*Supplementary Material*

# Application of Quantitative Systems Pharmacology to guide the optimal dosing of COVID-19 vaccines.

Mario Giorgi, Rajat Desikan, Piet H. van der Graaf and Andrzej M. Kierzek

## 1. INTRODUCTION

The purpose of this short Perspective article is to introduce questions in COVID-19 vaccine development, which in our view can be tackled by QSP approach. The purpose of this article is not to provide actionable predictions. The plots presented in main text show example simulations conducted for illustration purposes only. Therefore, we do not publish the entire model with a full set of rate law parameters and their documentation, which would be required in a full research article. The purpose of this Supplementary Material is to provide additional detail introducing the main assumptions of our modelling approach and comparison of virtual trials with clinical data, as well as type and magnitude of effort required for the development of QSP platform models applicable to vaccine development.

## 2. PREVIOUS WORK ON MECHANISTIC MODELLING OF IMMUNE RESPONSE AND APPLICATIONS TO VACCINES

Mechanistic modelling of the immune system has a been long-standing topic of interest in mathematical biology. Seminal models of B-cell clonal selection and antibody production date back to 1970s[1]. In our previous publication, we have provided a comprehensive review[2] of immune system modelling and identified ~130 models relevant to cytotoxic T-cell response and immune-oncology. These models did not include antibody responses and were not directly applicable to vaccine modelling. The biology covered in these models is more relevant to anti-viral response and inflammation. In fact, a QSP model of COVID-19 disease progression, covering cytotoxic response and inflammation but not B-cells and antibodies, has been recently published[3]. This model is applicable to development of therapeutics, but not vaccines. B-cell activation, antibody affinity maturation and other key processes occurring in germinal centres have been the subject of a number of detailed models. Garg, Desikan and Dixit introduce the state of the art in this field and present their own model[4]. These models however do not model vaccine administration and frequently use agent based modelling approaches, which would be too computationally expensive for virtual trial simulations involving hundreds of model instances simulated in timescales covering years. The ODE models of antibody response to viruses have been previously published. The model of Lee et al. 2009 describing immune response to influenza provides a classical model and review of previous work in this field[5]. This is a model of immune response to virus in mice, which does not include vaccine dose administration and translation to human.

Contrary to therapeutic drug development, where the support of dose selection by quantitative modelling is very well established, vaccine development still mostly follows long-standing empirical experience and practice. It has only recently been recognised that dose optimisation through approaches equivalent to PK/PD would be valuable to optimize vaccines. In their recent review article, Rhodes et al.[6] argue for the application of immunostimulation/immunodynamic modelling to dose finding for new vaccines. The earlier work from the same group presents such an IS/ID model with an

example application for dose selection in TB vaccines. They analyse mouse data on the number of IFN-gamma producing CD4 T-cells, fit these data with an empirical NLME model, allometrically scale mouse parameters to human, and make a prediction of the number of IFN-gamma producing T-cells observed in ex-vivo assays of human peripheral blood. The model has two variables for general populations of "transitional effector" and "resting memory" T-cells, does not represent B-cells and antibody production, and introduces dose as a dose group, rather than as a dose amount mechanistically linked to other biological processes. While the authors show how alternative doses established in animal studies could be prioritised for first in human trial, the model has, in our view, limited power to extrapolate. For example, without the dynamic, quantitative model translating dynamics of antigen concentration following specific quantity of a vaccine, antigen presentation and activation of B-cells, the model is unlikely to make predictions for dose intervals and dose amounts, which were not first tested in animal experiments. Empirical models fitted to specific animal dataset will have limited power in combining data from trials on different vaccines: below we show how the change of magnitude in IgG production in older influenza vaccine subjects, can be used to inform the QSP model of SARS-CoV-2 mRNA vaccines and predict changes observed in clinical data not used for calibration. Also, key human population parameters such as HLA genetics, body weight, compartment volumes, blood and lymph flows are missing. Thus, the model of Rhodes et al would not be applicable to make a prediction for a specific clinical population – the model could only be used to provide descriptive statistics for such a population, when clinical data becomes available.

While the distinction between PK/PD and QSP modelling is fluid and frequently debated, in general, the aim of the QSP approach to dose selection is to create a mechanistic model of underlying biology to expand the range of extrapolation with respect to the data used to inform the model. This may come at the expense of less precise fits of model predictions to data, often discovered after the data became available, than the fits obtained during calibration. However, if applied with caution and in the appropriate context, predictions of clinical biomarker changes before clinical trials are conducted may prove invaluable for drug/vaccine development programs. The benefits of top-down versus bottom-up mechanistic approach are subjects of long standing scientific if not philosophical debate, and it is not our aim to provide a resolution or even a comprehensive review here. Our article is a Perspective on the potential application of QSP approaches in SARS-CoV-2 vaccine dose optimisation.

