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 identifies 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. It was validated by predicting effectiveness against the B.1.167.2 (Delta) variant. The predictions, of 51.7% (34%, 69%) after one and of 88.6% (76%, 97%) after two vaccine doses, were close to the corresponding means, 49% and 85.4%, of observations in real-life effectiveness studies. We illustrate its applications to inform decisions, such as about the 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

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. The life-critical questions — What are the optimal doses and timings for vaccine boosters? How likely is a new variant to evade vaccine immunity? — would be answered a magnitude quicker. Unlike with influenza, no practically accepted correlates of protection frameworks yet exist for Covid-19.

Their research has been particularly active, with groundbreaking modelling frameworks proposed in Earle et al. (2021) and Khoury et al. (2021). These frameworks predict Covid vaccine efficacy using live-virus neutralisation titres, quantified in plasma collected from vaccinated volunteers in Phase 1/2 vaccine trials. A key strength of both frameworks lies in combining data across trials and vaccines. To overcome assays disparity between trials, the papers normalised vaccine-induced titres by mean titres from convalescent plasma used as a reference in a given trial.

There could, however, be challenges in applying these frameworks in some cases. For instance, what also matters is real-life effectiveness of vaccines, whereas the frameworks focus on vaccine efficacy from clinical trials. Both papers assumed a full-course vaccination (e.g. two doses of the approved mRNA vaccines), whereas predictions after a single dose are also important, given 5-12 weeks dosing intervals in practice. Predictions for emerging SARS-CoV-2 variants may be increasingly uncertain, since the Phase 1/2 neutralisation studies, used by the frameworks, were done with the virus Wild Type, whereas Phase 3 efficacy data were collected already with its successors. For instance, the predominant genotype during the Phase 3 Pfizer-BioNTech (Comirnaty) vaccine trials was D614G (Korber et al., 2020), which had a 2.3 fold-decrease in neutralisation versus the Wild Type (Wall et al., 2021). Predictions for the Delta and successor variants would be a further step away from the Wild Type.

Perhaps the key issue is how well early-stage neutralisation studies assumed within the frameworks could represent wider populations in Phase 3 trials and beyond. Indeed, those studies were often small: for example, 15 vaccinated and three convalescent subjects in the Moderna study (Jackson et al., 2020) used by Khoury et al. (2021). Different neutralisation studies with the same vaccine could also have disparate results even after normalisation. Although a correlates of protection model could fit neutralisation inputs well, if these are not representative, then so could be the model conclusions.

To address these challenges we propose a novel framework of immune correlates, which was built upon Earle et al. (2021) and particularly Khoury et al. (2021). Although the framework was developed for the Comirnaty mRNA vaccine, similar methodology could be applied to other vaccines. As with the other frameworks, generalisability to wider populations is a key concern, which we investigate in this paper.

2 Results

2.1 Key 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 Supplement for the studies details.)

The correlates of immune protection were based on data obtained at the Crick Institute within the ongoing Legacy Study (Wall et al., 2021). The study had one cohort with 149 UK healthcare workers who had a single dose of the Comirnaty vaccine, and another cohort, with $N = 159$, who had both doses. Blood plasma from each 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 same dose and variant. We fit these data by using the population protection function in (2), considered in the Supplement, which is common in correlates of vaccine 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 yet been applied in vaccine research. (More 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. There appears to be a substantial nonlinear association between neutralising titres from the Crick study and the observations from the effectiveness studies, with this association well represented by the assumed model of protection. 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 or on its boundary. The exception is the single-dose symptomatic effectiveness 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 additional reassurance about the model quality.

The variability in the effectiveness observations is greater after a single dose and peaks around the

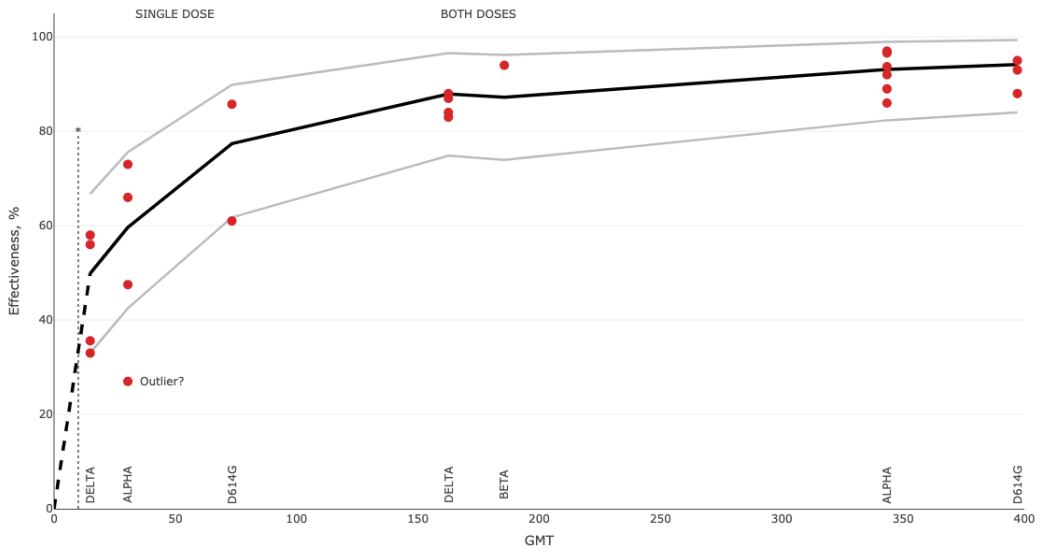


Figure 1: Protection model fit to symptomatic effectiveness observations. Each group of observations corresponds to the indicated variant and is plotted versus the geometric mean titre from the corresponding Crick sample. The 95% prediction bands are shown. The dotted line indicates the practical quantification threshold at $GMT = 10$. The dashed line extends the model fit beyond the data range towards zero.

