Antibody responses after SARS-CoV-2 infection and vaccination in the UK general population



Jia Wei

Nuffield Department of Medicine

Green Templeton College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity 2023

Abstract

Antibody responses after SARS-CoV-2 infection and vaccination in the UK general population

Jia Wei, Green Templeton College, a thesis submitted for the degree of Doctor of Philosophy in Clinical Medicine, Trinity 2023

Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has led to a global pandemic since its first identification in December 2019. In response to the pandemic, multiple vaccines have been developed, which showed good initial effectiveness, but different variants of concern caused by genetic mutations kept emerging and hampered vaccine protection. Assessment of the immunity acquired from both natural infection and vaccinations is important to optimize vaccination strategies and pandemic management.

This thesis aims to examine the nature and duration of antibody responses after SARS-CoV-2 infection and vaccination, analyse the association between antibody levels and protection (i.e. correlates of protection), and estimate the duration of protection following infection or vaccination in the UK general population.

Using data from the UK's COVID-19 Infection Survey (CIS), which is the largest household survey of COVID-19 infections and antibodies in the UK, I conducted five studies examining the antibody responses after Pre-Alpha and Alpha infections, after the first vaccination, after the second vaccination, and after the booster vaccination or breakthrough infection. Using latent class mixed models, I found that around 24% of people did not generate positive antibody response after infection, and 5-6% did not respond to the first vaccination, who were older and with more long-term health conditions. Using Bayesian linear mixed models, I estimated that the mRNA vaccines (BNT162b2 and mRNA-1273) elicited much higher antibody levels than the adenovirus vaccine (ChAdOx1) and natural infection, but antibodies after infection lasted longer. I also found that people with prior infection had much stronger antibody responses after the first and second vaccination. Using generalised additive models, I found that higher antibody levels were associated with a higher level of protection against (re)infections in the Delta and Omicron BA.4/5 epochs, and a hybrid immunity from both vaccination and infection led to much higher protection given the same antibody levels. Combining these results, I concluded that vaccine protection was relatively short-lived especially against Omicron infections, while natural infection induced longer protection.

This work has contributed to the understanding of immunogenicity and associated protection of SARS-CoV-2 natural infection and vaccination in the UK general population, which provides evidence to the UK's national vaccination strategies and policy making over the pandemic.

Table of Contents

Chapter 1 Preface *10*

1.1 Acknowledgement **10**

1.2 Funding **12**

1.3 Ethics statement **12**

1.4 Declaration of contribution to thesis **13**
1.4.1 Chapter 3 13
1.4.2 Chapter 4 13
1.4.3 Chapter 5 14
1.4.4 Chapter 6 14

1.5 Publications **15**
1.5.1 In relation to this thesis 15
1.5.2 Other papers during the course of this thesis 15

Chapter 2 Introduction and background *17*

2.1 The emergence of SARS-CoV-2 **17**

2.2 SARS-CoV-2 transmission and pathophysiology **18**

2.3 Coronavirus and SARS-CoV-2 **19**
2.3.1 Coronaviruses, SARS-CoV, and MERS-CoV 19
2.3.2 SARS-CoV-2 structure and variants of concern 21

2.4 SARS-CoV-2 vaccines **26**
2.4.1 Pfizer-BioNTech BNT162b2 vaccine 29
2.4.2 Moderna mRNA-1273 vaccine 30
2.4.3 Oxford-AstraZeneca ChAdOx1 vaccine 30

2.5 Vaccination programme in the UK **32**

2.6 Immunity against SARS-CoV-2 **34**

2.7 Thesis aims and outline **36**

Chapter 3 Thesis Methods *38*

3.1 UK's COVID-19 infection survey **38**

3.2 Vaccination data **40**

3.3 Laboratory data **40**
3.3.1 Antibody data 40
3.3.2 PCR data 44

3.3 SARS-CoV-2 definition **44**
3.3.1 Infection episode 44
3.3.2 Infection variant 45

3.4 Overall statistical methods **47**
3.4.1 Generalised additive models 47
3.4.2 Linear mixed models 50
3.4.3 Latent class mixed models 52
3.4.4 Bayesian statistics vs frequentist statistics 54
3.4.5 Bayesian linear mixed models 56

Chapter 4 Antibody responses to SARS-CoV-2 natural infection *62*

4.1 Introduction **62**

4.2 Methods **63**
4.2.1 Study population and data 63
4.2.2 Statistical analysis 65

4.3 Results **68**
4.3.1 Study population 68
4.3.2 Antibody trajectories following infection 72
4.3.3 Predictors of non-responders 76
4.3.4 Determinants of the peak and half-life of antibody responses 84
4.3.5 Duration of antibody responses and immune protection 94

4.4 Discussion **100**

Chapter 5 Antibody responses to first and second COVID-19 vaccination and correlates of protection against SARS-CoV-2 Delta variant *106*

5.1 Introduction **106**

5.2 Methods **109**
5.2.1 Population and data 109
5.2.2 Statistical analysis 115

5.3 Results **126**
5.3.1 Study population 126
5.3.2 Anti-spike IgG antibody positivity after vaccination and its predictors in the initial vaccination study (cohort 1) 133
5.3.3 Quantitative anti-spike IgG antibody responses after the first and second vaccination in the initial vaccination and second dose study (cohort 1 and 2) 145
5.3.4 Vaccine non-responders after the first and second vaccination and their predictors (cohort 1 and 2) 158
5.3.5 Determinants of anti-spike IgG peak levels and half-life after the first and second vaccination (cohort 2 and 3) 167
5.3.6 Antibody correlates of protection against infection in the Delta period (cohort 2) 187
5.3.7 Comparison of antibody trajectories and duration of protection (cohort 2 and 3) 191

5.4 Discussion **202**

Chapter 6 Correlates of protection against SARS-CoV-2 Omicron variant and antibody responses after booster vaccination or breakthrough infection *215*

6.1 Introduction **215**

6.2 Methods **217**
6.2.1 Correlates of protection analysis 217
6.2.2 Antibody trajectory analysis 223
6.2.3 Estimation of protection 233

6.3 Results **234**
6.3.1 Correlates of protection against Omicron BA.4/5 infection 234
6.3.2 Antibody trajectories after third/booster vaccination and breakthrough infection 241
6.3.3 Protection from the third/booster vaccination and breakthrough infection against new Omicron BA.4/5 infection 251

6.4 Discussion **256**

Chapter 7 Conclusions and future work *265*

Table of figures

Figure 2.1 Timeline of the key events of the COVID-19 outbreak in the first few months. 18
Figure 2.2 SARS-CoV-2 structure. 21
Figure 2.3 The life cycle of SARS-CoV-2 in the host cells. 22
Figure 2.4 Traditional and accelerated vaccine development pipelines. 27
Figure 2.5 Mechanism of action of mRNA vaccines. 29
Figure 2.6 Mechanism of action of ChAdOx1 vaccine 31
Figure 2.7 Priority groups for the UK COVID-19 vaccination programme. 33
Figure 2.8 Timeline of the key events in the UK COVID-19 vaccination programme. 34
Figure 3.1 Correlation between anti-spike IgG levels in mAb45 units (ng/ml) and WHO international units (BAU/mL) in 63 dilution replicates. 43
Figure 3.2 Gene positivity pattern by calendar time to determine Alpha, Delta, BA.1, BA.2, and BA.4/5 epochs. 46
Figure 3.3 Diagnostic plots for MCMC convergence. 60
Figure 4.1 Flowchart of the study cohort. 67
Figure 4.2 Individual trajectories for 7,256 participants infected with SARS-CoV-2 by class identified from latent class mixed models. 73
Figure 4.3 Distributions of continuous variables by classes identified from latent class mixed models. 74
Figure 4.4 Predicted mean anti-spike IgG trajectories in participants with natural SARS-CoV-2 infection from latent class mixed models in 7,256 participants (with standard error of the prediction). 75
Figure 4.5 Predicted probability of being in Class 1 (classical seroconversion in response to infection), 2 (possible late/reinfection), and 3 (seronegative non-responders). 77
Figure 4.6 Odds ratio with 95% confidence intervals from logistic regression comparing seronegative vs seroconverting (Class 3 vs Class 1) using demographic factors and individual symptoms that would be available without a positive test result. 79
Figure 4.7 Predicted probability of being in Class 3 (seronegative) compared with Class 1 (seroconverted) by age. 80
Figure 4.8 Determination of the ‘peak level’ for modelling anti-spike IgG trajectories. 85
Figure 4.9 Estimated mean trajectory of anti-spike IgG antibody levels and individual trajectories in 3,271 participants in Class 1. 86
Figure 4.10 Posterior estimations for individual anti-spike IgG half-lives and peak level 87
Figure 4.11 Posterior mean trajectories with 95% credibility intervals of anti-spike IgG after the peak level by age (a, b) and Ct values (c, d) in the multivariable model in 3,271 participants. 90
Figure 4.12 Posterior predictive check for the baseline model and the multivariable model. 92
Figure 4.13 MCMC trace plots for assessing convergence of chains. 93
Figure 4.14 Comparison of anti-spike IgG decline in 3,271 participants among the model assuming a linear decline in log2 scale (red line), the biphasic exponential model (blue line), and the model using splines (orange line). 95
Figure 4.15 Posterior predicted time (95% credibility interval) from the start of infection to three anti-spike IgG thresholds (23, 16, and 3 BAU/mL) by age (20, 40, 60, 80 years), sex, and ethnicity from the multivariable biphasic exponential model in 3271 participants. 96
Figure 4.16 Posterior predicted time (95% credibility interval) the start of infection to three anti-spike IgG thresholds (23, 16, and 3 BAU/mL) by age (20, 40, 60, 80 years), sex, and ethnicity from the multivariable linear model in 3271 participants 97

