# Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection

Shuo Feng 1, Daniel J. Phillips 1, Thomas White 2, Homesh Sayal 2, Parvinder K. Aley 1, Sagida Bibi 1, Christina Dold 1, Michelle Fuskova 3, Sarah C. Gilbert 3, Ian Hirsch 2, Holly E. Humphries 4, Brett Jepson 5, Elizabeth J. Kelly 6, Emma Plested1, Kathryn Shoemaker 5, Kelly M. Thomas 4, Johan Vekemans 7, Tonya L. Villafana 8, Teresa Lambe 3,8\*\*, Andrew J Pollard 1,9\*\*, Merryn Voysey 1,9\*\* and the Oxford COVID Vaccine Trial Group\*.

1. Oxford Vaccine Group, Department of Paediatrics, University of Oxford, UK
2. Late-stage development, Respiratory and Immunology (R&I), BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK
3. The Jenner Institute, Nuffield Department of Medicine, University of Oxford, UK
4. National Infection Service, Public Health England, Salisbury, UK
5. 5Late-stage development, Respiratory and Immunology (R&I), BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, US
6. 6Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, US
7. 7Late-stage development Respiratory and Immunology (R&I), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
8. Chinese Academy of Medical Science (CAMS) Oxford Institute (COI), University of Oxford, Oxford, UK.
9. NIHR Oxford Biomedical Centre, Oxford, UK

\*A list of authors and their affiliations appears at the end of the paper.

\*\* These authors contributed equally

Corresponding author:

Merryn Voysey [merryn.voysey@paediatrics.ox.ac.uk](mailto:merryn.voysey@paediatrics.ox.ac.uk)

Oxford Vaccine Group, Department of Paediatrics, University of Oxford,

CCVTM, Churchill Road, Churchill Hospital,

Oxford OX3 7LE United Kingdom

# Abstract

Global supply of COVID-19 vaccines remains limited. An understanding of the immune response predictive of protection could facilitate rapid licensure of new vaccines.Data from a randomised efficacy trial of ChAdOx1 nCoV-19 (AZD1222) vaccine in the UK was analysed to determine the antibody levels associated with protection against SARS-CoV-2. Binding and neutralising antibodies at 28 days after the second dose were measured in infected and non-infected vaccine recipients. Higher levels of all immune markers were correlated with a reduced risk of symptomatic infection. Vaccine efficacy of 80% against symptomatic infection with majority Alpha (B.1.1.7) variant of SARS-CoV-2 was achieved with 264 (95%CI: 108, 806) Binding Arbitrary Units (BAU)/ml and 506 (95%CI: 135, not computed (beyond data range) (NC)) BAU/ml for anti-spike and anti-RBD antibodies, and 26 (95%CI: NC, NC) International Unit (IU)/ ml and 247 (95%CI: 101, NC) normalised neutralisation titres (NF50) for pseudo- and live virus-neutralisation respectively. Immune markers were not correlated with asymptomatic infections at 5% significance level. These data can be used to bridge to new populations using validated assays, and allow extrapolation of efficacy estimates to new COVID-19 vaccines.

# Main text

## Introduction

Within 17 months of the identification of SARS-CoV-2 in Wuhan, China, in response to the pandemic, six COVID-19 vaccines were recommended for use by the WHO as of 16th June 2021.1 Vaccine efficacy ranging from 50% to 95% against symptomatic COVID-19 infections was reported using varying endpoint definitions.2-7 Real world evidence from vaccine rollout programmes has shown that COVID-19 vaccines are highly effective against severe disease, hospitalisation, and death, and reduce both asymptomatic infection and within household transmission.8-13

Global supply of COVID-19 vaccines remains limited despite intense production efforts. Authorization of new vaccines could help meet demand. As more countries implement vaccine programmes it will become increasingly difficult to conduct clinical efficacy studies of new vaccines. Understanding the relationship between immune responses to vaccines and protection against clinical outcomes is urgently needed to speed vaccine development. Knowledge of immune measures that are statistically associated with protection against disease (“correlates of protection”) may allow new vaccines to be authorised for use based on immunogenicity and safety data alone, when large efficacy trials are not feasible. In addition, understanding the immune response allows vaccines to be compared across cohorts of people who differ by age, race, ethnicity or other factors.

Both binding and neutralising antibodies are thought to be potential correlates of protection against COVID-19 and are correlated with each other.3,14-16 Previous human challenge studies of seasonal coronaviruses reported high levels of baseline neutralising antibody in uninfected or asymptomatic persons.17 However protection from infection with seasonal coronaviruses is not long lasting.17,18

Early evidence from a fishery vessel outbreak of SARS-CoV-2 suggested higher pre-existing neutralising antibodies were potential correlates of protection.18,19 A longitudinal cohort study of healthcare workers highlighted the association between baseline anti-spike and anti-nucleocapsid IgG and decreased risk of SARS-CoV-2 infection in the following 6 months.19,20

Evidence that antibodies may play a role in mediating protection against overt disease has come from vaccination and challenge studies in animals. Both neutralising antibody titres and Fc-dependent functional antibody responses correlate with protection induced by DNA and adenoviral vectored vaccines in rhesus macaques.21,22 Additionally, higher doses of passively transferred monoclonal antibodies were more protective than lower doses in golden hamsters and rhesus macaques challenged with the SARS-CoV-2 virus containing D614 in its spike protein.23

A meta-analysis modelling the relationship between vaccine efficacy (VE) reported from phase 3 vaccine clinical trials and neutralisation titres in convalescent patients showed a significant association at the study level between VE and neutralising antibody levels .24 Nevertheless, no study to date has defined a correlate of protection against SARS-CoV-2 infection or disease that can be used by regulators and vaccine developers.

The ChAdOx1 nCoV-19 vaccine (AZD1222) is a chimpanzee adenoviral vectored vaccine with full length SARS-CoV-2 spike insert which was developed at the University of Oxford and is in widespread global use and produced by AstraZeneca and their manufacturing partners. Using data from the UK and Brazil, we previously estimated an overall vaccine efficacy of 66.7% (95% CI: 57.4 to 74.0) against symptomatic infection and 27.3% (95% CI: −17.2 to 54.9) against asymptomatic infection.2,3 We previously showed that estimates of vaccine efficacy against symptomatic COVID-19 infection were higher in subgroups with higher pseudovirus neutralisation antibody titres, or higher anti-spike IgG levels in vaccine clinical trials of ChAdOx1 nCoV-19 in adults using summarised data.3 Here we report the relationship between a continuous measure of the humoral immune responses to vaccination and protection afforded by this vaccine which may facilitate further vaccine development. Specifically, we used individual data from the UK and identified the thresholds for four immune markers associated with protection against symptomatic infection. The WHO international standard units were reported for all assays, to allow comparisons across studies and platforms.

## Results

Using the COV002 data from the UK, we assessed the correlation between immune markers at 28 days post the second dose (PB28) of ChAdOx1 nCoV-19 and symptomatic and asymptomatic infections. Participants were reminded weekly to contact their study site if they experienced any of the primary symptoms of COVID-19 (fever ≥ 37.8oC; cough; shortness of breath; anosmia or ageusia) and were assessed in clinic, with a nose and throat swab taken for nucleic acid amplification testing (NAAT). Additionally, participants were asked to complete a nose and throat swab at home each week, which was used to detect asymptomatic infections. NAAT+ participants who had symptoms other than the main five COVID-19 symptoms were categorised as non-primary symptomatic and were not included in correlates analysis.

Table 1 summarises baseline characteristics for the defined Correlates Population, Control Population, and Correlates Cohort by cases and non-cases status. Extended Data Fig.1 summarises the exclusions for each study group. Participants were followed for a median of 88 and 85 days counting from 7 days after the post-boost + 28 days (PB28) visit, among Correlates and Control Populations respectively. The follow-up time was censored at the earliest timing of infection, withdrawal, unblinding or cut-off date February 28th 2021. Among 4,372 Correlates Population participants, there were a total of 174 breakthrough nucleic acid amplification test positive (NAAT+) cases of SARS-CoV-2 infection. Data were available for at least one of four assay readouts (anti-spike IgG, anti-RBD IgG, pseudovirus neutralization assay and live virus neutralization) for 171/174 (98.3%) cases and 1404/4195 (33.5%) non-cases. Data were available for anti-spike and anti-RBD IgG from 1318 PB28 samples (163 cases and 1155 non-cases, Table S2). A smaller set of data was available for analysis for pseudovirus neutralisation titres (149 cases, 828 non-cases) and for live virus neutralisation (110 cases and 412 non-cases) (Table S2). Cases were younger, with 84.2% of cases being 18-55 years compared with 71.6% of non-cases, and more likely to be healthcare workers (62.0% of cases were healthcare workers compared with 57.5% of non-cases, Table 1). In our baseline exposure model developed among the MenACWY group, younger age and being a healthcare worker facing more than 1 COVID patient per day were associated with a higher risk of being NAAT+. Other variables were not significant (see model output in Table S3). The distribution of baseline risk was similar for cases and non-cases (Table 1).