To the best of our knowledge, the QSP approach has not yet been applied to vaccine dose selection. The most relevant QSP model of Chen, Hickling and Vicini[7, 8] (CHV model) mechanistically represents protein antigen administration and PK, antigen uptake, digestion and MHC II binding, naïve, memory and effector CD4 T-cells, naïve, memory and plasma B-cells, antigen specific IgG synthesis antibody synthesis and immune complex formation. Similar to the model of Rhodes and colleagues, the CHV model was first parameterised by mouse data and then translated to human. However, due to the mechanistic detail included, the model produced antibody concentration profiles, which could be analysed with clinical trial criteria to predict the incidence of immunogenicity. Instead of using models fitted to specific doses, introduced as factors in NLME fitting, the CHV model quantitatively translates dose amount and dosing interval to quantitative antibody dynamics. Moreover, specific predictions could be made for patient populations with known HLA allele distributions as well as for individual genotypes. At the same time, the model does include CD4 T-cells and thus can also be informed by data that Rhodes and colleagues used in their study. As described in more detail below, the Certara Immunogenicity Consortium further expanded the biological scope of the CHV model and validated it by clinical case studies to create the IG Simulator tool[9]. However, both the CHV model and IG Simulator were so far applied to prediction of the impact of anti-drug antibodies (ADA) on drug PK and were not used for vaccine dose selection.

Computational approaches to COVID-19 vaccines have been recently comprehensively reviewed by Hwang et al.[10]. The review shows that the field focusses on "static" analyses of protein sequences and structures to optimise antigens by protein engineering. Furthermore, statistics and machine learning are used to identify biomarkers of patient safety and protection. Authors state that PK modelling approaches are currently unexplored. Authors cite the IS/ID modelling perspective mentioned above and state that there is a scope for application of Physiologically Based Pharmacokinetics (PBPK) to investigate antigen pharmacokinetics following administration in different modalities.

Based on literature reviewed above we conclude that application of QSP approach to SARS-CoV-2 vaccine development has not yet been published in a peer-reviewed article. We expect that other research groups are also considering utility of QSP in this context and thus wish to share our perspective.

## 3. VACCINE SIMULATOR DEVELOPMENT

Since 2017, the Immunogenicity QSP Consortium has focussed on modelling formation of anti-drug antibodies, an unwanted immunological response to therapeutic proteins[9]. Since the basic biology of the humoral immune response is the same regardless of whether we simulate an unwanted ADA response to therapeutic proteins or desired immunogenicity to a vaccine antigen, we could quickly re-purpose this model for COVID-19 vaccines. Below we present the major modelling assumptions of re-purposed model and work involved in its development.

The starting point in the development of the IG Simulator was the CHV model[7, 8]. This is an Ordinary Differential Equation model describing the following biological processes involved in antibody response to protein antigen 1) activation of antigen presenting cell 2) antigen uptake and digestion to peptides 3) peptide binding to MHC II receptors 4) antigen presentation 5) activation, proliferation and death of naïve, functional and memory CD4 T-cells 6) activation, proliferation and death of naïve, memory and plasma B-cells 7) IgG Anti-Drug Antibody production 8) Binding of ADAs to an antigen and formation of immune complex 9) Two compartment Pharmacokinetic model of therapeutic protein (antigen). The model introduces one specific clone of T-cells for each T-cell epitope with binding constants specific to compound sequence and HLA gene encoding particular MHC II receptor. A polyclonal antibody response was modelled by 17 clones of B-cells spanning entire range of biologically plausible antibody affinity. The compound specific input of the model is based on bioinformatics prediction of T-cell epitopes and binding constants to MHC II genes, but experimentally determined epitopes can be used as well. The PK model parameters are also required either from extrapolated pre-clinical model or clinical data for ADA negative patients. The model is used for virtual trial simulation with population frequency HLA alleles and PK model clearance being used as sources of variability. While antibody concentration and compound pharmacokinetics are major outputs of interest, all mechanistic model variables can be output and examined. The CHV mode is available[7, 8] with full documentation and executable Matlab code and further detail will not be re-iterated here.

The first step of IG Simulator[9] development was to replace the two-compartment PK model with Physiologically Based Pharmacokinetic (PBPK) model for biologics published by Li et al 2014[11] to expand model applicability in pre-clinical stage of drug development, when compound PK data may not yet be available and need to be extrapolated from molecular properties. The CHV model contains all immune system processes in one compartment. We compartmentalised immune system model into lymph node, peripheral blood and vascular blood compartments to better represent key immune system events occurring in lymph node and biomarker concentrations usually measured in peripheral

blood. The baseline immune cell numbers in tissues were set based on literature data. Example publications illustrating types of data sources include widely accepted order-of-magnitude estimates of absolute tissue numbers from pathology data[12], flow cytometry data from excised lymph-nodes[13] and organ donor tissues[14], experimentally determined frequencies of pre-existing naïve T-cells[15]. Compartment volumes were set following PBPK model[11], and cell circulation and migration were assumed to be limited by blood and lymph flows also available in PBPK model.

We expanded the antibody part of the model. The IgM species was included, and we also incorporated affinity maturation. PBPK model was used to capture the antibody distribution between physiological compartments and elimination from the system.