Figure 5.1 Diagram summarising the time-period, population, and analyses in three different papers in Chapter 5. 110
Figure 5.2 Flowchart of the study cohort. 112
Figure 5.3 Flowchart of the study cohort. 114
Figure 5.4 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline. 119
Figure 5.5 Distribution of study visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection. 123
Figure 5.6 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline. 125
Figure 5.7 Predicted probability of anti-spike IgG positivity by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by vaccination and prior infection. 134
Figure 5.8 Predicted probability of anti-spike IgG positivity by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by age 135
Figure 5.9 Predicted probability of anti-spike IgG positivity by time from first vaccination and age, according to vaccine type and prior infection status (full model). 135
Figure 5.10 Unadjusted percentage (95% CI) anti-spike IgG positive by days after first vaccination in 23,368 participants receiving a single dose of ChAdOx1 vaccine and without evidence of prior infection. 136
Figure 5.11 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 3,767 participants receiving a single dose of ChAdOx1 vaccine and with evidence of prior infection. 137
Figure 5.12 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 14,894 participants receiving a single dose of BNT162b2 vaccine and without evidence of prior infection. 138
Figure 5.13 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 2,067 participants receiving a single dose of BNT162b2 vaccine and with evidence of prior infection. 139
Figure 5.14 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 1,869 participants receiving two doses of BNT162b2 vaccine and without evidence of prior infection. 140
Figure 5.15 Predicted probability with 95% confidence intervals of spike IgG positivity 14-60 days after first vaccination in 24,977 participants without evidence of prior infection from a multivariable logistic regression model. 145
Figure 5.16 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by vaccination type and prior infection. 147
Figure 5.17 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by age 147
Figure 5.18 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination and age based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, according to vaccine type and prior infection status (full model). 148
Figure 5.19 Mean anti-spike IgG levels by time from second vaccination in 20-, 40-, 60- and 80-year-olds according to dosing interval using generalised additive models adjusted for age and dosing interval. 150
Figure 5.20 Unadjusted mean anti-spike IgG levels (BAU/mL) and count of IgG measurements by week from second vaccination in those without prior infection. 152

Figure 5.21 Mean anti-spike IgG levels by time from second vaccination in 3-, 8-, 10-, and 12-week dosing interval groups according to age using generalised additive models adjusted for age and dosing interval. 153
Figure 5.22 Unadjusted mean anti-spike IgG levels (BAU/mL) and count of IgG measurements by week from second vaccination in those with prior infection. 155
Figure 5.23 Predicted mean anti-spike IgG levels in participants with 10-week dosing interval by time from second vaccination according to vaccine type, age and prior infection status using generalised additive models adjusted for age and dosing interval. 156
Figure 5.24 Mean anti-spike IgG levels by time from second vaccination according to vaccine type, prior infection status, and age using generalised additive models adjusted for age and dosing interval. 157
Figure 5.25 Predicted anti-spike IgG trajectory in 36,518 participants without prior infection by class identified from latent class mixed models, using data from 14 days before vaccination to the 90th percentile of the observed time points after vaccination. 159
Figure 5.26 Individual trajectories by classes identified from latent class mixed models for single dose ChAdOx1 and BNT162b2 vaccine. 160
Figure 5.27 Age distribution by classes identified from latent class mixed models for single dose ChAdOx1 (N = 22,424) and BNT162b2 (N = 14,094) vaccines. 161
Figure 5.28 Distribution of timing of antibody measurements relative to the time of the second vaccination for the Bayesian linear mixed model. 169
Figure 5.29 Estimated unadjusted mean trajectory of anti-spike IgG antibody levels and individual trajectories from 21 days after the second vaccination**.** 170
Figure 5.30 Posterior predicted mean trajectories (95% credible interval) of anti-spike IgG levels from 21 days post-second dose using Bayesian linear mixed interval censored models. 171
Figure 5.31 Comparison of effect of factors on antibody responses in participants who received two doses of ChAdOx1 or BNT162b2 or had natural SARS-CoV-2 infection. 175
Figure 5.32 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day after the first vaccination. 177
Figure 5.33 Mean anti-spike IgG levels (95% Cl) by time from first vaccination by age, vaccine type, and prior infection status. 178
Figure 5.34 Posterior predicted trajectories (95% credible interval) of anti-spike IgG levels from 28 days post-first dose by prior infection status. 180
Figure 5.35 Distribution of study visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection. 188
Figure 5.36 Association between anti-spike IgG levels and protection from SARS-CoV-2 infection using the most recent antibody measurement obtained 21-59 days prior to the current visit. 190
Figure 5.37 Association between anti-spike IgG levels and protection from SARS-CoV-2 infection using the maximum antibody measurement obtained ≥21 days prior to the visit. 191
Figure 5.38 Posterior predicted mean days (95% credible interval) from the second vaccination or first infection to the threshold level associated with 67% protection. 193
Figure 5.39 Posterior predicted mean days (95% credible interval) from the second vaccination or first infection to the positivity threshold of 23 BAU/mL. 194
Figure 5.40 Percentage of participants above the threshold level associated with 67% protection by time from second vaccination/infection. 196
Figure 5.41 Posterior predicted days (95% credible interval) from the second vaccination to the threshold level associated with 67% protection (ChAdOx1: 107 BAU/mL, BNT162b2: 94 BAU/mL) multiplied by 2, 3, 5, 8, according to prior infection status and vaccine type. 197
Figure 5.42 Posterior predicted mean days (95% credible interval) from the first vaccination to the positivity threshold of 23 BAU/mL. 199

Figure 5.43 Posterior predicted mean days (95% credible interval) from the first vaccination to the threshold level associated with 67% protection (107 BAU/mL for ChAdOx1 and 94 BAU/mL for BNT162b2). 201
Figure 6.1 Association between anti-spike IgG levels and protection from new SARS-CoV-2 infection in the Omicron BA.4/5 epoch using the most recent antibody measurement obtained 21–59 days before the current assessment by infection variant. 221
Figure 6.2 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day from 21 days after the second vaccination to 14 days after the booster/infection. 226
Figure 6.3 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day from 42 days after the booster/infection for antibody decline modelling. 227
Figure 6.4 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline. 228
Figure 6.5 Association between anti-spike IgG levels and protection from new SARS-CoV-2 infection using the most recent antibody measurement obtained 21–59 days before the current assessment. 236
Figure 6.6 Effect of age on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels 238
Figure 6.7 Effects from time since last vaccination or infection on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels. 239
Figure 6.8 Effects from time since last infection on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels. 240
Figure 6.9 Posterior predicted mean trajectories (95%Crl) of anti-spike IgG levels from second vaccination through third/booster vaccination or infection using Bayesian linear mixed interval censored models. 246
Figure 6.10 Comparisons of antibody levels 42 days post third/booster vaccination or infection, half-lives, and days from third/booster vaccination or infection to reaching antibody levels associated with 67% protection by primary vaccine course, third/booster vaccination or infection, and age. 247
Figure 6.11 Predicted mean anti-spike IgG levels (95%Crl) at 90, 180, 270, and 360 days after the third/booster vaccination or breakthrough infection by ages. 248
Figure 6.12 Posterior predicted mean trajectories (95%Crl) of anti-spike IgG levels from third/booster vaccination or infection. 249
Figure 6.13 Posterior predicted mean trajectories (95%Crl) of anti-spike IgG levels from third/booster vaccination or infection by sex, ethnicity, long-term health condition, and healthcare role. 250
Figure 6.14 Effects of SARS-CoV-2 variant (Delta vs Omicron BA.1) on antibody responses after infection. 251
Figure 6.15 Proportion of participants above the anti-spike IgG antibody threshold level associated with 67% protection by time from third/booster vaccination or infection 255
Figure 6.16 Median protection level by calendar time. 256
Figure 6.17 Percentages of participants' most recent previous infection by age using last assessment from 1 August 2022 to 12 September 2022. 261
Figure 6.18 Distribution of the intervals between survey assessments included in the study. 264

