medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Predicting the Effectiveness of the Pfizer-BioNTech BNT162b2 Vaccine from SARS-CoV-2 Variants Neutralisation Data

Oleg Volkov[1,*], Svetlana Borozdenkova[1], and Alexander Gray[2]

[1]Xitific LTD, London, UK
[2]IDEAPharma, Milton Keynes, UK
[*]Corresponding author: Oleg Volkov, oleg@xitific.blue

Abstract

Modelling frameworks for vaccine protection are sorely needed to fight the Covid-19 pandemic with vaccines. We propose such a framework for the BNT162b2 and potentially other vaccines. It utilises correlates of protection based on live SARS-CoV-2 variants neutralising antibody titres from vaccinated individuals. We applied it to predict vaccine effectiveness in overall populations and age subgroups, following partial or full vaccination. It was validated by predicting effectiveness against D614G, B.1.1.7 (Alpha) and B.1.167.2 (Delta) variants. The predictions for Delta were 44.8% (22%, 69%) after one and 84.6% (64%, 97%) after two vaccine doses, which were close to the corresponding means, 45.6% and 85.5%, of observations in real-life effectiveness studies. Predictions for the other variants were also accurate. We also consider methods of safeguarding predictions against biases. Finally, we discuss implications for vaccination policies, such as decisions about doses and timing of vaccine boosters.

Keyword: Covid-19 vaccines; BNT162b2 vaccine; Correlates of protection; Vaccine effectiveness predictions; SARS-CoV-2 variants neutralisation; Delta B.1.167.2

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

1 Introduction

Immune correlates of protection can be a game changer in the fight against the Covid-19 pandemic. They allow prediction of vaccine efficacy and effectiveness from levels of immune response, such as antibody levels. Unlike with months-long clinical trials and observational studies, correlates data can be rapidly obtained from blood sampling and automated in-vitro studies. Life-critical questions — What are the optimal doses and timings for vaccine boosters in a population? How likely is a new variant to evade vaccine immunity? — would be answered more quickly.

A correlates of protection *framework* combines immune correlates with statistical tools for making predictions. Its desirable aspects include

Ease of implementation A framework does not require large human trials. Existing in-vitro technologies are sufficient or easily adapted. Modelling and software tools are simple to use and adapt by practitioners.

Wide scope Crucial prediction tasks are covered, including those to support regimen modifications of established vaccines; accelerated development of new vaccines; tailored policies for population subgroups; and risk assessment of emerging variants.

Accuracy Predictions within a specific task are relatively unbiased and precise.

Improving upon any of these aspects could save lives. And despite the recently proposed frameworks for SARS-CoV-2 vaccines, there is still room for improvement. Moreover, given the gravity and complexity of the pandemic, a one-size-fits-all framework may be elusive, and tailored approaches more suitable.

One key framework was developed within a purpose-designed trial with a subgroup ($N = 1005$) of vaccinated participants in the mRNA-1273 clinical trials (Gilbert et al., 2021). Several correlates of protection were established based on symptomatic Covid-19 cases in this subgroup, and rigorous causality analyses were performed.

Despite being comprehensive, this study does not cover all practically relevant tasks. As pointed out in Gilbert et al. (2021), it did not consider virus variants, predictions after one dose, or predictions for other SARS-CoV-2 vaccines. While the study used pseudovirus neutralisation assays, evidence has been accumulating that live SARS-CoV-2 neutralisation assays could be more informative (Lu et al., 2021). The study also involved sophisticated modelling and computational code, available on GitHub, which may take time to implement in some cases.

Two other groundbreaking frameworks were proposed in Earle et al. (2021) and Khoury et al. (2021), with the latter extended in Cromer et al. (2021). The frameworks predicted efficacy in Phase 3 vaccine trials using live-virus neutralisation titres from vaccinated volunteers in Phase 1/2 immunogenicity trials. To overcome assay disparity between trials, vaccine-induced titres were normalised by mean titres from convalescent plasma, which served as a reference in a given trial. Key strengths of both frameworks include combining data across different trials and vaccines and making accurate predictions of vaccine efficacy.

These frameworks could be potentially improved for predicting real-life effectiveness of vaccines in some practical cases. For example, the frameworks considered full vaccination (e.g., with two doses of

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

mRNA vaccines), whereas prediction after one dose is also important. Their predictions may be increasingly uncertain for emerging SARS-CoV-2 variants, since the assumed Phase 1/2 neutralisation studies were done with the virus Wild Type, whereas Phase 3 efficacy data were collected already with its successors. For example, the predominant genotype during the Phase 3 Pfizer-BioNTech (Comirnaty) vaccine trials was D614G (Korber et al., 2020), which could markedly decrease neutralisation compared with the Wild Type (e.g., by 2.3 folds according to Wall et al. (2021b)). Delta and successor variants would be further steps away from the Wild Type. Another task is prediction for age groups, which was not addressed by the frameworks.

A key point is that although a correlates of protection model could fit well, if its inputs are not representative, so could be the inferences from the model. When, for instance, implementing the framework in Khoury et al. (2021), practitioners would need to heed several sources of error due to modelling inputs:

- Neutralisation assays can differ not only in how they measure (i.e., the technology) but also in what they measure (e.g., $PRNT_{80}$ versus $MN\,IC_{50}$; Khoury et al. (2021)). There are also numerous other sources of inter-laboratory variation.
- Differences among vaccines and particularly between vaccine platforms (e.g. mRNA versus adenovirus vector) could influence predictions.
- Convalescent plasma samples may not be representative of wider convalescent populations.
- Individuals in early-stage immunogenicity studies may not generalise well to wider populations in Phase 3 trials and beyond.

The last two points are exacerbated by the often small size of early immunogenicity studies, such as 15 vaccinated and three convalescent subjects in the Moderna study (Jackson et al., 2020) assumed in Khoury et al. (2021).

Normalisation by the mean convalescent titres is beneficial as it allows combining data from disparate studies (Khoury et al., 2021; Earle et al., 2021). However, it may not always suffice, with substantial differences remaining even for neutralisation studies with the same vaccine. It also brings an extra distribution (of convalescent titres) into analysis, and hence extra variability. It could also introduce biases. For instance, unvaccinated younger Covid-19 patients could produce lower antibody levels than older patients would, perhaps due to less severe disease (Li et al., 2021); in vaccinated age groups the opposite occurs. Such effects, as well as different age distributions of individuals in convalescent samples, could bias predictions, particularly for age groups. Furthermore, convalescent data may be not available in some SARS-CoV-2 neutralisation studies.

Building upon Earle et al. (2021) and particularly Khoury et al. (2021) we developed a novel modelling framework of immune correlates, which addresses some of the above issues. It could also complement the existing frameworks, particularly in predicting effectiveness, including that after partial vaccination; in age subgroups; and against SARS-CoV-2 variants. The assumed correlates are the neutralising titres against SARS-CoV-2 variants from a large-scale "Legacy" study of live-virus neutralisation at the Crick Institute (Wall et al., 2021b). Normalisation by convalescent titres was not required, thus its associated errors were eliminated.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Although the framework was developed for the Comirnaty mRNA vaccine, it could apply to other vaccines with sufficient neutralisation and effectiveness, or efficacy, data. This focus on one vaccine at a time eliminates error due to vaccine differences.

Generalisability of neutralisation study populations to wider populations is still a concern. We therefore investigated the impact of population differences and suggested mitigation strategies. Like Earle et al. (2021) and Khoury et al. (2021), but unlike Gilbert et al. (2021), our framework does not address causality. It does, however, add an indirect support for the use of live -SARS-CoV-2 neutralisation titres as correlates of protection.

Our framework makes several trade-offs. For instance, its current implementation assumes a single vaccine and a single neutralisation study, and therefore does not use data from other vaccines and studies. We are currently looking into enabling this and into other extensions.

2 Results

2.1 Effectiveness modelling

To model symptomatic vaccine effectiveness we used its estimates from the Pfizer-BioNTech Phase 3 vaccine trial (FDA, 2020; Polack et al., 2020) and from seven large observational studies with the Comirnaty vaccine post-authorisation. We included studies that either reported effectiveness for specific variants, or in which the single dominant variant could be established from study dates and locations. These studies were conducted in Canada (Nasreen et al., 2021), France (Charmet et al., 2021), Scotland (Sheikh et al., 2021) and two in England (Bernal et al., 2021; Pouwels et al., 2021) and Israel (Haas et al., 2021; Mor et al., 2021) each. (See the Methods Supplement for details of the studies.)

The correlates of immune protection were based on data obtained at the Crick Institute within the ongoing Legacy Study (Wall et al., 2021b). The study ran one cohort with 149 UK healthcare workers who had had a single dose of the Comirnaty vaccine, and another cohort, with $N = 159$, who had had both doses. Blood plasma from cohort participants was tested in automated live-virus neutralisation assays against the D614G, B.1.1.7 (Alpha), B.1.351 (Beta) and B.1.167.2 (Delta) variants. The features of the resulting neutralisation titres are summarised in the Supplement.

Each distribution of neutralisation titres per dose and variant was paired with the effectiveness observations for the dose and variant. These data were fitted using the population protection function (2), considered in the Supplement, which is common in correlates of protection research (Khoury et al., 2021; Dunning, 2006; Nauta et al., 2009).

Unlike Khoury et al. (2021), who assumed a Normal distribution for vaccine efficacy, we assumed a beta distribution for vaccine effectiveness. The latter distribution is more suitable for outcomes on a bounded interval — such as between 0% and 100% in our case — than a normal distribution with an unbounded domain. Given our assumption, we fit the population model using nonlinear beta regression. To the best of our knowledge beta regression modelling has not been applied in vaccine research. (Modelling and methodology details are in the Supplement.)