**Table S1** summarises biological scope of IG Model.

| Variables | Lymph compartment | Vascular blood compartment | Peripheral blood compartment |
|---|---|---|---|
| **T cells** | Naïve | Naïve | Naïve |
| | Activated from naïve | X | X |
| | Memory | Memory | Memory |
| | Activated from memory | X | X |
| | Functional | Functional | Functional |
| **B cells** | Naive | Naive | Naive |
| | Activated from naive | X | X |
| | Memory | Memory | Memory |
| | Activated from memory | X | X |
| **DC cells** | Matured | X | X |
| | Immature | Immature | Immature |
| **ADA** | IgG | IgG | IgG |
| | IgM | IgM | IgM |
| **Immune complex** | IgG-Ag | IgG-Ag | IgG-Ag |
| | IgM-Ag | IgM-Ag | IgM-Ag |
| **Plasma cells** | X | Long-lived | X |
| | X | Short-lived | X |
| **Main modules of IG simulator** | | | |
| • Minimal PBPK model<br>• Bioinformatics input<br>• Antigen presentation module (MHCII)<br>• Cells circulation across compartments<br>• Antibodies circulation across compartments<br>• Antibody affinity maturation | | | |

**Table S1. Biological scope of IG Simulator.**

The IG Simulator is used to run virtual trial and predict incidence of immunogenicity and its impact on PK. The virtual trials and application of IG Simulator are introduced in our previous publications[2, 9]. The IG Simulator has been calibrated and validated using 15 case studies of compounds for which clinical data on pharmacokinetics and immunogenicity were available. Currently it is applied to ~20 ongoing drug development projects.

In the wake of the SARS-CoV-2 pandemic we re-purposed the IG Simulator to predict antibody responses to COVID-19 vaccines. The general strategy was to implement the model of vaccine administration and antigen expression, which was subsequently connected to IG Simulator already including all processes describing antibody production in response to protein antigen. Our initial focus was on Lipid Nanoparticle (LNP) mRNA vaccines. Figure S1 shows a diagram of LNP mRNA model, which we connected to IG Simulator.

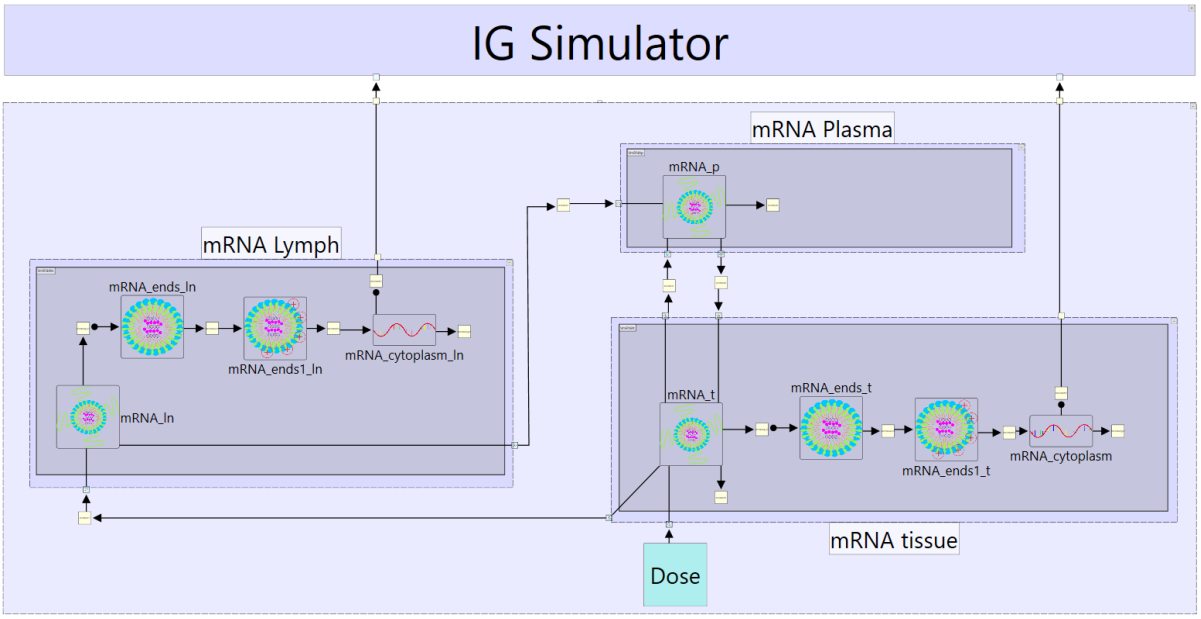


**Figure S1. Connectivity of LNP mRNA administration model.** Dose – the dose of LNP mRNA [ug mRNA]. mRNA_t, mRNA_p, mRNA_ln – mRNA concentrations [ng mRNA / mL] in tissue, plasma and lymph node compartments. mRNA_ends_t, mRNA_ends_ln – mRNA concentrations [ng mRNA / mL] in endosomes in lymph and tissue compartments, following uptake of LNP mRNAs by cells. mRNA_ends1_t, mRNA_ends1_ln – mRNA concentrations [ng mRNA / mL] in acidified endosomes. mRNA-cytoplasm and mRNA_cytoplasm_t denote translating mRNA concentrations [ng mRNA / mL] in the cytoplasm. Translation reactions produce antigen. Antigen variables are located within tissue and lymph node compartments of IG Simulator and updated accordingly. IG Simulator models antigen distribution, clearance and antigen specific immune response. While the model tracks mRNA concentration only, graphical symbols indicate mechanistic steps when mRNA is coated in LNPs.