Table of tables

Table 4.1 Priors used in the Bayesian linear mixed interval censored models. 68
Table 4.2 Characteristics of participants in classes identified from latent class mixed models for 7,256 participants infected with SARS-CoV-2.....70
Table 4.3 Odds ratio with 95% confidence intervals from multinomial logistic regression, using Class 1 as the reference. 78
Table 4.4 Odds ratio with 95% confidence intervals from logistic regression comparing seronegative vs seroconverting (Class 3 vs Class 1) using demographic factors and individual symptoms that would be available without a positive test result. 81
Table 4.5 Odds ratio with 95% confidence intervals from a logistic regression model comparing seronegative vs seroconverting participants (Class 3 vs Class 1) using demographic factors, individual symptoms, and comorbidities. 82
Table 4.6 Sensitivity analysis using different starting point (28, 84 days) and different censoring threshold (225, 280 BAU/mL) for anti-spike IgG trajectory modelling (peak and half-life estimation).....87
Table 4.7 Posterior mean and 95% credibility intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the univariable and multivariable models in 3,271 participants in Class 1. 89
Table 4.8 Model coefficients and MCMC diagnostics for the univariable and multivariable models. 90
Table 4.9 Posterior mean and 95% credibility intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in multivariable model including the time period factor (pre-Alpha vs Alpha) in 3,271 participants in Class 1. 93
Table 4.10 Posterior predicted days of protection from the start of infection at different thresholds. 98
Table 5.1 Priors used in the Bayesian linear mixed interval-censored models. 120
Table 5.2 Frequency of visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection. 123
Table 5.3 Participants' characteristics by cohort in the initial vaccination study. 127
Table 5.4 Characteristics of participants with two vaccine doses and at least one antibody measurement from 91 days before the first vaccination through to 4th October 2021 in the second dose study.....129
Table 5.5 Characteristics of participants with a single ChAdOx1, BNT162b2, or mRNA-1273 vaccination in the first dose study.....131
Table 5.6 Predicted probability of anti-spike IgG seropositivity with 95% confidence interval (CI) according to age, vaccine type and dose, and prior infection status by weeks after vaccination. ..141
Table 5.7 Predictors of antibody positivity 14-60 days post first vaccination in participants without evidence of prior infection from univariable and multivariable logistic regression models. 144
Table 5.8 Predicted anti-spike IgG levels (BAU/mL) with 95% confidence interval (CI) from generalised additive models. 151
Table 5.9 Characteristics of classes identified from latent class mixed models for single dose ChAdOx1 and BNT162b2 vaccines in those without evidence of prior infection. 162
Table 5.10 Previous health conditions identified from linked primary care records in 29,575 English participants without evidence of prior infection. 163
Table 5.11 Percentages of non-responders to first or second dose of ChAdOx1 or BNT162b2 using a heuristic rule defined as all antibody measurements being <16 BAU/mL and having at least one antibody measurement 21 days after the first or second dose.....166

Table 5.12 Comparison of the characteristics of responders with non-responders to first or second dose of ChAdOx1 or BNT162b2. 167
Table 5.13 Frequency of antibody measurements per participant, and frequency of antibody measurements relative to the time of the second vaccination, for the Bayesian linear mixed model. 168
Table 5.14 Posterior median and 95% credible intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the unadjusted models and multivariable models in 100,639 participants with two ChAdOx1 doses and 55,053 participants with two BNT162b2 doses. 172
Table 5.15 Posterior median and 95% credible intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the multivariable models. 181
Table 5.16 Posterior predicted peak levels (BAU/mL), half-lives (days), and time from first/second dose to the positivity threshold (days) with 95% credible intervals. 183
Table 5.17 Posterior predicted median peak levels (BAU/mL) and half-lives (days) with 95% credible intervals in participants who received one vaccination with prior infection, two vaccinations without prior infection, and one vaccination without prior infection, by vaccine type. 185
Table 5.18 Frequency of visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection. 188
Table 5.19 Simulation results to demonstrate the Bayesian interval-censored (log)linear mixed model is robust to the upper quantification limit. 214
Table 6.1 Proportion of study assessments with positive PCR results in the correlates of protection analysis. 222
Table 6.2 Settings for each Bayesian linear mixed model. 230
Table 6.3 Priors used in the Bayesian linear mixed interval-censored models on estimating antibody decline 42 days after third/booster vaccination or infection. 231
Table 6.4 Model coefficients and MCMC diagnostics for the Bayesian linear mixed models estimating antibody decline 42 days after the third/booster vaccination or infection. 231
Table 6.5 Characteristics of all participants included in the antibody trajectory analyses. 242
Table 6.6 Characteristics of participants included in the antibody decline analyses post third/booster vaccination or infection. 243
Table 6.7 Posterior predicted anti-spike IgG levels at 42 days from booster/infection (BAU/mL), half-lives after booster/infection (days), and time from booster/infection to reaching antibody levels associated with 67% protection. 253

Chapter 1 Preface

1.1 Acknowledgement

First, I would like to express my deepest appreciation to my supervisors, David Eyre, Sarah Walker, and Koen Pouwels, for their invaluable help and guidance during my DPhil. Their encouragement and support gave me confidence and determination during my research journey. Without them, this thesis would never have been possible. I consider myself fortunate enough to have them as my supervisors.

Second, I want to say thank you to all my friends in Oxford: Yingxi, Danyao, Sufen, Shuhan, Qingze, Jiaqi, Jiayu, Hanwen, and many other wonderful people I met here. I felt lucky to have them around me, whose support and company helped me pass some lonely moments during my time studying abroad, and made my life in Oxford colourful.

Third, I had the pleasure of working with some passionate and talented people in our lab at the Big Data Institute: Femke, Kevin, Chang, Jiandong, and Anders. They always have some good ideas that inspired me. I also enjoyed having lunch together and exchanging formal dinners with each other.

Finally, special thanks to my boyfriend Yifei, for being right by my side all the time. To my parents and family, for their love, care, and support throughout my life.

This research is based on data collected from the Office for National Statistics (ONS) COVID-19 Infection Survey. I would like to thank all COVID-19 Infection Survey participants and study workers for their selfless contribution to this study. The data collected have provided valuable evidence for COVID-19 pandemic control and vaccination strategies.

Also, many thanks to all those in the COVID-19 Infection Survey team:

Office for National Statistics: Sir Ian Diamond, Emma Rourke, Ruth Studley, Tina Thomas, Duncan Cook, Nick Taylor.

Office for National Statistics COVID Infection Survey Analysis and Operations teams, in particular Daniel Ayoubkhani, Russell Black, Antonio Felton, Megan Crees, Joel Jones, Lina Lloyd, Esther Sutherland.

University of Oxford, Nuffield Department of Medicine: Ann Sarah Walker, Derrick Crook, Philippa C Matthews, Tim Peto, Emma Pritchard, Nicole Stoesser, Karina-Doris Vihta, Alison Howarth, George Doherty, James Kavanagh, Kevin K Chau, Stephanie B Hatch, Daniel Ebner, Lucas Martins Ferreira, Thomas Christott, Brian D Marsden, Wanwisa Dejnirattisai, Juthathip Mongkolsapaya, Sarah Cameron, Phoebe Tamblin-Hopper, Magda Wolna, Rachael Brown, Sarah Hoosdally, Richard Cornall, David I Stuart, Gavin Screaton, Jaison Kolenchery, E Yvonne Jones.

University of Oxford, Nuffield Department of Population Health: Koen Pouwels.

University of Oxford, Big Data Institute: David W Eyre, Katrina Lythgoe, David Bonsall, Tanya Golubchik, Helen Fryer.

University of Oxford, Radcliffe Department of Medicine: John Bell.

Oxford University Hospitals NHS Foundation Trust: Stuart Cox, Kevin Paddon, Tim James.

University of Manchester: Thomas House.

Public Health England: John Newton, Julie Robotham, Paul Birrell.

IQVIA: Helena Jordan, Tim Sheppard, Graham Athey, Dan Moody, Leigh Curry, Pamela Brereton.

National Biocentre: Ian Jarvis, Anna Godsmark, George Morris, Bobby Mallick, Phil Eeles.

Glasgow Lighthouse Laboratory: Jodie Hay, Harper VanSteenhouse.

Department of Health and Social Care: Jessica Lee.

Welsh Government: Sean White, Tim Evans, Lisa Bloemberg.

Scottish Government: Katie Allison, Anouska Pandya, Sophie Davis.

Public Health Scotland: David I Conway, Margaret MacLeod, Chris Cunningham.

1.2 Funding

The COVID-19 Infection Survey is funded by the Department of Health and Social Care with in-kind support from the Welsh Government, and the Department of Health on behalf of the Northern Ireland Government and the Scottish Government.

This DPhil thesis is supported by a joint-scholarship from the Nuffield Department of Medicine, University of Oxford and the China Scholarship Council.

1.3 Ethics statement

The COVID-19 Infection survey received ethical approval from the South Central Berkshire B Research Ethics Committee (20/SC/0195).

1.4 Declaration of contribution to thesis

I, Jia Wei, designed and conducted all the analyses presented in this thesis with the support from my supervisors Prof. David W. Eyre, Prof. Ann Sarah Walker, and Dr. Koen B. Pouwels.

The COVID-19 infection survey was designed and planned by Ann Sarah Walker, Jeremy Farrar, John I Bell, John N Newton, Ian Diamond, and Koen B. Pouwels, and was conducted by Ann Sarah Walker, Ruth Studley, Duncan Cook, Emma Rourke, Alison Howarth, Brian D. Marsden, Tim E. A. Peto, Philippa C. Matthews, Nicole Stoesser, Sarah Hoosdally, E Yvonne Jones, David I Stuart, Derrick W .Crook, and David W. Eyre.

Additional assistance and acknowledgement in each chapter is outlined below.

1.4.1 Chapter 3

Researchers and study workers from the COVID-19 Infection survey team and University of Oxford performed laboratory testing including the PCR testing, antibody testing, and antibody conversion.

1.4.2 Chapter 4

Prof. Ann Sarah Walker performed data cleaning and preparation in STATA MP 17.

Daniel Ayoubkhani performed the linkage to the General Practice Extraction Service (GPES) Data for Pandemic Planning and Research and Hospital Episode Statistics (HES) via the NHS number and the regression model to analyse associations with specific comorbidities.

