Accelerating model-informed decisions for COVID-19 vaccine candidates using a model-based meta-analysis approach

Bhargava Kandala,1,2 Nele Plock,1 Akshita Chawla,1 Anna Largajolli,1 Seth Robey,1,2 Kenny Watson,2 Raj Thatavarti,2 Sheri A. Dubey,1 S.Y. Amy Cheung,2 Rik de Greef,2 Julie Stone,1 and Jeffrey R. Sachs2

1 Merck & Co., Inc.; Rahway, NJ, USA
2 Certara; Princeton, NJ, USA

Summary

Background The COVID-19 pandemic has increased the need for innovative quantitative decision tools to support rapid development of safe and efficacious vaccines against SARS-CoV-2. To meet that need, we developed and applied a model-based meta-analysis (MBMA) approach integrating non-clinical and clinical immunogenicity and protection data.

Methods A systematic literature review identified studies of vaccines against SARS-CoV-2 in rhesus macaques (RM) and humans. Summary-level data of 13 RM and 8 clinical trials were used in the analysis. A RM MBMA model was developed to quantify the relationship between serum neutralizing (SN) titres after vaccination and peak viral load (VL) post-challenge in RM. The translation of the RM MBMA model to a clinical protection model was then carried out to predict clinical efficacies based on RM data alone. Subsequently, clinical SN and efficacy data were integrated to develop three predictive models of efficacy—a calibrated RM MBMA, a joint (RM-Clinical) MBMA, and the clinical MBMA model. The three models were leveraged to predict efficacies of vaccine candidates not included in the model and efficacies against newer strains of SARS-CoV-2.

Findings Clinical efficacies predicted based on RM data alone were in reasonable agreement with the reported data. The SN titre predicted to provide 50% efficacy was estimated to be about 21% of the mean human convalescent titre level, and that value was consistent across the three models. Clinical efficacies predicted from the MBMA models agreed with reported efficacies for two vaccine candidates (BBV152 and CoronaVac) not included in the modelling and for efficacies against delta variant.

Interpretation The three MBMA models are predictive of protection against SARS-CoV-2 and provide a translational framework to enable early Go/No-Go and study design decisions using non-clinical and/or limited clinical immunogenicity data in the development of novel SARS-CoV-2 vaccines.

Funding This study was funded by Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA.

Copyright © 2022 Merck Sharp & Dohme LLC., a subsidiary of Merck & Co., Inc., Rahway, NJ, USA and The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: COVID-19; Vaccines; SARS-CoV-2; Systematic literature search; Model-based meta-analysis (MBMA); Modelling; Modelling and simulation

Introduction

The COVID-19 pandemic has caused significant morbidity and mortality across the world and is unlikely to end until effective vaccines are administered globally to reduce transmission, prevent hospitalization, severe disease, and death. Several COVID-19 vaccines have currently been authorized, and/or approved. However, with the emergence of newer variants of concern and waning immunity in the currently vaccinated population, there is still a need for newer vaccines to be developed, and for optimization of existing vaccines. Conduct of efficacy trials is lengthy and challenging to perform in a progressively more immunized population, necessitating reliable methods for prioritizing...
Evidence before this study

With the emergence of newer strains of SARS-CoV-2, there is a need for booster shots of existing vaccines and/or newer vaccines to be developed rapidly while maintaining adequate safety and efficacy. Establishing immune correlates of protection (CoP) and the availability of quantitative models predictive of clinical efficacy could substantially help address that need. A (non-systematic) literature review of PubMed (LitCovid), and preprint servers (MedRxiv, BioRxiv) between June, 2020 and October, 2021 revealed 3 publications1-3 that established neutralizing titres as potential CoP in RM, and 4 publications4-7 that propose neutralizing titres as a CoP in humans, with two establishing a 50% protective titre against SARS-CoV-2 through quantitative models of immune protection.8 However, the RM publications to date do not integrate available vaccine protection data across different platforms, and none of the publications describe the translation of non-clinical vaccine data and its integration with early clinical immunogenicity data to inform early Go/No-Go and study design decisions for novel vaccine development.

Added value of this study

We leverage a model-based meta-analysis (MBMA) approach using emerging non-clinical and clinical data; this enables decisions to be made based on the totality of evidence8 through the integration of prior information across different vaccine mechanisms, dose-levels, regimens, endpoints, assay methods, and study designs. We demonstrate that neutralizing titres are a potential CoP in RM through the MBMA model developed with aggregate level RM protection data (protection for RM is defined here as reduction in viral load). Further, via translation of the RM MBMA to a clinical protection model we demonstrate the successful prediction of clinical efficacy of novel vaccine candidates against SARS-CoV-2 using only RM immunogenicity data (SN titres). Subsequently, we illustrate how clinical efficacy predictions can be improved with the additional integration of early clinical data via the calibration approach and with a joint MBMA model. Thereafter, once clinical efficacy data from additional candidates became available, we developed a clinical MBMA model that allowed for clinical efficacy predictions based on clinical data alone. We corroborate previous findings with the estimation of a 50% protective titre in humans using all three MBMA models. Finally, we demonstrate the predictive ability and the future application of these models through efficacy predictions of BBV152 and CoronaVac, whose data weren’t used in the development of the models, and with predictions of vaccine efficacies against delta and omicron variants of SARS-CoV-2.

Implications of all the available evidence

To effectively control the COVID-19 pandemic and to meet the unprecedented global need for vaccines, there is a need for newer vaccines and/or optimization of existing vaccines (expanding indication to other age groups, booster shots for waning immunity, increased efficacy against newer variants, etc). Vaccine efficacy trials are lengthy and expensive, and therefore quantitative tools that can be leveraged to inform early Go/No-Go decisions are invaluable. The MBMA models presented here are predictive of protection against SARS-CoV-2 and provide a powerful translational framework to enable early Go/No-Go decisions and to accelerate study design decisions using only non-clinical and/or limited clinical immunogenicity data in the development of novel vaccines.

Methods

We provide a framework to use animal (non-clinical) data available early in vaccine development to predict vaccine efficacy, to improve vaccine efficacy predictions by enabling incorporation of clinical (and non-clinical) immunogenicity data when available, and, thus, to support model-informed decision making in vaccine development. To enable development of such a framework rapidly enough to support vaccine discovery and
Development of a predictive model-based meta-analysis (MBMA) in rhesus macaques (RM)

Systematic literature search. A systematic search was performed to identify non-clinical trials of vaccines against SARS-CoV-2 and of prior infection with SARS-CoV-2 (vaccine and re-challenge) in rhesus macaque (RM). It included searches in PubMed (LitCovid), BioRxiv, and MedRxiv. Through an automated screening approach, titles and abstracts of all references related to SARS-CoV-2 were screened for a defined list of lexicon terms (a description of systematic search is in Supplementary Material, including Table S1). This resulted in approximately 1429 references of potential interest from screening of 104,472 references with data cut-off of 25 Jan 2021. Out of these, 111 references contained words related to RM, and were reviewed to check if the reference was reporting challenge study results after vaccination, monoclonal antibody administration or initial virus challenge, and provided information on

1) SN titre data following vaccine administration, monoclonal antibody administration, or initial challenge (for rechallenge studies), and
2) Viral load (VL) data post challenge in relevant tissue matrices.

The details of these studies were reviewed, and an exploratory analysis was conducted to assess information provided on vaccine type (mechanism), dose-level, regimen, challenge dose, time of challenge, SN assay description, SN titres, units of SN, and viral load in tissue matrices. 13 of the publications provided information enabling them to be used for the MBMA analysis.

Software. Data preparation and visualization was performed using Excel 2016 and R version 3.5.0 or higher. MBMA model development was done using the nlme package in R, and efficacy predictions were carried out using R version 3.5.0 or higher.

Immunogenicity and viral load data derivation rules. Summary level log10 SN titres and geometric mean log10 peak VL data were of interest. SN titre data from live virus assays were preferred over pseudo-virus assay results if both were reported in the same reference at the time point of interest (time point closest to and prior to the time of challenge). Geometric means were derived from individual data, if available from the source. If only median data were available and individual data were not reported, the median values were retained in the analysis dataset. SN titre and VL data were log-transformed using base two and base 10 (respectively) prior to analysis to align with the standard units used in the literature.

Data were either directly reported or derived from reported values. If the SN titres for control arms were clearly reported at (or at half of) the lower limit of quantification/detection (LLOQ/LLOD), or, if no LLOQ/ LLOD values were reported, log10 SN data for control arms were set to zero. Alternate imputation rules for SN titres (e.g., fixing to half LLOQ/LLOD) were tested during model development, however no substantial impact on model parameters was identified. If peak log10 VL was reported at LLOQ/LLOD, it was set to half this threshold level, and if no LLOQ/LLOD was available, log10 VL was set to zero. This approach ensures that, for sufficiently high titres, we can predict viral load corresponding to the maximum possible protection (as measured in each experiment and specimen type). The high efficacies in the clinic and low VLS in RM suggest that it is reasonable to assume that viral load can be reduced to zero. To minimize any possible impact of VL imputation on model parameters and to ensure consistency between measured and modelled minimum response, we matched the appropriate model parameter value (VL\textsubscript{min} defined below) to the corresponding imputed VL value (for a given experiment and specimen type). To avoid numerical problems during model fitting, a constant with a value small enough to not impact results (determined through a sensitivity analysis, << 0.1) was added to VL values.