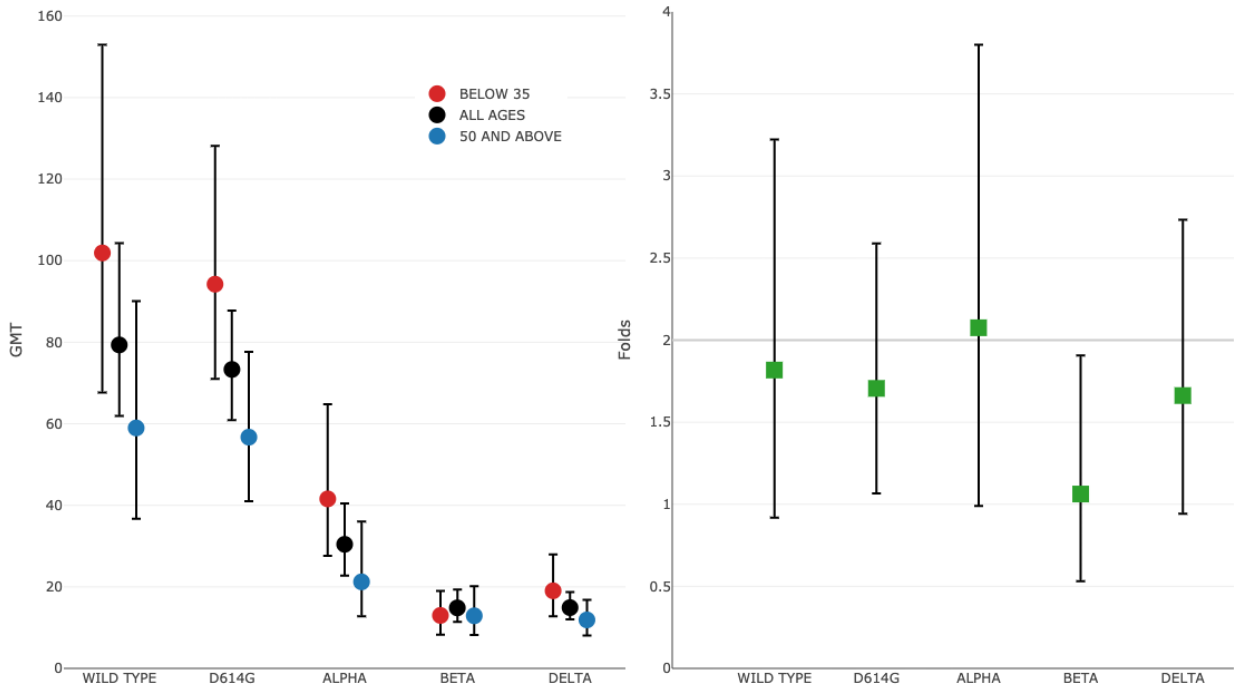
50% effectiveness, which is also reflected in the prediction bands in Figure 1. This pattern is unsurprising given that the uncertainty of a beta random variable is maximal in the middle of its domain. This provides additional empirical evidence for the assumed beta distribution.

2.2 Age subgroups

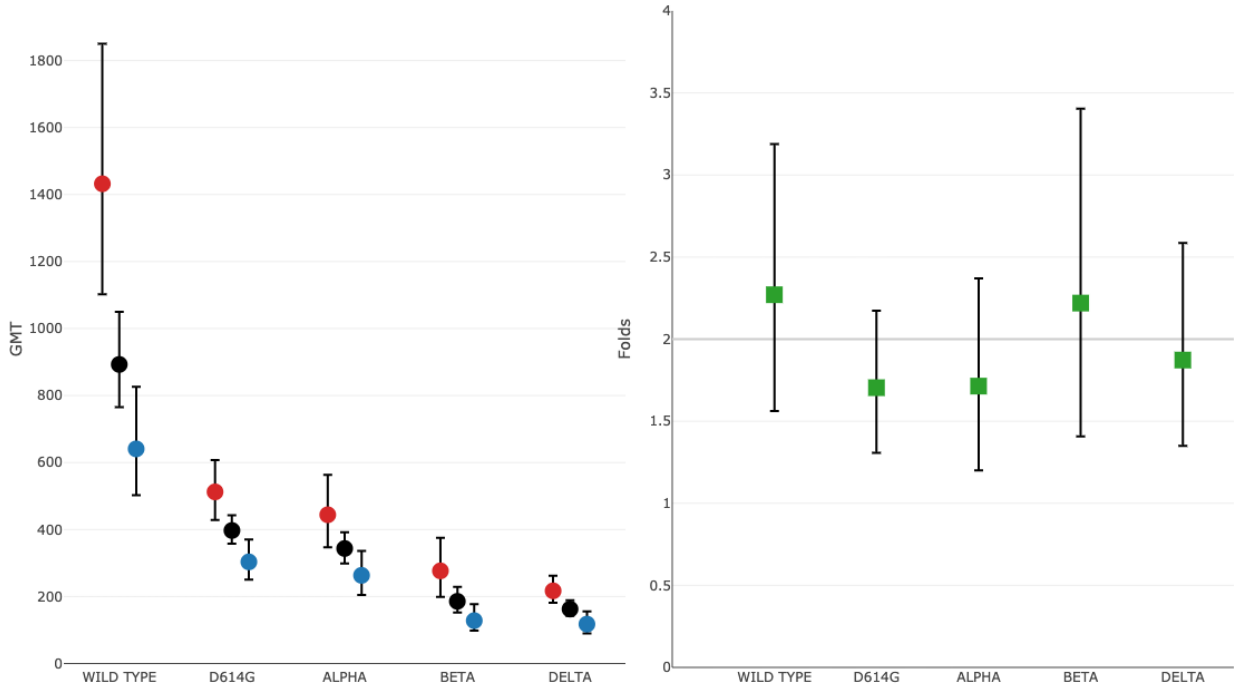
Differences in immune response between age groups could lead to substantial differences in vaccine effectiveness and thus could be relevant to vaccination policies.

We first compared the Crick neutralisation data per dose cohort in those below 35 with those 50 and over and with the overall 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 2.2, 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 the Supplement for its definition).