Prof. Ann Sarah Walker, Prof. David W. Eyre, and Dr. Koen B. Pouwels contributed to the study design, data analysis, and the manuscript preparation. Prof. Philippa C. Matthews, Dr. Nicole Stoesser, and Prof. Tim E.A. Peto contributed to the discussion of study results.

1.4.3 Chapter 5

Prof. Ann Sarah Walker performed data cleaning and preparation in STATA MP 17.

Daniel Ayoubkhani performed the linkage to the General Practice Extraction Service (GPES) Data for Pandemic Planning and Research and Hospital Episode Statistics (HES) via the NHS number and the regression model to analyse associations with specific comorbidities.

Dr. Koen B. Pouwels performed the analyses and wrote the code for the model on correlates of protection in the 'second vaccination study'.

Prof. Ann Sarah Walker, Prof. David W. Eyre, and Dr. Koen B. Pouwels contributed to the study design, data analysis, and the manuscript preparation. Prof. Philippa C. Matthews, Dr. Nicole Stoesser, and Prof. Tim E.A. Peto contributed to the discussion of study results.

1.4.4 Chapter 6

Prof. Ann Sarah Walker performed data cleaning and preparation in STATA MP 17.

Prof. Ann Sarah Walker, Prof. David W. Eyre, and Dr. Koen B. Pouwels contributed to the study design, data analysis, and the manuscript preparation. Prof. Philippa C. Matthews, Dr. Nicole Stoesser, and Prof. Tim E.A. Peto contributed to the discussion of study results.

1.5 Publications

1.5.1 In relation to this thesis

Wei, J., Stoesser, N., Matthews, P. C., Ayoubkhani, D., Studley, R., Bell, I., ... & Walker, A. S. (2021). Antibody responses to SARS-CoV-2 vaccines in 45,965 adults from the general population of the United Kingdom. *Nature Microbiology*, *6*(9), 1140-1149.

Wei, J., Matthews, P. C., Stoesser, N., Maddox, T., Lorenzi, L., Studley, R., ... & Eyre, D. W. (2021). Anti-spike antibody response to natural SARS-CoV-2 infection in the general population. *Nature Communications*, *12*(1), 6250.

Wei, J., Pouwels, K. B., Stoesser, N., Matthews, P. C., Diamond, I., Studley, R., ... & Eyre, D. W. (2022). Antibody responses and correlates of protection in the general population after two doses of the ChAdOx1 or BNT162b2 vaccines. *Nature Medicine*, *28*(5), 1072-1082.

Wei, J., Matthews, P. C., Stoesser, N., Diamond, I., Studley, R., Rourke, E., ... & Pouwels, K. B. (2022). SARS-CoV-2 antibody trajectories after a single COVID-19 vaccination with and without prior infection. *Nature Communications*, *13*(1), 3748.

Wei, J., Matthews, P. C., Stoesser, N., Newton, J. N., Diamond, I., Studley, R., ... & Eyre, D. W. (2023). Protection against SARS-CoV-2 Omicron BA. 4/5 variant following booster vaccination or breakthrough infection in the UK. *Nature Communications*, 14(1), 2799.

1.5.2 Other papers during the course of this thesis

Lumley, S. F., Wei, J., O'Donnell, D., Stoesser, N. E., Matthews, P. C., Howarth, A., ... & Eyre, D. W. (2021). The duration, dynamics, and determinants of severe acute respiratory

syndrome coronavirus 2 (SARS-CoV-2) antibody responses in individual healthcare workers. *Clinical Infectious Diseases*, *73*(3), e699-e709.

Eyre, D. W., Lumley, S. F., Wei, J., Cox, S., James, T., Justice, A., ... & Jeffery, K. (2021). Quantitative SARS-CoV-2 anti-spike responses to Pfizer–BioNTech and Oxford–AstraZeneca vaccines by previous infection status. *Clinical Microbiology and Infection*, *27*(10), 1516-e7.

Ward, V., Wei, J., Gordon, W., Barnes, E., Dunachie, S., Jeffery, K., ... & O'Donnell, A. M. (2022). SARS-CoV-2 antibody responses post-vaccination in UK healthcare workers with pre-existing medical conditions: a cohort study. *BMJ open*, *12*(12), e066766.

Eyre, D. W., Futschik, M., Tunkel, S., Wei, J., Cole-Hamilton, J., Saquib, R., …& Fowler, T. (2023). Performance of antigen lateral flow devices in the UK during the alpha, delta, and omicron waves of the SARS-CoV-2 pandemic: a diagnostic and observational study. *Lancet Infectious Diseases*, 23(8), 922-932.

Wei, J., Stoesser, N., Matthews, P. C., Khera, T., Gethings, O., Diamond, I., ... & Eyre, D. W. (2023). Risk of SARS-CoV-2 reinfection during multiple Omicron variant waves in the UK general population. *medRxiv*, 2023.06.29.23292043.

Chapter 2 Introduction and background

2.1 The emergence of SARS-CoV-2

Coronavirus disease 2019 (COVID-19) is a respiratory infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus[1]. It was first identified in the city of Wuhan, China. In December 2019, several hospitals in Wuhan identified patients with pneumonia of an unknown cause. These patients showed symptoms of viral pneumonia, including cough, fever, and dyspnoea in severe cases[2]. On 31 December, Wuhan Municipal Health Commission notified the public of a pneumonia outbreak and reported it to the World Health Organization (WHO)[3]. The WHO declared the outbreak a Public Health Emergency of International Concern on 30 January 2020[4]. On 11 February, the International Committee on Taxonomy of Viruses named the novel coronavirus 'SARS-CoV-2', and the WHO named the resulting disease 'COVID-19'[5].

The number of COVID-19 cases continued to increase sharply in China in January and February 2020, in response to which the Chinese government implemented strict health regulations and the city of Wuhan was shut down. However, being highly transmissible, this novel coronavirus disease spread rapidly across the world. A large number of infections were then reported in many other countries. On 11 March 2020, the WHO officially announced the COVID-19 outbreak as a pandemic[6] (**Figure *2.1***).

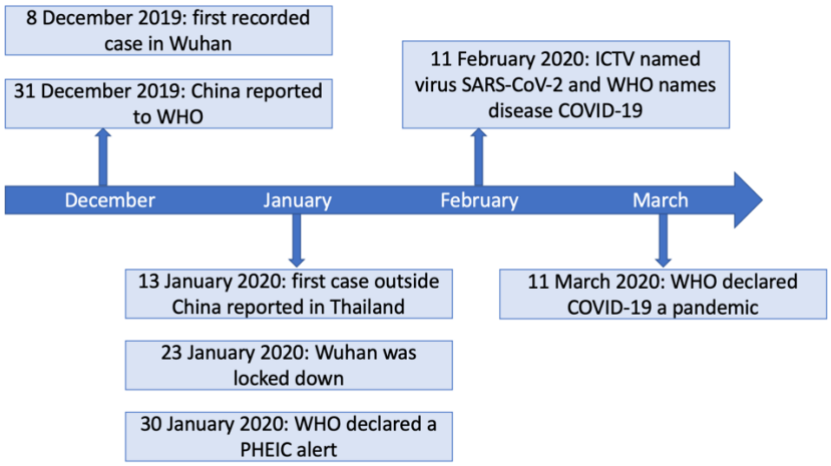


Figure 2.1 Timeline of the key events of the COVID-19 outbreak in the first few months.

Note: Adapted from [7]. PHEIC: Public Health Emergency of International Concern. ICTV: International Committee on Taxonomy of Viruses.

COVID-19 infection was first detected at the end of January 2020 in the United Kingdom (UK), and the virus was spreading rapidly. In March 2020, the UK government announced a national lockdown, asking people to stay at home and social distancing. These measures were gradually lifted since May 2020.

2.2 SARS-CoV-2 transmission and pathophysiology

SARS-CoV-2 is primarily transmitted via fomites and droplets during close direct contact with infected individuals or indirect contact through contaminated hands. While droplet transmission is likely the dominant route, aerosol transmission and faecal-oral transmission are also possible[8].

SARS-CoV-2 infections can be symptomatic and asymptomatic. For symptomatic infections, the most common symptoms of COVID-19 are dry cough and fever. Prior to the emergence of Omicron variant (described below), anosmia (loss of smell) and ageusia (loss of taste) were also strongly associated with COVID-19. Other symptoms include

fatigue, headache, sore throat, shortness of breath, muscle pain, and diarrhoea. In severe cases, patients can develop dyspnoea, low oxygen saturations, respiratory failure, shock, and multiple organ functional failure. Humans are believed to be universally susceptible to SARS-CoV-2, but the clinical manifestations differ with age and any underlying health conditions. Older adults and people with underlying health conditions have more severe symptoms and worse prognosis, and are more likely to require hospitalisation, or even die[7,8].

While most people recover within a few days or weeks, some people may have persistent symptoms for longer than 12 weeks. This has been termed ‘long-COVID’ or ‘post COVID-19 syndrome’. The most common symptoms of long-COVID are fatigue, shortness of breath, loss of smell, and muscle aches, but other symptoms may also appear[9].

The rapid spread of this highly contagious disease has posed the severest threat to global health across the world. As of July 2023, there have been a total of 768 million confirmed SARS-CoV-2 infections and 6.9 million related deaths around the world[10]. In the UK, the total number of confirmed cases were over 24 million and over 200,000 people died with COVID-19 as one of the causes on the death certificate[11].

