatment of aHCC using PD-1 and CTLA-4 immune checkpoint inhibitors, nivolumab and ipilimumab, using a QSP model of aHCC. To our knowledge this is the first QSP model of HCC. We also presented our procedure for virtual patient generation which extends from the work of Allen *et al*.[40] We then applied our virtual patient generation procedure to retrospectively predict the clinical trial results of the combination of nivolumab and ipilimumab for the treatment of aHCC. Additionally, we used a machine learning approach to quantify the importance of potential biomarkers.

The FDA has approved nivolumab and ipilimumab combination therapy for the treatment of several cancers including metastatic melanoma, advanced renal cell

carcinoma (RCC), colorectal cancer (CRC), HCC, non-small cell lung cancer (NSCLC) and, most recently, malignant pleural mesothelioma. This combination immunotherapy had ORR of 50% in melanoma (Check-Mate 067[46]) 42% in RCC (CheckMate 214[47]), 49% in CRC (CheckMate 142[48]), 32% in HCC (CheckMate 040); and 36% in NSCLC (CheckMate 227[49]). Simulating the treatment with our mechanistic QSP model of HCC, we predicted that the combination therapy would have an ORR of 27.8% (arm A) which falls within the confidence interval (21.8%–48.9%) of the results of the CheckMate 040 clinical trial[16] for HCC providing partial validation that modeling approach is predictive and cancer-specific. To further test the predictability of our model, we also conducted a preliminary analysis for monotherapy of pembrolizumab on 1,000 proposed patients in accordance with the clinical trial, Keynote-224.[14] The simulated per cent change in tumor diameter over the course of treatment in each class of patients (based on RECIST V.1.1) is shown in online supplemental figure S7. Our prediction for overall response rate was 16.6% which is close to the response rate of 17% as observed in Keynote-224 (online supplemental table S3). This shows that such detailed mechanistic models can be used to design new combination therapies for HCC and the optimum dose by predicting the efficacy of different dosing strategies.

Virtual patient generation has become an active area of research as QSP becomes a more prominent tool in the pharmaceutical industry. Despite the mechanistic detail
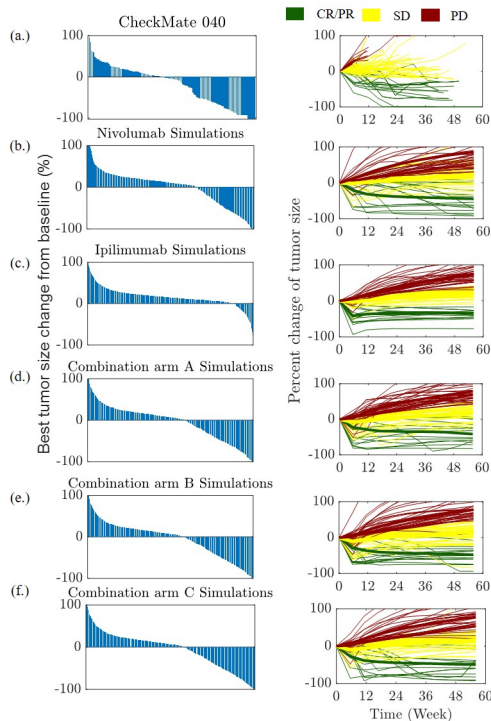
Figure 5 Waterfall (left) and spider plots (right) are shown for clinical trial CheckMate 040 and as predicted by the proposed QSP model for HCC. (A) CheckMate 040, (B) monotherapy of nivolumab, (C) ipilimumab monotherapy, combination treatments for (D) arm A (E) arm B (F) arm C. The waterfall plots show the best percent change in tumor diameter from baseline for each patient in the trial. The spider plots show the per cent change in tumor diameter at each measurement time point (every 6 weeks) for 100 randomly selected virtual patient. CR, complete response; HCC, hepatocellular carcinoma; PD, progressive disease; PR, partial response; QSP, quantitative systems pharmacology; SD, stable disease.

included in the models, the population level statistics are highly sensitive to virtual patient generation. There are various approaches for virtual patient generation. A comprehensive study of methodological considerations for virtual population in QSP model was done by Cheng et al.[50] A simple approach is to select input parameters from a known distribution.[21 23–26 51] This approach has the benefit of both being simple to implement, and computationally inexpensive. The drawback being that the distribution of outputs may not reflect the clinical population. Another approach is to assign weight to virtual patients such that the distribution reflects the clinical population.[52–54] These methods tend to be more computationally expensive, but have the benefit of accurately representing the population level statistics. One potential drawback is that this approach limits the diversity of the virtual patient population.[55] A third approach is to generate physiologically plausible patients and use rejection sampling to fit the population level statistics.[40 55] This approach has the benefit of both being computationally efficient and reduces the likelihood that unlikely patients get overrepresented in the virtual population.[40 55]