MBMA of serum neutralizing titres vs peak viral load. Non-linear mixed-effects modelling was implemented
to characterize the relationship between geometric mean SN titres and VL data from vaccine studies against SARS-CoV-2 conducted in RM. During modelling, study-level parameters could either be shared across specimens or be estimated separately. The model relationship is of the form

$$VL_{i,j,k,l} = VL_{\text{max},i} + e^{\ln\left(VL_{\text{max},i}/\text{LLOQ}\right)} \cdot \text{suppr} + e_{ijkl}$$

where $VL_{i,j,k,l}$ is the post-challenge log$_{10}$ peak viral load for study $i$, arm $j$, specimen type $k$, and challenge inoculum dose $l$ with SARS-CoV-2. This functional form was originally developed to describe Poisson distribution data such as respiratory syncytial virus (RSV) incidence rates (IR) and was subsequently shown to also describe nonclinical RSV viral load data successfully.\(^{13}\)

$VL_{\text{max},i}$ is the maximum log$_{10}$ peak viral load for study $i$, specimen type $k$, and challenge (inoculum) dose $l$ (which had more than one value in at least one study). $VL_{\text{max},i}$ was estimated in the MBMA model using a study-specific exponential random effect on $VL_{\text{max}}$ to account for the variability in peak viral load because of different experimental conditions between and within studies:

$$VL_{\text{max},i} = VL_{\text{max}} + \exp(n_{i,l})$$

Here $n_{i,l}$ is normally distributed with a mean $\mu$ and variance $\sigma^2$. $VL_{\text{min},i}$ is the minimum log$_{10}$ peak viral load for study $i$ and specimen type $k$ and was fixed to half of the respective reported LLOQ/LLOD values; and suppr represents the suppression of the peak viral load with increasing SN titre values and is described by

$$\text{suppr} = \left(\frac{\log_{10}(SN_{ij})}{\log_{10}(SN_{ij})} + \frac{\log_{10}(IT_{so})}{\log_{10}(IT_{so})}\right)^{-1}$$

Here $\log_{10}(SN_{ij})$ is the log$_{10}$ serum neutralizing titre closest and prior to the time of SARS-CoV-2 virus challenge in study $i$, arm $j$; $\log_{10}(IT_{so})$ is the SN titre for each specimen type $k$ at which log$_{10}$ VL has been reduced by 50% of the difference between $VL_{\text{min}}$ and $VL_{\text{max}}$. At this log$_{10}$, SN, VL is the geometric mean of $VL_{\text{min}}$ and $VL_{\text{max}}$. While study-specific differences between $\log_{10}(IT_{so})$ may be present (e.g., due to assay differences), no study-specific exponential random effect was included on the $\log_{10}(IT_{so})$ parameter to not overparameterize the model. It was assumed that this would not introduce substantial bias into the estimated typical value parameter, $\log_{10}(IT_{so})$, i.e., that there are sufficient data and that they are appropriately representative of assays used. The parameter $\gamma$ is the Hill coefficient that describes the slope of the curve, and $e_{ijkl}$ is the within-arm residual variability estimated with an additive normal distribution. The residual error was weighted by the inverse of the sample size in each trial arm. In this model, the predictor variable (SN titer), covariate effect (tissue type on IT$_{so}$; see Results), and the random effect on $VL_{\text{max}}$, together with the residual variability, account for any heterogeneity between studies.

**Covariate analysis.** A visual inspection of the covariate relationships was carried out before testing covariate effects in the model. If no clear trend was observable, then no formal model testing was performed. If a trend was observable from visual inspection, the covariate was evaluated in terms of model fit and statistical significance using the criteria described in the following section. The list of covariates deemed potentially impactful on the SN titre vs. peak VL relationship and explored during the MBMA model development and their relevance are shown in Table S2.

**Model assessments and qualification.** Nested models (e.g., covariate models) were compared through the likelihood ratio test that assumes that minus twice the difference of the log-likelihood (-2LL) values of the nested models is approximately chi-square distributed with the number of degrees of freedom (dof) equal to the number of new parameters. A drop of 3.84 (p-value<0.05, likelihood ratio test) caused by an introduction of one dof was considered a significant improvement and retained in the model. Non-nested models were compared by means of the Akaike Information Criterion (AIC), which is the likelihood of the model corrected for the number of parameters estimated; the lower the AIC values, the better the model fit. In addition, the precision of the parameter estimates was assessed by the confidence intervals (CIs) calculated using a normal approximation to the distribution of maximum likelihood estimators: the narrower the CIs, the lesser the uncertainty on model parameters. Furthermore, Pearson residual plots and $V^2$ACHER\(^{14}\) plots were produced to evaluate goodness-of-fit. ($V^2$ACHER is a method that enables intuitive visualization of data overlayed on the covariate-containing models fit to them).

**Translation of RM MBMA to a clinical protection model**

The translation of the RM MBMA model to a clinical protection model was carried out in two steps. First, the peak VL vs. log$_{10}$ SN titre relationship in RM (Eq. (1)) was transformed to obtain a % IR of clinical disease vs log$_{10}$ SN titre relationship in humans. Clinical disease was defined as symptomatic disease of any severity with PCR-confirmed SARS-COV-2. This transformation was made using the assumption that VL in RM is a translational surrogate for the IR of clinical disease and assuming that the shape of these two relationships is the same (Figure S2A-B). The equation for the transformation is a simple linear scaling as shown below:

$$\frac{IR_i - IR_{\text{min}}}{IR_{\text{max}} - IR_{\text{min}}} = \frac{VL_i - VL_{\text{min}}}{VL_{\text{max}} - VL_{\text{min}}}$$

\(^{(4)}\)
IRi is the log10 peak post-challenge viral loads measured for vaccine i; VLmin is the mean of minimum log10 peak viral loads across all studies in the RM-MBMA; VLmax is the estimated maximum log10 peak viral load from the RM MBMA model; IR is the predicted incidence rate of clinical disease for vaccine i; IRmin and IRmax are the minimum and maximum attainable incidence rates of clinical disease. IRmin was prospectively assumed to be 0, based on findings from RM data that robust neutralizing titres resulted in complete protection against detectable viral load (qPCR copy number). When IRmin is set to 0 IR scales proportionally with IRmax, and therefore IRmax can be arbitrarily fixed to 1. This assumption resulted in the following simplification of Eq. (4) above:

\[
IR_i = \frac{VL_i - VL_{\text{min}}}{VL_{\text{max}} - VL_{\text{min}}} \quad (5)
\]

Next, the % vaccine efficacy vs. log2 SN titre relationship in humans was derived from the predicted % incidence rate vs log2 SN relationship (Figure S2C) using:

- IRvaccine – incidence rate of clinical disease for vaccinated subjects, predicted using the reported Phase 1 or 2 geometric mean (GM) SN, and
- IRplacebo – incidence rate of clinical disease for placebo subjects. We are predicting the efficacy of a vaccine in COVID-19 naive individuals with baseline SN titres assumed to be near the LLLOD, corresponding to an incidence rate approaching IRmax, here fixed to 1.

The % Vaccine Efficacy is defined as the relative reduction in incidence rate of the vaccine compared to placebo arm in each trial:

\[
\text{% Vaccine Efficacy} = 100 \cdot \left(1 - \frac{IR_{\text{vaccine}}}{IR_{\text{placebo}}} \right) \quad (6)
\]

Curation of clinical immunogenicity and efficacy data of COVID-19 vaccine candidates. A systematic literature review was performed to identify clinical data for those vaccine candidates with reported Phase 1/2 immunogenicity data and Phase 3 efficacy data (a cut-off date of July 2021 was employed, at which a total of 71 references were selected for curation based on criteria outlined in the Supplementary Material, including Table S1). Additional data which were available after this date were included on an ad hoc basis if deemed to add sufficient value to the analysis (e.g., preliminary efficacy data from CoronaVac and BBV152 vaccines to facilitate external model validation). Summary level data from eight vaccines that met the criteria were curated and were included in the clinical database. Consistent with curation of the RM data, SN titre data obtained from live virus assays were preferred over pseudo-virus assay data. When available at more than one timepoint, SN titres obtained closest to (but after) the start of the efficacy assessment period (after the last dose) were selected. If not reported along with the published vaccine efficacy, SN titre values were obtained from alternative publications (e.g., Phase 1 safety and immunogenicity reports). However, SN titres were always matched with estimates of vaccine efficacy obtained from identical dosing regimens. SN titres from population strata (e.g., age) were selected to be from strata commensurate with the population in which efficacy was measured. Geometric mean Human Convalescent Serum (HCS) titres using the same assay were also included to quantitatively account for assay differences across trials. The respective definitions in each trial of efficacy endpoints and of HCS are described in Table S3. Data from 6 vaccine candidates (BNT162b2, mRNA-1273, Ad26.COV2.S, 2019nCoV, Gam-COVID-Vac, ChAdOx1 nCoV-19) were leveraged in the development of the MBMA models. The efficacy measure employed was the relative reduction of COVID-19 IR of all disease severities, confirmed by PCR, relative to control. Efficacy selected for inclusion was assessed at the time defined as the primary endpoint, at least 7 days after the final dose. The data from CoronaVac and BBV152, which became available after the models were built, were used as part of the external validation by comparing their published efficacies to the VE predicted using their respective SN GMT data as input to the three MBMA models. Additional SN titre data were also collected from a non-systematic ad hoc literature search to inform simulations to predict vaccine efficacy against the delta and omicron variants.