At the time of Vaccine Simulator model construction only pre-clinical data relevant to LNP mRNA administration, bio-distribution and expression were available. Moreover, there were no data expression of SARS-CoV-2 spike protein. However, the LNP mRNA technology is constructed as a "platform", suitable for administration of any protein. Therefore, we could capitalise on pre-clinical data for other protein antigens and create general model, with protein specific parameters, which can be deduced from sequence. The mouse bio-distribution data published by Moderna[16] for their mRNA

vaccines against H10N8 and H7N9 influenza viruses indicate that following intramuscular injection large amount of mRNA distributes from injection site to liver and lymph nodes. Majority of mRNA is distributed to tissues within 2 hours after administration, distribution to lymph compartment is delayed with Tmax between 2 and 8h. Based on this information we decided to create a model where LNP mRNA is distributed to tissue, lymph and plasma compartments. Expression of mRNA takes place in tissues and lymph nodes. The intracellular of the model was based on the detailed, quantitative experimental study of LNPs expressing human EPO in adipocyte and hepatocyte cell cultures[17]. Authors published separate time profiles for LNP uptake and protein expression demonstrating 10 hours delay before first protein molecules were produced. Thus minimal mechanistic model explaining these data includes separate uptake and expression processes. Moreover, to reproduce observed delay we had to introduce a "delay chain" step. This step can be mechanistically interpreted as endosome acidification and mRNA release process. The translation rate is dependent on transcript and product lengths following general gene expression models accepted in Systems Biology field[18]. The model was parameterised to simultaneously reproduce time profiles of cell culture mRNA and protein concentrations[17] and mouse mRNA AUC, Cmax, Tmax and half life data for lymph, plasma and tissues[16]. In cell culture simulations only uptake and intracellular processes were included. Protein lengths of hEPO and influenza vaccine antigens were used in cell culture and mouse data fitting respectively.

The LNP mRNA administration model was allometrically scaled to human and connected to IG Simulator. Mouse BW of 28g was assumed. In virtual trial simulation scaling was conducted for each virtual patient separately, using individual body weight. Following well accepted allometric scaling formulas, distribution rates were scaled with body weight using exponent -0.25. Compartment volumes were scaled with exponent 1. The rates of intracellular processes of endosome uptake, acidification and protein translation and mRNA degradation were not scaled. Connection to IG Simulator was created by adding protein production by translation to ODEs describing antigen concentrations in lymph and tissues. Molecular weights of transcripts and antigens were used for unit conversion to picomole.

The QSP model described above was used for virtual trial simulations of mRNA vaccines mRNA-1273 and BNT162b2 presented in main text. The virtual population file containing model input (parameters and initial states) for each virtual patient was created following procedures established for IG Simulator[9]. The NetMHCIIPan software was used to predict T-cell epitopes and MHC II receptor binding constants in SARS-CoV-2 S protein[19]. The HLA allele frequencies for North American population were obtained from Allele Frequency net database[20] and used to randomly generate individual patients genotypes and assign MHC II affinities accordingly. Individual physiological parameters (body weight, blood and lymph flows, compartment volumes were created using biologics model[11] implemented in Simcyp simulator (https://www.certara.com/software/simcyp-pbpk/).

## 4. CALIBRATION AND VALIDATION OF VACCINE SIMULATOR.

At the time of the Vaccine Simulator development, results of Phase II/III trial of mRNA-1273[21] provided the only available clinical data for COVID-19 mRNA vaccine. We used antibody titers reported in this trial to refine parameters of the model. We adjusted model parameters within biologically plausible ranges to reproduce relative change of antibody level between primary and booster dose. We used booster to primary dose ratio to make our comparison independent of a particular way of converting titers to concentrations. Furthermore, we

created version of the model representing older adult population (>65 years of age). At the time this work was conducted there were no data from COVID-19 mRNA vaccines for these age group. Therefore, we used clinical data from other vaccines to adjust antibody response in older age groups. The clinical work of Herati et al[22]. provides data on a change of IgG antibody level between younger and older adults in response to influenza vaccine. Data show 33% decrease in antibody amount in older population. Without any further evidence we have calibrated Vaccine Simulator to reproduce this effect. We varied parameters describing carrying capacity for a functional T-cell to stimulate the activation and proliferation of target naïve/memory B cells. We acknowledge that this is empirical, rather than mechanistic approach. Having examined literature on aging immune system, we did not find mechanism, which could be implemented within current model framework.

When Phase II/III trial of BNT162b1 and BNT162b2 vaccines was published[23], it constituted an ideal data set for validation, since none of the scientists refining the model has seen the data. We chose BNT162b1 vaccine for validation, as it is more difficult case study. The BNT162b1 uses SARS-CoV-2 S-protein Receptor Binding Domain (RBD) as an antigen, which is different from mRNA-1273 and BNT162b2, which use full S-protein. Other differences between both BNT162b and mRNA-1273 trials were shorter interval between primary and booster dose (21 days, rather than 28 days) and different dose amounts (25, 100, 250 microgram in mRNA-1273 and 10, 20, 30 microgram in BNT162b). Crucially, BNT162b1/2 trial contains data for older adults, thus allowing validation of our calibration based on influenza vaccine data.