Antibody levels at PB28 in cases and non-cases across four immune markers are shown in Extended Data Fig.2. Anti-spike IgG and anti-RBD IgG were highly correlated with each other (Pearson correlation coefficient *r* = 0.926) while the correlation between pseudovirus neutralisation titre and normalised live virus neutralisation titre (NF50) was moderate (*r*= 0.572). Anti-spike IgG values were also correlated with pseudovirus neutralisation titres (*r* = 0.657) and normalised live virus neutralisation titres (NF50) (*r* = 0.600) (Extended Data Fig.3). Non-normalised live virus neutralisation titres (ND50) were less highly correlated with anti-spike IgG (*r* =0.411) and pseudovirus neutralisation titres (*r* =0.305).

The risk of symptomatic COVID-19 decreased with increasing levels of anti-spike IgG (p=0.003), anti-RBD IgG (p=0.018), pseudovirus neutralisation titre (p=0.005) and live virus neutralisation titre (p<0.001) (Figure 1, Figure 2, Table 2). In contrast, there were no significant associations between any of the assays and protection against asymptomatic infection including for sensitivity analysis restricting to high viral load (all p>0.05, Figure 3, Extended Data Fig.4, Extended Data Fig.5, Table S4). When primary symptomatic COVID-19 cases were classified according to the presence of shortness of breath, we observed a similar trend with increasing immune marker levels associated with lower risk of infection (all p<0.05, Table S4, Extended Data Fig.6), but not for those with no shortness of breath (all p>0.05, Table S4, Extended Data Fig.7). Higher pseudovirus and live virus neutralisation titres were associated with lower risk of infection for those who had 3 or more COVID symptoms (Table S4, Extended Data Fig.8). The number of cases and non-cases included for correlates analysis by each immune-marker and outcome has been summarised in Table 2 and Table S4.

The antibody level associated with 80% VE against primary symptomatic COVID-19, was 40923 (95% CI: 16748, 125017) arbitrary units (AU)/ml for anti-spike IgG, equivalent to 264 binding antibody units (BAU)/ml (95% CI 108, 806) using the WHO international standard (NIBSC code 20/136). For anti-RBD IgG, 80% efficacy was achieved with median antibody level of 506 (95% CI: 135, not computed (NC)) BAU/ml (Figure 2, Figure 4, Table 2).

For pseudo- and live virus-neutralising antibody titres, values of 26 (95% CI: NC, NC) IU/ml and 247 (95% CI: 101, NC) normalised neutralisation titres (NF50) respectively were associated with 80% VE against symptomatic infection (Table 2). No values from these assays were associated with protection against asymptomatic infection (Table S4)

For all assays, when the analysis was restricted to symptomatic cases with shortness of breath, 80% VE was achieved at lower levels of immune markers than for symptomatic cases in general. Higher baseline exposure risk of SARS-CoV-2 infections predict higher probability of all outcomes (all p < 0.05, Table 2, Table S4) except for asymptomatic infections (p>0.05) in generalised additive models.

## Discussion

Here, we report an analysis of potential correlates of protection using data from 171 SARS-CoV-2 infection cases and 1404 non-cases, showing that higher anti-spike, anti-RBD IgG, and neutralising antibody titres are all associated with lower risk of symptomatic disease. We used immune responses in a phase 2/3 clinical trial to derive a model to predict absolute risk of infection, with appropriate adjustment for bias, assigning estimates for each level of antibody in the dataset. The relative risk of infection was then derived by reference to risk of infection in the control group. This is a robust approach to derive population estimates and adapted from recently described methods.25,26

We previously published overall aggregate-level summaries of antibody levels in participants with different prime-boost intervals. Vaccination prime-boost intervals were associated with varying levels of vaccine efficacy and there are some intriguing similarities between the aggregate level data with the estimate provided from analysis of individual participant level data in this report. The estimated anti-spike IgG level of 40923 AU/ml and the pseudo neutralising antibody titre of 185 associated with 80% VE in our models, were similar to the GMTs of 48961 AU/ml and 237.0 respectively, previously reported in the subgroup of participants vaccinated with ChAdOx1 nCoV-19 with a dose interval of at least 12 weeks between their 1st and 2nd dose – a regimen that provided 80.0% (95% CI 65.2 to 88.5) vaccine efficacy in the pooled analysis of clinical trialdata from the UK, Brazil and South Africa.3 The aggregate level results previously published included all eligible participants in the assessment of vaccine efficacy but only those with available antibody data were included in the summaries of immunogenicity meaning that direct comparisons of efficacy with immunogenicity were not in the same populations. Our current approach analyses the relationship between infections and antibody levels at the individual level in a single set of participants, with appropriate adjustment for confounding, providing robust outputs. In addition, the current work provides outputs in WHO standard units which are necessary for comparisons with data from other laboratories with different assays.

In a preprint by Gilbert et al., correlates of protection derived from the Moderna phase 3 efficacy trial are reported using similar methodology.27 Although overall binding and pseudovirus neutralising antibody titres after vaccination were higher in that study than those measured after the ChAdOx1 nCoV-19 vaccine, the correlates of protection findings appear similar to those we report here

No serological measurements in our data were shown to correlate with protection against asymptomatic infection or against symptomatic illness with only mild upper respiratory symptoms. This is consistent with our interim analysis that vaccine efficacy against asymptomatic infection was 27.3% (95% CI: −17.2 to 54.9) and was not significant at the 5% level.2 These results are consistent with the real world observation that infection remains possible in fully vaccinated individuals, despite high effectiveness against severe disease.

Antibody correlates presented in this report, relate to protection against mild disease, defined as a NAAT positive test with at least one symptom present. Weekly self-swabbing in the trial enabled detection of many mild cases. At these antibody titres, efficacy against more severe endpoints, used in other trials, would be higher than the estimates in this analysis. Notably, this has been confirmed in the analysis of real world effectiveness, in which the milder cases are not detected, after two doses of the vaccines in older adults in England where VE was 90% for Pfizer and 89% for ChAdOx1 nCoV-19 against symptomatic disease using the same case definition for both vaccines,11 while lower efficacy estimates were measured in our previously reported efficacy analysis with a milder disease endpoint.2

The correlates of vaccine efficacy reported here could be used to extrapolate efficacy to immunogenicity data for novel vaccines where clinical efficacy results are unavailable. A trial of a new vaccine that works through similar immune mechanisms and which produces antibody responses that are above the correlate values reported here, in at least 50% of participants (i.e. has a similar or higher median), might be expected to have similar efficacy against the clinical endpoints used in our UK trial, and higher efficacy against more severe endpoints. We provide correlates for vaccine efficacy estimates ranging from 50% to 90% to allow flexibility in the way these estimates are utilised by the regulators and policy-makers.

It has previously been shown that protection against lower respiratory tract infection (LRTI) may be easier to achieve than against upper respiratory tract infection (URTI) as challenge studies in rhesus macaques have shown stronger correlation between neutralising titres and the level of subgenomic mRNA in bronchoalveolar lavage samples than in nasal swab samples.28

Similarly, ChAdOx1 nCoV-19 [vaccinated hamsters, with low neutralising titres against B.1.351, were fully protected against LRTI following challenge with B.1.351, despite no evidence of protection against shedding of virus from the upper airway.29 Protection against upper respiratory tract or asymptomatic infections may be more closely associated with the presence of secretory IgA on the mucosal surface which was not measured in this study.30

These observations indicate that reduced neutralising capacity against B.1.351, and other variants of concern, might drive reduced protection against initial infection, and perhaps transmission, but protection against severe disease is maintained. Clinical trials of SARS-CoV-2 vaccines have consistently shown higher efficacy against more severe forms of disease such as hospitalisation or death, than against mild infections.2-5,15,31 We are unable to assess correlates of protection against severe disease or hospitalisation as there were no vaccinated participants hospitalised in the COV002 study.