The model fit is plotted in Figure 1 against the geometric mean titre (GMT) per variant sample. (As

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

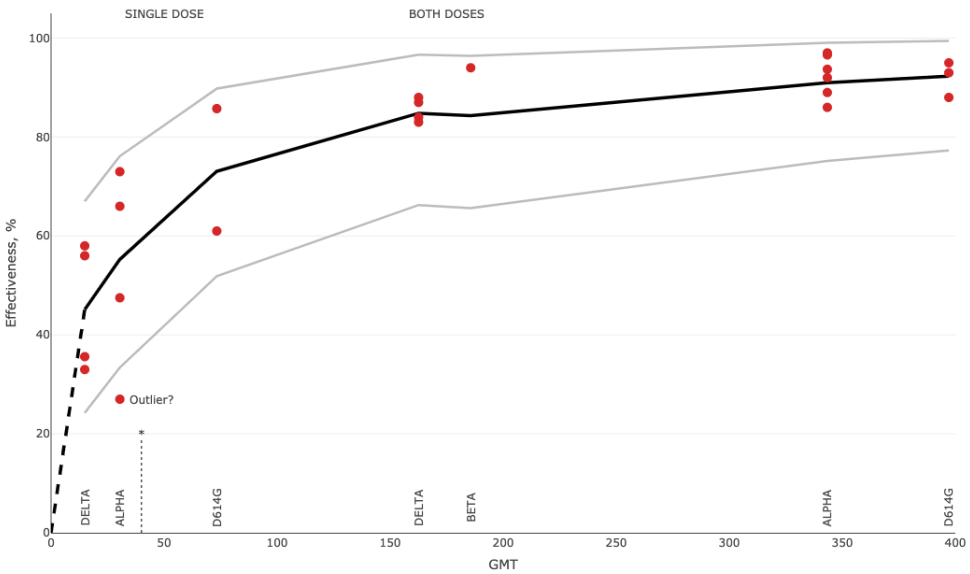


Figure 1: Protection model fit to symptomatic effectiveness observations. Each group of observations corresponds to the variant indicated and is plotted versus the geometric mean titre from the corresponding Crick sample. The 95% prediction bands are shown. The dotted line indicates the lower quantification threshold at $GMT = 40$. The dashed line extends the model fit beyond the data range towards zero.

our model function inputs empirical distributions, this figure is a simplified 2D representation of a multi-dimensional problem. Caution is therefore required in interpretation. More details on this, along with an approximation of the model, are considered in Appendix B.)

Figure 1 suggests a substantial nonlinear association between neutralising titres from the Crick study and the observations from the effectiveness studies. The association is well represented by the assumed model. This is despite the observational studies being performed in different countries, at different times post-vaccination and in diverse populations.

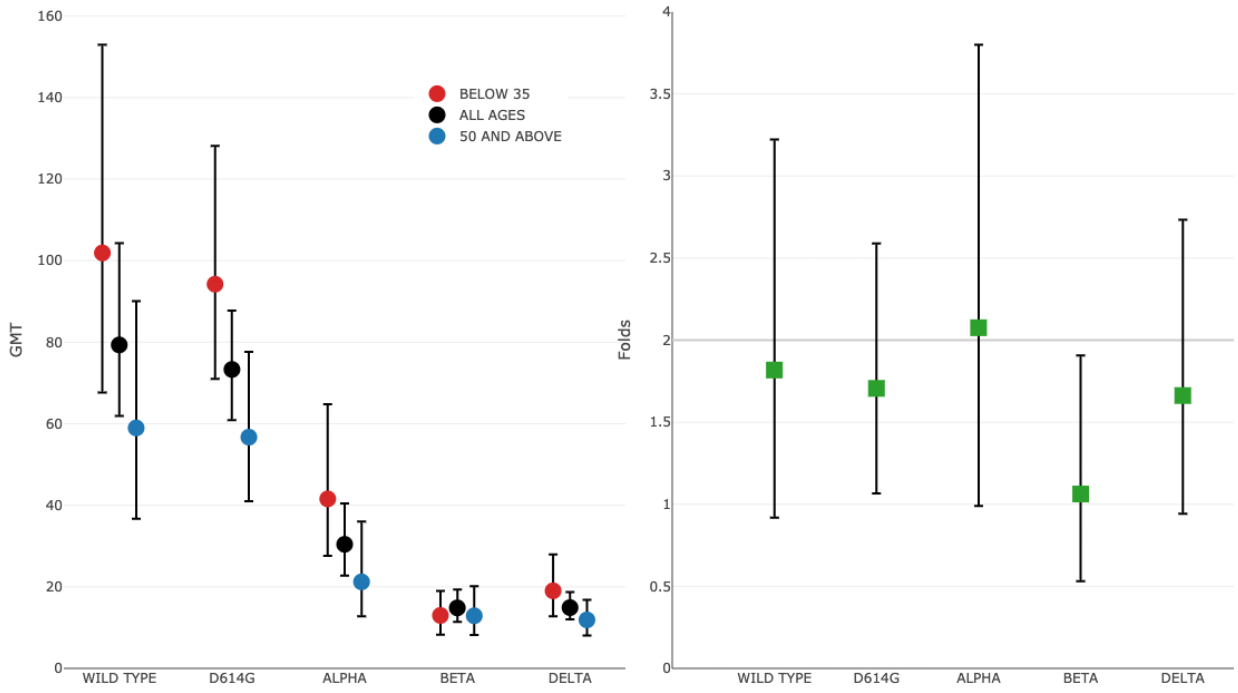
The model also characterised the prediction uncertainty well, with 23 out of 24 observations lying within 95% prediction interval. The exception is the single-dose symptomatic effectiveness, in %, of $27(13, 39)$ against the Alpha variant from the Scotland study in Sheikh et al. (2021). This is a potential outlier, as it is smaller than every symptomatic effectiveness estimate against the Delta variant in the studies considered here. It is also smaller than the effectiveness against infection caused by either the Alpha $38(29, 45)$ or Delta $30(17, 41)$ variant in Sheikh et al. (2021). That our model identified this outlier gives an additional reassurance about its quality.

Variability in the effectiveness observations is greater after a single dose and peaks in the region of 50% effectiveness, a pattern echoed by the prediction bands in Figure 1. This is unsurprising, since the uncertainty of a beta random variable is maximal in the middle of its domain. Hence there is additional empirical evidence for assuming a beta distribution.

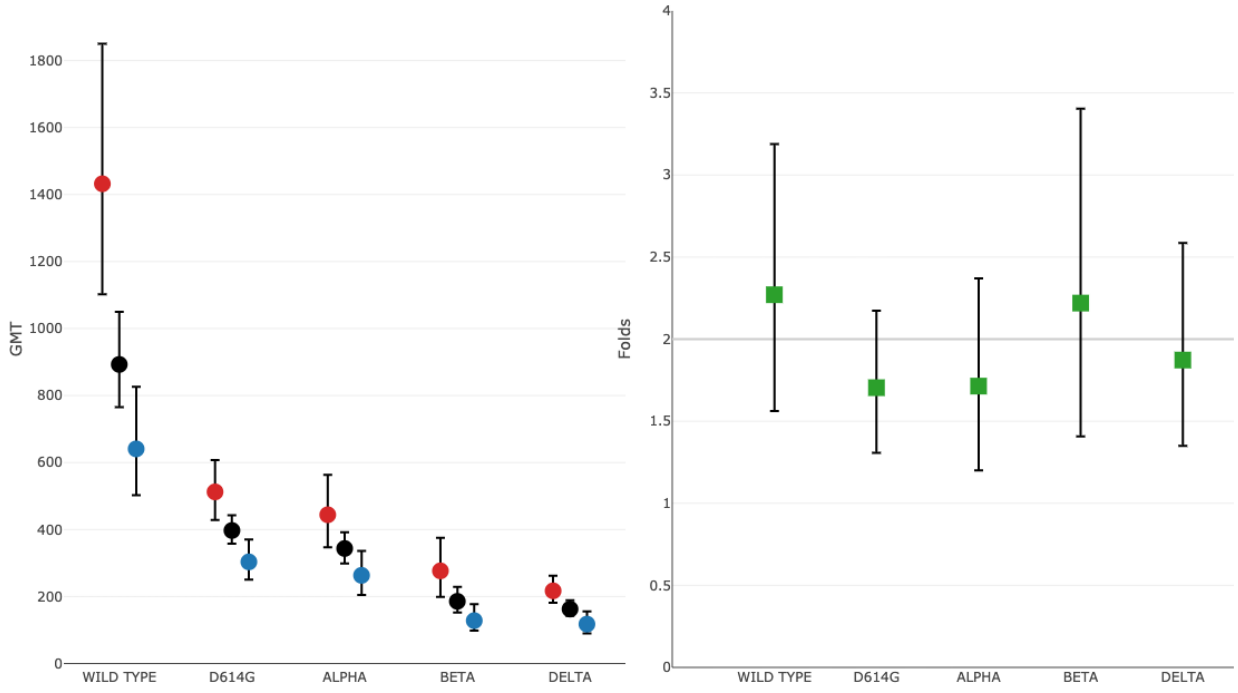
Age subgroups Differences in immune response between age groups could lead to substantial differences in vaccine effectiveness and thus could be relevant to vaccination policies.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .



(a) Single dose cohort.



(b) Both doses cohort.

Figure 2: Geometric mean titres and their folds in age groups. Left: GMT values for those below 35; all ages; and for 50 and above. Right: GMT folds are for those below 35 versus those 50 and above. Also shown are 95% bootstrap confidence intervals. Data source: Wall et al. (2021b)

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

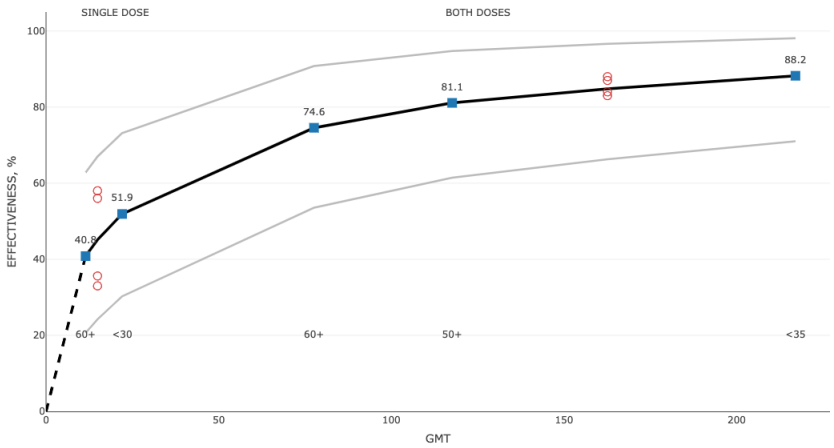


Figure 3: Predicted effectiveness against the Delta variant for selected age groups. The model fit, also in Figure 1, for all ages was used for this prediction. Also shown are effectiveness observations against Delta obtained in all-age groups.

We first compared the Crick neutralisation data per dose cohort in those below 35 with those 50 and over and with the overall study population. (The participants' ages were roughly between 20 and 70.) We used non-parametric bootstrap to estimate the GMT per each of these groups and the fold ratio of GMTs for the younger versus older groups. As illustrated in Figure 2, in almost all cases, the younger group had 1.6-2.3 times higher GMT than the older group. The exception was for the Beta variant after a single dose, possibly due to most individual titres being near the lower detection threshold (see Figure 10 in the Supplement).