Estimation of SN calibration factor

The clinical SN titres were normalized using the geometric mean HCS titres from their respective assays to account for the assay differences across trials. The same approach was not carried out for the RM SN titre data, as convalescent titres using the corresponding assays were either not reported or the definition of convalescence was not clearly stated in these RM studies.

Generalized least squares (GLS) was used to estimate a calibration factor to align the RM MBMA (scaled to 0-100% efficacy range) with the HCS-normalized clinical SN. The calibration factor, \( F_c \), was estimated by minimizing the sum of squared residuals between the reported point estimates for clinical efficacy and the corresponding HCS-normalized titre with the RM-MBMA curve predicted from the maximum likelihood estimates of the MBMA using the following equation:

\[
\log(SN_{\text{cal}}) = \log\left( \frac{SN_{\text{human}}}{HCS_{\text{human}}} \right) + \log(F_c) \quad (7)
\]

where \( SN_{\text{human}} \) is the observed SN titre of the \( i \)th vaccine candidate at timepoint closest to (but after) the start of
the efficacy assessment (after the last dose), $HCS_{human}$, is the observed HCS titre using the same assay and the same trial as the $i^{th}$ vaccine, and $SN_{cal}$ is the SN titre of the $i^{th}$ vaccine candidate when it is used for prediction of clinical results from the RM MBMA. $F_i$ was then used for calibration to enable prediction of clinical efficacy from clinical immunogenicity data and the non-clinical MBMA.

Development of joint MBMA model
A joint MBMA using RM data and clinical data was conducted to establish the relationship between SN titres and protection in both species. To model the non-clinical and clinical data jointly, the following steps were taken:

**Step 1. Transformation of the NHP data**

The viral load data in RM was transformed to an efficacy scale (0, 1) using Eqs. (8)-(9) so the response variables were the same for both species

\[
Scaled \ V_{li,j,k,l} = \frac{VL_{li,j,k,l}}{VL_{li,j-control,k,l}} \quad (8)
\]

where $VL_{li,j,k,l}$ is the $\log_{10}$ peak viral load for study $i$, arm $j$, specimen type $k$, and challenge inoculum dose $l$ post challenge with SARS-CoV-2. Scaled $VL_{li,j,k,l}$ was set to 0 if $VL_{li,j,k,l}$ was equal to reported LLOQ or LLOD, corresponding to an assumed full response. $VL_{li,j-control,k,l}$ refers to the $\log_{10}$ peak viral load corresponding to the placebo/control arm in study $i$, specimen type $k$, and challenge inoculum dose $l$ post challenge with SARS-CoV-2.

Efficacy (Scaled response) was calculated as

\[
Scaled \ response_{li,j,k,l} = 1 - Scaled \ VL_{li,j,k,l} \quad (9)
\]

**Step 2. Collate the efficacy and immunogenicity data from both species (RM and humans(h))**

\[
VE = (Scaled \ response_{li,j,k,l}; \ Clinical \ Efficacy_{rm}) \ \forall \ i,j,k,l,m \quad (10)
\]

\[
Imm = (\log_2 [SN_{RM}])_{ij} \left[ \frac{SN_{h,m}}{HCS_{h,m}} \right] \ \forall \ i,j,k,l,m
\]

Here, for the $m^{th}$ clinical vaccine candidate, Clinical Efficacy$_{rm}$ is the reported efficacy, $SN_{h,m}$ is the human serum neutralizing titre obtained at time closest to (but after) the start of efficacy assessment (after last dose), and $HCS_{h,m}$ is the corresponding human convalescent serum titre reported. ($\log_2 [SN_{RM}]$)$_{ij}$ is the $\log_{10}$ serum neutralizing titre closest prior to the time of SARS-CoV-2 virus challenge in study $i$, arm $j$, in the RM experiments.

**Step 3. Build the joint MBMA model (with data indexes (i, j, k, l, and m) omitted for readability):**

\[
VE \sim VE_{max} + \frac{\log_2 (SN_{RM})^{\gamma_{RM}}}{SN_{RM}} + \frac{\log_2 (IT_{50})^{\gamma_{RM}}}{IT_{50}} I[Species = RM]
\]

\[
+ \frac{SN_{v,clin}}{SN_{v,ref}} \cdot \frac{SN_{v,ref}}{\gamma_{RM}} I[Species = Human] \quad (11)
\]

$VE_{max}$ is the maximum efficacy estimated in the MBMA model. $SN_h$ is the serum neutralizing titre obtained at time closest to start of efficacy assessment (after last dose), and $HCS$ is the human convalescent serum titre reported for the corresponding clinical vaccine candidates. $\log_2 [SN_{RM}]$ is the $\log_{10}$ serum neutralizing titre closest to and prior to the time of SARS-CoV-2 virus challenge in the RM experiments. $log_2 [IT_{50}]$ is the $\log_{10}$ peak viral load required to obtain 50% protection in RMs for specimen type $k$ = BAL (& lung tissue), and nasal; $\gamma_{RM}$ is the ratio of vaccine SN titre to HCS titres required to obtain 50% efficacy in humans, $\gamma_{RM}$ and $\gamma_{clin}$ are the Hill coefficients that control the slope of the curve for the RM and human data, respectively. (The log of the hill coefficient was modelled to ensure non-negative parameter values.) $I[Species = RM]$ is an indicator variable that takes a value of 1 for RM data, and is otherwise 0, and analogously for $I[Species = Human]$. A GNLS approach was used for model building. When fitting the model parameters, the vaccine efficacy observations were weighted by the number of participants in each trial, and the RM studies were, separately, weighted by the number of animals. The weights for the clinical data ranged from 0.129-1, and for RM data from 0.1-1. A normally distributed residual error term was used in the joint MBMA model. A species-specific random effect was also tested: its variance was estimated to be close to zero, and it resulted in a worse model fit based on AIC criteria.

Development of clinical MBMA model
A clinical MBMA was conducted to establish the relationship between HCS-normalized SN titres (SN/HCS ratio) and clinical efficacy. The clinical SN titres were normalized to the geometric mean HCS titres as described above. A sigmoidal relationship was used:

\[
VE_{ij}(\%) = \frac{VE_{max} \cdot SN_{RM}}{SN_{RM} + (ET_{50, human})^\gamma} + e_{ij} \quad (12)
\]

Here, $VE_{max}$ is the maximum vaccine efficacy (%), $SN_{RM}$ is the HCS-normalized SN titre of the $i^{th}$ vaccine candidate and $j^{th}$ age group, $ET_{50, human}$ is the SN/HCS titre ratio corresponding to 50% vaccine efficacy in humans, $\gamma$ is the Hill coefficient that controls the slope of the curve,
and $e_{ij}$ is the additive residual variability estimated with a normal distribution. When fitting the model parameters, the vaccine efficacy observations were weighted by the number of participants in each trial.

Methods for prediction/simulation

Efficacy predictions were generated from each model using available clinical or non-clinical immunogenicity data. First, sets of 1,000 model parameters were sampled from the multi-variate normal variance-covariance matrix to account for parameter uncertainty. To account for uncertainty in the reported GMT in vaccinated and convalescent samples, each set of model parameters was paired with a set of GMTs from vaccinated or convalescent samples drawn from the log-normally distributed uncertainty (standard error). For predictions of efficacy against the newer variants (delta, omicron), we also accounted for uncertainty in the reported ratio of SN for WT vs. newer variant when available. While predicting from the RM-MBMA model using clinical calibration factor $F_C$, each parameter set was also paired with a calibration factor sampled from the uncertainty in the mean estimate for $F_C$. Clinical efficacy was predicted using each model for each sampled parameter set, calibration factor (where appropriate), and clinical GMT. Predicted efficacies were summarized as the median and 95% confidence interval. Root mean squared prediction error (RMSPE) was used to compare the predictive power of different modelling approaches.

