Supplementary Information for the manuscript

Integration of omics data sources to inform mechanistic modelling of immune-oncology therapies: a tutorial for clinical pharmacologists.

Georgia Lazarou*, Vijayalakshmi Chelliah*, Ben G. Small, Michael Walker, Piet H. van der Graaf, Andrzej M. Kierzek**

Certara QSP, Certara UK Limited, Sheffield, S1 2BJ, United Kingdom

*Equal contribution

** Corresponding author, andrzej.kierzek@certara.com

MINIMAL MODEL CALIBRATION USING OMICS DATA.

The minimal CIC model presented in this work can be parameterised for a particular human disease and therapy of interest. As our example application, we chose Non-Small Cell Lung Carcinoma (NSCLC) treatment with Pembrolizumab, an anti-PD1 monoclonal antibody. The aim was to calibrate the minimal CIC model to describe four individual NSCLC patients from the KEYNOTE-001 clinical trial for which tumour growth was measured during treatment with 10 mg/kg Pembrolizumab administered every 3 weeks (10 mg/kg Q3W)[1]. The calibrated QSP model could be used to predict variability of patient responses in full clinical trial of 489 patients, effects of other dosing regimens and to identify biological processes which can be pharmacologically targeted to enhance the effect of anti-PD1 treatment. We note that the purpose of our study is to demonstrate how omics data can be used to calibrate a QSP model, rather than to qualify a model for application in drug development.

Figure 3 in main manuscript summarises the experimental data used for model calibration. First, we used data from the "Immune Landscape of Cancer" (Figure 1 A,B in main manuscript) to constrain the relative amounts of antigen-presenting cells (APCs) and cytotoxic T-cells present in the

TME compartment (model variables APC_tme and Tc_tme). Since it is not possible to query the database for NSCLC, we averaged cell fractions available for two types of lung cancers: Lung Adenocarcinoma (LUAD) and Lung Squamous Cell Carcinoma (LUSC). For these cancers, the fraction of leukocytes in the TME is 0.14 (Leuk_frac parameter in our model); this quantity is further fractionated into 22 leukocyte types, for which the fraction of cytotoxic CD8 T-cells (Tc_frac parameter) was directly available (0.11), and to obtain the fraction of antigen-presenting cells (APC_frac parameter) we summed the fractions of M1 macrophages (0.055) and dendritic cells (0.042). Since we do not consider the remaining 20 cell types reported, we recalculate the ratio of cytotoxic cells to APCs assuming that the whole leukocyte fraction of the TME is composed only of these cell types. Therefore, the fractional compositions of cells relative to the TME were identified as 0.074 for cytotoxic T-cells (Tc_max parameter) and 0.066 for APCs (APC_max parameter). The proportion of the TME compartment composed of tumour cells (model variable Tumor) was therefore calculated as the remainder fraction of 0.86. These fractions were used in the model as maximal capacities in logistic rate law terms limiting both the migration of APCs and cytotoxic T-cells to the TME compartment (parameters Tc_max and APC_max, reactions v4 and v12, FigureS1), as well as the recruitment of TME capacity (parameter tumor_max, reaction v1, Figure S1). To set the numbers at the start of simulations, we multiplied the maximal TME fractions by the total number of cells corresponding to the specified initial tumour volume, where the total number of cells was calculated assuming all cell species have the same volume. Our method for defining initial states for given tumour volumes assumes that the tumours at the start of therapy are already infiltrated by the leukocytes up to maximal fractional TME capacity. Given that we do not simulate tumour development from a single cell and that solid tumours at the start of therapy would already be well developed, we believe that our assumption of maximal infiltration is biologically plausible.