We used the model fit (shown in Figure 1), obtained for all-age neutralisation data, to predict effectiveness against the Delta variant in selected age groups; see Figure 3. Note that not only did the effectiveness decrease with age also, but also the prediction uncertainty increased.

2.2 Predicting effectiveness against variants

Predicting vaccine effectiveness against a new SARS-CoV-2 variant could be required once it starts spreading, but no studies of vaccine effectiveness have been completed with it. (We assume that rapid neutralisation studies with the variant can be done.)

We first applied our framework to predict effectiveness against the Delta variant. The model was fitted using neutralisation and effectiveness data on the assumed variants except Delta. Next, the neutralisation titres against Delta were input into the fitted model to predict effectiveness against this variant. These predictions were compared with estimates of symptomatic effectiveness against the variant in Nasreen et al. (2021), Sheikh et al. (2021), Bernal et al. (2021) and Pouwels et al. (2021), detailed in Appendix D. The model fit and observations are plotted in Figure 4.

Similar predictions were done for D614G and Alpha (but not for Beta, as only one effectiveness estimate was available.) The results are illustrated in Figure 5. The predictions were remarkably close to the corresponding mean effectiveness values from effectiveness studies. For instance, for the Delta variant there was underprediction of only 0.8% after one dose, and of only 0.9% after two doses. The largest error —

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

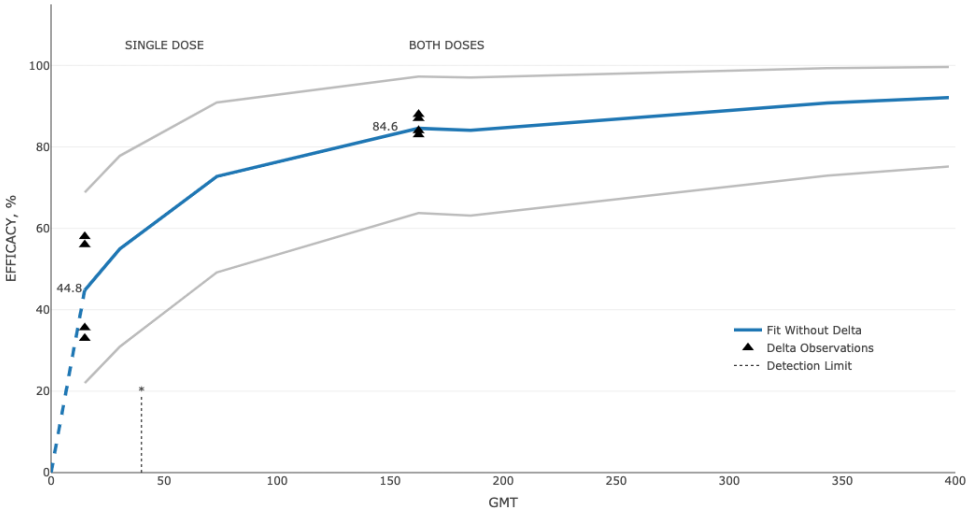


Figure 4: Predicted effectiveness against a variant. The model was fit using the effectiveness and neutralisation data on all the variants except the indicated variant. Prediction was evaluated at the neutralisation titres of the variant.

overpediction of 3% — was for the Alpha variant after one dose.

The prediction uncertainty appears realistic; see Figure 5. In all but one cases the 95% prediction interval included the corresponding estimates from the effectiveness studies. The exception was the effectiveness estimate after one dose from the Scotland study (Sheikh et al., 2021). (Recall that this estimate is a suspected outlier; see Section 2.1.) There was also good coverage of the confidence intervals for those estimates.

Notably, prediction is more difficult after a single dose than after both doses:

- There is larger uncertainty in effectiveness values around 50%, as evidenced by a greater spread of effectiveness observations.
- The neutralisation titres against Delta after one dose were outside the neutralisation data range used in the fitting of the other variants.
- The corresponding GMT was 13.3, well below the lower detection threshold of 40; see Figure 4. Left censoring could be particularly impactful, and predictions sensitive to the assumed nominal values, 10 and 5, representing weak and no neutralisation signal, respectively (see the Supplement for details).
- As the model function is relatively steep at small titres, any difference between the true and assumed titres is magnified disproportionately.

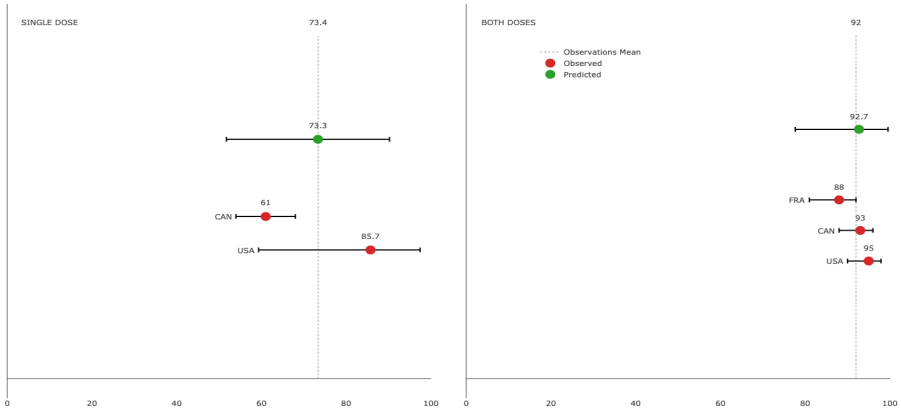
However, as more individuals are fully vaccinated, prediction uncertainty after one dose becomes less practically important.

2.3 Population differences

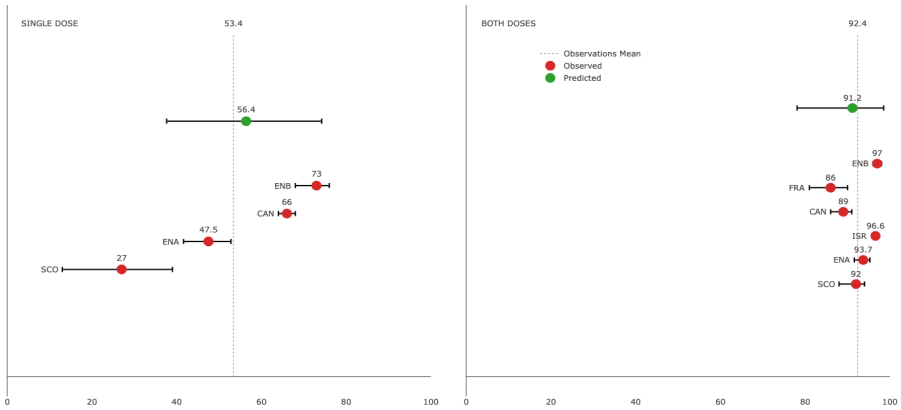
Until now we assumed that the titres distributions in the Crick study were representative of such distributions in the Comirnaty-vaccinated populations in the effectiveness studies. Here we explore how our modelling

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

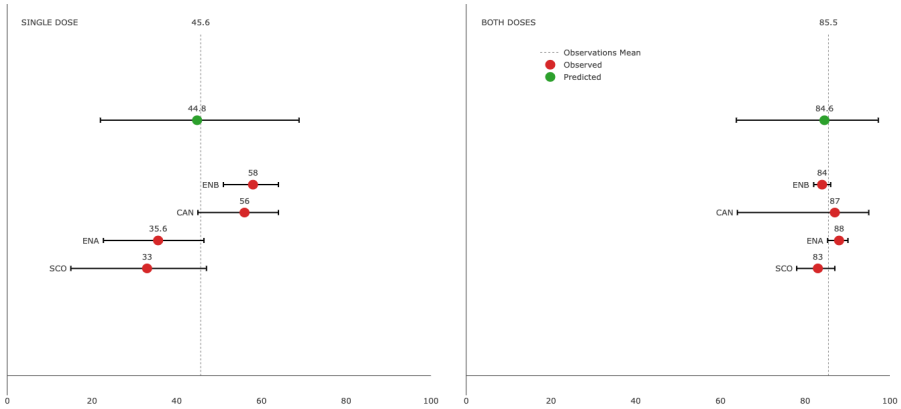
It is made available under a CC-BY 4.0 International license .



(a) D614G



(b) Alpha



(c) Delta

Figure 5: Predicted effectiveness against a variant. Estimates from effectiveness studies versus predictions from a model fitted without the indicated variant. 95% prediction bands and means of observations are shown. Study countries are indicated, with ENA and ENB standing for England study A and B, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

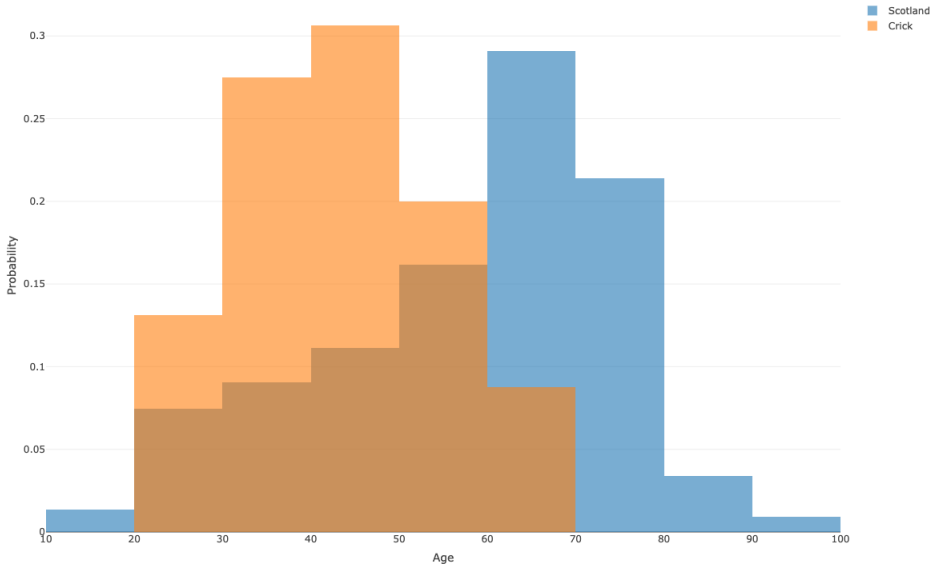


Figure 6: Age distribution for Comirnaty-vaccinated subjects in the Scotland study (Sheikh et al., 2021) versus that in the two-dose cohort in the Crick study (Wall et al., 2021b)

could be impacted by violations of this assumption due to population differences.