Role of the funders

The authors from Merck & Co., Inc., Rahway, NJ, USA participated in study design, data collection, analysis, interpretation, and writing of the report. The funders reviewed the penultimate draft of the manuscript. All authors had full access to the data in the study and approved the decision to submit for publication.

Results

Analogously to the Methods section, we first describe the RM dataset obtained from the systematic literature search, and the resulting relationship between SN titres and protection data in RM quantified by the RM MBMA model. We then show the predictions of vaccine clinical efficacies with RM data alone through the translation of the RM MBMA to a clinical protection model. Subsequently, we illustrate how clinical efficacy predictions can be made with the integration of early clinical data and RM data via the calibration approach, and also with the joint MBMA model. Thereafter, we describe the relationship between HCS-normalized SN titres and clinical efficacy via the clinical MBMA model. Finally, to illustrate the predictive ability of the three MBMA models, we show efficacy predictions of 2 vaccine candidates whose data were not used in the development of the models. In addition, predictions of vaccine efficacies against the delta and omicron variants of SARS-CoV-2 are predicted as this is an important application of the models.

Exploratory analysis suggests SN titres predictive of protection in RM

In total, summary level data from 13 studies including 21 SARS-CoV-2 vaccine candidates across various platforms were included in the RM MBMA analysis dataset (Table 1). To evaluate the potential of SN titres as an immune correlate of protection in the RM model, post-challenge peak VL in BAL, lung and nasal tissue were used as protection endpoints to indicate severity of infection, interpreting them as potential translational surrogates for clinical disease. Sub-genomic RNA (sgRNA) was assumed to best represent replicating virus. Therefore, only sgRNA peak VL was included as the “efficacy” (protection) endpoint in the MBMA model. The SN level closest to and prior to the time of challenge was assumed to be predictive of efficacy and hence included as the predictor. Exploratory visual analysis indicated a potential decrease in peak VL (increasing protection) with increase in SN in BAL, lung, and nasal tissues (Figure S1), supporting the evaluation of SN as a potential immune correlate of protection and the development of an MBMA model.

RM MBMA characterizes protective SN levels in rhesus macaques

A model-based meta-analysis (MBMA) enables decisions to be made based on the totality of evidence through the integration of prior information across different vaccine mechanisms, dose-levels, regimens, endpoints, and study designs. A non-linear mixed effect modelling approach was used to build an MBMA model (as described in Methods) to characterize the sigmoidal relationship between SN titres after vaccination and peak VL post-challenge in RM. The post-challenge peak sgRNA VL data in BAL, lung, and nasal tissue were modelled jointly. For $\text{IT}_{50}$ (defined to be the log$_2$ SN titre at which ln(VL$_\text{max}$) has been reduced by 50% of the difference between ln(VL$_\text{min}$) and ln(VL$_\text{max}$), model development indicated that estimating this parameter for BAL and lung tissue, and, separately, for nasal tissue ($\text{IT}_{50\text{, nasal}}$), was statistically superior to using a single $\text{IT}_{50}$ value across all tissue matrices. The final model therefore contained two $\text{IT}_{50}$ estimates (i.e., having tissue type as a covariate, as expected), while all other parameters were shared across all tissue matrices.

Challenge studies varied in terms of the type of neutralization assays used (plaque reduction neutralization test (PRNT), microneutralization (MN), and cytopathic effect (CPE) assays), and type of virus (live vs. pseudo).
| Study          | Study type | Number of arms | Mechanism of vaccine | Interventions                                                                 | Dose and regimen                                                                 | Adjuvant | Live/ Pseudo | SN assay (% Inhibition) | Specimen          | Reported efficacy in clinic |
|---------------|------------|----------------|----------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------|----------|---------------|------------------------|-------------------|---------------------------|
| Chandrashekar 2020 | Challenge  | 3              | NA                   | Control, none                                                                  | NA                                                                              | No       | lvnt          | 50                     | Nasal/naso-pharyngeal | NA                        |
| Corbett 2020   | Vaccine    | 3              | RNA                  | mRNA-1273, placebo                                                            | IM, week 0 and 4, 10 or 100 µg                                                  | No       | lvnt          | 50                     | Bal fluid, nasal/naso-pharyngeal | Yes                       |
| Yang 2020      | Vaccine    | 4              | Subunit              | SARS-CoV-2 vaccine, control                                                   | IM, day 0 and 7, 20 or 40 µg                                                    | Yes      | lvnt          | unknown                | lung tissue         | No                        |
| Mercado 2020   | Vaccine    | 8              | Adenovirus-vectored  | Placebo, S.PP, IPAS, IPAS, S.PP, S, S.dCT, IPA WT.S, 5dTM, FP            | IM, day 0, 10^3 viral particles                                                | No       | Psnv	     | 50                     | Bal fluid, nasal/naso-pharyngeal | Yes                       |
| Guébre 2020    | Vaccine    | 4              | Subunit              | SARS-CoV2373, placebo                                                         | IM, day 0 and 21, 5 or 25 µg                                                    | Yes      | lvnt          | 100                    | Bal fluid, nasal/naso-pharyngeal | Yes                       |
| Patel 2020     | Vaccine    | 2              | DNA                  | INO-4800, control                                                             | IDEP, week 0 and 4, 1 mg                                                        | No       | Psnv	     | 50                     | Bal fluid, nasal/naso-pharyngeal | No                        |
| Yu 2020        | Vaccine    | 7              | DNA                  | Placebo, S, S.dCT, S.dTM, S1, RBD peptide, S.dTM, FP                        | IM, week 0 and 3, 5 mg                                                          | No       | lvnt          | 50                     | Bal fluid, nasal/naso-pharyngeal | No                        |
| van Doremalen 2020 | Vaccine  | 3              | Adenovirus-vectored  | ChAdOx1 GFP, ChAdOx1 nCoV-19                                                  | IM, -56 and/or only -28 dpi, 2.5 × 10^10 VP/animal                             | No       | lvnt          | 100                    | Bal fluid, nasal/naso-pharyngeal, lung tissue | Yes                       |
| Rauch 2020     | Vaccine    | 3              | RNA                  | Control, CVnCoV                                                               | IM, day 0 and 28, 0.5 µg or 8 µg                                               | No       | lvnt          | 50                     | Lung tissue, bal fluid, nasal/naso-pharyngeal | Yes                       |
| Furuyama W 2021 | Vaccine    | 3              | VSV-EBOV             | VSV-EBOV, VSV-SARS2-EBOV                                                        | IM or IN, day10, 1 × 10^7 PFU                                                   | No       | lvnt          | 100                    | Nasal/naso-pharyngeal | No                        |
| Brouwer 2020   | Vaccine    | 2              | SARS-CoV-2 subunit    | Control, SARS-CoV-2 S1-513-50NP                                                | IM, week 0, 4 and 10, 50 µg                                                      | Yes      | lvnt          | 50                     | Bal fluid, nasal/naso-pharyngeal | No                        |
| McMahan 2020   | Antibody   | 4              | NA                   | Sham, IgG                                                                     | IV, day 0, 250 mg kg^-1, 25 mg kg^-1, 2.5 mg kg^-1                             | No       | Psnv	     | 50                     | Bal fluid, nasal/naso-pharyngeal | NA                        |
| van Doremalen 2021 | Vaccine  | 2              | Adenovirus-vectored  | ChAdOx1 GFP, ChAdOx1 nCoV-19                                                  | IN, -56 and -28 DPI, 1 ml                                                        | No       | lvnt          | 100                    | Nasal/naso-pharyngeal | Yes                       |

Table 1: Characteristics of studies in the RM database used in the MBMA of SN vs VL.

Bal=Bronchoalveolar lavage; IN=intranasal; IM=intramuscular; IV=intravenous; Lvnt=live-virus neutralization titres; Psnv=Pseudo-virus neutralization titres.

* Mercado et al.: pseudo-virus SN titres were preferred to the live-virus SN titres as only the pseudo-virus titres were available at the time of challenge.

b References 13, 20, 23, 24 and 25 are preprint/non peer reviewed articles.
used in those assays. Unlike clinical trials, convalescent titres using the corresponding assays were either not reported or the definition of (clinical) convalescence was not clearly stated in these RM studies. Hence the approach to normalize the SN titres to their corresponding convalescent titres was not carried out for the RM data. Therefore, after developing the base MBMA model, an exploratory analysis was conducted to assess the potential impact of these assay differences on the relationship between SN titres and peak VL. This visual assessment (Figure S4) showed notable consistency in the data from different assays, with no apparent trends suggesting systematic differences between the assays. As a result, no covariates on assay type were included in the final model. (In general, a covariate was considered unlikely to be important if the CIs of two V^Aucher plot curves corresponding to different values of that covariate overlap).