The approach presented in this work follows that of Allen et al[40] with some modifications. The main differences are in the way our plausible patients are generated. In their study, Allen et al randomly chose a physiologically realistic point in parameter space, and then optimized that point so that the outputs are also physiologically realistic. In our approach, we randomly choose a physiologically realistic point in parameter space, then either accept or reject it based on the outputs of the solution. Although, we end up rejecting just over 35% of our simulations, we found that, for application of QSP models, our approach is computationally more efficient than using an optimization to generate plausible patients.

Rieger et al[55] expanded on the work by Allen et al[40] by exploring three other means of generating the plausible patient population by using either nested SA (NSA), a modified genetic algorithm (GA) for optimization or a modified Metropolis-Hastings (MH) algorithm to replace the original SA optimization step from Allen et al. In their study, the authors compared the algorithms on the basis of the efficiency of going from the plausible population to the virtual patient population (ie, number of proposed patients needed to reach a specified goodness-of-fit), the computational cost, the convergence of the virtual population to the clinical population and the diversity of the virtual patient population. They concluded that their NSA approach is the most efficient, but also the most computationally costly and the MH approach was computationally fastest, however, the final virtual population was less diverse than the other approaches.

Although we did not make direct comparisons to their approaches, our approach converged to the clinical population and had low computational cost. We also expect our approach to be less efficient since we used a rejection-based approach rather than an optimization routine to generate the plausible patient population from the proposed population.

One of the biggest challenges is the availability of clinical data that can be used to obtain our final virtual patient population. In this study, we used the best per cent change in tumor diameter, which is often provided in the form of a waterfall plot. Despite its availability, using tumor diameter poses a problem for predicting novel therapies. However, other population-level information could be extremely valuable, such as distribution of the densities of different T cell populations in the tumor, T cell densities in the blood, or the densities of other cell types in the tumor immune microenvironment.[56 57] Using pretreatment distributions could potentially be an effective way to create virtual patients for predictive studies.