As an example we suppose that the "true population" in all the effectiveness studies was the Comirnaty-vaccinated population in the Scotland study (Sheikh et al., 2021). A difficulty for our analysis is that neutralisation data from this population were absent. Fortunately, the Crick study looked into individual covariates for neutralisation titres, including Sex, BMI and Age; see Wall et al. (2021b). Of these, only Age was highly statistically significant, whereas the others were not statistically significant. To represent the true population, we reweighted the titres from the Crick study according to the age distribution, shown in Figure S2 of Sheikh et al. (2021), of Comirnaty vaccinees in Scotland. The corresponding probabilities were applied to both dosing cohorts in the Crick study. According to Figure 6, the Scotland study had a noticeable skew towards older age groups.

The model fits of weighted and unweighted (original) Crick titres to effectiveness observations are illustrated in Figure 7. The weighted fit is shifted to the left, since lower titres, associated with older participants, have higher weights than in the original Crick data. If the effectiveness observations were generated by the Scotland population, but the Crick study population were used in modelling, we would infer that higher than actual titres were required for a given level of vaccine effectiveness and potentially underpredict it.

While both weighted and unweighted model fits in Figure 7 appear reasonable, their predictions could differ substantially. Unless potential biases are heeded, a "good" model fit could be misleading.

Some adjustments that could reduce biases due to population differences are noted in Section C. Section C.1 considers a special case when effectiveness predictions per dose/variant from a neutralisation study would equal corresponding predictions based on the true population. In this case, the true and "neutralisation" distributions need not be equal but only proportional to each other (as defined in Section C.1), which could simplify population matching.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

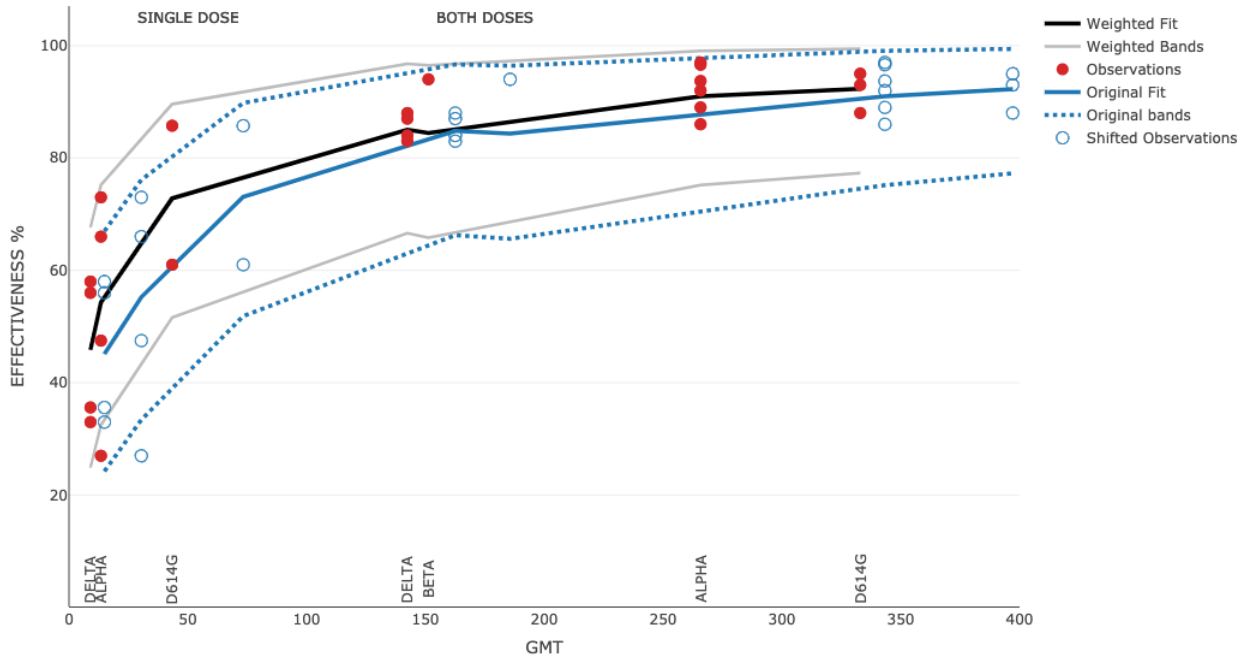


Figure 7: Model fits to age-weighted versus to original Crick neutralisation data. The observations are plotted versus the GMT per variant, for the corresponding titres distribution used in each fitting.

2.4 A sensitivity investigation

We now consider simulation results for a case where such proportionality does not hold. A detailed description of this simulation is in Section C.2. In a nutshell, we first take the age-weighted titre distributions, per variant, from the previous section as a baseline for true distributions. We next simulate five scenarios of how these baseline distributions could shift:

- DOWN: all distributions decrease (i.e., shift to the left)
- Down/up: distributions are more likely to decrease than to increase (i.e., shift to the right)
- Down/Up: distributions equally likely to decrease and increase
- Up/down: distributions are more likely to increase
- UP: all distributions increase

All the titres per variant/dose are shifted by the same multiple, with these multiples different across variant/dose combinations.

We simulated 1000 sets of "true" titre distributions per scenario. Each set was fitted to the effectiveness observations on all the variants. The resulting fit was used to evaluate the expected "true" effectiveness against Delta after one and two vaccine doses.

These true values were compared with the corresponding predictions, against Delta, based on the fit of the original (unweighted) neutralisation titres from the Crick study. These predictions are the same as in Figures 4 and 5. As in Section 2.2, model fitting was done without effectiveness observations on Delta.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

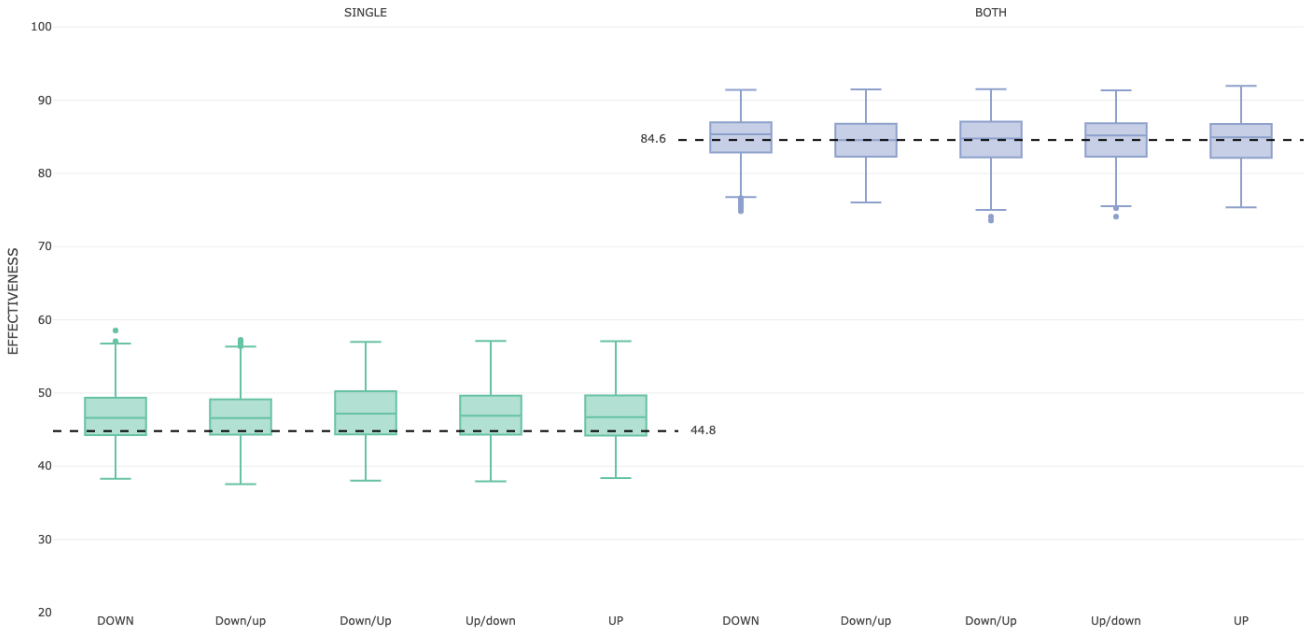


Figure 8: Robustness of effectiveness predictions against Delta: Crick study population ve Shifted Age-Weighted "true" distributions per dose. Box plots show effectiveness values evaluated for the Delta variant at a model fitting all effectiveness observations to the simulated true titres. The dashed line shows predictions from the original Crick study neutralisation data fitted to the effectiveness observations without Delta.

Figure 8 compares the true effectiveness values simulated under the assumed scenarios to the predictions from unweighted Crick distributions. After one dose, the medians of true values per scenario were in the range of $[46.6, 47.2]$, compared to 44.8 for the unweighted Crick data. After both doses, they were in $[84.6, 85.3]$ compared to 84.6. The predictions thus had a small bias after two doses and a slightly larger bias after one dose.

When we used the age-weighted Crick distributions instead of the original distribution, the predictions were 46.5 and 84.9, for one and two doses, respectively. Both values are closer to the corresponding medians of simulated predictions, which illustrates a benefit from population adjustments.

We also note a relatively small variability in true predictions, despite allowing large shifts (decreases or increases) per titre distribution for a variant. (The uniform distributions for these shifts are defined in Section C.2.) If we examine the interquartile ranges in Figure 8, we notice that the smallest lower quartile over the scenarios is 82.1%, and the largest, 87.1%. After one dose, these values are 44.2% and 50.3%.