We conducted virtual clinical trials of BNT162b1 vaccine using Vaccine Simulator calibrated as described above. The T-cell epitope prediction was conducted for RBD sequence alone. HLA allele frequencies for North American population were used. Simulations were run for both for younger and older subjects using two variants of the model described above. Distributions of PBPK model parameters were also made age group specific, using appropriate distributions from Simcyp simulator. We run 18 virtual trials corresponding to 18 measurements of antibody titers reported in clinical data: 3 timepoints (21, 28, 35 days), 3 doses (10, 20, 30 microgram) and 2 age groups. For each trial we simulated 250 subjects with 21 day interval between primary and booster dose. The number of subjects was sufficiently large to assure stability of the results and sufficiently small to be practical given computational cost of the model. For each clinical timepoint we calculated Geometric Mean Concentration of IgG [ng/mL]. We acknowledge that simulated GMC values in [ng/mL] and clinical observations expressed in arbitrary units based on titer [U/mL] cannot be directly compared. However, we can assume that concentration and titer values are correlated. Similarly to clinical situation our model produces polyclonal response with wide range of affinities. The GMC differences between timepoints, doses and age groups are thus determined by antibody amount.

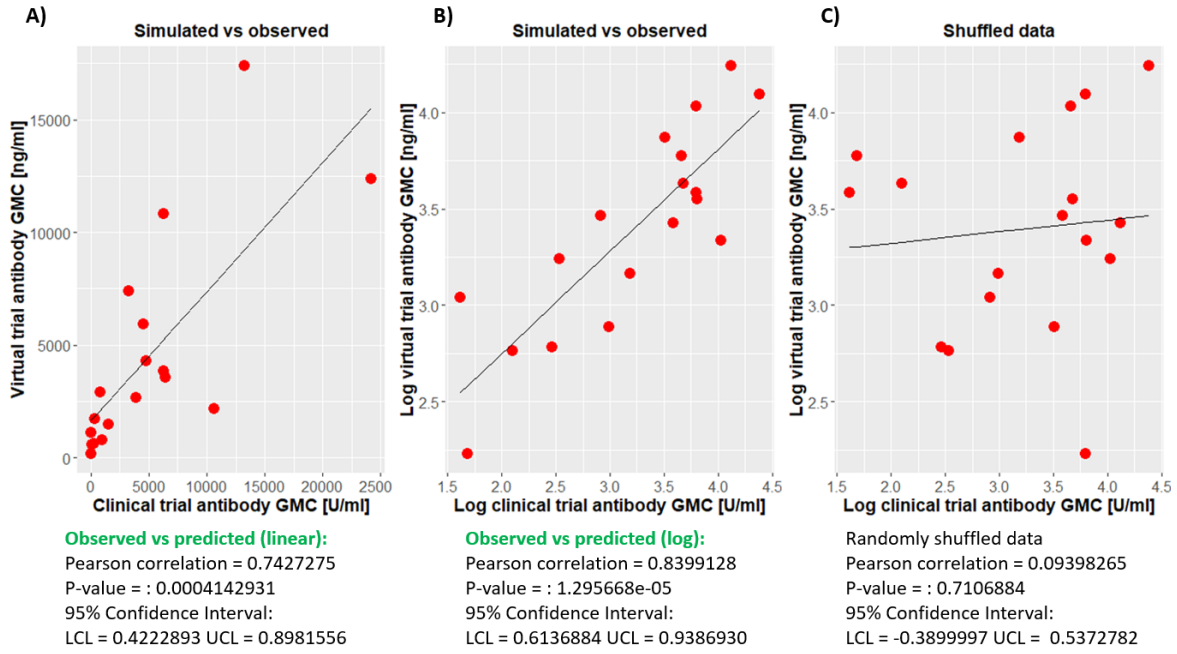


**Figure S2. Comparison of predicted and observed data for BNT162b1 vaccine.** Scatter plots show observed and predicted values for 18 measurements. A) linear scale B) Base 10 logarithm values of GMC C) Example of uncorrelated data: predicted and observed columns were randomly permuted.

Figure S2 shows that there is statistically significant correlation between clinical data and predicted values for BNT162b1 vaccine. Correlation coefficients for both GMC and log10 of GMC are 0.74 and 0.83 respectively with P-values of 4.1e-4 and 1.29e-5. In both cases 95% confidence interval does not include 0. As a reference, we created randomly permuted version of the dataset, by shuffling order of observed and predicted values in columns of the data file. The correlation coefficient is 0.09, with p-value of 0.71 and 95% confidence interval of correlation coefficient including 0. We conclude that predicted and observed anti-RBD IgG geometric mean concentrations are significantly correlated across different doses, time points and age groups. Therefore, Vaccine Simulator is applicable to prediction of changes in antibody levels at different timepoints following administration of different doses of COVID-19 mRNA vaccines to different age groups of clinical subjects.

Figure 2 presents validation of vaccine simulator for prediction of a specific biomarker – serum anti-RBD IgG concentration. As shown on Figure 2 of main manuscript, the QSP model can be used to generate hypotheses about time profiles of other biomarkers. We believe that these hypotheses are useful in directing experimental work in discovery of new biomarkers reflecting vaccine action and possibly new correlates of protection. The measurement of these biomarkers in Phase I/II stage of vaccine development would provide data for validation of Vaccine Simulator for prediction of these biomarkers. Validated QSP model would then become valuable tool for interpretation of biomarker results in larger clinical trials and large scale vaccination programs.