Using subgroup analysis and machine learning on our simulation results, we identified six important biomarkers for the prediction of immunotherapy response in HCC. Based on our analyses, patients with high tumor infiltrating lymphocyte (cytotoxic T cells, Tregs and exhausted T cells) densities and high T cell to Treg ratio are more likely to respond to this treatment. Increased CD8+T cell infiltration has been associated with response
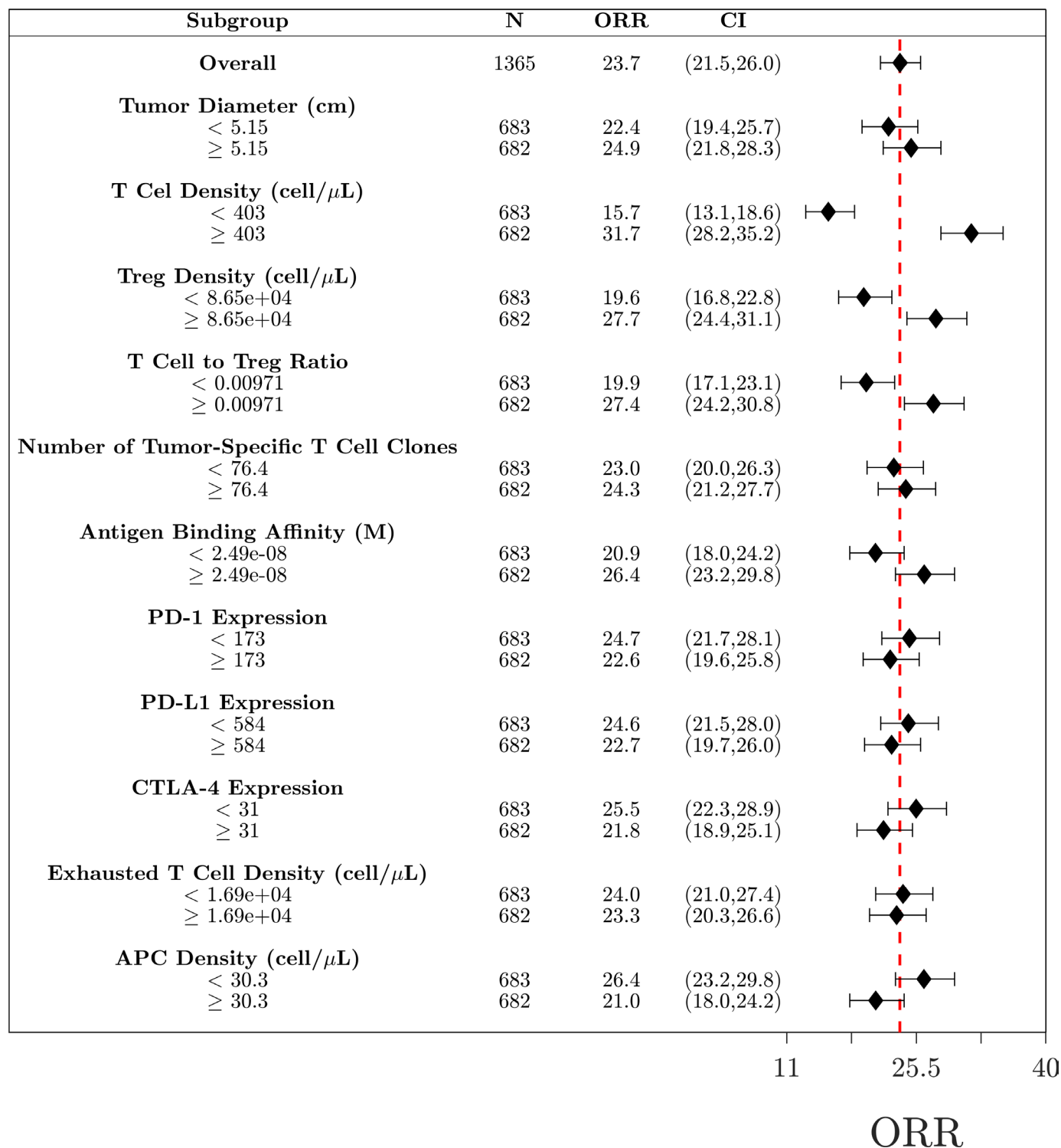
Figure 6 Subgroup analysis of the combination therapy. Virtual patients (N=1365) are divided into 22 subgroups based on the pretreatment values of selected biomarkers. Objective response rates (ORR) for each group are given along with the 95% CI estimated by the Agresti-Coull interval. The red dashed line indicated the ORR for the total population. APC, antigen-presenting cell.

to immune checkpoint inhibitor treatment in HCC[58–60] and in other cancers.[61 62] Infiltration of CD8+T cells and CD4+T cells has been reported in HCC in a recent study of communication landscape in tumor microenvironment.[57] PD-L1 expression may affect response to this therapy and their expression can be controlled by epigenetic modifications, genomic alterations, regulation at transcription, post-transcription and post-translation levels for better response rates.[63] Our subgroup analysis also shows that patients with high antigen binding affinity are more likely to respond to this treatment. Although no single biomarker has been confirmed for immune
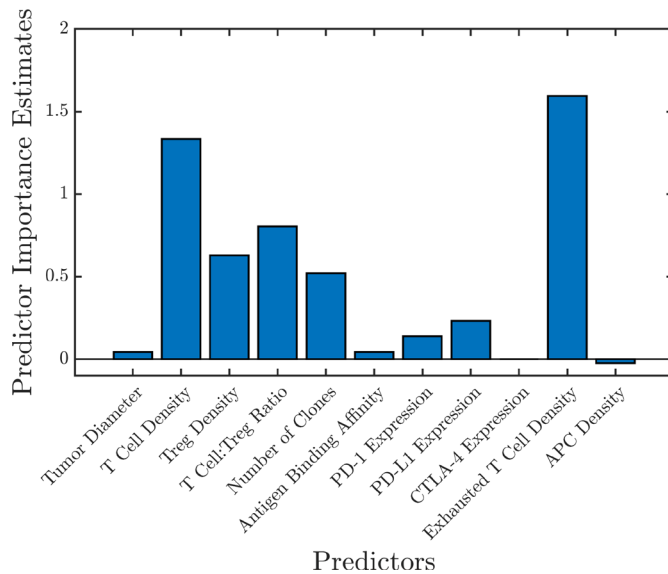
Figure 7 Predictor importance estimated by random forest using the mean increase in out-of-bag error caused by permuting the observations of each predictor. This shows the important predictive biomarkers for the combination therapy (arm A). APC, antigen-presenting cells.

checkpoint inhibitors in HCC,[64 65] several studies report positive correlation with response to immune checkpoint inhibitors in melanoma and lung cancer.[61 66–68]