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



(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. (2021)

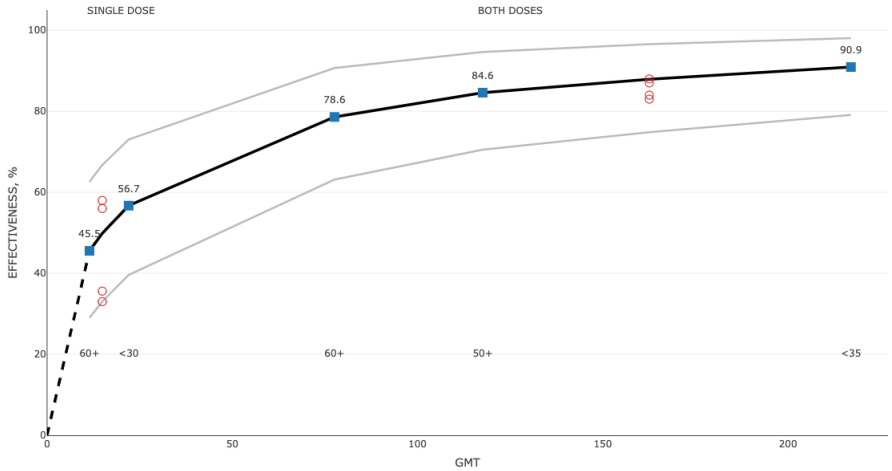


Figure 3: Predicted effectiveness against the Delta variant for selected age groups. The model fit, also in Figure 1, for all ages was used for this prediction. Also shown are effectiveness observations against Delta obtained in all- ages groups.

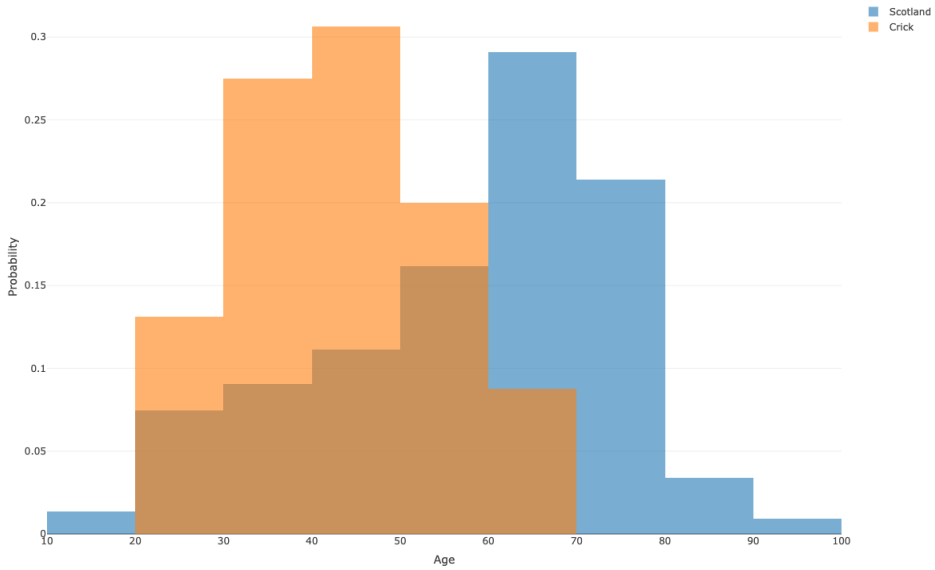


Figure 4: Age distribution for Comirnaty-vaccinated subjects in the Scotland study (Sheikh et al., 2021) versus that in the two-dose cohort in the Crick study (Wall et al., 2021)

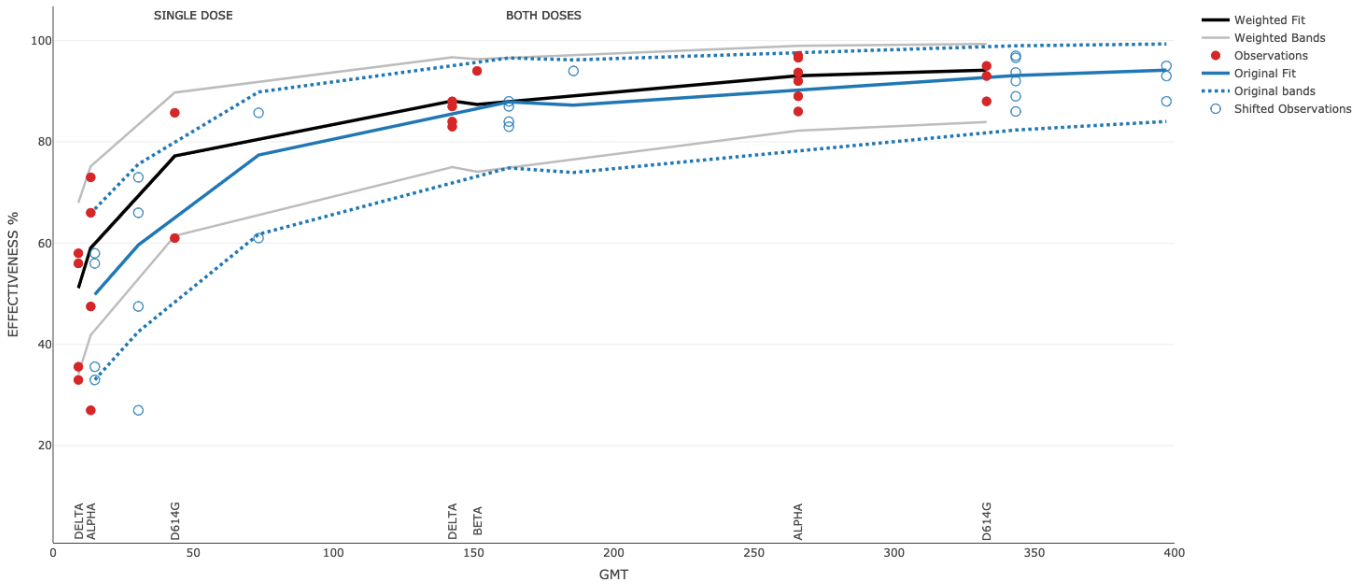


Figure 5: Model fits to age-weighted versus to original Crick neutralisation data. The observations are plotted versus the GMT per variant, for the corresponding titres distribution used in each fitting.