2.3 Coronavirus and SARS-CoV-2

2.3.1 Coronaviruses, SARS-CoV, and MERS-CoV

The name ‘coronavirus’ is derived from the ‘corona’-like or crown-like morphology of these viruses in the electron microscope first reported in 1960's[12]. Coronaviruses (CoVs) are a diverse family of enveloped single-stranded RNA viruses which include four genera: alpha, beta, gamma, and delta. Alpha and beta coronaviruses are capable of infecting humans. To date, there are seven coronaviruses identified that infect humans, including

human coronavirus (HCoV)-229E, HCoV-NL63, HCoV-OC43, or HCoV-HKU1 which are the main seasonal coronaviruses that cause upper respiratory tract infections[13], and the severe acute respiratory syndrome virus (SARS-CoV-1) and the Middle Eastern respiratory syndrome virus (MERS-CoV) both of which are associated with relatively high mortality. 229E and NL63 are alpha coronaviruses and OC43, HKU1, SARS-CoV, and MERS-CoV are beta coronaviruses. SARS-CoV-2 is also a new type of beta coronavirus[14].

SARS-CoV is a lower respiratory tract infection with atypical pneumonia and progressive lung damage. It was first discovered in November 2002 in China and had spread to over 30 countries with 8,096 reported cases and 774 deaths by July 2003. The case fatality rate is about 10%[13]. However, the pandemic ended after that with no additional infections detected since 2004. This could be because SARS-CoV was less transmissible and most contagious after symptom onset, giving time to implement public health restrictions such as travel bans and quarantine. SARS-CoV affects the renin–angiotensin system (RAS) system by suppressing ACE2, leading to the onset of symptoms.

MERS-CoV is a zoonotic virus transmitted to humans from camels and can cause a wide range of infections. It was first identified in 2012 in Saudi Arabia, and since over 2,000 cases have been reported in other 27 countries to date with a case fatality rate of 35%. Different from SARS, MERS continued to be a problem, causing sporadic cases and small clusters but not a widely spread pandemic. MERS-CoV causes symptoms by producing inflammatory cytokines and invading T cells[13,15].

SARS-CoV and MERS-CoV have similar symptoms including fever, cough, shortness of breath, and myalgia. Human-to-human transmission of SARS-CoV and MERS-CoV occurs mainly in health care settings[16].

2.3.2 SARS-CoV-2 structure and variants of concern

The SARS-CoV-2 genome contains approximately 29.9k nucleotides. The positive-stranded RNA genome has a 5'-cap and a 3'-poly-A tail. At the 5'-end is the open reading frame (ORF) 1ab, which encodes 16 non-structural proteins (Nsp1-16). At the 3'-end are four structural proteins: spike protein (S), envelope protein (E), membrane protein (M), and nucleocapsid protein (N)[17] (**Figure *2.2***).

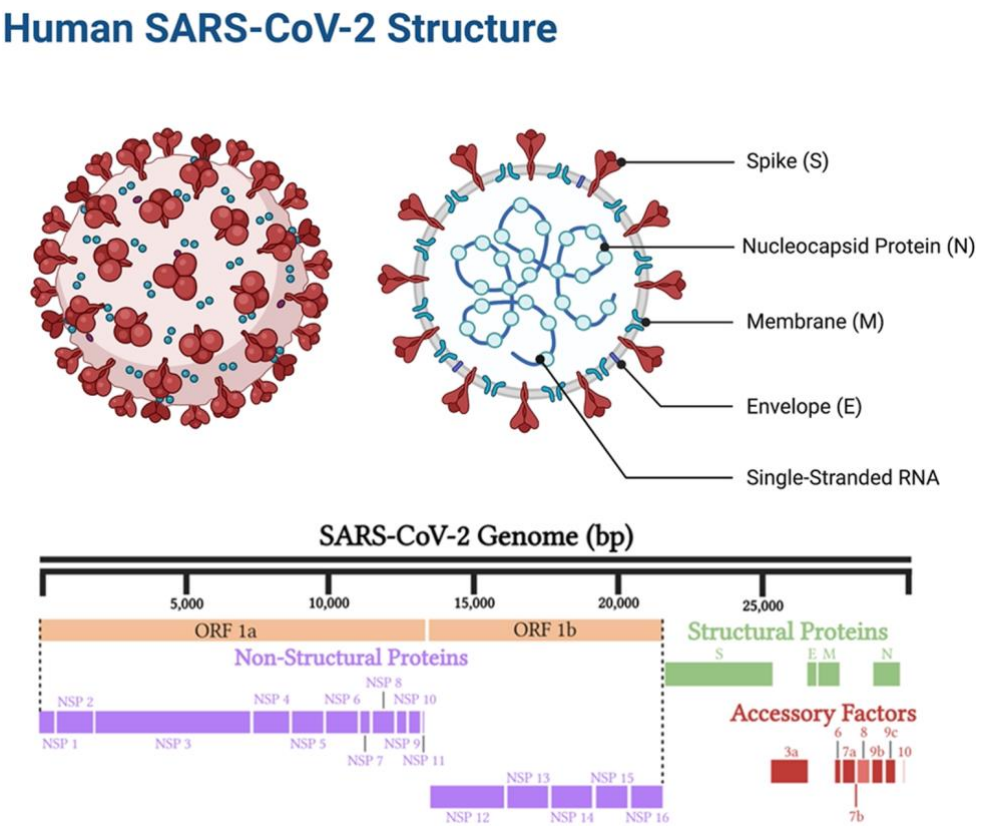


Figure 2.2 SARS-CoV-2 structure.

Note: Taken from [18].

SARS-CoV-2 infection starts with the virus particle binding its spike protein to the cellular angiotensin-converting enzyme 2 (ACE2), which is not only an enzyme, but acts as a functional receptor, leading to viral entry via membrane fusion or endocytosis. During the intracellular life cycle, viruses release the viral genome, express and replicate the genomic RNA, followed by subgenomic transcription and translation of viral structural proteins

including S, M, and E proteins at the endoplasmic reticulum (ER) membrane. Nucleocapsids are assembled in the cytoplasm and followed by budding into the lumen of the ER–Golgi intermediate compartment. Virions are then released from the infected cell through exocytosis[19,20] (**Figure *2.3***).

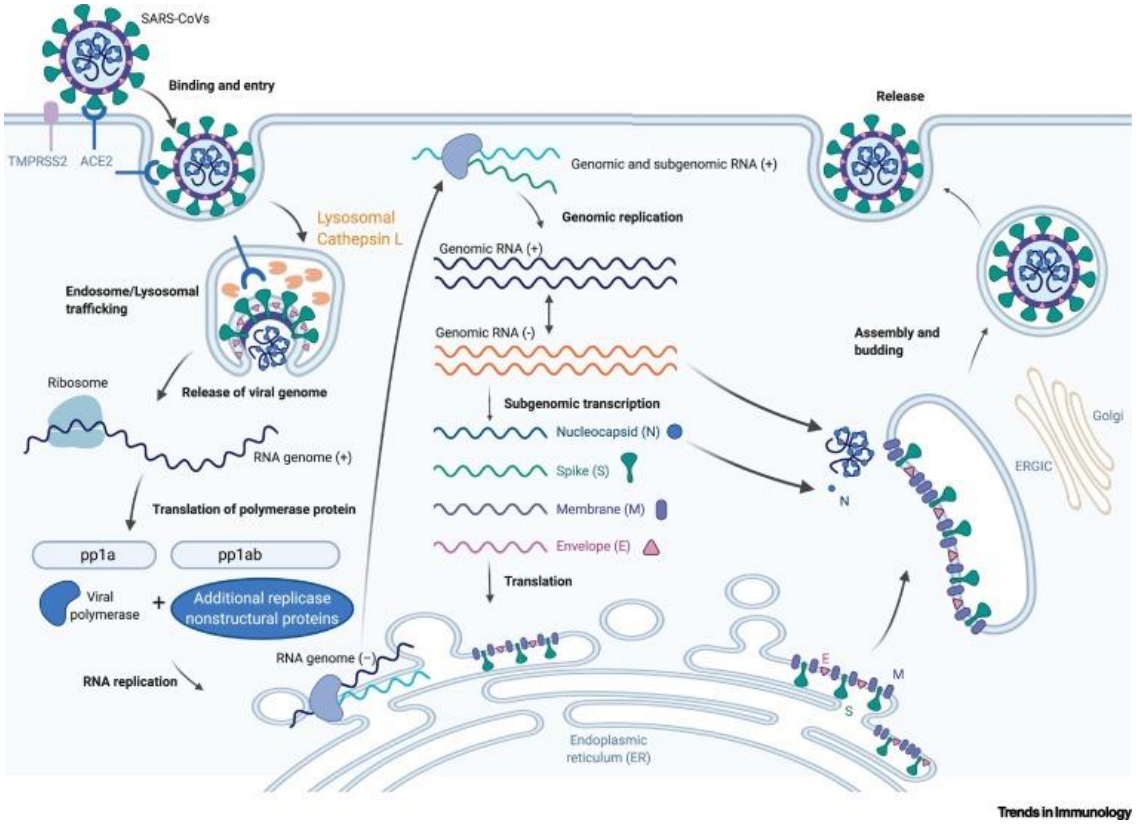


Figure 2.3 The life cycle of SARS-CoV-2 in the host cells.