There have been many efforts recently to incorporate machine learning into mechanistic modeling.[69–72] For example, Ribba *et al* used reinforcement learning with a mechanistic model of tumor inhibition in glioma response to chemotherapy and radiotherapy in application to precision dosing[69] Despite the lack of clinical data availability, QSP model simulations produce a large amount of data that can be difficult to analyze by using traditional techniques. In this study, we have demonstrated one way to incorporate machine learning to quantify the importance of potential biomarkers in the prediction of patient response. The approach presented here is only one possibility, however, with the availability of complex data analysis tools, there is huge potential for machine learning to adequately understand the complex state space and parameter space of these large mechanistic models.

QSP models are powerful, but their limitation is that they do not take into account spatial information, such that is provided by digital pathology or spatial transcriptomics.[57] Spatial heterogeneity is regarded a hallmark of cancer, and taking it into account would further enhance
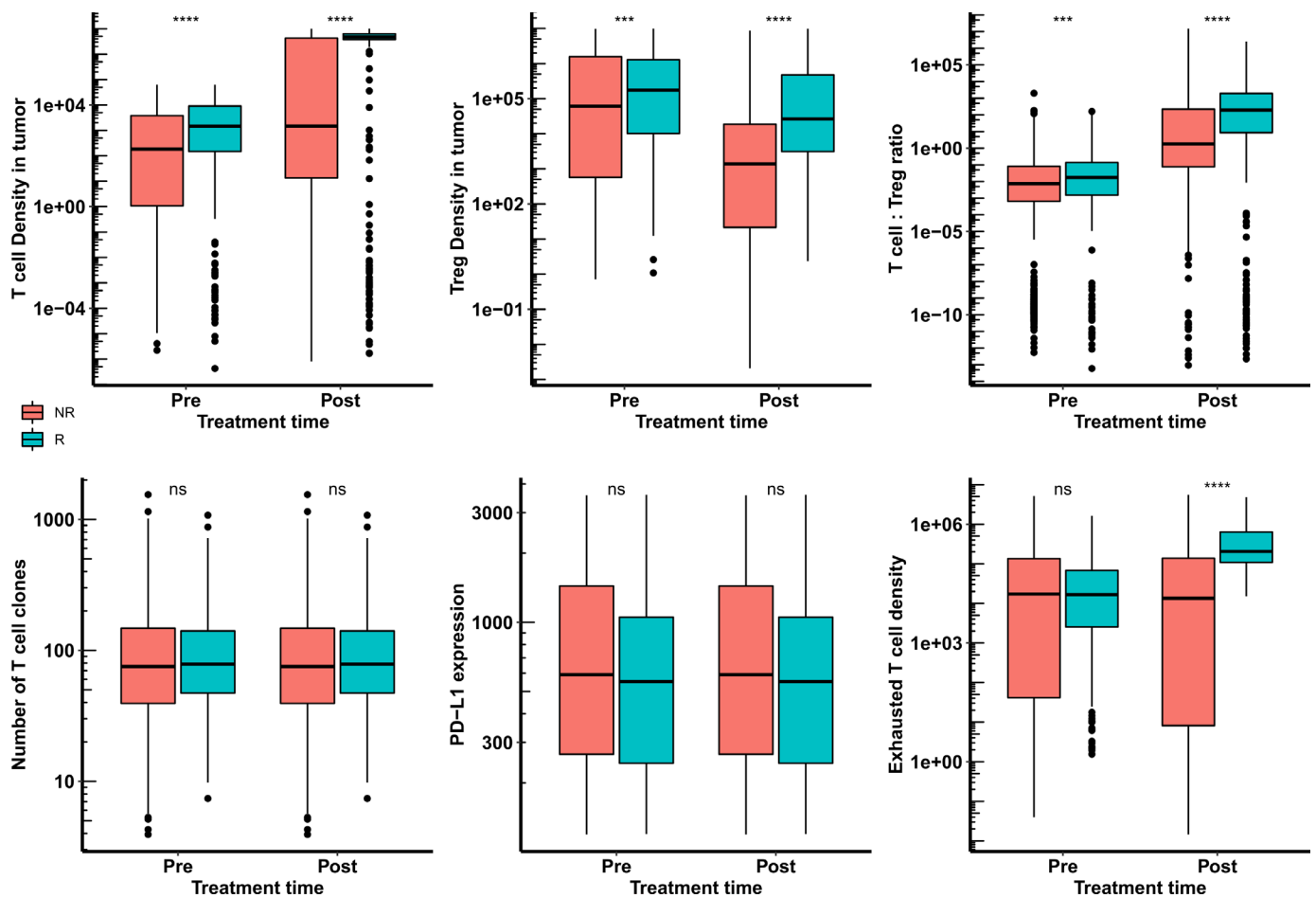