2.3 Population differences

An exploration

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 could be impacted by violations of this assumption due to population differences. (This presentation is a conceptual illustration rather than a quantitative prediction of real-life impacts.)

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. (2021). Of these, only Age was highly statistically significant, whereas the others were not statistically significant.

To represent the true population, we used the titres values from the Crick study, but reweighed them according to the age distribution from the Scotland study. Specifically, we used the age distribution for the Comirnaty vaccine in Figure S2 of Sheikh et al. (2021) defined for "those vaccinated in Scotland with one dose by 1 April 2021". The corresponding probabilities were applied to both dosing cohorts in the Crick study. According to Figure 4, 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 5. 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 indeed generated by the Scotland population, but the Crick population were used in modelling, our conclusions would be biased in the following sense: We would infer that higher than actual titres were required to attain a given

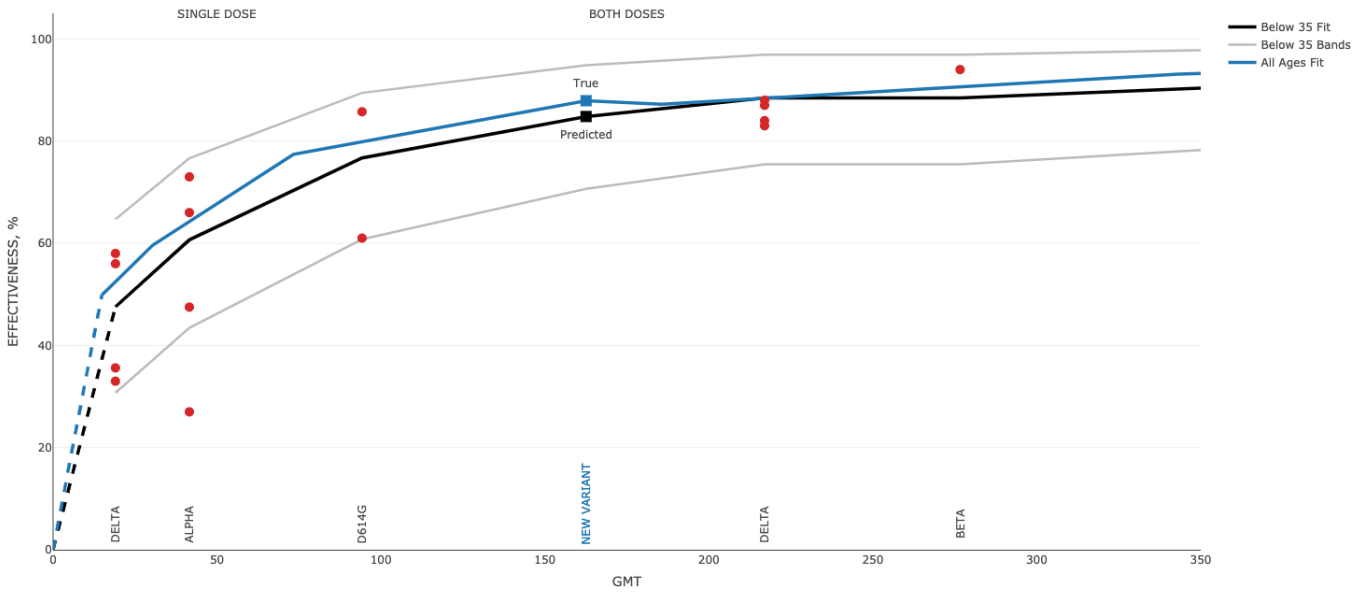


Figure 6: Cautious Prediction. Observations were fit to the below 35 age group; the all-ages fit is shown for comparison. Effectiveness for a new variant was predicted for all-age titres but assuming the below-35 fit.

level of vaccine effectiveness.

While weighted and unweighted models both appear to fit the corresponding effectiveness observations reasonably well, the predictions could be rather different. Unless we take care to address potential biases, such "good" model fits could be misleading. This might also be the case with the frameworks in Earle et al. (2021) and Khoury et al. (2021). (The discussion section reviews some adjustments for reducing such biases.)

Cautious prediction

Suppose that the overall Crick study population — possibly after an adjustment — has similar or slightly higher neutralisation titres than a true population does. There still would be uncertainty about how well the populations are matched and how well the "true" titres are approximated. There are also modelling and other uncertainties. Here we propose an approach to mitigate the impact of these uncertainties by including a safety margin in effectiveness predictions.

We first defined a sub-population of the Crick participants that, on average, has higher neutralisation titres. Here we simply chose all the participants aged under 35. We next assumed that they had generated effectiveness observations and so we fit the protection model to their neutralisation titres. Since these titres are higher than in the all-age group, this fit is to the right of the all-age fit; see Figure 6. In other words, we deliberately "raised the bar" to attain a given effectiveness level.

Suppose we aim to predict effectiveness against a new variant. For this we would collect neutralisation titres against the variant in the all-age population. To make a cautious prediction we would input these titres into the under-35 model fit.

In a numerical example, illustrated in Figure 6, the cautious prediction is 84.8%, which is 3.1% less than