Another key parameter of any QSP model of CIC is the number of pre-existing cancer neoantigen-specific naïve cytotoxic T-cells in the lymph node compartment (Tc_ln variable in our model). This can now be estimated using two recent datasets resulting from application of high-throughput

flow cytometry and high-throughput ex-vivo screening of cancer antigen-specific T-cells. In a series of studies from the Farber lab, high-throughput flow cytometry has been applied to organ donor tissues to quantitatively describe the state of the immune system in healthy individuals [2,3,4]. As flow cytometry studies are typically confined to peripheral blood whilst most of the key immune processes occur in tissues, these recent studies constitute an invaluable resource for many areas of basic and applied science, including mechanistic modelling of the immune system. From Thome et al. [5] we have determined the fraction of all cytotoxic T-cells in the lymph node which are naïve (0.44). Finding clinical data on the proportion of naïve cytotoxic T-cells which would also be specific to NSCLC neoantigens proved more difficult. We ultimately used a recent clinical study where high-throughput ex-vivo T-cell proliferation assays were used to estimate the frequency of naïve tumour neoantigen-specific CD4 T-cells in melanoma patients[6]. In doing so, we have assumed that the fundamental biological processes for creating the diversity of pre-existing naïve antigen specific T-cells are similar in CD4 and CD8 types, and similar across tumours. This assumption is corroborated by the fact that these frequencies were found to be similar across the antigens included in the study[6]. Moreover, our minimal model does not explicitly include CD4 cells as a state variable so their availability and function is implicitly accounted for in the cytotoxic T-cell parameters. This further justifies the use of CD4 data to constrain parameters representing combined CD4/CD8 activities. Bearing in mind these assumptions, we estimated the frequency of tumour neoantigen-specific naïve T-cells in the lymph node compartment as 1.85×10^{-5} (18.5 per million).

State variables describing cell populations in most of QSP models in IO, including the minimal model presented here, are expressed as absolute cell numbers rather than relative amounts. Given the sophistication of experimental methods used to obtain relative data, it is surprising that data on reference absolute leukocyte counts in human tissues are still sparse. Here, we used the estimates of absolute numbers of leukocytes in human tissues presented by Trepel in 1974[7], which have also been accepted as a reference in other work in quantitative immunology[8]. Trepel estimated the absolute number of cells in lymph nodes as 1.9×10^{11}. The fraction of CD8 events in T-cells isolated from excised

healthy human lymph nodes was found by Scott et al.[9] to be 0.13. By multiplication of the total number of lymphocytes in lymph nodes by all frequencies above, we estimate the number of pre-existing neoantigen-specific cytotoxic T-cells in the lymph node compartment to be 195,624.

Finally, we used clinical data for the individual growth trajectories of NSCLC tumours of patients treated with 10 mg/kg Pembrolizumab Q3W (shown in Figure 3) to calibrate parameters in the model that could not be defined using omics data[1]. These parameters correspond to the processes of tumour growth, its killing by cytotoxic T-cells, and drug action via the PD1 checkpoint. Therefore, our model describes—rather than predicts—the behaviour of the system for these four tumours subjected to this dosing regimen, and could be used to predict the behaviour of other tumours in full clinical trial, the effects of other dose regimens applied in the clinical trial (10 mg/kg Q2W, 2mg/kg Q3W), as well as sensitivity of dose response to perturbation of individual steps of the CIC, thus informing the discovery of combination targets. However, we emphasise that the model presented here has been built for the purpose of demonstrating data sources that can inform mechanistic modelling of immune-oncology therapies.

PARAMETERS VARIED IN VIRTUAL TRIAL SIMULATION

Figure 4 in main manuscript shows a virtual trial simulation for a cohort of 489 subjects treated with 19 doses of 10 mg/kg anti-PD1 Pembrolizumab administered every 3 weeks (simulation duration of t=399[days] with the first dose given at t=0; all patients are assigned a weight of 70 kg). We varied 9 model parameters representing the biological variability of tumours: tumour growth rate (mu), the rate of TME remodelling (kv1), basal death rate of tumour cells (beta_tumor), maximal rate of T-cell exhaustion by PD1 checkpoint (kPD1), the rate of tumour killing by cytotoxic T-cells (kkill), T-cell proliferation rate in the lymph node (Pro_Tc_ln), initial tumour size (Initial_tumour_volume), and the fractional composition of the TME compartment with respect to cytotoxic T-cells (Tc_max) and antigen-presenting cells (APC_max). These model inputs are presented in a high-resolution map (Figure