Figure 8 Distribution of potential biomarkers at pretreatment and post-treatment in responders and non-responders for the treatment of combination of nivolumab and ipilimumab (arm A) in HCC patients. Statistical significance is calculated using Wilcoxon test. HCC, hepatocellular carcinoma; ns, not significant; ***, p<=0.001; ****, p<=0.0001.

model capabilities, including biomarker identification. Stochastic discrete agent-based models and deterministic continuum partial differential equations-based models are the most common framework used to understand the heterogeneity in tumors.[73–76] We have developed a spatial platform, spQSP, that merges whole-patient QSP models with agent-based models of tumor with stochastic cell-level representation of the tumor microenvironment that is capable of incorporating results of digital pathology and spatial transcriptomic.[77–79]

In summary, QSP is a quickly developing field that has been shown to play a crucial role in drug development and QSP models are increasingly becoming a standard part of regulatory submissions.[80–82] The predictive power of QSP models is continually being demonstrated and constantly improving.[21 24–26] Further, the importance of virtual patient generation is increasingly becoming evident.[30] Finally, machine learning has a huge potential in the field of QSP, enabling detailed analyses and interpretation of complex state or parameter spaces typical of large QSP models.

ORCID iDs
Richard J Sové http://orcid.org/0000-0003-1491-616X
Hanwen Wang http://orcid.org/0000-0001-5480-431X

REFERENCES

1 Kaya E, Daniel Guillermo M, Yilmaz Y. Prevention of hepatocellular carcinoma and surveillance of high risk patients. *Hepatology Forum* 2022;3:33–8.
2 Rumgay H, Ferlay J, de Martel C, *et al*. Global, regional and national burden of primary liver cancer by subtype. *Eur J Cancer* 2022;161:108–18.
3 Sung H, Ferlay J, Siegel RL, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.
4 Boland P, Wu J. Systemic therapy for hepatocellular carcinoma: beyond sorafenib. *Chin Clin Oncol* 2018;7:50.
5 Finn RS, Qin S, Ikeda M, *et al*. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med* 2020;382:1894–905.
6 Casak SJ, Donoghue M, Fashoyin-Aje L, *et al*. FDA approval summary: Atezolizumab plus bevacizumab for the treatment of patients with advanced unresectable or metastatic hepatocellular carcinoma. *Clin Cancer Res* 2021;27:1836–41.
7 Abou-Alfa GK, Meyer T, Cheng A-L, *et al*. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. *N Engl J Med* 2018;379:54–63.
8 Bruix J, Qin S, Merle P, *et al*. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2017;389:56–66.
9 Zhu AX, Kang Y-K, Yen C-J, *et al*. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol* 2019;20:282–96.
10 Hodi FS, O'Day SJ, McDermott DF, *et al*. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010;363:711–23.
11 Topalian SL, Taube JM, Pardoll DM. Neoadjuvant checkpoint blockade for cancer immunotherapy. *Science* 2020;367:eaax0182.
12 Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science* 2018;359:1350–5.
13 Makaremi S, Asadzadeh Z, Hemmat N, *et al*. Immune checkpoint inhibitors in colorectal cancer: challenges and future prospects. *Biomedicines* 2021;9:1075.
14 Zhu AX, Finn RS, Edeline J, *et al*. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. *Lancet Oncol* 2018;19:940–52.
15 El-Khoueiry AB, Sangro B, Yau T, *et al*. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* 2017;389:2492–502.
16 Yau T, Kang Y-K, Kim T-Y, *et al*. Efficacy and safety of nivolumab plus ipilimumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib: the CheckMate 040 randomized clinical trial. *JAMA Oncol* 2020;6:e204564.
17 Crocenzi TS, El-Khoueiry AB, Yau TC, *et al*. Nivolumab (nivo) in sorafenib (sor)-naive and -experienced pts with advanced hepatocellular carcinoma (HCC): CheckMate 040 study. *Journal of Clinical Oncology* 2017;35:4013.
18 Larkin J, Chiarion-Sileni V, Gonzalez R, *et al*. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015;373:23–34.
19 AR H, Yau T, Hsu C. Nivolumab (NIVO) + ipilimumab (IPI) combination therapy in patients (PTS) with advanced hepatocellular carcinoma (aHCC): subgroup analyses from CheckMate 040. *Journal of Clinical Oncology* 2020;512.
20 Lemaire V, Shemesh CS, Rotte A. Pharmacology-based ranking of anti-cancer drugs to guide clinical development of cancer immunotherapy combinations. *J Exp Clin Cancer Res* 2021;40:1–17.
21 Wang H, Milberg O, Bartelink IH, *et al*. In silico simulation of a clinical trial with anti-CTLA-4 and anti-PD-L1 immunotherapies in metastatic breast cancer using a systems pharmacology model. *R Soc Open Sci* 2019;6:190366.
22 Milberg O, Gong C, Jafarnejad M, *et al*. A QSP model for predicting clinical responses to monotherapy, combination and sequential therapy following CTLA-4, PD-1, and PD-L1 checkpoint blockade. *Sci Rep* 2019;9:1–17.
23 Jafarnejad M, Gong C, Gabrielson E, *et al*. A computational model of neoadjuvant PD-1 inhibition in non-small cell lung cancer. *Aaps J* 2019;21:1–14.