Although live virus and pseudovirus neutralisation assays were modestly correlated with each other, the live virus assay was more closely associated with protection against symptomatic COVID-19 than the pseudovirus assay. This may reflect the sensitivity and dynamic range of the assays.

Protection against symptomatic COVID-19 is not absolute with any vaccine, and the results presented here show that there is no single threshold value for any of the assays investigated that was indicative of sterilising immunity in our data. Instead, the probability of infection decreases on average with higher immune responses but substantial variation exists between individuals. This is similar to studies of respiratory syncytial virus where risk of infection decreased with higher antibody levels, although infections were still observed at high levels of antibody, suggesting a definitive individual threshold of protection does not exist.32 We provide antibody estimates that correspond with 50% to 90% VE however the wide confidence intervals around these estimates should be noted.

These estimates represent the antibody level observed 28 days after a *second* dose of vaccine that provide protection during the subsequent 4-6 month period among UK COV002 efficacy and immunogenicity cohorts. This is different from the antibody level that would protect an individual at the time of exposure to the virus. Further work is needed to determine the durability of antibody and long term protection after vaccination.

High levels of protection were noted after vaccination with one dose of a lipid nanoparticle RNA vaccine, despite modest levels of neutralising antibody, strongly supporting the concept that other mechanisms are at play as co-correlates of protection.5,33 We have previously shown that a wide range of Fc-mediated antibody functions are induced by vaccination, and it is possible that these functions may be important in the absence of neutralising antibody.34 Furthermore, strong T cell responses induced by ChAdOx1 nCoV-19 may contribute to protection14,16 and have been associated with recovery from COVID-19 disease.35-37 The relationship between antibody and T cell responses may differ depending on the type of vaccine used, and care should be taken in interpreting data from clinical testing of different vaccine technologies.

There are some limitations to our analysis. These analyses are based on cases of COVID-19 detected in a mainly white population in the UK, which were mostly due to B.1.177 and B.1.1.7 variants. In settings where these are not the dominant variants causing disease, or where neutralisation assays use different strains of the virus, the modelled relationships between immune markers and disease outcomes shown here may not apply. In addition, we have conducted a large number of analyses and therefore some caution should be taken when drawing conclusions based on single p values alone as these are presented unadjusted for multiple comparisons. Furthermore, these analyses have been conducted on samples taken after 2 doses of ChAdOx1 nCoV-19 and might not apply to protection afforded by a single dose of the same vaccine or other COVID-19 vaccines. Correlates may also vary according to age profile, but this not explored in our study due to the small number of older adults recruited. The potential role of T cells and interaction between humoral and cellular immunity has not been evaluated in this study. It is not possible to determine in this study if our results represent mechanistic or non-mechanistic correlates of protection, as many immune responses are highly correlated.

Correlates of protection can be used to bridge to new populations and new vaccines using validated assays. These data can be used to extrapolate efficacy estimates for new vaccines that use similar immune mechanisms and where efficacy data is unavailable.

**Acknowledgments**

This Article reports independent research funded by UK Research and Innovation (MC\_PC\_19055: SG, AJP, TL), Engineering and Physical Sciences Research Council (EP/R013756/1: SG, AJP, TL), National Institute for Health Research (COV19 OxfordVacc-01: SG, AJP, TL), Coalition for Epidemic Preparedness Innovations (Outbreak Response To Novel Coronavirus (COVID-19: SG, AJP, TL)), National Institute for Health Research Oxford Biomedical Research Centre (BRC4 Vaccines Theme), Thames Valley and South Midland’s NIHR Clinical Research Network, and AstraZeneca.  The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the UK Department of Health and Social Care. We acknowledge support from Thames Valley and South Midland’s NIHR Clinical Research Network and the staff and resources of NIHR Southampton Clinical Research Facility and the NIHR Oxford Biomedical Research Centre. AJP is a NIHR senior investigator.

The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the UK Department of Health and Social Care. AstraZeneca reviewed the final manuscript but the academic authors retained editorial control. Other Funders had no role in study design, data collection and analysis, or decision to publish.

We thank the volunteers who participated in this study. We thank Professor Peter B. Gilbert and Dr Peter Dull for their advice and contributions to the methodology.

**Author Contributions Statement**

MV and SF designed the study. SF, DJP, TW, HS, BJ, KS, MV and IH contributed to the data analysis and methods. PKA, SB, CD, MF, HEH, EJK, EP, KMT, TLV, and JV contributed to implementation of the study and/or laboratory experimentation. SF, MV,  TL, and AJP contributed to the preparation of the report. All authors critically reviewed and approved the final version.

**Competing Interests Statement**

Oxford University has entered into a partnership with AstraZeneca for further development of ChAdOx1 nCoV-19. SCG is co-founder of Vaccitech (collaborators in the early development of this vaccine candidate) and named as an inventor on a patent covering use of ChAdOx1 vectored vaccines and a patent application covering this SARS-CoV-2 vaccine (PCT/GB2012/000467). TL is named as an inventor on a patent application covering this SARS-CoV-2 vaccine and was a consultant to Vaccitech for an unrelated project during the conduct of the study. AJP is chair of the UK Department of Health and Social Care (DHSC) Joint Committee on Vaccination & Immunisation (JCVI) but does not participate in discussions on COVID-19 vaccines, and is a member of WHO SAGE. AJP is a National Institute for Health Research (NIHR) Senior Investigator. The views expressed in this article do not necessarily represent the views of the DHSC, JCVI, NIHR, or WHO. TW, HS, IH, JB, EJK, KS, JV, TLV are employees of AstraZeneca. The other authors declare no competing interests.

**Oxford COVID Vaccine Trial Group** Authorship List

Syed Adlou 1, Lauren Allen 4, Brian  Angus 3, Rachel Anslow 1, Marie-Claude  Asselin 4, Natalie Baker 4, Philip Baker 1, Thomas  Barlow 4, Amy Beveridge 1, Kevin R. Bewley 4, Phillip Brown 4, Emily  Brunt 4, Karen R. Buttigieg 4, Susana Camara 1, Sue Charlton 4, Emily Chiplin 4, Paola Cicconi 3, Elizabeth A. Clutterbuck 1, Andrea M. Collins 10,11, Naomi S. Coombes 4, Sue Ann Costa Clemens 1,12, Melanie Davison 4, Tesfaye Demissie 1, Tanya Dinesh 1, Alexander D. Douglas 3, Christopher J. A.  Duncan 13,14, Katherine R. W. Emary 1, Katie J. Ewer 3, Sally  Felle 1, Daniela M. Ferreira 7, Adam  Finn 15,16, Pedro M. Folegatti 3, Ross Fothergill 4, Sara  Fraser 4, Harriet Garlant 4, Laura Gatcombe 4, Kerry J. Godwin 4, Anna L. Goodman 17,18, Christopher A. Green 19,20, Bassam  Hallis 4, Thomas C. Hart 1, Paul T. Heath 21, Helen  Hill 10, Adrian V. S. Hill 3, Daniel  Jenkin 3, Mwila Kasanyinga 1, Simon Kerridge 1, Chanice Knight 4, Stephanie  Leung 4, Vincenzo  Libri 22,23, Patrick J.  Lillie 24, Spyridoula Marinou 1, Joanna McGlashan 4, Alastair C. McGregor 25,26, Lorna  McInroy 4, Angela  M. Minassian3, Yama F. Mujadidi 1, Elizabeth J. Penn 4, Katrina M. Pollock 27, Pamela C. Proud 4, Samuel Provstgaard-Morys 1, Durga Rajapaska 4, Maheshi N. Ramasamy 1, Katherine Sanders 1, Imam Shaik 4, Nisha Singh 1, Andrew Smith 28, Matthew D. Snape 1, Rinn Song 1, Sonu Shrestha 1, Rebecca K. Sutherland 29, Emma C. Thomson 30, David P. J. Turner 31, Alice Webb-Bridges1, Christopher J. Williams 32