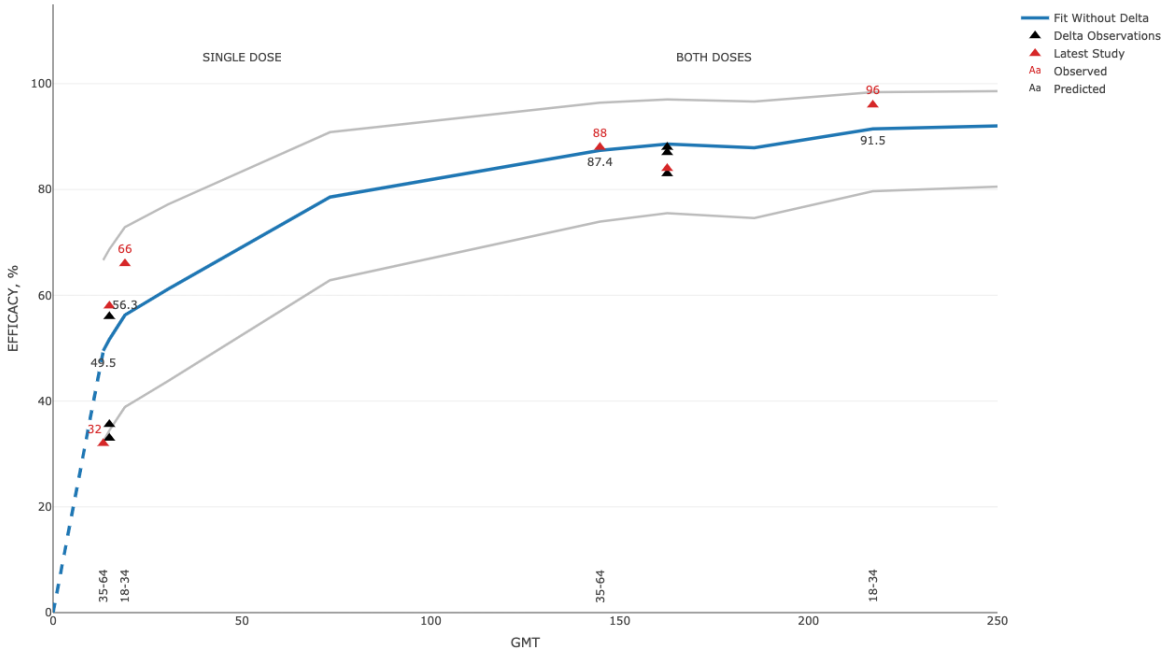


Figure 7: Predicted effectiveness against Delta for all-age groups and age subgroups. The model was fit using the effectiveness and neutralisation data on all the variants except Delta. The neutralisation titres were from all Crick study participants. Effectiveness values for all-age groups are stated in Figure 8.

87.9% with an all-ages fit. By using the cautious prediction we can infer that the all-age effectiveness would be at least 84.8% and likely higher. By extension, we infer the effectiveness in the true population likely to be approximately equal to and possibly higher than this number.

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

For the example in the previous section we note that the unweighted Crick population could already enable cautious predictions for the assumed true population. We could use the Crick population as is or apply the approach in this section for an extra margin of safety. The cautious approach could also be useful when combining effectiveness data from disparate populations in different effectiveness studies.

Excessive caution, however, is undesirable, as it may, for instance, lead to boosting too early.

2.4 Predicting vaccine effectiveness against the Delta variant

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 applied our framework to predict effectiveness against the Delta variant in the all-age, 18–34 and

35–64 groups. We first fit the all-age neutralisation titres to effectiveness observations on the assumed variants except Delta. We next input the neutralisation titres against Delta into the fitted model to predict effectiveness against this variant. We compared these predictions with effectiveness estimates from Nasreen et al. (2021), Sheikh et al. (2021), Bernal et al. (2021) and Pouwels et al. (2021) in Figure 7. These references reported all-age symptomatic effectiveness, whereas the last reference also for the age subgroups that we consider.

To assess the cautious approach, we also fit the all-age effectiveness data (without Delta) to the neutralisation titres obtained in those under 35. Figure 8 compares these "standard" and "cautious" predictions for all-age groups.

The predictions were particularly close after two vaccine doses. The standard approach over-predicted the mean of effectiveness observations only by 3.2% for the all-age group. For the age subgroups in Pouwels et al. (2021), the under-prediction was 4.5% for ages 18-34 and 0.5% for ages 35-64. The cautious approach under-predicted the mean of effectiveness observations only by 0.2% in the all-age group. (The cautious age threshold was chosen before making predictions.)

The predictions after a single dose were also close. The standard approach over-predicted the observations mean by 2.7%, and the cautious approach under-predicted by 6%; see Figure 8. Gauged against the effectiveness estimates in Pouwels et al. (2021), the prediction errors were noticeably larger for the age subgroups: an under-prediction by 9.7% for ages 18-34 and over-prediction by 17.5% for ages 35-64.

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 observations. Also, the confidence interval for the effectiveness of 32% in the 35-64 group was $(0, 53)$, the largest confidence interval reported in Pouwels et al. (2021) for the Comirnaty vaccine.

- The neutralisation data for the 35-64 group were outside the neutralisation data range modelled for the other variants

- As the model function is relatively steep at small titres, a difference between the true and assumed titres is magnified disproportionately.

- The Crick titres in the 35-64 group had a GMT of 13.3, which is close to the lower detection thresholds of 10 and 5 (see the Supplement for their definitions). Therefore the impact of left censoring could be more pronounced. Also, prediction could be sensitive to the choice of the threshold values.

Indeed, the largest prediction error (17.5%) was for ages 35-64, whose assumed neutralisation titres were the smallest. Once older members of this group get vaccinated with both doses, such prediction errors become less important in practice.

3 Discussion

Our proposed framework takes a step towards establishing correlates of protection for the Comirnaty and more broadly for SARS-CoV-2 vaccines. Our modelling produced close predictions of vaccine effectiveness

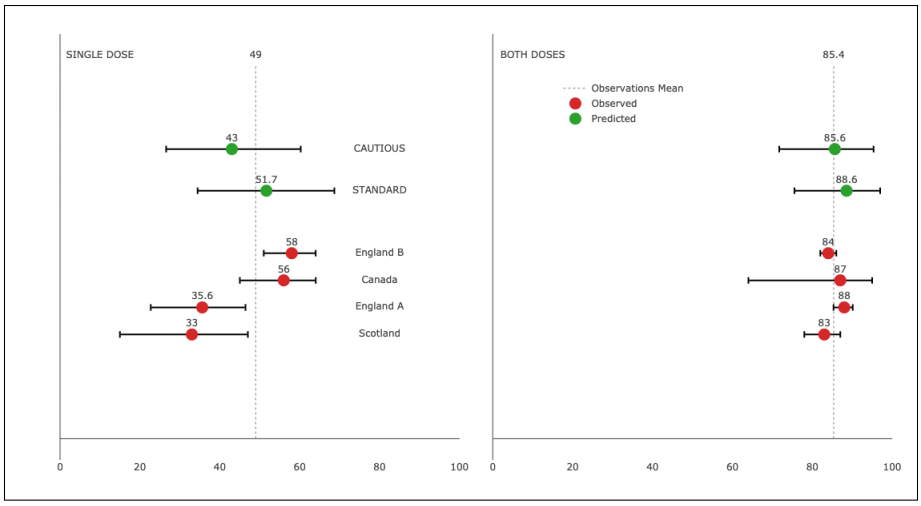


Figure 8: Predicting effectiveness against Delta. Estimates from effectiveness studies versus predictions from a model fit without Delta effectiveness observations. 95% prediction bands and weighted means of observations are shown.