24 Wang H, Sové RJ, Jafarnejad M, *et al*. Conducting a virtual clinical trial in HER2-negative breast cancer using a quantitative systems pharmacology model with an epigenetic modulator and immune checkpoint inhibitors. *Front Bioeng Biotechnol* 2020;8:141.

25 Ma H, Wang H, Sové RJ, *et al*. Combination therapy with T cell engager and PD-L1 blockade enhances the antitumor potency of T cells as predicted by a QSP model. *J Immunother Cancer* 2020;8:e001141–11.

26 Wang H, Ma H, Sové RJ, *et al*. Quantitative systems pharmacology model predictions for efficacy of atezolizumab and nab-paclitaxel in triple-negative breast cancer. *J Immunother Cancer* 2021;9:e002100–14.

27 Kosinsky Y, Dovedi SJ, Peskov K, *et al*. Radiation and PD-(L)1 treatment combinations: immune response and dose optimization via a predictive systems model. *J Immunother Cancer* 2018;6:1–15.

28 Wang H, Zhao C, Santa-Maria CA, *et al*. Dynamics of tumor-associated macrophages in a quantitative systems pharmacology model of immunotherapy in triple-negative breast cancer. *iScience* 2022;25:104702.

29 Abrams RE, Pierre K, El-Murr N, *et al*. Quantitative systems pharmacology modeling sheds light into the dose response relationship of a trispecific T cell engager in multiple myeloma. *Sci Rep* 2022;12:10976.

30 Chelliah V, Lazarou G, Bhatnagar S, *et al*. Quantitative systems pharmacology approaches for Immuno-Oncology: adding virtual patients to the development paradigm. *Clin Pharmacol Ther* 2021;109:605–18.

31 Ashworth WB, Davies NA, Bogle IDL. A computational model of hepatic energy metabolism: understanding Zonated damage and steatosis in NAFLD. *PLoS Comput Biol* 2016;12:e1005105.

32 Liao Y, Davies NA, Bogle IDL. Computational modeling of fructose metabolism and development in NAFLD. *Front Bioeng Biotechnol* 2020;8:762.

33 Verma BK, Subramaniam P, Vadigepalli R. Modeling the dynamics of human liver failure post liver resection. *Processes* 2018;6:115.

34 Wang Y, Brodin E, Nishii K, *et al*. Impact of tumor-parenchyma biomechanics on liver metastatic progression: a multi-model approach. *Sci Rep* 2021;11:1–20.

35 Generaux G, Rieger TR, Howell BA. QSP modeling of liver AMPK activation using NAFLDsym is predicted to reduce steatosis in NAFLD patients. *Clin Pharmacol The* 2017;101:S65.

36 Siler SQ, Generaux GT, Howell BA. Using quantitative systems pharmacology modeling to understand the effects of acetyl CoA carboxylase (ACC) inhibition on liver and plasma triglycerides in a simulated population. *In Hepatology* 2018;68:448A.

37 Shoda LKM, Woodhead JL, Siler SQ, *et al*. Linking physiology to toxicity using DILIsym®, a mechanistic mathematical model of drug-induced liver injury. *Biopharm Drug Dispos* 2014;35:33–49.

38 Sové RJ, Jafarnejad M, Zhao C, *et al*. QSP-IO: a quantitative systems pharmacology toolbox for mechanistic multiscale modeling for Immuno-Oncology applications. *CPT Pharmacometrics Syst Pharmacol* 2020;9:484–97.