Note: Taken from [20].

The original virus detected in Wuhan is regarded as the wild-type virus. After that, different variants of SARS-CoV-2 have been discovered. Variants come from the alteration of the SARS-CoV-2 genome through mutation and recombination, leading to changes in the viral proteins, and further leading to differences in disease severity, transitivity, and symptoms[21].

If a specific mutated variant is considered to be concerning because of rapid spread and unknown impact on clinical outcomes, it will be designated as Variant Under Investigation

(VUI) with a year, month, and number. Following a risk assessment, the variant may be designated as Variant of Concern (VOC)[22]. These variants usually have mutations in the spike protein receptor binding domain (RBD) which increase the binding affinity to the ACE2 receptor[23]. According to WHO's definition, a VOC is associated with one or more of the following global public health outcomes: 1) an increase in transmissibility or detrimental change in COVID-19 epidemiology; 2) an increase in virulence or change in clinical disease presentation; 3) a decrease in the effectiveness of public health and social measures[24].

The naming of the variants is done by using letters of the Greek Alphabet. Chronologically, up to March 2023, VOCs circulating around the world were Beta variant (B.1.351) (first identified in South Africa), Alpha variant (B.1.1.7) (first identified in the UK), Gamma variant (P.1) (first identified in Brazil), Delta variant (B.1.617.2) (first identified in India), and Omicron variant (B.1.1.529) variant (first identified in Botswana). The Omicron variant continues to evolve which has led to different lineages that show transmission advantages, including BA.1, BA.2, BA.3, BA.4, BA.5 and descendent sublineages[24].

Up to March 2023, the main VOCs circulating in the UK by chronological order were Alpha, Delta, Omicron BA.1, BA.2, BA.4/5, BQ.1, CH.1.1 and XBB.1.5, leading to multiple infection waves in the UK. I will briefly summarise the characteristics of these variants below.

2.3.2.1 Alpha variant (B.1.1.7)

The Alpha variant was first detected in Kent, United Kingdom in November 2020, and spread rapidly across England between November 2020 and January 2021. The WHO designated it as a VOC on 18 December 2020. The UK imposed lockdown measures and

domestic travel restrictions in reaction to the spread of the virus from December 2020. Many countries also banned passenger travel from the UK.

This variant has many non-synonymous substitutions of immunological importance, including the N501Y replacement on the spike protein to increase ACE2 binding and cell infectivity[25,26]. One major characteristic of the variant is a deletion at positions 69 and 70 of the S protein (Δ69–70), resulting in a diagnostic failure on the S gene target in a positive PCR test using the ThermoFisher TaqPath assay. This is called S-gene target failure (SGTF)[27]. This deletion also increased the spike infectivity and made the virus replicate more efficiently[28]. The Alpha variant was estimated to be 40-90% more transmissible than the wild-type virus[27,29], but there was no evidence of a change in disease severity[29,30].

2.3.2.2 Delta variant (B.1.617.2)

The Delta variant was first detected in India. It was designated as a VOC on 11 May 2021 given its higher transmissibility and reduced neutralisation by antibodies, and was indicated as the dominant strain globally in June 2021 by the WHO. It was partly responsible for India's deadly second wave of the pandemic[31], and contributed to the rapid increase in cases in July 2021 in the UK[32].

Delta variant has mutations in the genes encoding the SARS-CoV-2 S protein causing the substitutions D614G, T478K, P681R, and L452R, which are associated with enhanced transmissibility and decreased vaccine efficacy[33]. It was estimated that the Delta variant was at least 50% more transmissible than the previously dominant Alpha variant. Significantly reduced sensitivity to neutralising antibodies from sera derived following infection or vaccination compared with preceding variants was widely reported[34–36]. A few

studies also reported a higher hospitalisation rate associated with Delta than Alpha variant, potentially indicating an increased virulence with Delta variant[37–39].

2.3.2.3 Omicron variant (B.1.1.529)

Highly transmissible Omicron variants have been the dominant variants since the end of 2021. Omicron was first detected in Botswana and was designated as VOC on 26 November 2021 by the WHO. Omicron has a significant growth advantage over the Delta variant and spreads more rapidly than Delta. On 6 January 2022, it had been confirmed in 149 countries[40]. Compared with other variants, loss of taste and smell were less common. The risk of long-COVID was also lower compared with Delta variant[41].

The Omicron variant has more than 50 mutations, which is more than any previous SARS-CoV-2 variant. 32 mutations occurred in the spike protein, which caused its greater transmissibility and infectivity, and help it escape from neutralising antibody responses from prior non-Omicron infection or vaccination[42–44], resulting in a large number of breakthrough infections and reinfections. Although Omicron has rapidly spread in the community with higher incidence levels than previously seen in this pandemic, there has been no evidence that the Omicron variant was more severe than other variants, and some evidence suggests a lower risk of disease severity and death with Omicron[45]. Nevertheless, the very high level of transmission and risk of reinfection may still pose a threat to public health, especially in vulnerable populations.

In the UK, the dominant Omicron subvariants include BA.1 (original strain), BA.2, BA.4, BA.5, BQ.1, CH.1.1, and XBB.1.5. BA.2 subvariant became dominant in February 2022, which had a growth advantage over BA.1. It was further replaced by the more rapidly spreading BA.4/5 subvariant in June 2022. BA.4 and BA.5 are reported to be more transmissible and resistant to immunity generated by previous variants including Omicron

BA.1[46]. After November 2022, sequencing data in the UK showed a mixture of different sub-lineages including the BQ.1 variant (a sub-lineage of BA.5), CH.1.1 variant (a sub-lineage of BA.2.75), and XBB variant (a recombinant lineage derived from two BA.2 sub-lineages)[47].

2.4 SARS-CoV-2 vaccines

Traditional vaccine development is lengthy and can take 8-15 years, including four steps[48]:

1) Pre-clinical trial: design, vaccine production, formal toxicology studies;
2) Phase I clinical trial: conduct a safety assessment and obtain preliminary immunogenicity data, usually in fewer than 100 individuals;
3) Phase II clinical trial: further investigate immunogenicity and determine appropriate dosage and optimal vaccine regimens, usually in a few hundred individuals;
4) Phase III clinical trial: evaluate vaccine efficacy and safety, usually in a few thousand individuals.

After these steps, the vaccine will go through a licensing process and finally be approved by the regulatory agency, which will take another 1-2 years (**Figure *2.4***).

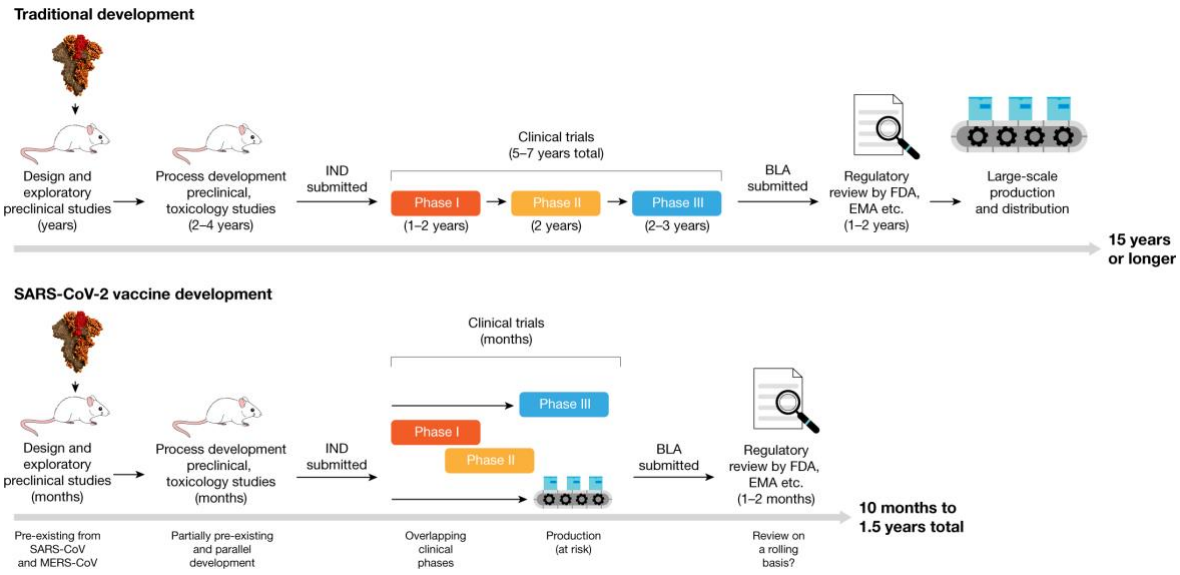


Figure 2.4 Traditional and accelerated vaccine development pipelines.

Note: Taken from [48].

Since the start of the pandemic, clinical development of SARS-CoV-2 vaccines has been conducted in an accelerated way. Researchers used data from the preclinical development of SARS-CoV-1 and MERS-CoV vaccine candidates to omit the initial exploratory step, and the production processes were adapted from those of existing vaccines, which saved a considerable amount of time. The clinical trials were conducted in an overlapping way, with initial phase I/II trials followed by rapid progression to phase III trials after an interim analysis of phase I/II data (**Figure *2.4***). What is more, pharmaceutical companies such as Pfizer, Moderna, and AstraZeneca invested enormous resources and research funds for SARS-CoV-2 vaccine development. Governments from the United States and several other countries also provided a large amount of funding for the scale-up of potentially effective vaccines[49,50]. Through these efforts, the process of SARS-CoV-2 vaccine development and approval was significantly shortened to 10 months, making it possible to vaccinate people within only one year after the start of the pandemic. From a mathematical modelling study, the first year of COVID-19 vaccination saved tens of millions of lives globally[51].