against the Delta variant, especially after two doses. This is despite potentially large differences between the Crick study population and the populations in the effectiveness studies, and also despite using the Delta variant only for prediction, but not for model fitting.

(An exception is prediction at neutralisation titres near the lower detection thresholds, which may benefit from additional investigation.)

3.1 Populations matching

To address potential biases due to populations differences, some of the following adjustments could be used

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 in Section 2.3 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 and multimodel inference, could be also beneficial.

Using unadjusted Crick study data

The Crick study could benefit vaccine effectiveness assessments even without such adjustments. Indeed, its population could be representative of larger populations of healthcare workers and of general working-age populations.

Despite populations disparity, we could still obtain accurate predictions of vaccine effectiveness, as illustrated for the Delta variant in Section 2.4. We could also obtain useful lower bounds on vaccine effectiveness in wider populations, as considered in Section 2.3. In particular, data on immunity waning or on response to a third booster in the Crick population could be informative about wider vaccinated populations.

As illustrated in Figure 5 the shapes of the effectiveness curves were similar between the two populations considered. If this holds more generally, then modelling for the Crick study population could provide useful insights into vaccination policies.

3.2 Informing vaccination policies

We consider some examples of how vaccination policies could be informed by our proposed framework.

Effectiveness regions

Our framework suggests the following regions of vaccine effectiveness.

Flat region includes effectiveness against every variant after two vaccine doses in the all-age population (see, e.g., Figure 1). This region allows maximum flexibility for policy-making, since the impact on effectiveness due to moderate changes in neutralising titres would be small. For example, Payne et al. (2021) reported greater neutralisation for the Comirnaty vaccine with an eight-week dosing interval than with the three-to-five week intervals used in the clinical trials. However, a longer interval could expose vaccinees 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 to using smaller vaccine doses. Policy-makers could confidently decide on dose reduction so as to address supply shortages, high costs and dose-dependent side effects. This could also free up doses for vulnerable populations in low-income countries.

In the long run, waning vaccine protection can imply that immune responses in vaccinated individuals — particularly the elderly and vulnerable — are getting outside a flat effectiveness region 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 doses than the current doses.

Given the world-wide vaccine shortages, higher-income countries could implement boosting with reduced doses, sufficient for protecting domestic vulnerable populations and free-up vaccine supply for other countries.

Steep region corresponds, for example, to effectiveness against Alpha or Delta after a single dose. The 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 effectiveness protection of a country's population.

Intermediate region is located between the other two. It is exemplified by the single-dose effectiveness against D614G for all ages (Figure 1) and perhaps the two-dose effectiveness against Delta in the 60+ age groups (Figure 3). Its distinguishing feature is an increased uncertainty of effectiveness, as modelled by the beta distribution. Policies for an intermediate region could combine those for the other two, while also accounting for the extra uncertainty.

Notably, these regions and their interpretation could change if we consider vaccine effectiveness against hospitalisation and death.

Age differences There are considerable virus neutralisation and vaccine effectiveness differences between age groups, which could support age-tailored vaccination policies. For example, those under 35 had almost twice as high GMT against the Delta variant as those over 60 did, and would still be within a flat region; see Figure 3. Therefore, reduced doses in younger groups may be justified so as to expand vaccine coverage and reduce side effects along with vaccine hesitancy, which are prevalent in younger groups.

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 work. Unsurprisingly, older groups may need a booster sooner, and perhaps more frequently than the younger group would. Our framework could help quantify how much sooner, to which age groups and how more frequently.

Recent effectiveness estimates Our framework could also give insights into various aspects of the pandemic. One example is a surprisingly low symptomatic effectiveness, of 64%, of the vaccine against Delta announced in Israel Ministry of Health (2021); Zimmer (2021). (At the moment of writing, few details were available, including on the uncertainty interval.)

Our modelling implies that this effectiveness might not be completely surprising for older people (see Figure 3), given that vaccine effectiveness wanes over time. As an intermediate effectiveness region is approached, upcoming effectiveness estimates could become more dispersed. 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. (2021).) 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 sorely needed.

Our modelling suggests a compromise strategy to provide a full first dose but reduce the initial booster (second) dose in younger age groups, and also the second booster in all age groups. It could particularly make sense to reduce the second and even first doses in those under 18, given the strong immune response and concerns about toxicities in this group.

Investigating emerging variants earlier Earlier worldwide attention to the Delta variant might have allowed prompt action and have prevented many deaths. Virus neutralisation assays for the emerging variants could perhaps be done earlier, still at their variant-of-interest stage. Drops in vaccine effectiveness could then be predicted, and resources allocated to variants accordingly. Likewise, neutralisation titres could be monitored and vaccine effectiveness predicted for mutations of an existing variant of concern, such as for the Delta Plus.

Human challenge studies It could be possible to validate immune correlates of protection models in human challenge studies with the Comirnaty 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 (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.

Future work

More realistic adjustments for population differences could be incorporated within our framework. The modelling could be further validated with detailed data (e.g., on age groups) from the current and future effectiveness studies.

Refinements of our modelling could include incorporation of titres measurement errors; of additional covariates; and of within-study correlations of effectiveness estimates. Crick study data could also be combined with data from other SARS-CoV-2 neutralisation studies.