The goodness of fit plots (Figure S5) and the V^Aucher plot (Figure 1) revealed that the final model provided a robust fit to the VL and SN titre data across all specimens and for all vaccine platforms. The model-estimated parameters are provided in Table 2. The $I_{50}$ estimate for BAL and lung tissue of 5.84 (5.2 – 6.5) [mean (95% CI)] log₂ units is lower than the $I_{50,nasal}$ for nasal tissue which was 8.3 (7.1 – 9.5) [mean (95% CI)] log₂ units, suggesting higher SN levels required for protection in the nasal tissue compared to the lower
respiratory tract. Data from 4 vaccine candidates whose Phase 3 clinical efficacy is reported (2019nCoV, Ad26.COV2.S, ChAdOx1 nCoV-19, and mRNA-1273) are coloured (non-gray) in Figure 1. It is apparent that the relative order of model-predicted and observed vaccine protection for these four vaccines is similar to that of their Phase 3 clinical efficacies (Table 3). Overall, higher SN titres resulted in lower peak viral loads across all specimens in RM.

Immunogenicity in RMs is broadly predictive of clinical vaccine efficacy

The translation of the RM MBMA model enabled prediction of VE in humans using only immunogenicity data in RM. This translation required scaling viral load to clinical incidence rate using the assumption that relative changes in RM VL vs SN titre relationship correspond to the same relative changes in incidence rate vs SN titre relationship (same shape) as described under Methods above. The translated RM MBMA model was then leveraged to predict clinical vaccine efficacy using RM SN titres as input. Five vaccine candidates were included in the predictions. (These are the only ones with reported clinical efficacy and corresponding SN titres in RM. Although the clinical efficacy was reported for BNT162b2, a suitable VL endpoint in RM was not available. Hence those data were used in efficacy predictions but not in the estimation of RM MBMA.) Predictions were made using the RM MBMA model for protection measured in BAL fluid/lung (i.e., with IT50) and then for the nasal swab (i.e., using IT50;nasal) parameters (Table 2). The predictions using the BAL/lung tissue model parameters performed better than the predictions from nasal swab model parameters (root mean squared prediction error (RMSPE) = 12% vs 21%). The VE predictions with the nasal swab model parameters were lower than those predicted with the BAL/lung model parameters given the higher estimated protective titres in nasal swab (IT50;nasal > IT50) as shown in Table S4. Figure 2 shows that the predicted vaccine efficacies were generally consistent with the observed clinical efficacies.

Calibration enables prediction of clinical vaccine efficacy using clinical SN titres and the RM MBMA

As phase 1 clinical immunogenicity data of COVID-19 vaccine candidates became available, the data were

| Parameter | Estimate | 95% confidence interval |
|-----------|----------|------------------------|
| VLmax     | 4.86     | 4.13 – 5.60            |
| γ         | 4.14     | 2.83 – 5.45            |
| IT50      | 5.84     | 5.21 – 6.47            |
| IT50;nasal| 8.32     | 7.11 – 9.53            |
| IIV on VLmax | 0.24   | 0.15 – 0.39            |
| Residual standard error | 1.91 | 1.65 – 2.22 |

Table 2: Parameter estimates of RM MBMA model.

VLmax is the maximum log_{10} peak viral load; IT_{50} is the log_{2} SN titre at which ln(VL_{max}) has been reduced by 50% of the difference between ln(VL_{max}) and ln(VL_{min}) in BAL and lung tissue, IT_{50;nasal} is defined analogously for nasal tissue; γ is the Hill coefficient that controls the slope of the curve; IIV is the inter-study/challenge inoculum dose variability expressed in standard deviation units.

Figure 2. Clinical vaccine efficacy predictions via translation of the RM MBMA Model. Each point represents the predicted vaccine efficacy using the RM MBMA translation versus the corresponding reported clinical vaccine efficacy. The error bars represent the 95% confidence intervals for the vaccine efficacy. The black dotted line is a reference (the line of identity along which predicted VE = observed VE), RMSPE – root mean square prediction error.
| Sponsor       | Vaccine          | Dosage      | Regimen   | Second Dose Day | SN Assay | Age Strata       | SN Title (GM) | SN Title 95% CI low | SN Title 95% CI high | SN Data Source | HCS Data Source | Efficacy 95% CI low | Efficacy 95% CI high | Efficacy 95% CI high |
|---------------|------------------|-------------|-----------|-----------------|----------|-----------------|---------------|----------------------|----------------------|-----------------|-----------------|---------------------|---------------------|---------------------|
| Pfizer        | BNT162b2         | 30 µg       | 2 Dose    | 21              | SN       | Younger adults  | 360.9         | 237.3                | 544.4                | Fig 6           | 93.7            | 61.8                | 143.1               | Figure 4B           |
| Pfizer        | BNT162b2         | 30 µg       | 2 Dose    | 21              | SN       | Older adults    | 155.7         | 90.5                 | 307.1                | Fig 7           | 95.7            | 61.8                | 143.1               | Figure 4B           |
| Moderna       | mRNA-1273        | 100 µg      | 2 Dose    | 28              | SN       | Adults          | 1179.4        | 1130.5               | 1230.5               | Table 15        | 321             | 237.1               | 440.3               | Figure 3B           |
| Janssen       | Ad26.COV2.S      | 0.5*10^11   | 2 Dose    | 28              | SN       | Younger adults  | 223           | 158                  | 318                  | Table 3         | 522             | 301.3               | 911                  | Figure 28           |
| Janssen       | Ad26.COV2.S      | 0.5*10^11   | 1 Dose    | 28              | SN       | Older adults    | 258           | 154.6                | 387.4                | Figure 3        | 522             | 301.3               | 911                  | Figure 28           |
| Novavax       | 2019nCoV         | 5 µg        | 2 Dose    | 21              | SN       | Adults          | 1433          | 978.2                | 2099.4               | Table S8        | 453             | 259                  | 790                  | Figure 38           |
| Gamaleya      | Gam-COVID-Vac    | 50 µg       | 1 Dose    | 21              | SN       | Matrix-M        | 44.5          | 31.8                 | 62.2                 | Table S2        | 33.0            | 31.5                 | 34.5                 | Table S6           |
| Astra-Zeneca  | ChAdOx1 nCoV-19  | 5*10^10     | 2 Dose    | 28              | SN       | Younger adults  | 347.4         | 94.8                 | 599.9                | Table 5         | 509.6           | 386.3               | 672.3                | Table 5            |
| Sinovac       | CoronaVac        | 3 µg        | 2 Dose    | 14              | SN       | Adults          | 23.8          | 20.5                 | 27.7                 | Table 56-1      | 163.7           | 128.5               | 208.6                | Fig S6-3           |
| Bharat Biotech| Bbv152           | 6 µg        | 2 Dose    | 28              | SN       | Adults          | 125.6         | 111.2                | 141.8                | Table 4         | 170.2           | 113.2               | 255.9                | Figure 2C          |

Table 3: Clinical Immunogenicity and Efficacy Data used in calibration, validation, and modelling.

(a) 18-55 years & 16-55 years for SN titre and efficacy data, respectively.
(b) 65-85 years & ≥65 years for SN titre and efficacy data, respectively.
(c) ≥18 years.
(d) 18-55 years & 18-59 years for SN titre and efficacy data, respectively.
(e) 65 years & ≥60 years for SN titre and efficacy data, respectively.
(f) 18-84 years & ≥18 years for SN titre and efficacy data, respectively.
(g) 18-55 years.
(h) 18-59 years & ≥18 years for SN titre and efficacy data, respectively.

References 33 and 41 are preprint/non peer reviewed articles.
integrated with the RM MBMA translation to predict their clinical efficacy. The immunogenicity data from phase 1 and 2 trials and the corresponding phase 3 clinical efficacies (Table 3) of six vaccine candidates (BNT162b2, mRNA-1273, Ad26.COV2.S, 2019nCoV, Gam-COVID-Vac, ChAdOx1 nCoV-19) were available at the time of model development and included in the analysis. (This analysis was originally done when only data from mRNA-1273 and BNT162b2 were available, and the resulting calibration (result not shown) was similar). The vaccine SN titres (all obtained from live virus assays, except for ChAdOx1 nCoV-19 for which only pseudo virus data were available) were normalized to their respective GMT HCS titres, as described in Methods, to account for inter-assay differences. The GNLS approach (see Methods) was used to estimate a calibration factor to align the RM-MBMA curve (using the BAL/lung model parameters) with the clinical immunogenicity and efficacy data of the six vaccine candidates. The calibration factor, $F_c$, was estimated to be 143.9 (95% CI: 90.9, 196.9): this number can be interpreted as the average SN titre (across these assays and their corresponding HCS data sets) of a typical HCS sample. An example of calibration with specific numbers can be found below Fig S6. The calibration shows good alignment between the clinical data and RM-MBMA prediction. (Figure S6)

Simulations from the calibrated RM-MBMA model were conducted to predict the HCS-normalized SN titre level expected to provide 50% vaccine clinical efficacy. As shown in Figure 3, the 50% protective titre was predicted to be 24.3% (95% CI: 14.1%, 46.5%) of HCS. To assess the predictive power of this approach, the calibrated RM-MBMA was used to predict (retrospectively) the clinical efficacy of each of the six vaccine candidates based on their respective Phase 1 or 2 HCS-normalized SN titres. Figure 3 shows reasonable agreement between the predicted and observed clinical efficacies for all the vaccine candidates. Predictions from the calibrated RM-MBMA model had a RMSPE of 7% (n=8 data points: 6 vaccines – Ad26.COV2.S and BNT162b2 each had separate efficacies for subjects aged 18-65 and 65-85 years).