Finally, it is important to stress, that application of Vaccine Simulator to any other platform than mRNA vaccines would require implementation of platform specific administration

module, followed be validation of entire model. While the mechanistic model of immune response to protein antigen is general, rather than mRNA vaccine specific, one cannot exclude possibility that other modalities may trigger processes, which were either not included or which calibration was not sufficiently tested. For example, adenovirus vector vaccines (e.g. AZD1222) may trigger immune response to the vector. Therefore, implementation of new administration module should be followed by similar validation to what is presented above for mRNA vaccines. This could be done with Phase I/II data for specific doses and validated model could then be used for prediction of optimal dosing regime in larger efficacy trials and long-term vaccination programs.

## 5. DIRECT COMPARISON OF MODEL OUTPUT AND ANTIBODY TITER.

Antibody response is usually reported as a titer rather than concentration. This is because bioassay signal depends both on antibody amount and affinity. While the standard curve can be created with spiked in amounts of antibody and used to convert assay signal to absolute concentration, the choice of reference antibody will always involve assumptions. For example, if polyclonal rabbit antibodies are used, the criticism may be that they may have different affinity range than patient's antibodies. If specific human monoclonal antibody is used it will not represent full affinity range in human polyclonal response. Because of these complexities antibody amounts in clinical trials are usually reported as titers, where titer is defined as the reciprocal of the highest dilution of the sample that yields a positive result (e.g., dilution of 1/100=titer of 100), i.e., a result above a predetermined "cut point" value of assay signal[24]. Depending on further normalisation, titers may also be reported as arbitrary units.

While antibody titers can be directly used as dependent variables in statistical or empirical models fitted to data, the lack of precise measurement of absolute antibody concentrations poses challenge for mechanistic quantitative modelling. In our model the dose in microgram RNA is translated into antibody concentration in ng/mL through simulation of a model involving mechanistic steps, described by absolute cell numbers or molecular concentrations in compartment volumes. This allows us to integrate a wide range of biology literature, where mechanistic steps are characterised in isolation (cell proliferation rates, antibody synthesis rates). However, the challenge specific to antibody response modelling is that we need to compare absolute concentrations with clinical outcomes reported in arbitrary unit or titers.

One way of addressing this challenge is to focus on relative changes between timepoints, for example the ratio of antibody amount after primary and booster dose. The ratios of reported titers and simulated concentrations can be directly compared. Figure S2 above is an example of model and clinical outcome comparison which assumes linear relationship between polyclonal response simulated by our model and observed titer, within physiological parameter range, but does not require specific mapping between titers and concentrations. Significant correlation between predictions and clinical outcomes indicates that we can predict relative change of antibody level between different dose amounts, dosing intervals, timepoints and age groups. Credible prediction of relative change in response to dose amount and interval of a particular vaccine in specific patients population, available before clinical

trial is conducted, is in itself highly valuable for dose selection. However, given that scientific effort triggered by the pandemic resulted in experimental data of unprecedented scope and quantitative detail, we believe that we can map mechanistic model output and clinical data more directly.

We identified three recent publications reporting convalescent serum levels of antibodies specific to RBD domain of anti-SARS-CoV-2 S-protein in absolute concentration units[25, 26, 27]. Chen and colleagues[25] used human monoclonal antibody CR3022[28] as a positive control. This antibody, originally discovered in Sars-Cov research, binds SARS-CoV-2 RBD[28]. The Geometric Mean Plasma concentration determined in 92 subjects who recovered from SARS-CoV-2 infection was 4422.11 ng/mL. This corresponds very closely to convalescent serum measurement in 34 subjects reported by Ibarrondo and colleagues[27]. While these authors published only the plots and did not include supplementary data, the concentration reported by Chen et al. (3.65 on log10 scale) is very close to the median of data shown on Figure 1. Ibarrondo et al. also publish their standard curve obtained using CR3022 antibody as a standard. Thus, two independent groups, analysing independent samples of convalescent subjects and using CR3022 as a standard obtained consistent estimates of convalescent serum concentration. Crucially, very close estimate of convalescent serum concentration was obtained by the group using different positive control. Hartley and colleagues[26] created standard curve by serial dilutions of Ritixumab, biding its specific antigen in separate wells on the same plate. The Geometric Mean Concentration of anti-RBD IgG in 25 convalescent subjects was 4676.62 ng/mL. Given that antibody concentrations change over many orders of magnitude and are generally reported on logscale, this is remarkably close to the value of 4422.11 ng/mL obtained by Chen and collegues, who used CR3022 as positive control. Instead of averaging these two values, we chose to use value reported by Chen et al. as it was obtained for larger number of subjects. We conclude that anti-RBD IgG convalescent serum concentration of 4422.11 ng/mL is supported by three independent studies using two different positive controls for assay calibration.

The clinical trial of mRNA-1273[21] included measurement of anti-RBD convalescent serum titer in 38 subjects with the same assay which was used for assessment of anti-RBD antibody response to vaccine (mRNA-1273 is full S-protein vaccine, but anti-RBD antibodies were assayed as well). The Geometric Mean Titer reported for 38 convalescent subjects was 37,857 (95% CI 19,528–73,391). The data reported by Widge et al.[29] show a subset of patients from the same trial. To directly compare clinical and simulated data on one plot, we divided each data set by relevant convalescent serum level. Simulated anti-RBD concentrations were divided by 4422.11 ng/mL and clinical titers were divided by 37,857 at each time point. **Figure 1A** in main manuscript shows the ratios of IgG to convalescent serum for clinical and simulated data on one plot.