Up to 2022, ten COVID-19 vaccines have been approved by WHO for global use. These vaccines involve four distinct vaccine platforms[49,52]:

1) Inactivated virus vaccines: use inactivated or weakened virus which does not cause disease but still generates an immune response, including Sinopharm's Covilo, Sinovac's CoronaVac, and Bharat Biotech's Covaxin.
2) Messenger RNA (mRNA) vaccines: use genetically engineered RNA to generate a protein that can trigger an immune response, including Moderna's Spikevax mRNA-1273 and Pfizer–BioNTech's BNT162b2. mRNA vaccines have the advantage of rapid development due to their nature that circumvents the need for cell culture or fermentation of viruses, low cost for manufacturing, and high potency[53].
3) Adenovirus vector-based vaccines: use a genetically engineered virus that cannot cause disease itself but generate a protein that triggers an immune response, including AstraZeneca's Vaxzevria and Covishield ChAdOx1 and Johnson & Johnson–Janssen's Ad26.COV2.S.
4) Adjuvanted protein vaccines: use harmless fragments of proteins or protein shells to generate immune response safely, including Novavax's Nuvaxovid and Covovax NVX-CoV2373.

The most used vaccine types in the UK are the Pfizer-BioNTech BNT162b2, Oxford-AstraZeneca ChAdOx1, and Moderna mRNA-1273 vaccines. Their characteristics, mechanisms, and reported efficacy from early clinical trials are summarised below. It should be noted that the efficacy was only based on the initial wild-type virus and the efficacy against further variants could be much lower.

2.4.1 Pfizer-BioNTech BNT162b2 vaccine

The BNT162b2 vaccine was developed by Pfizer in collaboration with the German company BioNTech. The BNT162b2 vaccine is a lipid nanoparticle-formulated, nucleoside-modified mRNA vaccine. The antigen is a full-length spike protein with two proline substitutions, and the dose per vaccine is 30 µg. The mechanism of action is summarised in **Figure *2.5***. The initially suggested dosage is two doses with a dosing interval of 21 days. The storage temperature for BNT162b2 is -70 °C, making it difficult to transport and be used in low- and middle-income countries[54].

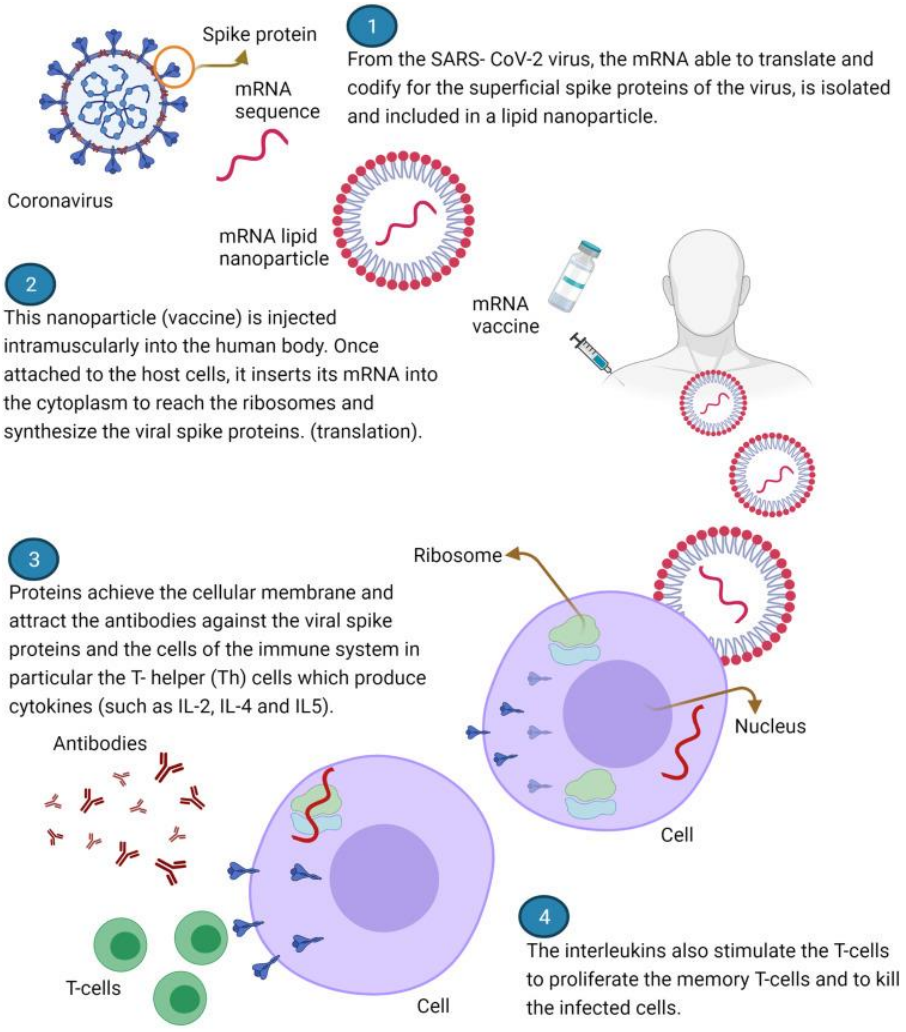


Figure 2.5 Mechanism of action of mRNA vaccines.

Note: Taken from [49].

Preliminary phase 2/3 clinical trials reported efficacy against infection of 52% (95% confidence interval [CI], 29.5-68.4%) after the first dose, and 94.6% (95% CI 89.9-97.3%) after the second dose. The efficacy against the severe disease was 88.9% (95% CI 20.1-99.7%). Local and systemic reactogenicity are common, including short-term, mild-to-moderate pain at the injection site, fatigue, and headache. The incidence of serious adverse events was low and was similar in the vaccine and placebo groups[55].

2.4.2 Moderna mRNA-1273 vaccine

The mRNA-1273 vaccine was developed by Moderna and the Vaccine Research Center at the National Institute of Allergy and Infectious Diseases (NIAID) in the United States. Similar to the BNT162b2 vaccine, mRNA-1273 vaccine is also a lipid nanoparticle–encapsulated mRNA-based vaccine that encodes the prefusion stabilized full-length spike protein of the SARS-CoV-2 virus, but with a higher dose of 100 µg. The mechanism of action is also similar to that presented in **Figure *2.5***. The suggested dosage is two doses with a dosing interval of 28 days, and the storage temperature is -20°C[54].

From an initial phase III randomised controlled trial, the overall efficacy after receiving two doses was estimated to be 94.1% (95% CI, 89.3-96.8%) against infection and 100% against the severe disease. Moderate, transient reactogenicity after vaccination was common, but serious adverse events were rare, and the incidence was similar in the vaccine and placebo groups[56].

2.4.3 Oxford-AstraZeneca ChAdOx1 vaccine

The ChAdOx1 vaccine was developed by AstraZeneca in collaboration with the University of Oxford. It is a viral vector vaccine using a replication-deficient chimpanzee DNA

adenoviral vector with the SARS-CoV-2 spike protein. Human populations have not been exposed to the adenovirus and it does not generate an immune response to the adenovirus itself, but only to the viral protein encoded in the host DNA. The mechanism is summarised in **Figure *2.6***. The initially suggested dosage is two doses with a dosing interval of 28 days. It can be stored at refrigerator temperature (2°–8°C), so it is more easily deployed in low- and middle-income countries than mRNA vaccines[54].

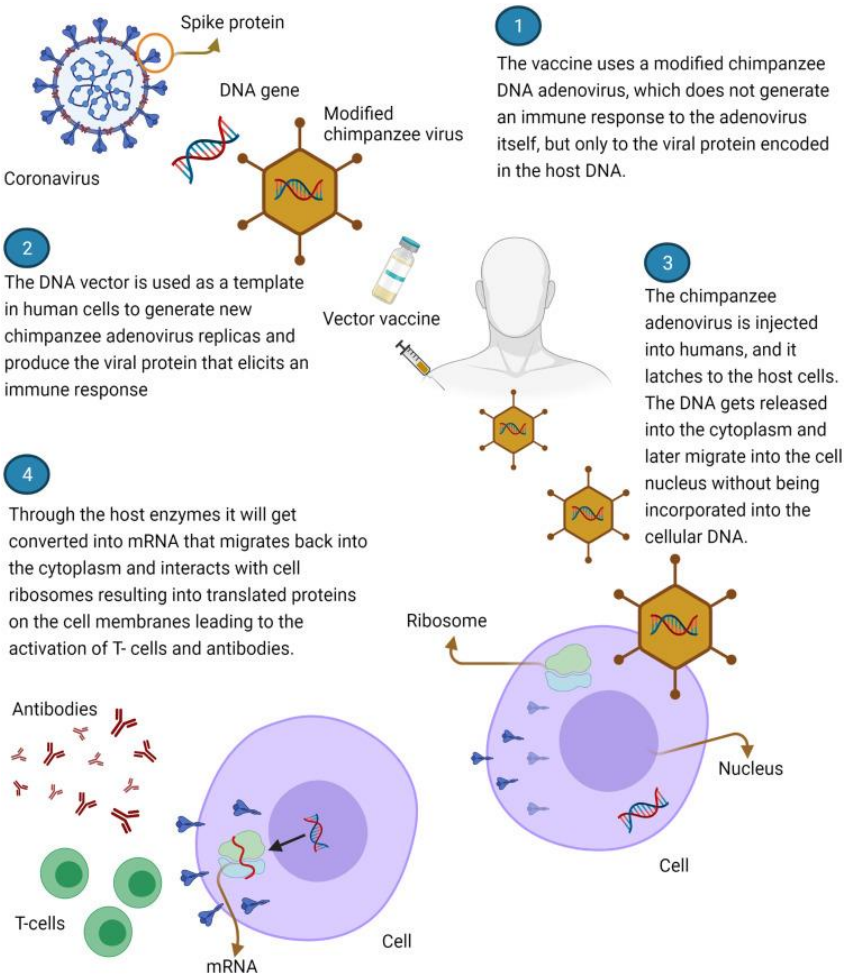


Figure 2.6 Mechanism of action of ChAdOx1 vaccine.