Joint MBMA allows simultaneous estimation of protective titres in RM and humans

A joint MBMA model (i.e., one using both RM and clinical data, with species as a covariate) was developed to explore an alternative approach to integrating clinical and RM data in developing a quantitative model predictive of clinical vaccine efficacy. The joint model uses clinical SN titres scaled by HCS GMT titres (see Methods) and assumes that the form of the sigmoidal function describing the efficacy vs. SN titre relationship in humans is the same as that for protection (VL reduction) versus titre in RM, leveraging all the data.
Figure 4. Joint MBMA of vaccine efficacy vs. SN titres in humans and RM informs the 50% efficacious titre levels in both species. The vaccine SN titres in humans are normalized to the mean HCS titre (SN/HCS) using the corresponding assay to account for the assay differences across trials. The vaccine SN titres in RM are reported in log2 units. Vaccine efficacy is expressed as a % on the y-axis. The solid line and the pink shaded region correspond to the predicted median and 95% Confidence Interval in humans from the joint model. The points are the reported SN/HCS titre ratio and the corresponding vaccine efficacy in humans. The error bars represent the 95% confidence interval of vaccine efficacy. The dashed line with the green shaded region and the double dashed line with the blue shaded region correspond, respectively, to the predicted median and 95% Confidence Interval in BAL fluid and nasal swab specimens in RM.

$VE_{max}$ is the maximum vaccine efficacy (parameter shared across species); $\gamma$ is the Hill coefficient that describes the slope of the curve (either for RM or human); $log_2([IT_{50}]_{RM,BAL})$ is the log2 SN titre corresponding to 50% VE in BAL and lung tissue; $log_2([IT_{50}]_{RM,nasal})$ is the log2 SN titre corresponding to 50% VE in nasal tissue; $\frac{SN}{HCS}_{50}$ is the SN/HCS titre ratio corresponding to 50% VE in humans.

| Parameter                               | Estimate (95% CI)                      |
|-----------------------------------------|----------------------------------------|
| $VE_{max}$                              | 1 (fixed)                              |
| $\gamma_{RM}$                           | 2.23 (1.50, 3.33)                      |
| $\gamma_{citin}$                        | 1 (fixed)                              |
| $log_2([IT_{50}]_{RM,BAL})$             | 4.92 (4.06, 5.78)                      |
| $log_2([IT_{50}]_{RM,nasal})$           | 8.57 (6.86, 10.30)                     |
| $\frac{SN}{HCS}_{50}$                  | 0.24 (0.07, 0.77)                      |
simultaneously and enabling simultaneous estimation of (human and RM) model parameters. As described in Methods, the viral load data in RM was transformed to an efficacy scale, so that the response variables were the same for both species. Summary level data from 13 RM studies and clinical data corresponding to six vaccine candidates (BNT162b2, mRNA-1273, Ad26.COV2.S, 2019nCoV, Gam-COVID-Vac, ChAdOx1 nCoV-19) were used in the development of the joint model.

Using the same model function for both species allowed flexibility in the number of shared parameters between them. During model development, initially all the model parameters were shared between the two species except for the 50\% protective titres (\( \log_2 [\text{IT}_{50,\text{RM,BAL}}] \), \( \log_2 [\text{IT}_{50,\text{RM,nasal}}] \frac{\text{SN}}{\text{HCS}} \)) in RM and humans. While this model was able to describe the data used for model building reasonably well, it did not perform well in predicting additional clinical efficacy data used for model validation (results not shown). Therefore, additional species-specific Hill coefficients (\( \gamma \)) were identified. Human \( \gamma (\gamma_{\text{clin}}) \) was estimated close to 1, but was imprecise, and was therefore fixed to a value of 1 (AIC reduced by 2 units). A similar approach of fixing \( VE_{max} \) to a value of 1 further reduced AIC by 2 units. The Hill coefficients for human and RM were estimated as 0.21 and 0.62, respectively.

**Figure 5. Clinical MBMA reveals a strong relationship between HCS-normalized titres and VE in humans.** The vaccine SN titres for each trial are normalized to (divided by) their assay’s respective mean HCS titre (SN/HCS) to account for the assay differences across trials. The solid line and the shaded region correspond to the predicted median and 95\% Confidence Interval of clinical MBMA prediction. The points are the reported SN/HCS titre ratio and the corresponding vaccine efficacy in humans. The error bars represent the 95\% confidence interval of vaccine efficacy. The vaccine efficacy observations are weighted by the number of participants in each trial. \( VE_{max} \) is the maximum vaccine efficacy (%). \( ET_{50,\text{human}} \) is the SN/HCS titre ratio corresponding to 50\% vaccine efficacy in humans.

| Parameter       | Estimate (95\% CI) |
|-----------------|--------------------|
| \( VE_{max} \)  | 1.00 (0.94, 1.07)  |
| \( ET_{50,\text{human}} \) | 0.21 (0.15, 0.30) |

The vaccine SN titres for each trial are normalized to (divided by) their assay’s respective mean HCS titre (SN/HCS) to account for the assay differences across trials. The solid line and the shaded region correspond to the predicted median and 95\% Confidence Interval of clinical MBMA prediction. The points are the reported SN/HCS titre ratio and the corresponding vaccine efficacy in humans. The error bars represent the 95\% confidence interval of vaccine efficacy. The vaccine efficacy observations are weighted by the number of participants in each trial. \( VE_{max} \) is the maximum vaccine efficacy (%). \( ET_{50,\text{human}} \) is the SN/HCS titre ratio corresponding to 50\% vaccine efficacy in humans.
coefficient for RM data was estimated to be (2.23, 95% CI [1.50 – 3.33]). The joint MBMA model has a shared error structure and parameters between the two species.

The overlay of the observed vaccine efficacy with the prediction from the joint MBMA (Figure 4, Figure S7) reveals that the final model fit the vaccine efficacy and titre data for all vaccine candidates. The ET50;human parameter estimate (Figure 4) revealed the 50% protective SN titre level to be 24% (95% CI (7%, 77%)) of the geometric mean human convalescent titre level. The 50% protective SN titres in RM were consistent with those estimated in the RM MBMA model.

Clinical MBMA informs protective SN titre levels in humans

While the RM MBMA model could be used for VE predictions with RM data alone, and the calibrated-RM MBMA model enables predictions with the integration of clinical data from one or two vaccine candidates, a stand-alone clinical MBMA required more clinical data. As soon as sufficient phase 3 clinical efficacy (and immunogenicity) data were reported, a clinical MBMA was conducted. This quantified the relationship between HCS-normalized SN titres (SN/HCS ratio) and clinical efficacy using only clinical data. The observed SN titre data from phase 1 and 2 trials and the corresponding phase 3 clinical efficacies (Table 3) of six vaccine candidates against SARS-CoV-2 (BNT162b2, mRNA-1273, Ad26.COV2.S, 2019nCoV, Gam-COVID-Vac, ChAdOx1 nCoV-19) were used in the model development. The sigmoidal relationship form described in Eq. (12) was used to quantify the relationship. The 95% CI of the log of the Hill coefficient estimate contained 0 (0.29; 95% CI (-0.88, 1.48)), and hence was fixed to 0 in the final model.

The overlay of the observed vaccine efficacy with the prediction from the clinical MBMA (Figure 5) revealed that the final model fit the % vaccine efficacy and SN/HCS titre data for all vaccine candidates. The model-estimated parameters are also provided in Figure 5 and the ET50;human parameter estimate revealed the 50% protective SN titre level to be 21% (95% CI – (15%, 30%)) of the geometric mean human convalescent titre level.

Independent validation of vaccine efficacy predictions

An independent validation of the three MBMA models — RM MBMA and its translation via the calibration method, the Joint MBMA, and the Clinical MBMA — was carried out to assess their predictive ability. The SN/HCS titre data of 2 vaccine candidates, BBV152 and CoronaVac, were used as input for the three models to predict the corresponding clinical efficacies. Table 4 reveals that the predicted vaccine efficacies for the two vaccine candidates using the three approaches are consistent with the reported efficacies.