S1) and described in more detail in the model documentation. They reflect variability of individual tumours arising from a variety of biological factors such as growth rates, PD1 and MHCI expression levels, molecular nature of antigens, and expression of factors affecting vascularisation and immune infiltration of the TME. The first 6 parameters listed above are those that were adjusted during model calibration; for the purpose of demonstration, we set the lower and upper bounds of these parameters as 0.5× and 1× the nominal value and assume that the values of these parameters in a virtual patient population would be uniformly distributed. We use the inclusion criteria from the clinical trial to define the lower bound for the initial tumour size as 0.01 mL, and choose an upper bound of 10 mL, again assuming that the initial size of tumour in a cohort of virtual patients would be uniformly distributed. The 2 latter parameters listed above are those which we are able to derive with the input of omics data informing parameters Tc_frac and APC_frac. Therefore, we account for variability in Tc_max and APC_max via the distributions of Tc_frac and APC_frac which emerge from the data available for 1156 individual NSCLC (LUAD and LUSC) patients: Tc_frac has a lognormal distribution with mean=-2.43600 and SD=0.62335; and APC_frac has a lognormal distribution with mean=-2.58098 and SD=0.69689. For each of the 9 model parameters which we use to define biological variability, we used computer-generated random numbers following the corresponding distributions within these ranges to generate VPs.

MODEL DOCUMENTATION AND EXECUTABLE CODE

The model has been built in Certara in-house QSP Platform software with modular Biological Process Map interface. Figure S1 shows the map exported with all modules expanded. The user can zoom the map in PDF and examine model structure in detail. The map with modules collapsed was used as on overview on Figures 2 and 3. Executable code and model documentation, exported from the platform is available in the CIC_v2.2.5.zip file. The ZIP file contains three directories:

CIC_v2.2.5_CSV – Model variables, rate laws, algebraic assignments parameters and documentation in CSV files.

CIC_v2.2.5_Matlab – Matlab code of the model used for simulation shown on Figure 4. Run Figure4.m to reproduce Virtual Clinical Trial shown in Figure 4.

CIC_v2.2.5_R – R code of the model. The RUN.R contains example execution of the model. We recommend this code for users who wish to access, read or modify model equations and doses. In this version of the code variable and parameter arrays are indexed with names corresponding to model documentation in CSV files.

REFERENCES

1. Chatterjee M, et al. Systematic evaluation of pembrolizumab dosing in patients with advanced non-small-cell lung cancer. Ann Oncol 27 1291-1298. (2016)
2. Thome JJ, et al. Spatial map of human T cell compartmentalization and maintenance over decades of life. Cell 159 814-828. (2014)
3. Granot T, et al. Dendritic Cells Display Subset and Tissue-Specific Maturation Dynamics over Human Life. Immunity 46 504-515. (2017)
4. Carpenter DJ, et al. Human immunology studies using organ donors: Impact of clinical variations on immune parameters in tissues and circulation. Am J Transplant 18 74-88. (2018)
5. Thome JJ, et al. Spatial map of human T cell compartmentalization and maintenance over decades of life. Cell 159 814-828.
6. Costa-Nunes C, et al. High-throughput Screening of Human Tumor Antigen-specific CD4 T Cells, Including Neoantigen-reactive T Cells. Clin Cancer Res 25 4320-4331. (2019)
7. Trepel F. Number and distribution of lymphocytes in man. A critical analysis. Klin Wochenschr 52 511-515. (1974)
8. Bains I, Antia R, Callard R, Yates AJ. Quantifying the development of the peripheral naive CD4+ T-cell pool in humans. Blood 113 5480-5487. (2009)
9. Scott GD, Atwater SK, Gratzinger DA. Normative data for flow cytometry immunophenotyping of benign lymph nodes sampled by surgical biopsy. J Clin Pathol 71 174-179. (2018)