2.5 Cautious prediction

Suppose that the Crick study population — possibly after an adjustment — has similar or slightly higher neutralisation titres than a true population does. There could still be uncertainty about how closely the populations match and how well the true titres are approximated. There are also modelling and other uncertainties.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

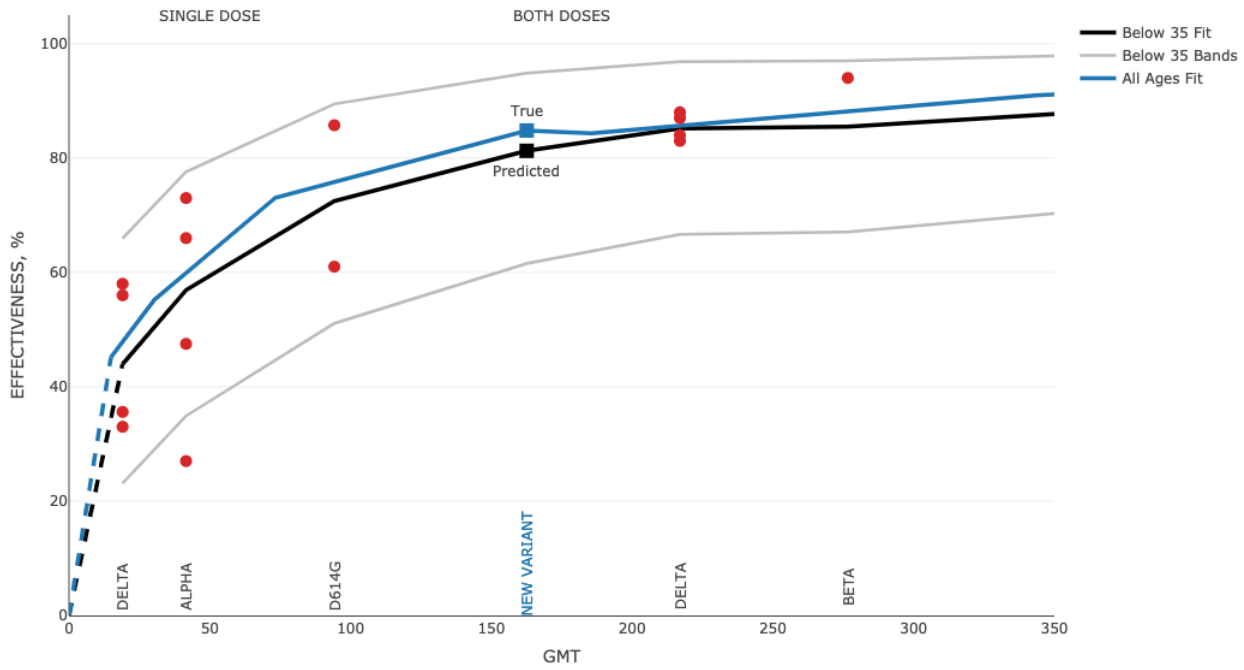


Figure 9: Cautious Prediction. Observations were fit to the below 35 age group; the all-ages fit is shown for comparison. Effectiveness for a new variant was predicted for all-age titres but assuming the below-35 fit.

One way to mitigate these uncertainties is to include a safety margin for effectiveness predictions, as follows. First we select a sub-population of the Crick participants with higher, on average, neutralisation titres. Here we simply choose all the participants aged under 35. We next fit their neutralisation titres to the effectiveness observations; see Figure 9. The resulting fit — based on higher titres — is to the right of the all-age fit.

To predict effectiveness against a new variant, for example, we would collect neutralisation titres against it in the all-age population. To make a cautious prediction we input these titres into the under-35 model fit.

In a numerical example, illustrated in Figure 9, the cautious prediction, 81.3%, is 3.5% less than 84.8%, the all-age fit. By using the cautious prediction we can infer that the all-age effectiveness would be at least 81.3% and likely higher. By extension, we infer that the effectiveness in the true population would be approximately equal to and possibly higher than 81.3%.

This cautious approach could be practically beneficial. It requires just enough information to assume that a true population has similar or lower titres than a neutralisation study population does. The true distribution is not required, and the impact of population mismatch can be reduced. If a cautious prediction, say, of 85% effectiveness after a half-booster, is still high enough in practice, then a sound decision could be made.

We note that the unweighted fit in Figure 7 could provide enable cautious predictions for the assumed true population. The same conclusion can be made from Figure 8, especially for the first dose. We could therefore use Crick study population as is or apply the approach in this section for an extra margin of safety.

The cautious approach could be also useful when combining effectiveness data from disparate populations in different effectiveness studies. Excessive caution, however, is undesirable, as it may, for instance,

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

lead to boosting too early.

3 Discussion

Our modelling established a strong nonlinear association between the neutralisation titre distributions and the effectiveness observations assumed. It also produced accurate predictions of vaccine effectiveness against the three variants after one and after two vaccine doses. This is despite potentially large differences between the Crick study population and the populations in the effectiveness studies, as well as the likely disparities between effectiveness definitions in different studies, and a generally high uncertainty after one dose. Although we did not examine causality, the high accuracy of predictions gave an indirect support for the use of neutralising titres as correlates of protection.

We pinpointed potential biases due to population differences between neutralisation and effectiveness studies. Effectiveness predictions could benefit from population adjustments, such as those overviewed in Section C, as well as from cautious prediction, considered in Section 2.5.

Future work It can be beneficial to incorporate further adjustments for populations into our framework. Our modelling could be additionally validated with detailed data from past and future effectiveness studies.

Modelling refinements could include incorporation of: titre measurement errors; additional covariates; between-study variability; imputation and other methods addressing censoring.

A valuable extension of our framework would be to combine data across neutralisation studies. Another, relatively straightforward extension, is to accommodate various types of vaccine effectiveness and efficacy; and also effectiveness against a mix of SARS-CoV-2 variants. A key extension, which we are currently working on, is to effectiveness predictions across Covid-19 vaccines.

3.1 Policy implications

Monitoring neutralisation titres

Our modelling supports conclusions in Wall et al. (2021a) about the benefits of monitoring neutralisation titres in various populations. For instance, the Crick study already enabled accurate predictions of vaccine effectiveness for the three SARS-CoV-2 variants (in Section 2.2). It could be representative of larger populations of healthcare workers and of general working-age populations. Any follow-up data from the study, for instance, on waning immunity or on response to a third booster could be highly informative and valuable for policy decisions.

Expanding the Crick study or running additional studies of SARS-CoV-2 variants neutralisation could bring further benefits, including early warnings of waning immunity in a subpopulation and early identification of immunity-evading variants.

Human challenge studies It could be possible to validate correlates of protection models in human challenge studies with a vaccine. To emulate lower antibody levels in older populations (and to provide information on dose reduction potential) a lower vaccine dose could be administered to the study participants (who

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

are usually aged between 18 and 30). As effectiveness could be modelled across variants, safer challenges with more benign variants, such as D614G, may be sufficient.

Effectiveness regions

Our modelling suggests the following regions of vaccine effectiveness, which could be relevant to decision-making:

Flat region includes effectiveness against every variant after two doses in the all-age population (e.g., Figure 1). This region allows a maximum flexibility for policy-making, since moderate changes in neutralising titres will have only a small impact on effectiveness. For example, Payne et al. (2021) reported greater neutralisation for the Comirnaty vaccine with an eight-week dosing interval than with three-to-five week intervals used in the clinical trials. However, a longer interval could expose partially vaccinated persons to higher infection risk, since single-dose protection is lower. Within a flat region, either longer or shorter intervals could be chosen on epidemiological grounds, without risking vaccine effectiveness.

There is also little effectiveness cost from reduced vaccine doses. Policy-makers could confidently decide on dose reduction so as to address supply shortages, high costs and dose-dependent toxicities. This could also free up doses for vulnerable populations worldwide.

Over time, waning vaccine protection could drive neutralisation titres– particularly in the elderly and vulnerable — outside the flat region, especially for the Delta variant. Still, a flat region could be reached again after another booster (i.e., a third dose). Boosting with reduced doses would have little impact on effectiveness, compared to full-dose boosters. Furthermore, a Comirnaty vaccine update for the Delta variant or a new universal vaccine could probably be administered in smaller than the current doses.

Steep region corresponds, for example, to effectiveness against Alpha or Delta after a single dose. This region gives less leeway for vaccination policies: for instance, reducing the dose may cause a relatively steep drop in effectiveness. First-dose reduction could still be beneficial, if implemented with due caution. For instance, if the pandemic is under control by non-pharmaceutical measures, reduced first doses could accelerate reaching full vaccination and protection of a country's population.

Intermediate region is located between the other two regions. It is exemplified by the single-dose effectiveness against D614G for all ages (Figure 1) and possibly by the two-dose effectiveness against Delta in the elderly (Figure 3). Its distinguishing feature is an increased uncertainty about effectiveness, as reflected by a beta distribution. Policies for an intermediate region could combine those for the other two, while accounting for the extra uncertainty.

These regions and their interpretation would be different for vaccine effectiveness against hospitalisation and death.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

We note that predictions within a flat region are less sensitive to modelling assumptions and population differences. Conversely, predictions within a steep region could be more sensitive due to the steepness of the model function and proximity to the lower detection limit.

Age differences We found considerable differences in virus neutralisation and vaccine effectiveness between age groups, which could warrant age-tailored vaccination policies. For example, those under 35 had almost twice as high GMT against the Delta variant as those over 60 had, and were very much inside the flat region; see Figure 3.

With the Delta variant, the twice-vaccinated 60+ age group appears to be within a region of reduced effectiveness and increased uncertainty. The situation could be even worse for wider elderly populations, since the Crick study participants were all under 70 and healthy enough to be in employment. Unsurprisingly, older groups may need a booster sooner, and perhaps more frequently than the younger group would.

Dosing strategies Our modelling suggests a compromise dose-sparing strategy to provide a full first dose but reduce the initial booster (second dose) in younger age groups, and also the second booster (third dose) in all age groups. Reduced doses in younger groups may be justified — even despite Delta — not only to expand vaccine coverage but also to reduce toxicity and vaccine hesitancy, which are prevalent in these groups. This could particularly make sense for those under 18.

Recent effectiveness estimates Our framework could give insights into various aspects of pandemic policy. One example is a surprisingly low estimate, of 64%, for symptomatic effectiveness of the Comirnaty vaccine against Delta (Israel Ministry of Health, 2021; Zimmer, 2021).

Our modelling implies that this estimate might not be completely surprising for older populations (see Figure 3), given that vaccine effectiveness wanes over time. Another point is that as an intermediate effectiveness region is approached, upcoming effectiveness estimates can become increasingly uncertain. Hence, vaccine boosting decisions may be suboptimal unless based on several effectiveness studies.

Addressing current and future variants The Delta variant already pushes us towards diminished vaccine effectiveness, particularly for the older and vulnerable people. (This is also evident from Khoury et al. (2021) and Wall et al. (2021b).) A new variant reducing neutralising titres even further would push us more into the danger zone. Therefore measures such as expanding worldwide vaccine coverage to reduce mutations and variants emergence, as well as possible updating of existing vaccines are needed.

Virus neutralisation studies for emerging variants could be done earlier. This would predict effectiveness decreases sooner and enable better mobilisation of resources. Likewise, neutralisation titres could be monitored and vaccine effectiveness predicted for successors of existing variants of concern, such as the Delta Plus.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Competing interests

The authors declare no competing interests.

Funding

This work received no external funding.