### Table 4: Comparison of vaccine efficacy predictions using clinical SN data.

| Vaccine       | Reported Phase 3 VE, % (95% CI) | Calibration Method VE, % (95% CI) | Joint MBMA VE, % (95% CI) | Clinical MBMA VE, % (95% CI) |
|---------------|---------------------------------|----------------------------------|---------------------------|-----------------------------|
| BBV152        | 77.8 (65.2 – 86.4)              | 75.3 (60.3 – 87.9)               | 76.2 (48.2 – 91.3)        | 77.2 (67.1 – 84.0)          |
| CoronaVac     | 50.7 (35.7-62.2)                | 34.3 (15.5 – 51.8)               | 38.0 (15.6 – 67.7)        | 40.9 (30.6 – 51.3)          |

### Table 5: Predictions of vaccine efficacy against the delta variant.

| Vaccine                  | Fold Decrease in SN Titre (Delta relative to WT) GM (SE) | Calibrated RM-MBMA VE, % (95% CI) | Joint MBMA VE, % (95% CI) | Clinical MBMA VE, % (95% CI) | Reported efficacy VE, % |
|--------------------------|----------------------------------------------------------|----------------------------------|---------------------------|-----------------------------|--------------------------|
| ChAdOx1 nCoV-19          | 4.29 (1.13)                                              | 57.0 (18.2, 88.9)                | 41.1 (11.6, 78.0)         | 43.0 (18.8, 67.7)           | 59.8, 60                 |
| BNT162b2                 | 2.49 (1.13)                                              | 85.3 (67.8, 96.3)                | 87.4 (58.8, 98.0)         | 87.4 (75.2, 96.5)           | 79, 87.9                 |

(a) Geometric mean: calculated as exp(mean(ln(x))) where x indicates 25 values obtained from Table S4 of the original publication.

Geometric SE: calculated as exp(SE) where SE=SD(ln(x))/25 and x indicates 25 values obtained from Table S4 of the original publication.

(b) SN titres for the age group 18 – 65 years were used for efficacy predictions.

Reference 44 is a preprint/non peer reviewed article.

Prediction of vaccine efficacy against newer strains of SARS-CoV-2

To further validate and to investigate the relevance of MBMA-based efficacy predictions to emerging variants of concern, clinical efficacies against the delta variant were predicted from reported fold change in SN titre against the delta variant relative to wild-type. The three MBMA models provided reasonable efficacy predictions when compared with the reported efficacies as shown in Table 5.
Discussion

Quantitative models that use only nonclinical and or Phase 1 immunogenicity data to predict clinical vaccine efficacy against SARS-CoV-2 provide a powerful translational framework to inform critical vaccine development decisions. These decisions include informing dose selection, target population, early Go/No Go (preventing execution of Phase 3 efficacy trials where success is unlikely), efficacy against newer strains, and expected efficacy with booster vaccinations. Here, we present this translational framework through the development of MBMA models by integrating prior information across different vaccine mechanisms, species, doses/regimens, endpoints, assay methods, and study designs. This modelling was used to help inform decisions on vaccine candidates, as reported elsewhere.45,46

Early in the pandemic, as challenge studies demonstrating vaccine protection against SARS-CoV-2 in RM were published (and no clinical data were available), we developed a RM MBMA model using data from 13 vaccine studies that quantified the robust relationship between SN titres post vaccination and peak VL post challenge in RM. The estimated 50% protective titre in RM in the BAL/lung tissue and nasal swab tissues were 5.84 (5.2 - 6.5) [mean (95% CI)] log2 units, and 8.3 (7.1 - 9.5) [mean (95% CI)] log2 units respectively, suggesting higher SN levels required for protection (defined as reduction in peak VL relative to sham-treated animals) in the nasal tissue compared to the lower respiratory tract. Similar findings were reported by Corbett et al.72 while evaluating immune correlates of protection after mRNA-1273 vaccination they concluded that, in RM, protection in the lower respiratory tract was achieved at lower serum antibody concentrations than in the upper respiratory tract. This outcome was also supported by the data from challenge studies in RM which were used as input to the RM MBMA model, wherein only a few vaccine candidates eliciting the strongest immune responses demonstrated protection in the upper airways in addition to the lower airways.11,12,19 This trend has also been seen in other respiratory viruses such as respiratory syncytial virus.57 A potential explanation for the higher infectivity in nasal tissue may be the skewed receptor distribution towards the upper airways, which has been reported for SARS-CoV-2.43 This may lead to differences in reproductive ratio between tissues, warranting higher titres and thus higher \( \frac{RT_0}{n} \) in nasal specimen to achieve the same level of virus inhibition.49

Results further demonstrate that non-clinical immunogenicity data alone can be predictive of clinical efficacy through the translation of the RM MBMA protection model to a clinical efficacy model. The clinical efficacy predictions from the translational model using the BAL/lung tissue model parameters performed better than the predictions from nasal swab model parameters (RMSPE of 12% vs 21% respectively). While reduction of viral load in the lower airways is relevant clinically in reducing moderate to severe disease, reduction in both upper and lower airways could be relevant clinically in preventing mild disease and reducing risk of transmission.4 This potentially explains why the BAL/lung tissue model was more accurate in predicting clinical vaccine efficacies, given that the definition of the efficacy endpoint across vaccine trials is symptomatic disease of any severity confirmed by PCR. Therefore, the BAL/lung tissue parameters are more relevant to the definition of efficacy in the clinical trials, and hence were used for clinical efficacy predictions in the calibration and the joint MBMA models.

As phase 1/2 clinical immunogenicity and efficacy data of COVID-19 vaccine candidates became available, we integrated the data into the RM MBMA translation through the calibration method or combined them with the RM data to develop a joint MBMA model. These two approaches provide quantitative frameworks to predict clinical efficacy of vaccine candidates when there are insufficient clinical data to develop a standalone clinical MBMA model. In the calibration method, a calibration factor was estimated to align the RM-MBMA curve with the clinical immunogenicity and efficacy data. The leave-one-out cross validation approach50 (Figure S8) showed that the calibration factor was relatively insensitive to any trial data being excluded during estimation. Integration of phase 1/2 clinical immunogenicity data improved the predictions of VE compared to predictions from RM alone (RMSPE 7% vs 12%). Using species as a covariate in a joint MBMA model provides an alternative approach to integrating clinical and RM data for predicting (quantitatively) clinical vaccine efficacy. The joint model assumes the same curve (titres appropriately normalized) for both humans and RM, thereby enabling simultaneous modelling of all the data while providing flexibility in the number of shared parameters between the two species during estimation. The joint MBMA model also reasonably predicted VE (RMSPE 6.6%).

Once sufficient phase 3 clinical data were reported, clinical data alone enabled an MBMA to quantify the relationship between HCS-normalized SN titres (SN/HCS ratio) and clinical efficacy. The 50% protective titres predicted from the calibrated RM MBMA model, joint MBMA model, and the clinical MBMA model were consistent across the three approaches (24.3% (95% CI: 14.1%-46.5%), 24.1% (95% CI: 7.0%, 77.0%), and 21.0% (95% CI: 15.0%, 30.0%), respectively, of HCS). The consistency in the point estimate provides additional confidence in leveraging these approaches for VE predictions based on the data available at a given stage of development for a new vaccine candidate. The larger CI obtained with the joint MBMA model is perhaps due to the model describing the protection data in both species simultaneously.
We also assessed the predictive ability of these three models through the prediction of clinical efficacies of BBV152 and Coronovac, whose data were not included in the development of the models. The predictions were carried out using only their clinical SN/HCS ratio as the input to the MBMA models. All three models accurately predict the efficacy of BBV152 and the 95% CI. While the predicted 95% CI of Coronavac encompasses the observed efficacy point estimate, the point estimate is not accurately predicted. Coronavac has a reported clinical efficacy of 50% and therefore occurs in the most sensitive region of the efficacy vs. titre relationship. We attribute the lack of accuracy in the efficacy predictions of Coronavac to the limited data and the large uncertainty in normalized SN titres corresponding to a ~50% vaccine efficacy in the available clinical dataset. ChAdOxi nCoV-19, for which data were included in model building, had 5-fold higher normalized SN titres than Coronavac and reported efficacy similar to Coronavac’s. When Coronavac was included in the clinical MBMA model estimation as a sensitivity analysis, model parameters changed only marginally: the IT50 estimate was slightly lower, 0.186 (95% C.I.: 0.111-0.265) as compared to 0.21 when Coronavac was not included in the model estimation. In addition, when the model was refit including BBV152 and Coronavac in the clinical MBMA model and excluded ChAdOxi nCoV-19, and BNT162b2, the RMSPE for the new model was lower (3.6%) compared to the original model (6.9%), indicating an improvement in fit. However, the model parameters changed only marginally (about 15% for IT50), therefore suggesting the model is relatively robust to the difference in data sets. To highlight another important application, we predicted the efficacies of two vaccines (BNT162b2, and ChAdOxi nCoV-19) against the delta variant using the reported drop in in-vitro SN titres for that variant compared to the original strain. Although the efficacy predictions of BNT162b2 are more accurate than those for ChAdOxi nCoV-19, the data suggest that, overall, the three models have demonstrated reasonable predictive power, especially given that only the in vitro data of the fold decrease in SN titre of the variant relative to WT was leveraged in predicting clinical efficacies for the corresponding variant. The wider 95% CIs are due to the wider variability of the titre ratios that needed to be incorporated in the predictions. Therefore, the VE predictions consistent with the reported efficacies of these vaccines against the delta variant demonstrate the utility of these models to predict vaccine efficacies against variants. In addition, these results demonstrate SN titres could be predictive of efficacy regardless of vaccine mechanism or variant, and therefore suggesting that this approach may also be predictive of vaccine efficacy after a booster dose. Recently, the fold change decrease in SN titres against the omicron variant after three doses (relative to SN titres against the WT strain after two doses of the vaccine) have been reported for a few vaccines (Table S5).51-53 Based on these reports, the model-predicted VE for BNT162b2, mRNA-1273, and 2019nCoV vaccines against the omicron variant would be 83.4% (95% CI: 72.3% -90.4%), 89.8% (95% CI: 85.5% -93.0%), and 92.7% (95% CI: 86.7% -96.3%) respectively. (No quantitative estimate of efficacy against omicron was available currently.) Recent literature based on nonclinical species as well as human data support that SN titre time-course seems predictive of efficacy at a given time (i.e., it is the SN titre at the time of exposure that determines efficacy, not just the peak titre after vaccination), and efficacy after vaccine booster shots has been predicted elsewhere.54 [Reference 54 is a preprint/non peer reviewed article]