1. Department of Clinical Sciences, Liverpool School of Tropical Medicine, UK
2. Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK
3. Institute of Global Health, University of Siena, Italy
4. Department of Infection and Tropical Medicine, Newcastle upon Tyne Hospitals NHS Foundation Trust;
5. Translational and Clinical Research Institute, Immunity and Inflammation Theme, Newcastle University
6. School of Population Health Sciences, University of Bristol, UK
7. University Hospitals Bristol and Weston NHS Foundation Trust, UK
8. Department of Infectious Diseases, Guy's and St Thomas' NHS Foundation Trust, St Thomas' Hospital, London, UK and
9. MRC Clinical Trials Unit at University College London, UK
10. NIHR/Wellcome Trust Clinical Research Facility, University Hospitals Birmingham NHS Foundation Trust and
11. Institute of Microbiology & Infection, University of Birmingham, UK
12. St George's Vaccine Institute, St George's, University of London, UK
13. NIHR UCLH Clinical Research Facility and
14. NIHR UCLH Biomedical Research Centre, London, UK
15. Hull University Teaching Hospitals NHS Trust and Hull York Medical School, UK
16. London North West University Healthcare NHS Trust
17. Department of Medicine, Imperial College London, UK
18. NIHR Imperial Clinical Research Facility and NIHR Imperial Biomedical Research Centre, London, UK
19. College of Medical, Veterinary & Life Sciences, Glasgow Dental Hospital & School, University of Glasgow, UK
20. Clinical Infection Research Group, Regional Infectious Diseases Unit, NHS Lothian, Edinburgh, UK
21. MRC - University of Glasgow Centre for Virus Research & Department of Infectious Diseases, Queen Elizabeth University Hospital, Glasgow, UK
22. University of Nottingham and Nottingham University Hospitals NHS Trust, UK
23. Aneurin Bevan University Health Board, Newport, Wales

# Tables

**Table 1 Baseline characteristics of correlates population, control population, and cases and non-cases among correlates cohort.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ChAdOx1 nCoV-19 Correlates population  (N = 4,372) | MenACWY Control population  (N = 4,194) | ChAdOx1 nCoV-19 Correlates cohort\* | |
| Cases  (N = 171) | Non-cases  (N = 1404) |
| Age group |  |  |  |  |
| 18-55 years | 3240 (74.1%) | 3229 (77%) | 144 (84.2%) | 1005 (71.6%) |
| 56-69 years | 542 (12.4%) | 482 (11.5%) | 10 (5.8%) | 194 (13.8%) |
| ≥70 years | 590 (13.5%) | 483 (11.5%) | 17 (9.9%) | 205 (14.6%) |
| Sex (Female) | 2533 (57.9%) | 2526 (60.2%) | 102 (59.6%) | 780 (55.6%) |
| Ethnicity |  |  |  |  |
| White | 4036 (92.3%) | 3914 (93.3%) | 160 (93.6%) | 1293 (92.1%) |
| Asian | 220 (5.0%) | 184 (4.4%) | 8 (4.7%) | 71 (5.1%) |
| Black | 21 (0.5%) | 15 (0.4%) | 1 (0.6%) | 10 (0.7%) |
| Other† | 95 (2.2%) | 81 (1.9%) | 2 (1.2%) | 30 (2.1%) |
| BMI (mean (SD)) | 26.4 (5) | 26.5 (5.2) | 27 (5.2) | 26.5 (5.1) |
| BMI<30 | 3519 (80.5%) | 3347 (79.8%) | 130 (76.0%) | 1124 (80.1%) |
| BMI≥30 | 852 (19.5%) | 846 (20.2%) | 41 (24.0%) | 280 (19.9%) |
| Co-morbidities | 1088 (24.9%) | 1032 (24.6%) | 44 (25.7%) | 360 (25.6%) |
| Respiratory disease | 547 (12.5%) | 537 (12.8%) | 20 (11.7%) | 178 (12.7%) |
| Cardiovascular disease | 572 (13.1%) | 514 (12.3%) | 24 (14.0%) | 192 (13.7%) |
| Diabetes | 99 (2.3%) | 85 (2%) | 3 (1.8%) | 36 (2.6%) |
| Healthcare worker status |  |  |  |  |
| Non-healthcare worker | 1652 (37.8%) | 1456 (34.7%) | 65 (38.0%) | 597 (42.5%) |
| Healthcare worker facing no more than 1 COVID patient per day | 1904 (43.6%) | 1938 (46.2%) | 74 (43.3%) | 587 (41.8%) |
| Healthcare worker facing at least 1 COVID patient per day | 816 (18.7%) | 800 (19.1%) | 32 (18.7%) | 220 (15.7%) |
| Baseline risk probabilities^ |  |  |  |  |
| Mean (SD) | 0.0786 (0.0303) | 0.0794 (0.0296) | 0.0824 (0.0283) | 0.0774 (0.0306) |
| Dosage schedule |  |  |  |  |
| LD/LD | 125 (2.9%) | 69 (1.6%) | 7 (4.1%) | 114 (8.1%) |
| LD/SD | 1420 (32.5%) | 1361 (32.5%) | 46 (26.9%) | 320 (22.8%) |
| SD/SD | 2827 (64.7%) | 2764 (65.9%) | 118 (69%) | 970 (69.1%) |
| Prime boost interval |  |  |  |  |
| <6 weeks | 1078 (24.7%) | 931 (22.2%) | 28 (16.4%) | 456 (32.5%) |
| 6-8 weeks | 538 (12.3%) | 478 (11.4%) | 43 (25.1%) | 197 (14%) |
| 9-11 weeks | 1158 (26.5%) | 1236 (29.5%) | 42 (24.6%) | 398 (28.3%) |
| ≥12 weeks | 1598 (36.6%) | 1549 (36.9%) | 58 (33.9%) | 353 (25.1%) |
| Length of follow-up (days) from 7 days post PB28 until infection occurred or Feb 28 2021 (median (IQR)) | 88 (64, 113) | 85 (62, 108) | 53 (29, 81) | 105 (81, 135) |
| NAAT+ cases | 174 | 333 | 171 |  |
| Symptomatic | 55 (31.6%) | 196 (58.9%) | 54 (31.6%) |  |
| Asymptomatic | 99 (56.9%) | 112 (33.6%) | 97 (56.7%) |  |
| Non-primary symptomatic | 20 (11.5%) | 25 (7.5%) | 20 (11.7%) |  |

PB28: Study visit occurring 28 days after the second dose of vaccine; NAAT: nucleic acid amplification test; BMI: body mass index; LD: low dose; SD: standard dose

\* The correlates cohort is a subset of all eligible participants in the ChAdOx1 nCoV-19 correlates populations who have samples processed for at least one assay.

^ The baseline risk exposure score summarises predicted probability of having NAAT+ outcome from the risk model developed using the MenACWY Control Population

† Options included in ‘Other’ are as follows: ‘Mixed’, ‘Other – Free text’, or ‘prefers not to give’

**Table 2. Outputs from generalised additive models, with immune marker values associated with 50%, 60%, 70%, 80% and 90% vaccine efficacy against symptomatic infection**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Assay Units | p value immune marker | p value baseline risk score | No.  case | No.  Non case | 50% VE (95% CI) | 60% VE (95% CI) | 70% VE (95% CI) | 80% VE (95% CI) | 90% VE (95% CI) |
| Anti-Spike IgG | | AU/ml | 0.003 | <0.001 | 52 | 1155 | 4446 (NC, 12822) | 8413  (NC, 22232) | 17538  (NC, 37929) | 40923  (16748, 125017) | 139306  (57276, NC) |
| BAU/ml |  |  |  |  | 29  (NC, 83) | 54  (NC, 143) | 113  (NC, 245) | 264  (108, 806) | 899  (369, NC) |
| Anti-RBD IgG | | AU/ml | 0.018 | <0.001 | 52 | 1155 | 2193  (NC, 13614) | 6266  (NC, 29105) | 20700  (NC, 56620) | 63383  (16903, NC) | 295781  (90567, NC) |
| BAU/ml |  |  |  |  | 17  (NC, 109) | 50  (NC, 232) | 165  (NC, 452) | 506  (135, NC) | 2360  (723, NC) |
| Normalised live virus neutralisation assay | | NF50 | <0.001 | <0.001 | 36 | 412 | 68  (NC, 129) | 91  (NC, 175) | 135  (48, 267) | 247  (101, NC) | 938  (294, NC) |
| Pseudovirus neutralisation assay | | ID50 | 0.005 | <0.001 | 47 | 828 | NC | 22  (NC, 76) | 57  (NC, 183) | 185  (NC, NC) | 982  (303, NC) |
|  | IU/ml |  |  |  |  | NC | 3  (NC, 11) | 8  (NC, 26) | 26  (NC, NC) | 140  (43, NC) | |

ID50: neutralisation dilution for 50% virus inhibition; NF50: Normalised neutralisation titres; NC: not computed; IgG: Immunoglobulin G; RBD: receptor binding domain, VE: vaccine efficacy, CI: confidence interval. AU/ml: arbitrary units per millilitre; BAU/ml: Binding antibody units (WHO international standard 20/136), IU/ml: international units per millilitre (WHO international standard 20/136).