Data availability

All data and R code used for this paper will be freely available upon publication from the paper's GitHub repository

Acknowledgements

O.V. is grateful to Rosemary Bailey, Ian Bartlett and Andrew Langley for their support and encouragement during this research. The authors thank Rebecca Lodwick and Barbara Bogacka for their suggestions on earlier versions of the paper which lead to its improvements. We also thank David L. Bauer for making available additional data from the Crick Legacy Study.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

A Methods

A.1 Neutralisation Titres

The explanatory variable in our modelling was the antibody titres attaining 50% neutralisation of SARS-CoV-2 variants after one or two doses of the Comirnaty vaccine. The titres data were obtained at the Crick Institute within the ongoing Legacy study Wall et al. (2021b). These data are available at Wall et al. (2021b) GitHub repository. The study collected plasma from 149 healthcare workers in the one-dose cohort and 159 healthcare workers in the two-dose cohort. (Overall, there were 250 *unique* participants.) The samples were tested in high-throughput automated assays of live-virus neutralisation against the D614G, Alpha, Beta and Delta variants, and against the virus Wild Type.

The resulting distributions of log10-transformed titres are shown in Figure 10. After the first dose, these distributions violate the Normality assumption. As their shapes differ between the variants, an exponential or another standard distribution may not be appropriate to model all of them. After the second dose, these distributions appear more Normal.

According to Wall et al. (2021b), any value above the assay quantitative limit of detection, equal to 2560, was recorded as 5120. Any neutralisation signal below the quantitative limit of 40 was recorded as 10. If no neutralisation signal was observed, a titre of 5 was recorded. These thresholds are illustrated in Figure 10.

A.2 Protection models

The discussion in this section draws on Dunning (2006); see Qin et al. (2007) for a related approach.

Individual protection function The probability of an individual with a particular titre being protected against symptomatic Covid-19 was assumed to follow

$$p(t) = \frac{\exp(\alpha + \beta t)}{1 + \exp(\alpha + \beta t)}, \tag{1}$$

which is a scaled logistic function. The *individual protection* function has the log-titre t as the covariate and the constants, α and β, as the model parameters to be estimated. (As is common in vaccine research, we used $\log_{10}$-transformed titres. The same function was used in Dunning (2006) and Nauta et al. (2009), and, under a different parametrisation, in Khoury et al. (2021).) We assume this function with the same parameter values for every individual and across all the variants.

Population protection function The proportion protected in a population with N individuals is defined as

$$P(t) = \frac{1}{N} \sum_{i=1}^{N} \frac{\exp(\alpha + \beta t_i)}{1 + \exp(\alpha + \beta t_i)}, \tag{2}$$

where t_i is the log-titre obtained in subject i against a particular variant. This *population protection* function is the expectation over the individual protection functions with the same parameters α and β. Unlike

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

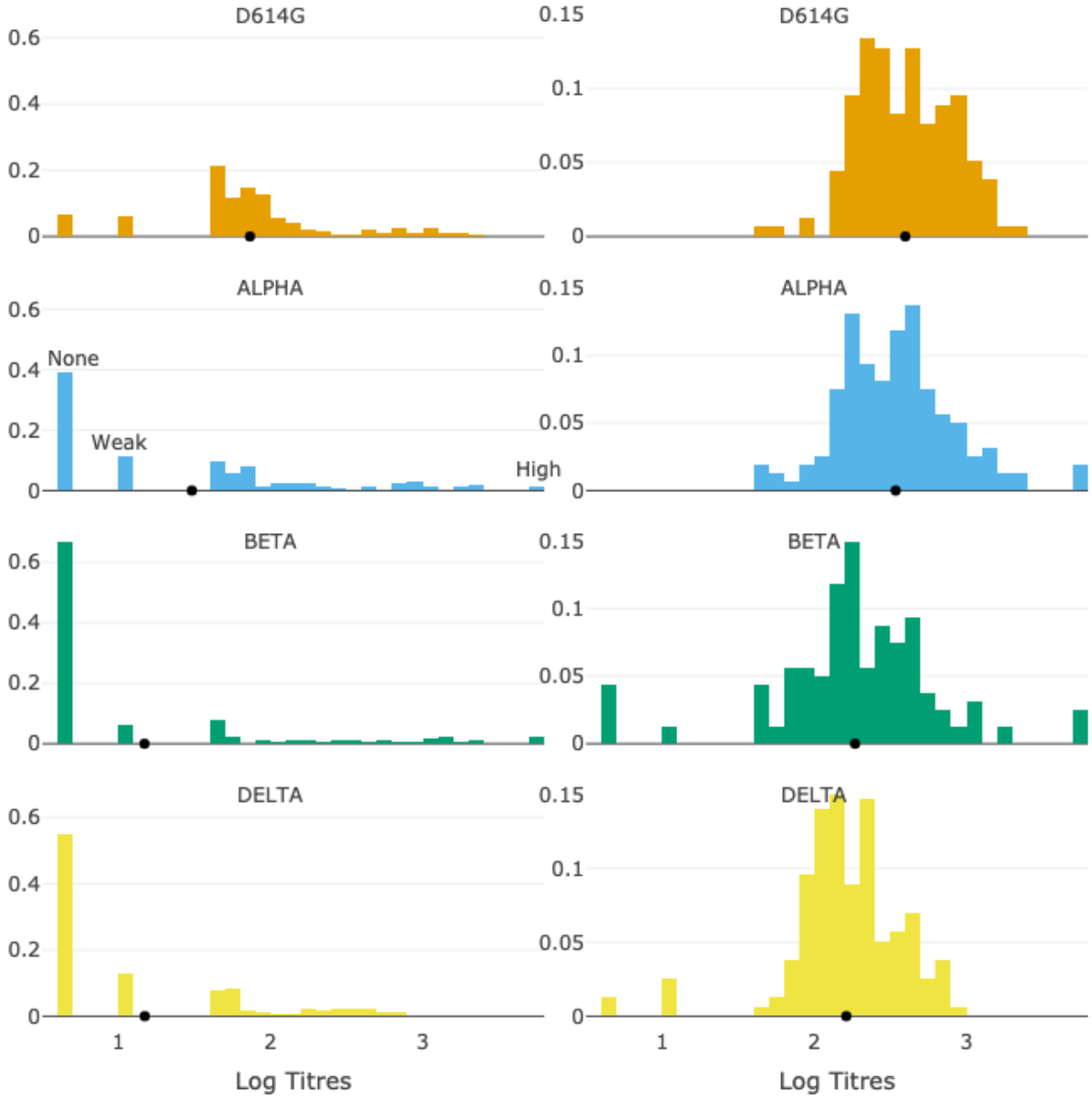


Figure 10: Distributions of log10 titres after one (Left) and two (Right) vaccine doses. Detection thresholds are illustrated for Alpha after one dose: "None", recorded as 5, means no signal; "Weak", recorded as 10, means a signal below 40; "High", at 5120, means a signal above 2560. Sources: Wall et al. (2021b).

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Khoury et al. (2021), who assumed normal distributions of log-titres, we evaluated the expectation using the empirical distributions from the Crick study, due to normality violations and non-standard censoring.

Relation to vaccine effectiveness Following Dunning (2006), we write vaccine efficacy as

$$VE = 1 - \frac{1/v \sum_i^v (1 + \exp(\alpha + \beta t_i))^{-1}}{1/u \sum_j^u (1 + \exp(\alpha + \beta t_j))^{-1}}; \tag{3}$$

for v vaccinated and u unvaccinated subjects in a vaccine effectiveness study.

We assume that the study subjects have negligibly small levels of neutralising antibodies and/or the number of seropositive subjects is negligibly small. Then, unvaccinated subjects' neutralisation titres would approach zero, and so each log-titre t_j would approach $-\infty$. Hence, each term $\exp(\alpha + \beta t_j)$ in the denominator of (3) would tend to zero, so that the denominator would approach unity. Therefore, we can write

$$VE = 1 - \frac{1}{v} \sum_i^v \frac{1}{1 + \exp(\alpha + \beta t_i)} = \frac{1}{v} \sum_i^v \left(1 - \frac{1}{1 + \exp(\alpha + \beta t_i)}\right) = \frac{1}{v} \sum_{i=1}^v \frac{\exp(\alpha + \beta t_i)}{1 + \exp(\alpha + \beta t_i)}$$

That is, assuming infection-naive individuals, vaccine effectiveness equals the population protection function.

The assumption is unlikely to hold in general for vaccine studies. However, we note that prior SARS-CoV-2 infection cases can cause large downward biases in effectiveness estimates, compared to effectiveness in SARS-CoV-2 naive populations (World Health Organisation, 2021, p. 27). Consequently, investigators in Covid-19 vaccine effectiveness studies take great care to remove these biases by identifying individuals with prior SARS-CoV-2 infections and excluding them from effectiveness analyses.

Accordingly, our modelling assumes (2) as the model function for (symptomatic) vaccine effectiveness.

A.3 Statistical modelling

Modelling the outcome Most of the observational studies that we assumed estimated vaccine effectiveness by multi-covariate logistic regression. The estimates so obtained were often far from crude estimates that could be computed using vaccinated and unvaccinated case ratios given in the papers reporting these studies. Rather than use case counts within a binomial infected/non-infected framework, as in Khoury et al. (2021), we fitted our model to the effectiveness estimates directly.

Since these estimates were percentages, we treated them similarly to probabilities. We therefore used a beta distribution, which is standard for modelling such outcomes. For simplicity, we did not include the endpoints 0% and 100%, which can be accommodated by an extended beta distribution.

Beta regression The model function (2) was fit using nonlinear beta regression (Ferrari and Cribari-Neto, 2004). The corresponding log-likelihood function was maximised numerically using the *optim* procedure in R.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

To obtain a starting guess for this maximisation we fit the individual protection function (1) to the geometric mean titres (GMTs) per variant/dose combination. Since a logit transformation of (1) is $\alpha + \beta t$, we applied a generalised linear model (GLM) framework of beta regression. Specifically, we used the *betareg* package (Cribari-Neto and Zeileis, 2010).

To fit either function, we used the beta distribution parametrisation proposed in Ferrari and Cribari-Neto (2004); Cribari-Neto and Zeileis (2010). The 95% prediction interval in either case was defined with 2.5% and 97.5% quantiles of the beta distribution.

To the best of our knowledge, beta regression — linear or nonlinear — has not yet been applied to model correlates of vaccine protection.

The data and R code used for this paper are freely available in a GitHub repository.