One straightforward extension of our modelling is to other types of vaccine effectiveness and efficacy; another is to effectiveness against a mix of SARS-CoV-2 variants. A key extension, which we are currently working on, is to effectiveness predictions for different Covid-19 vaccines.

Acknowledgements O.V. is grateful to Rosemary Bailey for her support and encouragement throughout this research. We thank Rebecca Lodwick and Barbara Bogacka for their suggestions on earlier versions of the paper which lead to its improvements. We are thankful to David L. Bauer for making additional data from the Crick Legacy Study available.

4 Supplement: Methods

4.1 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. (2021). These data are available at Wall et al. (2021) 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 9. 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. (2021), the observed titres were right censored at the upper detection limit of 5760 on the untransformed scale. Neutralisation values observed below 40 that did not have a strong enough signal to estimate the 50% inhibiting titre were recorded as 10. If no neutralisation signal was identified, a titre of 5 was recorded. These censoring thresholds are illustrated in Figure 9.

4.2 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

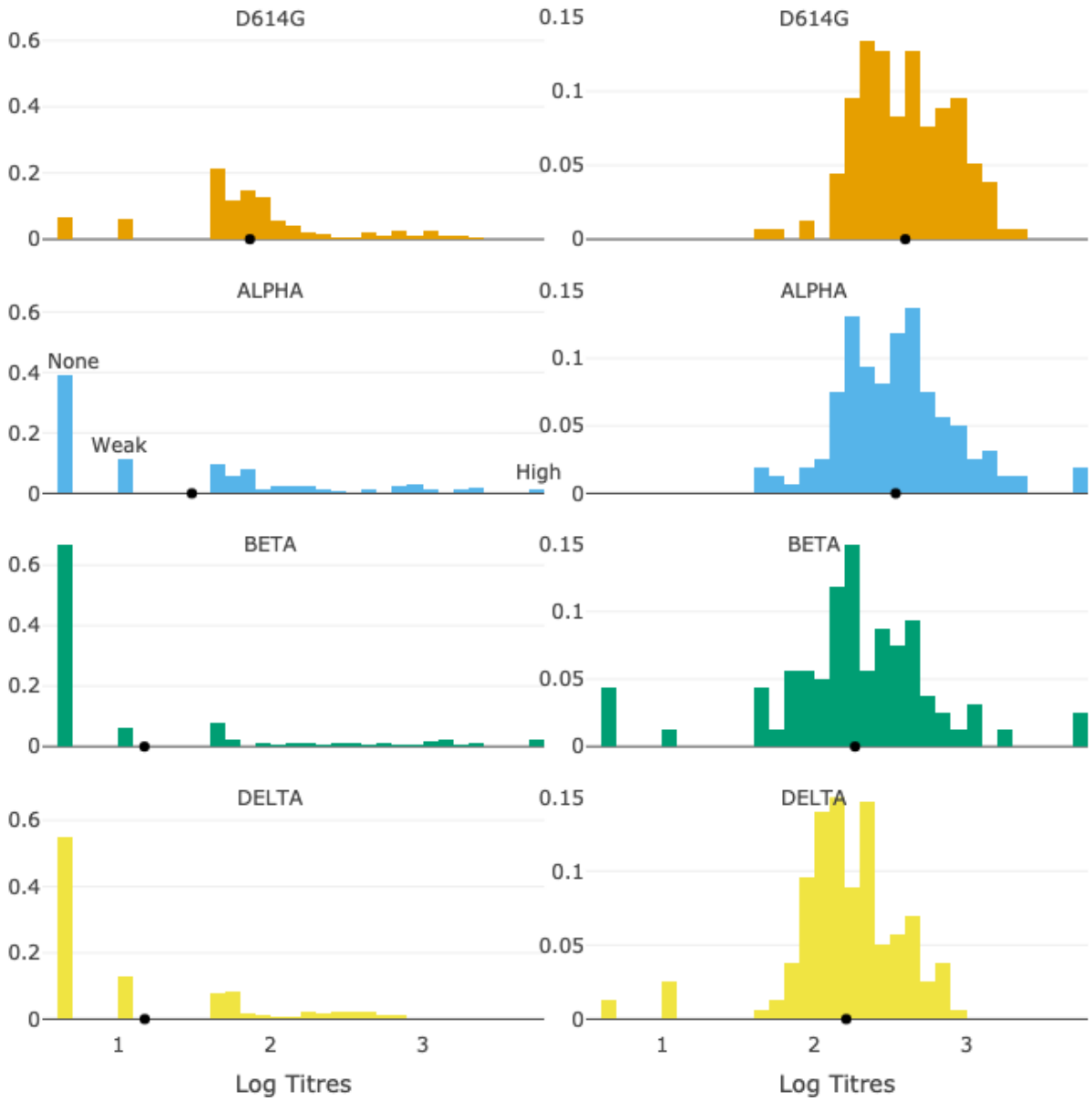


Figure 9: Distributions of log10 titres after one (Left) and two (Right) vaccine doses. Detection thresholds are illustrated for Alpha after one dose: "None", recorded as 5, means no signal; "Weak", recorded as 10, means the detected signal was not strong enough for fitting the neutralisation model; "High", at 5760, the upper limit of detection. The same thresholds are used across doses and variants. Sources: Wall et al. (2021).

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 a study with 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.

4.3 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.

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). We incorporated different uncertainties in effectiveness estimates by weighting each estimate by the reciprocal length of its reported confidence interval. 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.

5 Extended Data Supplement: Vaccine Effectiveness

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.

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)	
$*$	1	21	Delta	England B	66.0	(58.0, 72.0)	
$*$	2	14	Delta	England B	96.0	(93.0, 98.0)	
†	1	21	Delta	England B	32.0	(0.0, 53.0)	
†	2	14	Delta	England B	88.0	(78.0, 94.0)	

Table 1: Studies that provided symptomatic effectiveness data. Eff. - efficacy for Study 1, effectiveness for the other studies. Estimates were in all-age groups except $*$ for ages 18–34 and † for ages 35-64 in England B study. These age subgroups were only used with Figure 3 in Section 2.4.

6 Appendix: 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 10.

Figure 11 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.