Similar argument is applicable to clinical trial of BNT162b1 and BNT162b2 vaccines[23]. The Geometric Mean Concentration for convalescent serum of 38 subjects was 631 U/mL. The BNT162b1 and b2 vaccines use RBD and full spike antigens and it was not clear whether authors reported convalescent serum of anti-RBD or anti-S antibodies. However, in their previous paper, the same group reported convalescent serum GMC of anti-RBD IgG as 601

U/mL for a sample 38 subjects[30]. These values are close and likely represent the same measurement, but we chose to use the value of 601 U/mL as the reference. Therefore, to directly compare simulated and clinical data for BNT162b2 and calibrate the model for this vaccine, we divided simulated output by 4422.11 ng/mL and clinical data plots by 601 U/mL. Calibrated model was used for simulations shown on **Figures 1C, D** and following chapter discusses definition of responders.

## 6. ANTIBODY PROTECTION THRESHOLD.

Plasma antibody titer is relatively easy to measure biomarker reported in every vaccine development program. However, relationship between antibody level following vaccination and protection from infection and disease is complex, due to multiple different immune response pathways, which may potentially be recruited by the vaccine. In particular, vaccines invoke cytotoxic T-cell responses, where cells infected by pathogen are killed, thus limiting spread of infection within the body. Moreover, antibody titers in plasma may poorly reflect amount of antibodies at the site of infection. Thus, plasma antibody level may at best be a correlate of protection and the correlation needs to be established in clinical data collected for particular disease. Regardless of these complexities, the term "protective threshold" is used in vaccinology literature (e.g. Figure 4 and the discussion in Pollard and Bijker 2020[31]) and in the case that plasma antibodies are established as correlate of protection, clinical subjects may be stratified as "responders" based on antibody titer.

Clinical trial of mRNA-1273 vaccine[21] shows high correlation between antibody anti-RBD IgG titers and virus neutralisation titers in vitro ($r = 0.853$, 95% CI (0.811, 0.886)). Even stronger correlation was reported for BNT162b1 vaccine ($r = 0.9452$, $p<0.0001$)[30]. Therefore, all conclusions presented in main text, such as dosing interval corresponding to maximal response would remain unchanged if in-vitro neutralisation assay was a biomarker of interest. Still, virus neutralisation assay may be a correlate of protection but is not direct measure of vaccine efficacy.

With a growing number of Phase III trials completed for COVID-19 vaccines there is a mounting evidence that plasma antibody titer is indeed a correlate of protection. In recent medRxiv pre-print (https://doi.org/10.1101/2021.03.17.20200246), Earle et al. present a plot of vaccine efficacy vs Sars-Cov-2 spike IgG ELISA for 7 vaccines. While the manuscript is not yet peer-reviewed, authors convincingly demonstrate that the linear model can be fitted to these data with rank correlation coefficient of 0.93 and 94.2% of variance explained (Figure 1B). Therefore, dose recommendations (e.g. optimal interval) based on simulated or observed increase in plasma anti-RBD antibody level are very likely to translate to increase in efficacy.

In our analysis we focussed on identifying approximate antibody threshold which would be useful to stratify patients. As observed by United Kingdom Joint Committee on Vaccination and Immunisation (UK JCVI)[32] as a part of the justification for adjusted dosing interval, the cumulative incidence curves for vaccine and placebo arms of mRNA vaccines separate at between day 10 and 14. Assuming linear relationship between antibody level and protection, the antibody level observed between day 10 and 14 corresponds to the smallest

concentration required for protection. The day 14 antibody geometric mean titer for 100 ug mRNA-1273 dose is 34,073. This is within 95% CI of convalescent serum titer of 37,857. Therefore, we assume that convalescent serum level is useful threshold related to protection. This is further substantiated by the observation that Sars-Cov-2 re-infections are rare; immune response in convalescent patients is protective. Unfortunately, there are no day 14 antibody data for BNT162b2 vaccine. However, given similarity of both antibody responses and efficacy curves between mRNA-1273 and BNT1622b2, we expect that it is useful to use convalescent plasma concentration as a protection threshold for both vaccines.

We plotted convalescent plasma level on **Figure 1A** and **1B** in main manuscript as a guide-for-an-eye level threshold relevant for protection. Our conclusions regarding maximum antibody concentration as a function of dose interval and the shapes of antibody time profiles do not depend on the threshold and are not affected by assumptions discussed above. Given that i) antibody concentration is emerging as a correlate of protection ii) a timepoint where efficacy is first observed coincides with the timepoint where antibodies reach convalescent serum level, we believe that is useful to consider convalescent serum concentration as the threshold below which protection may be of concern. This is reflected in our discussion of response durability and protection after first dose. Furthermore, on **Figure 1C** and **1D** we use convalescent serum level to classify patients and plot concentration of responders. This is to demonstrate that our simulations can be used to simulate biomarker relevant to protection for different doses of a vaccine applied to different age groups. Users could of course use different thresholds in this type of analysis to examine consequences of different assumptions in design of long term vaccination programs. Finally, we also note that prediction of the time profile of antibody response to different doses in different patient populations is going to become more useful as the data on relationship between antibodies and protection accumulate. Based on Recent pre-prints of Earle et al. (https://doi.org/10.1101/2021.03.17.20200246) and Padmanabhan, Desikan and Dixit (https://doi.org/10.1101/2021.03.16.21253742) we expect that it will soon be possible to translate simulated antibody levels to more detailed estimates of efficacy.