B An approximation to the effectiveness model

Complexities Our assumed model function (2) computes the proportion of protected individuals in a vaccinated population, as common in vaccine research (Dunning, 2006; Nauta et al., 2009; Khoury et al., 2021). However, that this function — defined over distributions — is multivariate complicates regression visualisation and decision-making. Although we can plot predicted effectiveness against the geometric mean titre (GMT), as done in this paper, interpretation of the plots is not straightforward, since the underlying distribution of titres are hidden. Technically, the protection function is not a function of GMT: for instance, it could map the same GMT to multiple effectiveness values. Since we used empirical distributions of log-titres, rather than, say, a Normal distribution with a constant variance, as in Khoury et al. (2021), our problem has a high dimensionality. Another issue is that a distribution's shape could vary substantially between populations, which poses concerns about generalisability of our conclusions. A distribution's (geometric) mean, on the other hand, is usually more stable.

A simplified model for effectiveness To overcome these difficulties we define a simplified model as follows. We assume that all the titres in a sample for a dose/variant combination are equal to the GMT of the corresponding titres sample from the Crick study. Then, the population protection function (2) coincides with the individual protection function (1). The simplified model assumes the latter for vaccine effectiveness. (This appendix refers to the model assumed elsewhere in the paper as the full model.)

A logit transformation of (1) is $\alpha + \beta t$. Consequently, we can apply a framework akin to generalised linear models (GLM) and use a standard package (*betareg*) package of beta regression.

A simplified model fit to our assumed 24 symptomatic effectiveness observations is illustrated in Figure 11.

Figure 12 shows that the two model fits and their prediction bands are close. This is particularly true after two doses for all the variants and after one dose for Alpha and D614G. We could interpret the full model in Figure 1 as approximately a simplified model plotted as a function of GMT. Furthermore, our inferences could be robust to the shape of log-titre distributions per variant, provided their GMTs are stable.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

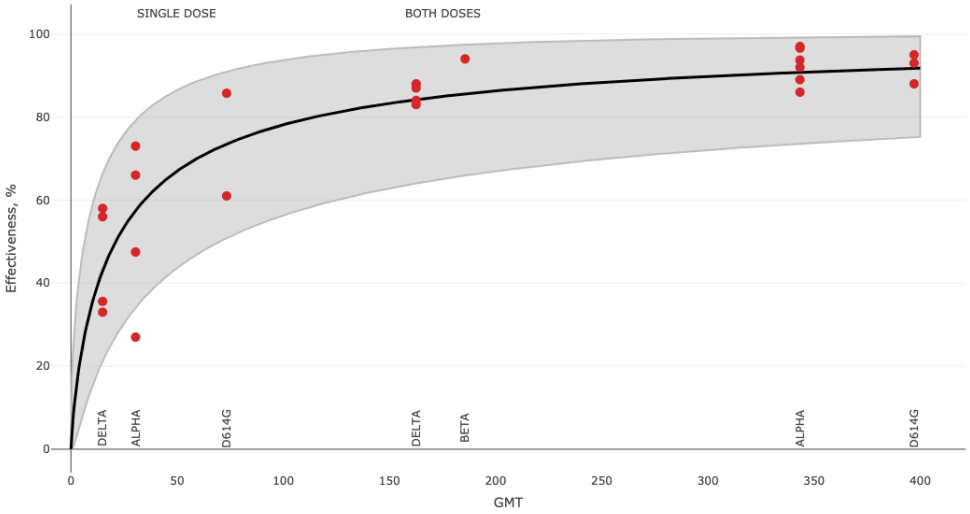


Figure 11: Simplified model fit to effectiveness observations, with 95% prediction bands shown.

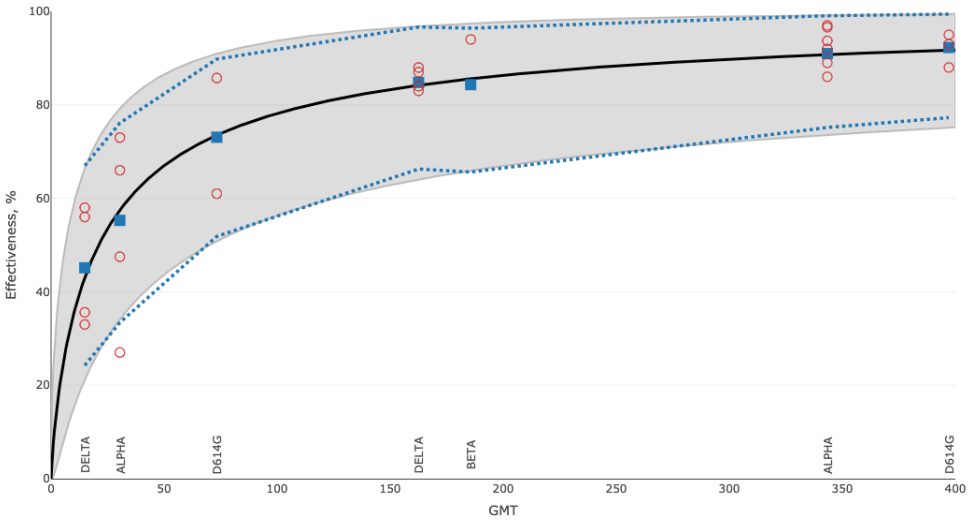


Figure 12: Simplified model fit and prediction bands are shown in a black line and grey shading. The full model fit and prediction bands are shown in blue squares and dashed lines.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

C Statistical adjustments of populations

This section overviews some adjustments that could overcome biases due to population differences.

Iterative proportional fitting (aka data raking) Neutralisation titres are iteratively weighted according to covariates so as to reflect participants in effectiveness studies more closely. The age-based reweighing above is in the spirit of this approach.

Propensity scores weighting Weights can be defined to account for selection biases in relevant populations.

Subjects matching Neutralisation study participants are matched by covariates to their closest vaccinated subjects in an effectiveness study. Each effectiveness study would then provide a matched subset of participants. Effectiveness values would be re-estimated for these subsets and fit to the neutralisation titres per variant.

This matching could also proceed in the opposite direction. Suppose that enough vaccinated subjects are selected for a neutralisation study to cover relevant covariates in wider vaccinated populations. Each effectiveness study population could then be matched to a subset of the neutralisation study population. The effectiveness values would be fit to neutralisation data from these matched subsets.

Other approaches Various other modelling approaches, such as sensitivity analyses (as illustrated below) and multimodel inference, could be also beneficial.

These adjustments rely on detailed information about both neutralisation and effectiveness studies populations. Such information is publicly available on (GitHub) for the Crick study but unfortunately not for effectiveness studies. Furthermore, some of the above adjustments may require additional, specially designed trials. Therefore the potential adjustments were not investigated for this current paper, but hopefully could be investigated in the future.

C.1 Equal multiples of titre distributions

We now consider a condition when effectiveness predictions per dose/variant from a neutralisation study would equal corresponding predictions based on the true population. The condition is that the family of true titre distributions per dose/variant is "proportional" to the family of corresponding titre distributions from the neutralisation study. In other words, each true distribution per dose/variant is obtained by multiplying the corresponding neutralisation study distribution by some constant $\kappa > 0$; see Figure 13. This constant must be the same for every dose/variant combination, but need not be known. This condition is unlikely to hold at the outset, but could hold approximately after adjustment of the populations (e.g., as described in Section C). It also implies that we need not strive to match titre distributions perfectly, but only to make them approximately proportional.

For illustration, suppose our goal is to ensure that a neutralisation study produced distributions representative of N typical individuals from a true effectiveness population. For the neutralisation study we would select N subjects similar to those individuals and then perform various statistical adjustments. Even

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

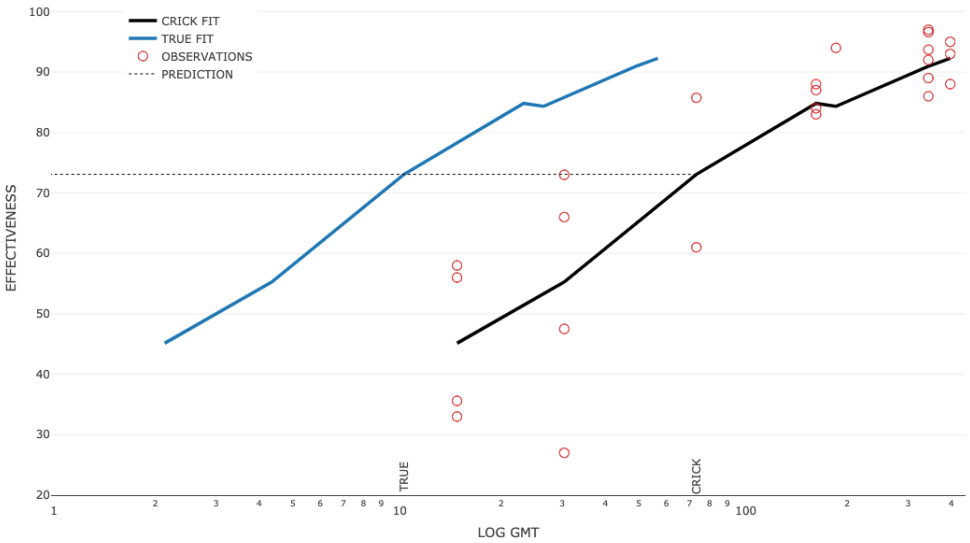


Figure 13: True distributions of titres are a constant multiple of Crick distributions

then, the true and neutralisation distributions may not be perfectly matched, for instance, due to unobserved confounders.

We may still in some cases obtain precise predictions about the true population. One such case is if each true distribution per dose/variant were a constant multiple, $\kappa > 0$, of the corresponding neutralisation distribution. Indeed, let $\log x_i$ and $\log \kappa x_i$ denote the log-titre for subject i in the neutralisation study and person i in the true population, respectively. The corresponding terms in model (2) are $\alpha + \beta \log x_i$ and $\alpha + \log \kappa + \beta \log x_i$, which differ by a constant independent from the titre.

Consequently, a model fitting neutralisation study titres to effectiveness observations would produce exactly the same predictions as a model fitting true titres to the same observations. This is illustrated in Figure 13, which has log-titres on the x-axis. Conveniently, the actual value of $\kappa > 0$ need not be known for such predictions.

C.2 Simulating distributions

This section provides a methodology background for the investigation in Section 2.4. Its purpose is to illustrate the impact of mismatched distributions of titres on effectiveness predictions. The assumed task is to predict effectiveness against the Delta variant, as in Section 2.2. The predictions are made using the titres distributions from the Crick study. To define a "true" distribution for dose/variant combination j we start with a corresponding age-weighted distribution, considered above. Then we multiply it by $\kappa_j > 0$, with these values possibly different between distributions. Consequently, both the distribution shapes and the multiples are different.

For each each dose/variant combination j, the fold κ_j was simulated under each of the following scenarios:

DOWN κ_j is a fold decrease, with values in $[0.25, 1]$

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Down/up κ_j is a fold decrease, from $[0.5, 1]$, with probability 0.67 and a fold increase, from $[1, 2]$ with probability 0.33.