Where CIs were outside the range of values of the assay the limits are reported as ‘not computed’ (NC). Vaccine efficacy estimates and confidence intervals are those shown in Figure 4, at every 10% increment in the y axis. The two-sided p value for each immune marker (column 3) is from the generalised additive models in Figure 1, showing the strength of the relationship between the antibody value and infection. The p-values were not adjusted for multiple comparisons.

**Figure Legends**

**Figure 1: Predicted absolute risk of primary symptomatic COVID-19 as a function of immune markers measured 28 days post second dose by generalised additive regression.**

Predicted absolute risk of primary symptomatic COVID-19 as a function of:

a: Anti-spike IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

b: Anti-RBD IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

c: Pseudovirus neutralisation antibody titres 28 days post boost (47 cases, 828 non-cases included in the analysis)

d: Live virus neutralisation antibody titres 28 days post boost (36 cases, 412 non-cases included in the analysis).

Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19) .

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**Figure 2: Relative risk of primary symptomatic COVID-19 among vaccine recipients compared with MenACWY control arm participants as a function of immune markers measured at day 28 post-second dose**

Results are shown for:

a: Anti-spike IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

b: Anti-RBD IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

c: Pseudovirus neutralisation antibody titres 28 days post boost (47 cases, 828 non-cases included in the analysis)

d: Live virus neutralisation antibody titres 28 days post boost (36 cases, 412 non-cases included in the analysis).

Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants, derived by dividing the output curve from Figure 1 by the overall risk of infection in the MenACWY control group. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are 95% bootstrapped confidence intervals for the relative risk. The arrows point to the immune marker values at 0.20 and 0.50 relative risk, i.e., 80% and 50% vaccine efficacy for illustrative purpose. The full range of VE estimates from 50 to 90% are shown in Table 2.

**Figure 3 Relative risk of asymptomatic SARS-CoV-2 infection among vaccine recipients compared with the MenACWY control arm participants as a function of immune markers measured at 28 days post second dose**

Results are shown for:

a: Anti-spike IgG measured at 28 days post boost (91 cases, 1155 non-cases included in the analysis)

b: Anti-RBD IgG measured at 28 days post boost (91 cases, 1155 non-cases included in the analysis)

c: Pseudovirus neutralisation antibody titres 28 days post boost (86 cases, 828 non-cases included in the analysis)

d: Live virus neutralisation antibody titres 28 days post boost (62 cases, 412 non-cases included in the analysis).

Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are bootstrapped 95% confidence intervals.

**Figure 4: Vaccine efficacy against primary symptomatic COVID-19 as a function of immune markers measured at day 28 post-second dose**

Results are shown for:

a: Anti-spike IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

b: Anti-RBD IgG measured at 28 days post boost (52 cases, 1155 non-cases included in the analysis)

c: Pseudovirus neutralisation antibody titres 28 days post boost (47 cases, 828 non-cases included in the analysis)

d: Live virus neutralisation antibody titres 28 days post boost (36 cases, 412 non-cases included in the analysis).

Blue shaded areas represent the immune marker density distribution. Green lines show the vaccine efficacy and green dotted lines are 95% bootstrapped confidence intervals for vaccine efficacy. Vaccine efficacy is computed as 1 minus the relative risks shown in Figure 2. These results are also shown in Table 2 at 10% increments on the y axis.

**Extended data figure legends**

**Extended data Figure 1. Participant flow chart showing inclusion in correlates models.** Eligible participants comprised the Correlates Population and those with samples processed comprised the Correlates Cohort.

**Extended data Figure 2. Immune markers measured at day 28 post-second dose, in primary symptomatic, asymptomatic, non-primary cases, NAAT positive cases and NAAT negative non-cases**

a: N = 1155 NAAT negative, 52 primary, 91 asymptomatic and 20 non-primary, b: N = 163 NAAT positive participants’ anti-spike IgG measured at 28 days post boost; c: N = 1155 negative, 52 primary, 91 asymptomatic and 20 non-primary, d: N = 163 NAAT positive participants’ anti-RBD IgG measured at 28 days post boost; e: N = 828 NAAT negative, 47 primary, 86 asymptomatic and 16 non-primary, f: N = 149 NAAT positive participants’ pseudovirus neutralisation titre measured at 28 days post boost; g: N = 412 negative, 36 primary, 62 asymptomatic and 12 non-primary, h: N = 110 NAAT positive participants’ live neutralisation titre measured at 28 days post boost

a-h: minima: smallest value; maxima: largest value; centre: median value; bounds of box: 25% and 75% quartile value; upper/lower whisker extends from the hinge to the largest/smallest value no further than 1.5 \* inter-quartile range from the hinge.

IgG: Immunoglobulin G; RBD: receptor binding domain.

Primary symptomatic cases: NAAT+ with at least one COVID symptom (cough, fever, shortness of breath, anosmia, aguesia). Asymptomatic cases: NAAT+ on weekly self-swab with no symptoms recorded. Non-primary cases: NAAT+ with only non-primary COVID symptoms (e.g. nausea, diarrhoea). P-value estimated by one-way ANOVA test comparing between primary, asymptomatic, non-primary cases and NAAT negative non-cases and by two sample t-test comparing between NAAT positive cases and NAAT negative non-cases (two-sided).

**Extended data Figure 3. Correlations between a, Anti-SARS-CoV-2 spike and RBD IgG. b, Anti-SARS-CoV-2 Spike IgG and pseudovirus neutralisation titre. c, Anti-SARS-CoV-2 Spike IgG and live virus neutralisation titre. d, pseudovirus neutralisation titres and live virus neutralisation titres.**

95% confidence ellipses assuming a t-distribution are shown for each outcome (primary symptomatic cases, asymptomatic cases and negative controls). Pearson correlation coefficients shown as r values using all available data. Primary symptomatic cases: NAAT+ with at least one COVID symptom (cough, fever, shortness of breath, anosmia, aguesia). Asymptomatic cases: NAAT+ on weekly self-swab with no symptoms recorded.

**Extended data Figure 4. Predicted absolute risk of asymptomatic SARS-CoV-2 infection as a function of immune markers measured 28 days post second dose.**

Predicted absolute risk of asymptomatic infection as a function of:

a: Anti-spike IgG measured at 28 days post boost (91 cases, 1155 non-cases included in the analysis)

b: Anti-RBD IgG measured at 28 days post boost (91 cases, 1155 non-cases included in the analysis)

c: Pseudovirus neutralisation antibody titres 28 days post boost (86 cases, 828 non-cases included in the analysis)

d: Live virus neutralisation antibody titres 28 days post boost (62 cases, 412 non-cases included in the analysis).

Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19).

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**Extended data Figure 5. Sensitivity analysis showing absolute and relative risk of asymptomatic SARS-CoV-2 infection as a function of immune markers measured at 28 days post second dose excluding cases with low viral load (Ct ≥ 30)**

Results are shown for:

a and b: Anti-spike IgG measured at 28 days post boost (39 cases, 1207 non-cases included in the analysis)

c and d: Anti-RBD IgG measured at 28 days post boost (39 cases, 1207 non-cases included in the analysis)

e and f: Pseudovirus neutralisation antibody titres 28 days post boost (37 cases, 877 non-cases included in the analysis)

g and h: Live virus neutralisation antibody titres 28 days post boost (29 cases, 445 non-cases included in the analysis).

**a, c, e, g**: Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19).

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**b, d, f, h**: Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are 95% bootstrapped confidence intervals for the relative risk.

**Extended data Figure 6. Sensitivity analysis showing absolute and relative risk of primary symptomatic SARS-CoV-2 infection in participants with symptoms of shortness of breath as a function of immune markers measured at day 28 post-second dose**

Results are shown for:

a and b: Anti-spike IgG measured at 28 days post boost (28 cases, 1155 non-cases included in the analysis)

c and d: Anti-RBD IgG measured at 28 days post boost (28 cases, 1155 non-cases included in the analysis)

e and f: Pseudovirus neutralisation antibody titres 28 days post boost (27 cases, 828 non-cases included in the analysis)

g and h: Live virus neutralisation antibody titres 28 days post boost (22 cases, 412 non-cases included in the analysis).