39 Ma H, Pilvankar M, Wang J, *et al*. Quantitative systems pharmacology modeling of PBMC-Humanized mouse to facilitate preclinical Immuno-oncology drug development. *ACS Pharmacol Transl Sci* 2021;4:213–25.

40 Allen RJ, Rieger TR, Musante CJ. Efficient generation and selection of virtual populations in quantitative systems pharmacology models. *CPT Pharmacometrics Syst Pharmacol* 2016;5:140–6.

41 Yarchoan M, Xing D, Luan L, *et al*. Characterization of the immune microenvironment in hepatocellular carcinoma. *Clin Cancer Res* 2017;23:7333–9.

42 Kalathil SG, Lugade AA, Miller A, *et al*. PD-1+ and Foxp3+ T cell reduction correlates with survival of HCC patients after sorafenib therapy. *JCI Insight* 2016;1:1–12.

43 Mi H, Gong C, Sulam J, *et al*. Digital pathology analysis quantifies spatial heterogeneity of CD3, CD4, CD8, CD20, and FOXP3 immune markers in triple-negative breast cancer. *Front Physiol* 2020;11:1–22.

44 Eisenhauer EA, Therasse P, Bogaerts J, *et al*. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.

45 Saung MT, Pelosof L, Casak S, *et al*. FDA approval summary: nivolumab plus ipilimumab for the treatment of patients with hepatocellular carcinoma previously treated with sorafenib. *Oncologist* 2021;26:797–806.

46 Larkin J, Chiarion-Sileni V, Gonzalez R, *et al*. Five-Year survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med* 2019;381:1535–46.

47 Motzer RJ, Tannir NM, McDermott DF, *et al*. Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *N Engl J Med* 2018;378:1277–90.

48 Overman MJ, Lonardi S, Wong KYM, *et al*. Durable clinical benefit with nivolumab plus ipilimumab in DNA mismatch Repair-Deficient/Microsatellite Instability-High metastatic colorectal cancer. *J Clin Oncol* 2018;36:773–9.

49 Hellmann MD, Paz-Ares L, Bernabe Caro R, *et al*. Nivolumab plus ipilimumab in advanced non-small-cell lung cancer. *N Engl J Med* 2019;381:2020–31.

50 Cheng Y, Straube R, Alnaif AE, *et al*. Virtual populations for quantitative systems pharmacology models. *Methods Mol Biol* 2022;2486:129–79.

51 Verma BK, Subramaniam P, Vadigepalli R. Model-based virtual patient analysis of human liver regeneration predicts critical perioperative factors controlling the dynamic mode of response to resection. *BMC Syst Biol* 2019;13:1–15.

52 Klinke DJ. Integrating epidemiological data into a mechanistic model of type 2 diabetes: validating the prevalence of virtual patients. *Ann Biomed Eng* 2008;36:321–34.

53 Schmidt BJ, Casey FP, Paterson T, *et al*. Alternate virtual populations elucidate the type I interferon signature predictive of the response to rituximab in rheumatoid arthritis. *BMC Bioinformatics* 2013;14:1–16.

54 Cheng Y, Thalhauser CJ, Smithline S, *et al*. QSP toolbox: computational implementation of integrated workflow components for deploying multi-scale mechanistic models. *Aaps J* 2017;19:1002–16.

55 Rieger TR, Allen RJ, Bystricky L, *et al*. Improving the generation and selection of virtual populations in quantitative systems pharmacology models. *Prog Biophys Mol Biol* 2018;139:15–22.

56 Ho WJ, Zhu Q, Durham J, *et al*. Neoadjuvant cabozantinib and nivolumab converts locally advanced HCC into resectable disease with enhanced antitumor immunity. *Nat Cancer* 2021;2:891–903.

57 Mi H, Ho WJ, Yarchoan M, *et al*. Multi-Scale spatial analysis of the tumor microenvironment reveals features of cabozantinib and nivolumab efficacy in hepatocellular carcinoma. *Front Immunol* 2022;13:1–16.

58 Ma L, Hernandez MO, Zhao Y, *et al*. Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* 2019;36:418–30.

59 Kaseb AO, Vence L, Blando J, *et al*. Immunologic correlates of pathologic complete response to preoperative immunotherapy in hepatocellular carcinoma. *Cancer Immunol Res* 2019;7:1390–5.

60 Haratani K, Hayashi H, Tanaka T, *et al*. Tumor immune microenvironment and nivolumab efficacy in EGFR mutation-positive non-small-cell lung cancer based on T790M status after disease progression during EGFR-TKI treatment. *Ann Oncol* 2017;28:1532–9.