Down/Up κ_j is a decrease from $[0.5, 1]$ and an increase from $[1, 2]$ with equal probabilities

Up/down κ_j is a decrease from $[0.5, 1]$ with probability 0.33 and an increase from $[1, 2]$ with probability 0.67

UP κ_j is a fold increase with values in $[1, 4]$

(In every case, the ratio of the endpoints equals four.)

We first drew the increase or decrease interval and then the actual value of κ_j from a uniform distribution over the interval. Our simulation ensured that the GMT after a first dose for a particular variant was less than that after two doses. For each scenario, we simulated 1000 vectors of κ_j, with 5000 vectors in total.

Hence, each vector of κ_j defines a set of true titre distributions. Each of these sets was used to fit our model to the effectiveness observations, including those for the Delta variant. The predicted effectiveness values, against the variant, after one and two doses were set as the true effectiveness values.

D Extended Data

We performed a literature search for studies of symptomatic effectiveness (and efficacy) with the Comirnaty vaccine. We sought studies either specifying the assumed SARS-CoV-2 variant or for which the predominant variant could be inferred from the study dates and location. We identified 24 effectiveness (and efficacy) observations in all-age groups, and four in age subgroups. The included studies and observations are listed in Table 1.

The first study is the Pfizer-BioNTech Phase 2/3 vaccine trial mostly conducted in the US (130 out of 152 study sites). The estimate of symptomatic efficacy after two vaccine doses was given in Polack et al. (2020) for the study period between May and November 2020. During that time, more than 90% of the SARS-CoV-2 infections in the US, and similarly world-wide, had the D614G mutation Korber et al. (2020). Hence, we assigned D614G to the efficacy estimates from this study.

Following Skowronski and Serres (2021), we also calculated the vaccine efficacy 14 days after the first dose, by using the revised efficacy data in the FDA submission (FDA, 2020). The estimate was made using the Bayesian approach used by Pfizer investigators to estimate efficacy after two doses (FDA, 2020).

We included seven observational studies as follows: one study in Canada (Nasreen et al., 2021), two in England (Bernal et al., 2021; Pouwels et al., 2021); one in France (Charmet et al., 2021), two in Israel (Mor et al., 2021; Haas et al., 2021) and one in Scotland (Sheikh et al., 2021). Note that (Charmet et al., 2021) reported combined effectiveness estimates for Pfizer-BioNTech and Moderna mRNA vaccines. We included this study as it stated that 87% of the participants had the former vaccine.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Study	Dose	Days	Variant	Country	Eff.	Confidence	Reference
1	1	14	D614G	US	85.7	(59.3, 97.5)	FDA (2020)
	2	7	D614G	US	95.0	(90.0, 97.9)	Polack et al. (2020)
2	1	14	D614G	Canada	61.0	(54.0, 68.0)	Nasreen et al. (2021)
	1	14	Alpha	Canada	66.0	(64.0, 68.0)	
	1	14	Delta	Canada	56.0	(45.0, 64.0)	
	2	7	D614G	Canada	93.0	(88.0, 96.0)	
	2	7	Alpha	Canada	89.0	(86.0, 91.0)	
	2	7	Delta	Canada	87.0	(64.0, 95.0)	
3	1	28	Alpha	Scotland	27.0	(13.0, 39.0)	Sheikh et al. (2021)
	1	28	Delta	Scotland	33.0	(15.0, 47.0)	
	2	14	Alpha	Scotland	92.0	(88.0, 94.0)	
	2	14	Delta	Scotland	83.0	(78.0, 87.0)	
4	1	21	Alpha	England A	47.5	(41.6, 52.8)	Bernal et al. (2021)
	1	21	Delta	England A	35.6	(22.7, 46.4)	
	2	14	Alpha	England A	93.7	(91.6, 95.3)	
	2	14	Delta	England A	88.0	(85.3, 90.1)	
5	2	7	D614G	France	88.0	(81.0, 92.0)	Charmet et al. (2021)
	2	7	Alpha	France	86.0	(81.0, 90.0)	
6	2	7	Alpha	Israel A	96.6	(95.8, 97.2)	Haas et al. (2021)
7	2	14	Beta	Israel B	94.0	(90.0, 98.0)	Mor et al. (2021)
8	1	21	Alpha	England B	73.0	(68.0, 76.0)	Pouwels et al. (2021)
	2	14	Alpha	England B	97.0	(96.0, 98.0)	
	1	21	Delta	England B	58.0	(51.0, 64.0)	
	2	14	Delta	England B	84.0	(82.0, 86.0)	

Table 1: Studies that provided symptomatic effectiveness estimates of the Comirnaty vaccine in all-age populations. Eff. stands for efficacy for Study 1, and effectiveness for the other studies.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

References

Bernal, J. L., N. Andrews, C. Gower, E. Gallagher, et al. (2021, August). Effectiveness of covid-19 vaccines against the b.1.617.2 (delta) variant. *New England Journal of Medicine 385*(7), 585–594.

Charmet, T., L. Schaeffer, R. Grant, S. Galmiche, et al. (2021, September). Impact of original, b.1.1.7, and b.1.351/p.1 SARS-CoV-2 lineages on vaccine effectiveness of two doses of COVID-19 mRNA vaccines: Results from a nationwide case-control study in france. *The Lancet Regional Health - Europe 8*, 100171.

Cribari-Neto, F. and A. Zeileis (2010). Beta regression in R. *Journal of Statistical Software 34*(2).

Cromer, D., M. Steain, A. Reynaldi, T. E. Schlub, et al. (2021, August). SARS-CoV-2 variants: levels of neutralisation required for protective immunity.

Dunning, A. J. (2006). A model for immunological correlates of protection. *Statistics in Medicine 25*(9), 1485–1497.

Earle, K. A., D. M. Ambrosino, A. Fiore-Gartland, D. Goldblatt, et al. (2021, May). Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine*.

FDA (2020). Pfizer-biontech covid-19 vaccine (bnt162, pf-07302048) vaccines and related biological products advisory committee briefing document. `https://www.fda.gov/media/144246/download`.

Ferrari, S. and F. Cribari-Neto (2004, August). Beta regression for modelling rates and proportions. *Journal of Applied Statistics 31*(7), 799–815.

Gilbert, P. B., D. C. Montefiori, A. McDermott, Y. Fong, et al. (2021, August). Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy trial.

Haas, E. J., F. J. Angulo, J. M. McLaughlin, E. Anis, et al. (2021, May). Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in israel: an observational study using national surveillance data. *The Lancet 397*(10287), 1819–1829.

Israel Ministry of Health (July 05, 2021). Decline in vaccine effectiveness against infection and symptomatic illness. available at: `https://www.gov.il/en/departments/news/05072021-03`. accessed on August 16, 2021.

Jackson, L. A., E. J. Anderson, N. G. Rouphael, P. C. Roberts, et al. (2020, November). An mRNA vaccine against SARS-CoV-2 — preliminary report. *New England Journal of Medicine 383*(20), 1920–1931.

Khoury, D. S., D. Cromer, A. Reynaldi, T. E. Schlub, et al. (2021, May). Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nature Medicine*.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Korber, B., W. M. Fischer, S. Gnanakaran, H. Yoon, et al. (2020, August). Tracking changes in SARS-CoV-2 spike: Evidence that d614g increases infectivity of the COVID-19 virus. *Cell 182*(4), 812–827.e19.

Li, C., D. Yu, X. Wu, H. Liang, et al. (2021, April). Twelve-month specific IgG response to SARS-CoV-2 receptor-binding domain among COVID-19 convalescent plasma donors in wuhan.

Lu, Y., J. Wang, Q. Li, H. Hu, J. Lu, and Z. Chen (2021, June). Advances in neutralization assays for SARS-CoV-2. *Scandinavian Journal of Immunology 94*(3).

Mor, O., N. S. Zuckerman, I. Hazan, R. Fluss, et al. (2021, July). BNT162b2 vaccination efficacy is marginally affected by the SARS-CoV-2 b.1.351 variant in fully vaccinated individuals.

Nasreen, S., H. Chung, S. He, K. A. Brown, et al. (2021, July). Effectiveness of COVID-19 vaccines against variants of concern in ontario, canada.

Nauta, J. J., W. E. Beyer, and A. D. Osterhaus (2009, August). On the relationship between mean antibody level, seroprotection and clinical protection from influenza. *Biologicals 37*(4), 216–221.

Payne, R. P., S. Longet, J. A. Austin, D. Skelly, et al. (2021). Sustained t cell immunity, protection and boosting using extended dosing intervals of BNT162b2 mRNA vaccine. *SSRN Electronic Journal*.

Polack, F., S. Thomas, N. Kitchin, et al. (2020). Safety and efficacy of the bnt162b2 mrna covid-19 vaccine. *New England Journal of Medicine 383*, 2603–2615.

Pouwels, K. B., E. Pritchard, P. C. Matthews, N. Stoesser, D. W. Eyre, et al. (2021, August). Impact of delta on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK.

Qin, L., P. B. Gilbert, L. Corey, M. J. McElrath, and S. G. Self (2007, November). A framework for assessing immunological correlates of protection in vaccine trials. *The Journal of Infectious Diseases 196*(9), 1304–1312.

Sheikh, A., J. McMenamin, B. Taylor, and C. Robertson (2021, June). SARS-CoV-2 delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *The Lancet 397*(10293), 2461–2462.

Skowronski, D. and G. D. Serres (February 2021). Correspondence to "safety and efficacy of the bnt162b2 mrna covid-19 vaccine".

Wall, E. C., M. Wu, R. Harvey, G. Kelly, et al. (2021a, July). AZD1222-induced neutralising antibody activity against SARS-CoV-2 delta VOC. *The Lancet 398*(10296), 207–209.

Wall, E. C., M. Wu, R. Harvey, G. Kelly, et al. (2021b, June). Neutralising antibody activity against SARS-CoV-2 VOCs b.1.617.2 and b.1.351 by BNT162b2 vaccination. *The Lancet 397*(10292), 2331–2333.

World Health Organisation (2021). Evaluation of covid-19 vaccine effectiveness. interim guidance. available at: `https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccine_effectiveness-measurement-2021.1`. accessed on July 23, 2021.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.06.21263160; this version posted September 24, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Zimmer, C. (2021). The world is worried about the delta virus variant. studies show vaccines are effective against it. *The New York Times*. Available at: https://www.nytimes.com/2021/07/06/science/Israel-Pfizer-covid-vaccine.htm.