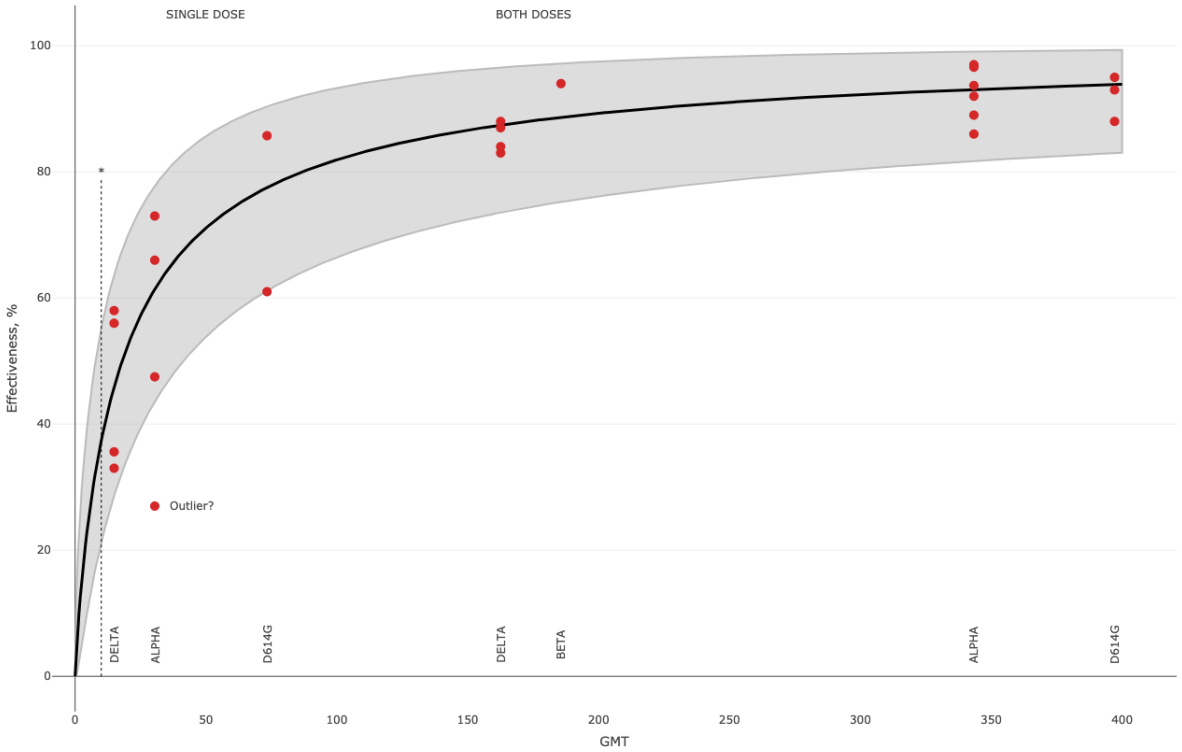


Figure 10: Simplified model fit to effectiveness observations, with 95% prediction bands shown. The dashed line indicates the quantification threshold at $GMT = 10$

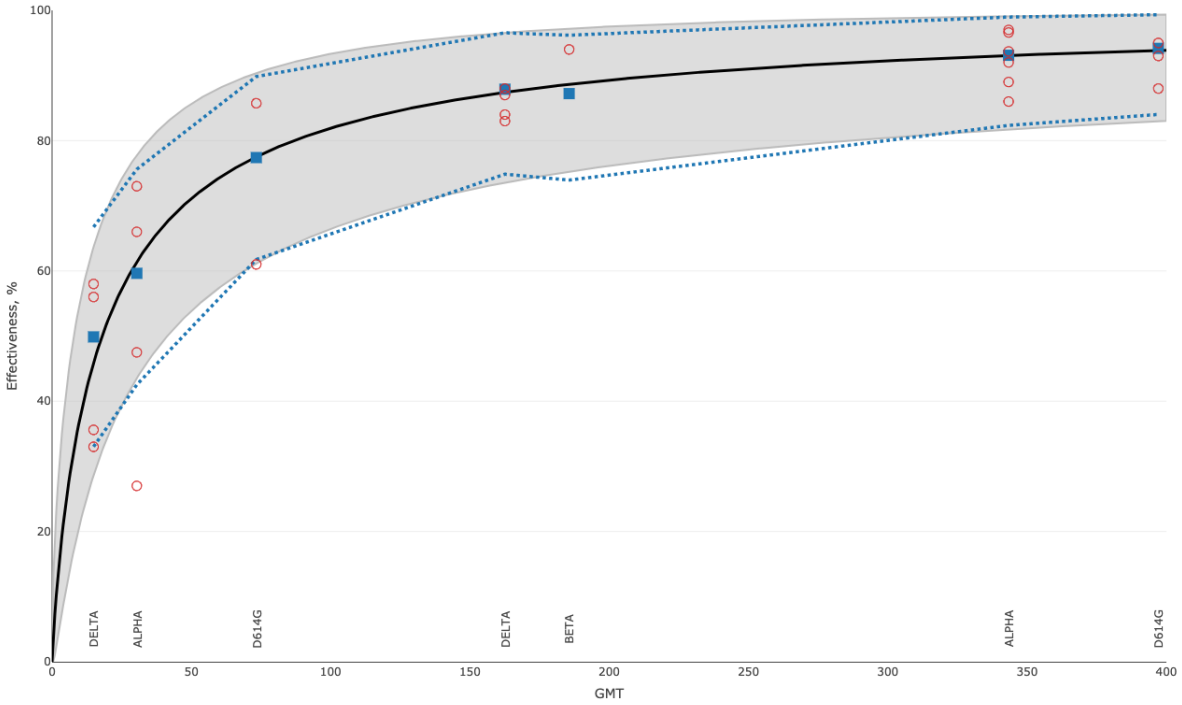


Figure 11: Simplified model fit and prediction bands are shown in a black line and grey shading. The full model fit and prediction bands are shown in blue squares and dashed lines.

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 are freely available from the paper's GitHub repository.

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).

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.

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*.

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.

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. (2021, 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.

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`.