61 Tumeh PC, Harview CL, Yearley JH, *et al*. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014;515:568–71.

62 Chen P-L, Roh W, Reuben A, *et al*. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov* 2016;6:827–37.

63 Yu X, Li W, Young KH, *et al*. Posttranslational modifications in PD-L1 turnover and function: from cradle to Grave. *Biomedicines* 2021;9:9.

64 Pinter M, Jain RK, Duda DG. The current landscape of immune checkpoint blockade in hepatocellular carcinoma: a review. *JAMA Oncol* 2021;7:113–23.

65 Rizzo A, Ricci AD. Pd-L1, TMB, and other potential predictors of response to immunotherapy for hepatocellular carcinoma: how can they assist drug clinical trials? *Expert Opin Investig Drugs* 2022;31:415–23.

66 Inoue H, Park J-H, Kiyotani K, *et al*. Intratumoral expression levels of *PD-L1*, *GZMA*, and *HLA-A* along with oligoclonal T cell expansion associate with response to nivolumab in metastatic melanoma. *Oncoimmunology* 2016;5:e1204507.

67 Riaz N, Havel JJ, Makarov V, *et al*. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* 2017;171:934–49.

68 Forde PM, Chaft JE, Smith KN, *et al*. Neoadjuvant PD-1 blockade in resectable lung cancer. *N Engl J Med* 2018;378:1976–86.

69 Ribba B, Kaloshi G, Peyre M, *et al*. A tumor growth inhibition model for low-grade glioma treated with chemotherapy or radiotherapy. *Clin Cancer Res* 2012;18:5071–80.

70 Ribba B, Dudal S, Lavé T, *et al*. Model-Informed artificial intelligence: reinforcement learning for precision dosing. *Clin Pharmacol Ther* 2020;107:853–7.

71 Gaw N, Hawkins-Daarud A, Hu LS, *et al*. Integration of machine learning and mechanistic models accurately predicts variation in cell density of glioblastoma using multiparametric MRI. *Sci Rep* 2019;9:1–9.

72 Zhang J, Petersen SD, Radivojevic T, *et al*. Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. *Nat Commun* 2020;11:1–13.

73 Gallaher JA, Massey SC, Hawkins-Daarud A, *et al*. From cells to tissue: how cell scale heterogeneity impacts glioblastoma growth and treatment response. *PLoS Comput Biol* 2020;16:e1007672.

74 Jenner AL, Smalley M, Goldman D, *et al*. Agent-based computational modeling of glioblastoma predicts that stromal density is central to oncolytic virus efficacy. *iScience* 2022;25:104395.

75 Massey SC, Rockne RC, Hawkins-Daarud A, *et al*. Simulating PDGF-Driven glioma growth and invasion in an anatomically accurate brain domain. *Bull Math Biol* 2018;80:1292–309.

76 Cess CG, Finley SD. Multi-scale modeling of macrophage-T cell interactions within the tumor microenvironment. *PLoS Comput Biol* 2020;16:e1008519.

77 Gong C, Ruiz-Martinez A, Kimko H, *et al*. A spatial quantitative systems pharmacology platform spQSP-IO for simulations of Tumor–Immune interactions and effects of checkpoint inhibitor immunotherapy. *Cancers* 2021;13:3751–33.

78 Zhang S, Gong C, Ruiz-Martinez A, *et al*. Integrating single cell sequencing with a spatial quantitative systems pharmacology model spQSP for personalized prediction of triple-negative breast cancer immunotherapy response. *Immunoinformatics* 2021;1-2:100002.

79 Ruiz-Martinez A, Gong C, Wang H, *et al*. Simulations of tumor growth and response to immunotherapy by coupling a spatial agent-based model with a whole-patient quantitative systems pharmacology model. *PLoS Comput Biol* 2022;18:e1010254.

80 Zineh I. Quantitative systems pharmacology: a regulatory perspective on translation. *CPT Pharmacometrics Syst Pharmacol* 2019;8:336–9.

81 Grimstein M, Yang Y, Zhang X, *et al*. Physiologically based pharmacokinetic modeling in regulatory science: an update from the U.S. food and drug administration's office of clinical pharmacology. *J Pharm Sci* 2019;108:21–5.

82 Bai JPF, Earp JC, Florian J, *et al*. Quantitative systems pharmacology: landscape analysis of regulatory submissions to the US food and drug administration. *CPT Pharmacometrics Syst Pharmacol* 2021;10:1479–84.