**a, c, e, g**: Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19).

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**b, d, f, h**: Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are 95% bootstrapped confidence intervals for the relative risk.

**Extended data Figure 7. Sensitivity analysis showing absolute and relative risk of primary symptomatic SARS-CoV-2 infection in participants with no symptoms of shortness of breath as a function of immune markers measured at day 28 post-second dose**

Results are shown for:

a and b: Anti-spike IgG measured at 28 days post boost (24 cases, 1155 non-cases included in the analysis)

c and d: Anti-RBD IgG measured at 28 days post boost (24 cases, 1155 non-cases included in the analysis)

e and f: Pseudovirus neutralisation antibody titres 28 days post boost (20 cases, 828 non-cases included in the analysis)

g and h: Live virus neutralisation antibody titres 28 days post boost (14 cases, 412 non-cases included in the analysis).

**a, c, e, g**: Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19).

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**b, d, f, h**: Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are 95% bootstrapped confidence intervals for the relative risk.

**Extended data Figure 8. Sensitivity analysis showing absolute and relative risk primary symptomatic SARS-CoV-2 infection with 3 or more COVID-19 symptoms as a function of immune markers measured at day 28 post-second dose**

Results are shown for:

a and b: Anti-spike IgG measured at 28 days post boost (32 cases, 1155 non-cases included in the analysis)

c and d: Anti-RBD IgG measured at 28 days post boost (32 cases, 1155 non-cases included in the analysis)

e and f: Pseudovirus neutralisation antibody titres 28 days post boost (28 cases, 828 non-cases included in the analysis)

g and h: Live virus neutralisation antibody titres 28 days post boost (21 cases, 412 non-cases included in the analysis).

**a, c, e, g**: Grey horizontal lines show the overall risk of primary symptomatic COVID-19 in the control group (MenACWY) and vaccine groups (ChAdOx1 nCoV-19).

Blue dots show the absolute risk predicted from the model across the range of antibody values included in the analysis, adjusting for baseline exposure risk to SARS-CoV-2 infection. Green shaded areas show the confidence interval around the predicted mean probability (green line)

**b, d, f, h**: Blue shaded areas represent the immune marker density distribution. Green lines show the relative risk of infection among vaccine recipients compared to the MenACWY control arm participants. The green lines are the median relative risk obtained from 10,000 bootstrap samples. Green shaded areas are 95% bootstrapped confidence intervals for the relative risk.

# References:

1. World Health Organisation. The COVID-19 candidate vaccine landscape. <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines> Last accessed on 13 May 2021.

2. Voysey, M.*, et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99-111 (2021).

3. Voysey, M.*, et al.* Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. *Lancet* **397**, 881-891 (2021).

4. Moderna. Moderna’s COVID-19 Vaccine Candidate Meets its Primary Efficacy Endpoint in the First Interim Analysis of the Phase 3 COVE Study. <https://investors.modernatx.com/news-releases/news-release-details/modernas-covid-19-vaccine-candidate-meets-its-primary-efficacy> Last accessed on 17 Nov 2020.

5. Polack, F.P.*, et al.* Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *New England Journal of Medicine* **383**, 2603-2615 (2020).

6. Centers for Disease Control and Prevention. Johnson & Johnson’s Janssen COVID-19 Vaccine Overview and Safety. <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/janssen.html> Last accessed on 13 May 2021.

7. Al Kaabi, N.*, et al.* Effect of 2 Inactivated SARS-CoV-2 Vaccines on Symptomatic COVID-19 Infection in Adults: A Randomized Clinical Trial. *JAMA* **326**, 35-45 (2021).

8. Tenforde, M.W.*, et al.* Effectiveness of Pfizer-BioNTech and Moderna Vaccines Against COVID-19 Among Hospitalized Adults Aged ≥65 Years - United States, January-March 2021. *MMWR Morb Mortal Wkly Rep* **70**, 674-679 (2021).

9. Haas, E.J.*, et al.* 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. *Lancet* **397**, 1819-1829 (2021).

10. Vasileiou, E.*, et al.* Effectiveness of first dose of COVID-19 vaccines against hospital admissions in Scotland: national prospective cohort study of 5.4 million people. in *SSRN* (2021).

11. Public Health England. COVID-19 vaccine surveillance report Week 20. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/988193/Vaccine_surveillance_report_-_week_20.pdf> Last accessed on 24 May 2021.

12. Lopez Bernal, J.*, et al.* Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ* **373**, n1088 (2021).

13. Pritchard, E.*, et al.* Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. *Nature Medicine* **27**, 1370-1378 (2021).

14. Folegatti, P.M.*, et al.* Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *Lancet* **396**, 467-478 (2020).

15. Emary, K.R.W.*, et al.* Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. *Lancet* **397**, 1351-1362 (2021).

16. Ewer, K.J.*, et al.* T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial. *Nat Med* **27**, 270–278 (2020).

17. Callow, K.A. Effect of specific humoral immunity and some non-specific factors on resistance of volunteers to respiratory coronavirus infection. *J Hyg (Lond)* **95**, 173-189 (1985).

18. Edridge, A.W.D.*, et al.* Seasonal coronavirus protective immunity is short-lasting. *Nat Med* **26**, 1691-1693 (2020).

19. Addetia, A.*, et al.* Neutralizing Antibodies Correlate with Protection from SARS-CoV-2 in Humans during a Fishery Vessel Outbreak with a High Attack Rate. *J Clin Microbiol* **58**, e02107-02120 (2020).

20. Lumley, S.F.*, et al.* Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. *N Engl J Med* **384**, 533-540 (2021).

21. Yu, J.*, et al.* DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* **369**, 806-811 (2020).

22. Mercado, N.B.*, et al.* Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583-588 (2020).

23. Baum, A.*, et al.* REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science* **370**, 1110-1115 (2020).

24. Khoury, D.S.*, et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nature Medicine* **27**, 1205-1211 (2021).

25. Qin, L., Gilbert, P.B., Corey, L., McElrath, M.J. & Self, S.G. A Framework for Assessing Immunological Correlates of Protection in Vaccine Trials. *The Journal of Infectious Diseases* **196**, 1304-1312 (2007).

26. Team UC-19 RT / CPN (CoVPN) B, Gilbert PB, Fong Y, et al. USG COVID-19 Response Team / CoVPN Vaccine Efficacy Trial Immune Correlates Statistical Analysis Plan [Internet]. Available from: <https://figshare.com/articles/online_resource/CoVPN_OWS_COVID-19_Vaccine_Efficacy_Trial_Immune_Correlates_SAP/13198595/11> Last accessed on 2021 May 17.

27. Gilbert, P.B.*, et al.* Immune Correlates Analysis of the mRNA-1273 COVID-19 Vaccine Efficacy Trial. *medRxiv*, 2021.2008.2009.21261290 (2021).

28. van Doremalen, N.*, et al.* ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **586**, 578-582 (2020).

29. Fischer, R.J.*, et al.* ChAdOx1 nCoV-19 (AZD1222) protects hamsters against SARS-CoV-2 B.1.351 and B.1.1.7 disease. *bioRxiv* (2021).

30. Wang, Z.*, et al.* Enhanced SARS-CoV-2 neutralization by dimeric IgA. *Sci Transl Med* **13**, eabf1555 (2021).

31. AZD1222 US Phase III trial met primary efficacy endpoint in preventing COVID-19 at interim analysis. Available from: <https://www.astrazeneca.com/media-centre/press-releases/2021/astrazeneca-us-vaccine-trial-met-primary-endpoint.html> Last accessed on 13 May 2021.

32. Walsh, E. & Falsey, A.R. Humoral and Mucosal Immunity in Protection from Natural Respiratory Syncytial Virus Infection in Adults. *The Journal of Infectious Diseases* **190**, 373-378 (2004).

33. Mulligan, M.J.*, et al.* Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature* **586**, 589-593 (2020).

34. Barrett, J.R.*, et al.* Phase 1/2 trial of SARS-CoV-2 vaccine ChAdOx1 nCoV-19 with a booster dose induces multifunctional antibody responses. *Nat Med* **27**, 279-288 (2020).