The quantitative relationship the model provides between clinical efficacy and titres enables dose selection using only a clinical dose-immunogenicity relationship (for a given vaccine candidate). The presented model relates immunogenicity to protection in RM and to relative incidence rate observed in clinical trials. Therefore, an initial prediction based on nonclinical data requires the assumption that the immunogenicity in RM of the dose used in RM represents the (calibrated) clinical immunogenicity in a manner consistent with the other vaccines tested. In the examples used here, the fit and prediction results generally seem to indicate that the assumption is supported when the RM and clinical doses are the same. If dose-ranging is performed in RM, then a dose which provides a “plateau” response would, based on the model, be assumed to provide maximal possible clinical efficacy for that vaccine. Translation of the dose-response itself requires a different kind of modelling than that used here, and this becomes substantially more complex when multiple doses are required so that inter-dose timing becomes an additional (and important) factor. As suggested above, using dose-immunogenicity modelling on clinical data can provide more direct evidence of a plateau in immune response (caveats about multi-dose regimens are similarly applicable). The framework presented here illustrates how such emerging clinical immunogenicity data can be used to predict efficacy and, once corresponding efficacy data are available, and how to refine predictions based on a learn-and-confirm paradigm.

There are limitations to our analysis. Different neutralization assays were used across studies. Therefore, normalization of SN titres with the corresponding convalescent serum titre using the same assay was carried out (even though the definition of convalescence was not uniform across studies), as has been established by other groups8,11 as a reasonable approach to harmonize SN titres across studies. Vaccine studies in RM also used different types of neutralization assays (PRNT, MN, and CPE assays), and the type of virus (live vs. pseudo) used in those assays also differed. Unlike clinical trials, convalescent titres using the corresponding
assays were either not reported or the definition of convalescence was not clearly stated in most of these RM studies. Hence the approach to normalize the SN titres to their corresponding convalescent titres was not carried out for the RM data. The RM MBMA analysis established a strong relationship between SN titres and viral load without normalizing the SN titres, probably because of the relatively large number of vaccine studies (data points) included in the RM MBMA model (essentially enabling the model to average out the impact of assay differences), a different circumstance than that for the relatively sparse clinical data. We found a notable consistency in the data that used different assays and no systematic impact of the assay differences was found on the relationship between SN titres and VL in RM. This seems supported by literature, indicating substantial variability in assay results, but no clear systematic differences between live and pseudo virus assay results.55–57

The translation of the RM MBMA model assumed that the relationship of log(viral load) to log (SN) has the same shape as that of incidence rate (IR) to log (SN) (an assumption about the mapping of viral load in RM to clinical incidence rate). The assumption was made based on previous investigations in RSV14 which successfully used the same type of mapping. The efficacy predictions resulting from the mapping were in reasonable agreement with the observed data. However, a reasonable fit does not constitute an independent validation of the translational assumption, and therefore this assumption is a limitation of the translational approach requiring support by additional (future) data. Another limitation of our approach is that we have only looked at neutralizing antibody data as predictor of efficacy given the strong evidence of its role in protection. Cellar responses (T and B cell memory) and non-neutralizing Fc effector functions could also potentially contribute to protection, and as additional evidence on their role in protection becomes available that could also be modelled.

In summary, the MBMA models presented here are predictive of vaccine-induced efficacy against SARS-CoV-2. Although there are reports in the literature demonstrating neutralizing antibodies as a CoP in RM,15–33 this study adds the translation of protection in RM to clinical efficacy based on SN titres. This translation provides a powerful quantitative framework, enabling early prioritization of (and Go/No Go decisions for) vaccine candidates in discovery before substantial time and resources are invested in the clinical development of those candidates. The predicted 50% protective titres from the MBMA models also corroborate previous findings by Khoury et al.,4 and are consistent with more recent work by Padmanabhan et al.,33 increasing the confidence in that number. The models also can predict VE against variants of concern (including delta and omicron variants). Overall, the correlates-of-protection models relating SN titre to clinical outcome presented here are also generalizable to vaccines beyond SARS-CoV-2, as recently shown by Maas et al. for RSV.34 The work shown here serves as an important example of a framework to predict VE using only protection data in animal models, and thereafter a framework to incorporate early clinical immunogenicity data to improve VE predictions driving model-informed decision-making during vaccine development.

Contributors

BK, NP, AC, AL, SR, SYAC, RDG, RJS, and JRS were involved in conceptualization of this study. BK, NP, AC, AL, KW, RT, SYAC, and JRS managed data curation for this study. Formal data analysis was performed by BK, NP, AC, AL, SR, SAD, SYAC, and JRS. Funding for this study was acquired by JRS. Investigation was performed by BK, AC, SR, SYAC, and JRS. BK, NP, AC, AL, SR, SYAC, JS, and JRS developed or designed methodology for this study. BK, SYAC, RDG, and JRS served as project administrators. AC, AL, SR, and SYAC provided software assistance. BK, KW, SYAC, RDG, JS, and JRS played a supervisory role. Validation of research outputs and verification of underlying data was provided by BK, NP, AC, AL, SR, SYAC, and JRS. BK, NP, AC, AL, SR, RDG, and JRS prepared or created visualization/data presentations for this study. BK, NP, AC, AL, KW, SYAC, and JRS wrote the original draft and all authors participated in reviewing and editing the manuscript. All authors read and approved the final version of the manuscript.

Data sharing statement

The data sharing policy, including restrictions, of Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA, is available at http://engagezone.msd.com/ds_documentation.php through the EngageZone site or via email to dataaccess@merck.com. (For the purposes of this manuscript, this also applies to code).

Declaration of interests

Certara received funding from Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA, for modelling work. BK, AC, SR, SAD, JS, and JRS are employees of Certara, Princeton, NJ, USA and hold stock and/or stock options in Merck & Co., Inc., Rahway, NJ, USA, and have received support for attending meetings and/or travel. NP, AL, KW, RT, and RDG are employees of Certara, Princeton, NJ, USA and holds Certara and AstraZeneca stock or stock options. SYAC is an employee of Certara, Princeton, NJ, USA and holds Certara and AstraZeneca stock or stock options.

Acknowledgements

We thank Maheswar Ganta, Marianne van Vugt, Nina Tsuneda, Niharika Gandhapuneti, and Neelima...
Kankanampati from Certara Data Science Group for literature search and review, digitization of literature data, database development, curation, augmentation, and QC. We also thank (both of Certara, Princeton, NJ) Jos Lommerse for valuable scientific input on MBMA model development, and Amy Porter for medical writing support. We thank (all of Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA at the time of this analysis); Amy Epseseth, Kalpit Vora, Andrew Bett, Antonios Aliprantis, Kevin Russell and many others from Vaccine Discovery, Clinical, Translational Medicine, Bioanalytical, and Quantitative Pharmacology and Pharmacometrics for careful review of the abstracts, for valuable input on assays and model assumptions, and for critical review of the manuscript. This study was funded by Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA.

Supplementary materials
Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jebiom.2022.104264.