**References.**


1. Bell GI. Mathematical model of clonal selection and antibody production. *J Theor Biol* **29** 191-232. (1970)

2. Chelliah V*, et al.* Quantitative Systems Pharmacology Approaches for Immuno-Oncology: Adding Virtual Patients to the Development Paradigm. *Clin Pharmacol Ther*. (2020)

3. Dai W*, et al.* A Prototype QSP Model of the Immune Response to SARS-CoV-2 for Community Development. *CPT Pharmacometrics Syst Pharmacol* **10** 18-29. (2021)

4. Garg AK, Desikan R, Dixit NM. Preferential Presentation of High-Affinity Immune Complexes in Germinal Centers Can Explain How Passive Immunization Improves the Humoral Response. *Cell Rep* **29** 3946-3957 e3945. (2019)

5. Lee HY*, et al.* Simulation and prediction of the adaptive immune response to influenza A virus infection. *J Virol* **83** 7151-7165. (2009)

6. Rhodes SJ, Knight GM, Kirschner DE, White RG, Evans TG. Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling. *J Theor Biol* **465** 51-55. (2019)

7. Chen X, Hickling TP, Vicini P. A mechanistic, multiscale mathematical model of immunogenicity for therapeutic proteins: part 2-model applications. *CPT Pharmacometrics Syst Pharmacol* **3** e134. (2014)

8. Chen X, Hickling TP, Vicini P. A mechanistic, multiscale mathematical model of immunogenicity for therapeutic proteins: part 1-theoretical model. *CPT Pharmacometrics Syst Pharmacol* **3** e133. (2014)

9. Kierzek AM*, et al.* A Quantitative Systems Pharmacology Consortium Approach to Managing Immunogenicity of Therapeutic Proteins. *CPT Pharmacometrics Syst Pharmacol* **8** 773-776. (2019)

10. Hwang W*, et al.* Current and prospective computational approaches and challenges for developing COVID-19 vaccines. *Adv Drug Deliv Rev* **172** 249-274. (2021)

11. Li L, Gardner I, Dostalek M, Jamei M. Simulation of monoclonal antibody pharmacokinetics in humans using a minimal physiologically based model. *AAPS J* **16** 1097-1109. (2014)

12. Trepel F. Number and distribution of lymphocytes in man. A critical analysis. *Klin Wochenschr* **52** 511-515. (1974)

13. 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)

14. Thome JJ*, et al.* Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell* **159** 814-828. (2014)

15. Jenkins MK, Moon JJ. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J Immunol* **188** 4135-4140. (2012)

16. Bahl K*, et al.* Preclinical and Clinical Demonstration of Immunogenicity by mRNA Vaccines against H10N8 and H7N9 Influenza Viruses. *Mol Ther* **25** 1316-1327. (2017)

17. Yanez Arteta M*, et al.* Successful reprogramming of cellular protein production through mRNA delivered by functionalized lipid nanoparticles. *Proc Natl Acad Sci U S A* **115** E3351-E3360. (2018)

18. Shamir M, Bar-On Y, Phillips R, Milo R. SnapShot: Timescales in Cell Biology. *Cell* **164** 1302-1302 e1301. (2016)

19. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48** W449-W454. (2020)

20. Gonzalez-Galarza FF*, et al.* Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res* **48** D783-D788. (2020)

21. Jackson LA*, et al.* An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. *N Engl J Med* **383** 1920-1931. (2020)

22. Herati RS*, et al.* Circulating CXCR5+PD-1+ response predicts influenza vaccine antibody responses in young adults but not elderly adults. *J Immunol* **193** 3528-3537. (2014)

23. Walsh EE*, et al.* Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *N Engl J Med* **383** 2439-2450. (2020)

24. Shankar G*, et al.* Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. *AAPS J* **16** 658-673. (2014)

25. Chen Y*, et al.* Quick COVID-19 Healers Sustain Anti-SARS-CoV-2 Antibody Production. *Cell* **183** 1496-1507 e1416. (2020)

26. Hartley GE*, et al.* Rapid generation of durable B cell memory to SARS-CoV-2 spike and nucleocapsid proteins in COVID-19 and convalescence. *Sci Immunol* **5**. (2020)

27. Ibarrondo FJ*, et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N Engl J Med* **383** 1085-1087. (2020)

28. Yuan M*, et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368** 630-633. (2020)

29. Widge AT*, et al.* Durability of Responses after SARS-CoV-2 mRNA-1273 Vaccination. *N Engl J Med* **384** 80-82. (2021)

30. Sahin U*, et al.* COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. *Nature* **586** 594-599. (2020)

31. Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. *Nat Rev Immunol* **21** 83-100. (2021)

32. Iacobucci G, Mahase E. Covid-19 vaccination: What's the evidence for extending the dosing interval? *BMJ* **372** n18. (2021)