35. McMahan, K.*, et al.* Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630-634 (2021).

36. Yang, P.-H.*, et al.* Increased circulating level of interleukin-6 and CD8+ T cell exhaustion are associated with progression of COVID-19. *Infectious Diseases of Poverty* **9**, 161 (2020).

37. Urra, J.M., Cabrera, C.M., Porras, L. & Ródenas, I. Selective CD8 cell reduction by SARS-CoV-2 is associated with a worse prognosis and systemic inflammation in COVID-19 patients. *Clin Immunol* **217**, 108486-108486 (2020).

38. Ramasamy, M.N.*, et al.* Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. *The Lancet* **396**, 1979-1993 (2020).

39. Sterne, J.A.C.*, et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393 (2009).

40. World Health Organization - National Institute for Biological Standards and Control. (17-December-2020). WHO International Standard First WHO International Standard for anti-SARS-CoV-2 immunoglobulin (human) – NIBSC code: 20/136: Instructions for Use, v2.0. Potters Bar, Hertfordshire, EN6 3QG: WHO International Laboratory for Biological Standards, UK Official Medicines Control Laboratory.

41. Lubin, J.H.*, et al.* Epidemiologic evaluation of measurement data in the presence of detection limits. *Environmental health perspectives* **112**, 1691-1696 (2004).

42. Chen, H., Quandt, S.A., Grzywacz, J.G. & Arcury, T.A. A Bayesian multiple imputation method for handling longitudinal pesticide data with values below the limit of detection. *Environmetrics* **24**, 132-142 (2013).

43. Brand, J., van Buuren, S., le Cessie, S. & van den Hout, W. Combining multiple imputation and bootstrap in the analysis of cost‐effectiveness trial data. *Statistics in medicine* **38**, 210-220 (2019).

44. Box, G.E. & Tiao, G.C. *Bayesian inference in statistical analysis*, (John Wiley & Sons, 2011).

45. R Core Team. R: A language and environment for statistical computing. in *R Foundation for Statistical Computing.* (2020).

46. Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**, 3-36 (2011).

# **Online Methods**

## **Study description**

The data included in this analysis comes from participants enrolled in COV002 (registration NCT04400838), a phase 2/3 randomised single blind vaccine efficacy trial conducted across 19 sites in the UK. A full description of the trial including immunogenicity, efficacy, and safety data, and the protocol has been previously published.2,3,14,15,38

This study was approved in the UK by the Medicines and Healthcare products Regulatory Agency (MHRA), reference 21584/0428/001 0001, and the South-Central Berkshire Research Ethics Committee, reference 20/SC/0179. All participants provided informed consent.

Briefly, participants in the study were randomised to receive ChAdOx1 nCoV-19 or a MenACWY control vaccine. The randomisation ratio (ChAdOx1 nCoV-19:MenACWY) differed by study cohort, and was either 1:1, 5:1, or 3:1. (see CONSORT diagram, Extended Data Fig.1). Open label groups and are not included in this report.

## **Study endpoints and outcomes**

Participants were reminded weekly to contact their study site if they experienced any of the primary symptoms of COVID-19 (fever ≥ 37.8oC; cough; shortness of breath; anosmia or ageusia) and were assessed in clinic, with a nose and throat swab taken for nucleic acid amplification testing (NAAT). Additionally, participants were asked to complete a nose and throat swab at home each week.

The outcomes for this analysis were 1) primary symptomatic COVID-19: a NAAT+ swab with at least one qualifying symptom, and 2) asymptomatic infections identified from weekly self-administered swabs and defined as a NAAT+ swab with no symptom reported. Sensitivity analysis of asymptomatic infections removed potential false-positive cases by restricting to those with higher viral load (Ct value < 30). NAAT+ participants who had symptoms other than the main five COVID-19 symptoms were categorised as non-primary symptomatic and were not included in correlates analysis.

Primary symptomatic COVID-19 outcomes were further classified according to whether a symptomatic participant reported shortness of breath or not, and whether 3 or more COVID-19 symptoms among 5 were present, indicators of more severe disease.

All endpoints were evaluated by a blinded independent clinical review committee.

## **Immune markers and time points**

A proportion of serum samples from vaccine recipients at the 28-day post-boost visit (PB28) were tested on three different assays with four assay readouts. All NAAT+ cases were tested if sample volume allowed, while a proportion of non-cases were tested. Samples were tested blinded to case status. The data from non-cases was obtained first, and consisted mainly of the samples processed for the initial application for emergency use which needed 15% of samples included in the efficacy cohort to be processed on validated assays. Subsequent to this NAAT+ cases were sent for testing as they occurred, if not already including the 15%. We assume the mechanism of missingness for samples that were not tested to be missing at random.39 To account for the missing data, factors associated with sample availability were controlled as weights in the analysis (see Correlates of risk and Inverse probability weighting below).

Anti-SARS-CoV-2 Spike and RBD IgG were measured by a multiplex immunoassay on the MSD platform at PPD. The assay sequences were based on the ancestral Wuhan sequences. Antigen information and sequence information are provided in Table S1. Assay validation included precision and ruggedness, dilutional linearity, selectivity, and relative accuracy for each SARS-CoV-2 antigens. Post validation studies for stability and for conversion to the WHO standard, as well as the establishment of a cut-point were performed. The LLOQs for anti-spike and anti-RBD are 33 and 204 AU/ml, respectively.

Antibody neutralisation was measured with a lentivirus-based pseudovirus particle expressing the D614 SARS-CoV-2 spike protein. The pseudovirus neutralizing antibody assay was validated at Monogram Biosciences (South San Francisco, CA, USA). Validation included accuracy, repeatability, intermediate precision, linearity, specificity/selectivity, sensitivity, and stability utilising pooled sera from high-titre, intermediate-titre, and low-titre pooled convalescent SARS-CoV-2 sera, as well as historical negative samples collected in the year 2017 (prior to SARS-CoV-2 circulation). The LLOQ for pseudovirus neutralizing antibody is 40 (ID50).

Antibody neutralisation was also measured by a live microneutralisation assay using the Victoria/01/2020 strain of the virus (Public Health England). Qualification of the assay included assessment of specificity, parallelism, dilutional linearity, repeatability, intermediate precision and assessment of the assay range. A formal validation has since been completed (after the testing of clinical study samples in this manuscript). Normalised values (NF50) were used for the main analyses as the normalisation process removes the plate to plate variability and normalised values are more highly correlated with binding antibody and pseudovirus neutralisation assays. However, normalised values cannot be converted into WHO standard units. A sensitivity analysis is provided in the Table S4 using non-normalised values (ND50) which are also presented as IU/ml using the WHO standard, but are less highly correlated with other assays. The LLOQ of the assay is 58 (ND50) and 8.6 (NF50).

Due to the limitations of laboratory capacity fewer samples were tested for virus neutralisation than were tested using the quicker multiplex assay.

### **Imputation on censored immune marker data in main analysis**

Immune marker values were log10-transformed prior to analysis. Values which were censored at the lower limit of quantification (LLOQ) were imputed with the value LLOQ/2. Approximately 10% of the pseudovirus neutralisation titre were censored at the LLOQ, and sensitivity analysis were conducted by imputing these values using a Gibbs sampler.

**Conversion to WHO International Standard (20/136)**

Each assay was analysed in its original scale and results were then converted to the WHO international standard units using the conversion factors supplied by each laboratory. WHO standard units are Binding Antibody Units per millilitre (BAU/ml) for anti-spike and anti-RBD IgG, and International Units per millilitre (IU/ml) for neutralisation titres.40 For PPD conversion factors are supplied with confidence intervals. These are not able to be applied to the converted data as it is a one-to-one conversion. For the Monogram assay multiple forms of the conversion factor were supplied and all three are implemented.

Conversion factors were as follows:

PPD: Conversion from Arbitrary Units (AU/ml) per ml to Binding Antibody Units per ml (BAU/ml)

* Anti-spike IgG 0.00645, 95% CI (0.00594, 0.00701)
* Anti-RBD IgG 0.00798, 95% CI (0.00735, 0.00866)

Monogram pseudovirus neutralisation assay (D614) conversion from ND50 to international units per ml (IU/ml)

* 0.1428 (mean
* 0.1458 (Geo.mean)
* 0.1534 (median)

PHE live virus neutralisation assay conversion from ID50 to IU/mL

* 0.2461 (1/4.064)

## **Study design and analysis populations**

We first defined the Correlates Population by restricting to participants who met the eligibility criteria and received ChAdOx1 nCoV-19: participants were eligible for inclusion if they were baseline seronegative to the SARS-CoV-2 N protein at first vaccination, had their PB28 visit within a 14 to 42 day window after the second dose, and were followed up to at least 7 days after PB28 with no prior evidence of infection. Participants were excluded from analysis if infection occurred before PB28. Participants who received two doses were included in the analysis, either standard dose followed by standard dose (SDSD), or low dose followed by low or standard dose (LDSD or LDLD). 9 participants who received mixed schedules (one dose of ChAdOx1 nCoV-19 and one dose of MenACWY control) in error were excluded from analysis (Extended Data Fig.1). The same eligibility criteria were applied to define a Control Population of MenACWY recipients.

Among the ChAdOx1 nCoV-19 Correlates Population, those who had biomarker data available comprised the Correlates Cohort. Participants who tested NAAT positive more than 7 days after PB28 were defined as cases while those who did not have a positive test were defined as non-cases. The 7 day window was implemented to exclude cases in which exposure is likely to have occurred before a blood sample was taken.

## **Statistical Analysis**

## **Baseline exposure risk to SARS-CoV-2 infections**

To control for potential confounding due to variation in exposure risk among participants with available immune marker data, a logistic regression risk model was developed among the Control Population of MenACWY recipients. Baseline factors associated with exposure risk were used to model the probability of being NAAT positive in this population. Baseline variables for the risk model included age in years, ethnicity (white and non-white), BMI (<30 kg/m2, ≥30 kg/m2), co-morbidities (having any of: respiratory disease; cardiovascular disease; or diabetes) and healthcare worker status (non-healthcare worker, healthcare worker exposed to no more than 1 COVID patient on an average day, healthcare worker exposed to 1 or more COVID patients on an average day). Output is shown in Table S3. The linear predictor from the risk model developed using the MenACWY Control Population was used to predict the baseline risk of exposure in the ChAdOx1 nCoV-19 Correlates Cohort.

## **Correlates or risk (CoR)**

The CoR analysis was conducted within the Correlates Cohort. Log-transformed immune marker values were analysed using generalised additive models (GAM) for binary data with a cubic spline smooth applied to immune marker values to allow a non-linear effect. The logit-transformed predicted baseline exposure risk was included as a linear covariate in the GAM model. A p value <0.05 from the approximate significance test from the smooth GAM was used to determine if an immune marker was associated with protection. There was no adjustment for multiple comparison. Separate models were fitted for each immune marker controlling for baseline exposure risk, and weighted by inverse probability weights as described below.

## **Inverse probability weighting**

Immune marker data were not available for everyone in the correlates population, and cases are over-represented in the immune marker datasets as these were preferentially processed over non-cases. Unadjusted estimates of absolute risk of infection will therefore be inflated and result in bias to correlates estimates. We used a logistic regression model to predict the probability that a participant will have immune marker data available to the analysis. The outcome variables were each immune marker, and predictors were age group (18-55 years, 56-69 years, 70 years or above), whether the participant is a case or non-case, the type of case (primary symptomatic, non-primary symptomatic, asymptomatic), prime-boost interval, and dosage (LD/LD, LD/SD, SD/SD). The inverse probability from this model was used to weight the correlates of risk models for each immune marker to remove this source of bias. (Table S3)

## **Correlates of vaccine efficacy (CoVE)**

For each outcome, to derive the relative risk (RR) and CoVE, an estimate of the absolute averaged predicted risk from the CoR model was computed. The averaged absolute risk was then compared to the overall risk among MenACWY Correlates Population, which was itself weighted by the randomisation ratio for study groups not randomised 1:1.

Vaccine efficacy (VE) was defined as 100% x (1 – RR). Mean estimate of VE at each level of antibody in the dataset, as well as 95% confidence intervals were calculated from 10,000 bootstrap samples.

Further analysis details are provided with the original trial statistical analysis plan (SAP) and the separate SAP developed for immune correlates analyses. The immune correlates SAP leant heavily on the methods proposed in the publicly available SAP by the Coronavirus Prevention Network (CoVPN) Biostatistics Team.26

## **Bootstrap**

We resampled from all participants enrolled in the study. For each bootstrap sample, we calculated the inverse probability weights to account for sampling bias. We then estimated the CoR by GAM, adjusting for the baseline risk exposure and weighting by inverse probability weights. We compared the predicted absolute risk from the GAM across the full range of antibody values, with the resampled MenACWY control population weighted overall risk. 10,000 bootstrap samples were used for each immune marker and outcome. The overall estimates for correlates of risk and correlates of vaccine efficacy were given by the median value in the bootstrap. 95% confidence intervals were calculated using the bootstrap percentile method, i.e., the 2.5% and 97.5% quantiles from the bootstrap.

Correlates and their CIs were not computed for assays in which the relationship between antibody and outcome was non-significant. Where CIs were outside the range of values of the assay these are reported as ‘not computed’ (NC).

## **Sensitivity analyses**

### **Viral load**

To account for potential of misclassification in asymptomatic infections, a sensitivity analysis was conducted excluding cases with lower viral loads (defined as those for whom all returned PCR positive tests had a Ct value ≥ 30) as these are potential false positives.

### **Imputation of censored antibody values**

Approximately 10% of the pseudovirus neutralisation antibody titre were below the LLOQ. We performed a sensitivity analysis to account for the potential bias caused by imputing LLOQ/2. Studies have shown that imputing LLOQ/2 can lead to bias and confidence intervals with poor coverage when a significant proportion of the data are censored41-43. When a bootstrap is required for missing data, Brand et al. (2019) found single imputation embedded inside a bootstrap showed better statistical properties than other methods43. We used an iterative Gibbs sampler proposed by Chen et al. (2013) to impute the censored log pseudovirus neutralisation antibody values42.

Not all participants with results from the pseudovirus neutralisation titre also have results from the anti-spike, anti-RBD and live neutralising antibody titres. For each bootstrap sample, we iteratively predicted the missing and censored values for each antibody titre in a Gibbs sampler, constraining the predictions for the censored values to be less than or equal to the LLOQ. The antibody titres were iteratively predicted by a sequence of Bayesian linear regressions. For each regression, the independent variables were the current prediction for all other titres, the baseline risk score and all variables used in the inverse probability weighting model.

Let , j = 1, 2, 3, 4 be the vector of the jth antibody titre values. Let and be the variance and vector of regression parameters for the jth linear regression update respectively. We chose a non-informative prior42,44 for and , namely

Then the Gibbs sampler proposed by Chen et al. (2013) is as follows42:

Initialise the missing and censored values for each , j = 1, 2, 3, 4.

For i = 1, … , N

For j = 1, 2, 3, 4

Update and from the posterior distribution given the current predictors for all other antibody values , k < j; , k > j and the fixed covariates.

Update from the posterior predictive distribution given and and the current values of the predictor variables.

We imputed a single value for each of the censored log pseudovirus neutralisation antibody values from the N = 100th iteration of the Gibbs sampler. Note participants with missing log pseudovirus neutralisation antibody titre values were excluded from the sensitivity analysis. The sensitivity analysis was then run on the imputed dataset for the bootstrap sample.

We initialised the Gibbs sampler by predicting the missing and censored values from a sequence of linear regressions on the non-missing data. This sequence was developed with the data structure in mind, aiming to initialise the chain as close to the posterior mode as possible.

We ran multiple chains on bootstrap samples and tested for convergence by inspecting trace plots of the censored log pseudovirus neutralisation titres. From these plots we determined the 100th iteration to be approximately converged.

## **Data cut-off**

The data cut-off date for inclusion in this analysis was Feb 28, 2021. Cases occurring after this date are not included in the analysis.

## **Software**

Data analysis was done using R version 3.6.1 .45 The GAM was coded using the mgcv package.46 Three knots were used for each GAM, and the smoothing parameter was estimated by generalized cross validation.

## **Data availability statement**

Anonymised participant data will be made available when the trials are complete, upon requests directed to the corresponding author. Proposals will be reviewed and approved by the sponsor, investigator, and collaborators on the basis of scientific merit. After approval of a proposal, data can be shared through a secure online platform after signing a data access agreement. All data will be made available for a minimum of 5 years from the end of the trial.

## **Code availability statement**

The R code for the main correlates estimates is available in supplementary file (item 5, page 8).