Note: Taken from [49].

From two studies in Brazil and the UK, the overall efficacy against symptomatic infection was 70.4% (95%CI 54.8–80.6) after the second dose. The efficacy against severe COVID-19 was 100% after the first dose[57]. However, although initial studies showed that this

vaccine was safe with only some mild side effects, there was some evidence emerging on the potential adverse effects of the vaccine including the rare development of immune thrombotic thrombocytopenia[58,59]. Therefore, ChAdOx1 was banned for use in many European countries in March 2021[60]. However, ChAdOx1 remains as the first-line vaccine in many low- and middle-income countries due to its cheaper price, greater stability, and no freezing requirement.

2.5 Vaccination programme in the UK

On 8 December 2020, the UK became the first country in the world to start its national vaccination programme, with the approval of the Pfizer-BioNTech BNT162b2 vaccine on 2 December 2020[61], followed by the approval of the Oxford-AstraZeneca ChAdOx1 nCoV-19 vaccine on 30 December 2020[62]. The Spikevax Moderna mRNA-1273 vaccine was approved on 8 January 2021, becoming the third COVID-19 vaccine approved by the UK regulator. All vaccines have been widely used in the UK since then.

In the UK, vaccines were initially administered to priority groups, including care home residents, people >80 years old, healthcare workers, and those clinically vulnerable (≥16 years), and then offered to the rest of the adult (≥18 years) population in decreasing age order[63] (**Figure *2.7***). In January 2021, the dosing interval was extended to 12 weeks for all vaccines, regardless of the licensed dosing schedule (3-4 weeks). This was because the government believed that initially vaccinating a greater number of people with a single dose would prevent more deaths and hospitalisations than vaccinating a smaller number of people with two doses[64].

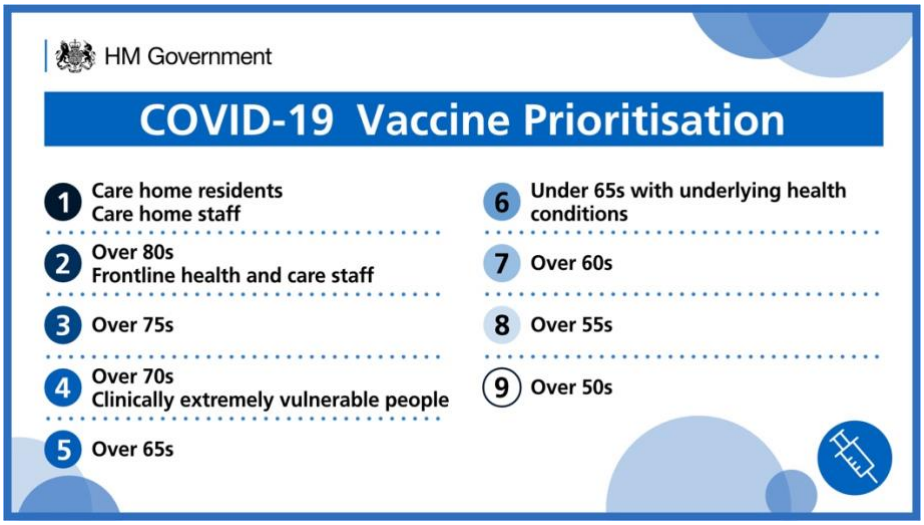


Figure 2.7 Priority groups for the UK COVID-19 vaccination programme.

Note: Taken from[65].

In September 2021, vaccines were further offered to younger people aged 12-15 years[66]. In November 2021, due to the evidence suggesting the potential of waning immunity from vaccination and the newly emerged Omicron variant, a booster campaign was announced by the government to offer a third/booster vaccination to all UK adults from three months after their second dose[67]. The booster vaccination was either Pfizer BNT162b2 or Moderna mRNA-1273 vaccine, regardless of the primary vaccine received. From February 2022, priority groups including people ≥75 years old, care home residents, and those clinically vulnerable were able to receive a fourth vaccination[68]. In September 2022, Joint Committee on Vaccination and Immunisation (JCVI) further announced an autumn booster programme that aimed to increase immunity in the priority groups from winter 2022 to 2023. All adults ≥50 years were offered a fourth dose, and those ≥75 years were able to receive a fifth dose[69] (**Figure *2.8***).

The vaccination rollout was successful compared with many other countries and vaccine take-up was high. By September 2022, 93.6% of the population aged ≥12 years had

received a first vaccination, 88.3% had completed a two vaccination course, and 70.2% had received a third/booster vaccination[70].

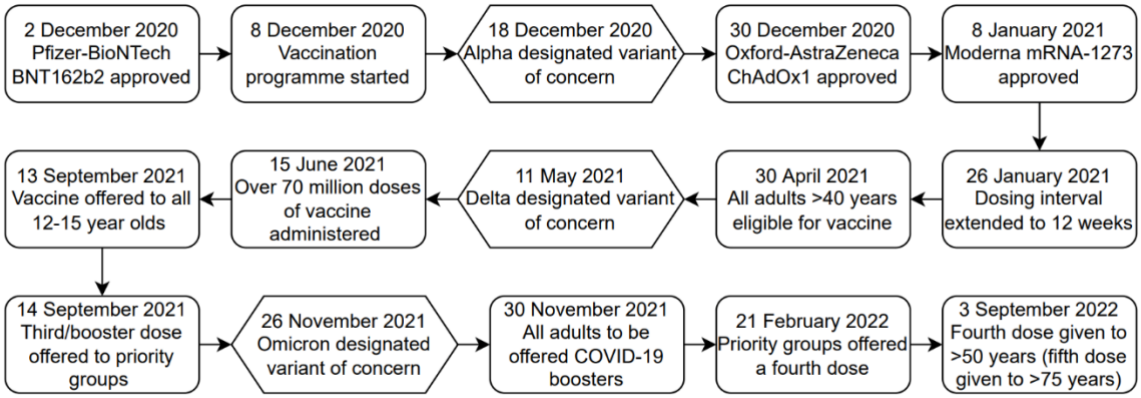


Figure 2.8 Timeline of the key events in the UK COVID-19 vaccination programme.

2.6 Immunity against SARS-CoV-2

The human immune system is divided into innate and adaptive immune systems. When the virus enters the human body, the innate immune response is the first line of defence, which secretes type I interferons, antiviral cytokines, and cellular responses are activated, including neutrophils, monocytes, macrophages, dendritic cells, and natural killer cells. The innate immune system lacks immunologic memory, thus remains unchanged however often the antigen is encountered[71].

Adaptive immunity is the second line of defence, which is generated in the lymph nodes, spleen, and mucosa-associated lymphoid tissue. It involves the proliferation of antigen-specific B and T cells which occurs when the surface receptors of these cells bind to viral antigens. Adaptive immunity includes humoral immunity (B-cell immunity) and cellular immunity (T-cell immunity). Humoral immunity generates binding antibodies to bind the SARS-CoV-2 virus spike protein and eliminate the virus, or neutralising antibodies to block the entry of the virus into cells. Cellular immunity includes virus-specific B cells and

T cells that can provide immunological memory. CD8+ T cells eliminate infected cells directly and CD4+ T cells help support immune responses[52,71].

Immunity can be acquired from both natural infection and vaccination. SARS-CoV-2 infection induces a robust humoral immune response, which plays a critical role in antiviral defence by providing newly produced antibodies from activated plasma cells, including anti-spike (anti-S), anti-receptor binding domain (anti-RBD) and anti-nucleocapsid (anti-N) antibodies. Apart from antibodies, the induction of SARS-CoV-2 specific memory T cells and B cells is also important for long-term protection[72]. SARS-CoV-2 vaccines are designed to elicit an immune response against the spike protein, so vaccination only induces anti-S and anti-RBD antibodies, but not anti-N antibodies. Antibodies have different types, and immunoglobulin G (IgG) is the most common type of antibody found in blood circulation. Other types include immunoglobulin M (IgM) and immunoglobulin A (IgA). Both SARS-CoV-2 IgM and IgG can be detected around the same time after infection, but IgM usually becomes undetectable after a few weeks to months after infection, while IgG can last longer[73].

Although SARS-CoV-2 infection induces antibody responses, antibody levels may be dependent on the disease severity and the virus inoculum. For example, IgG and IgM antibodies have been found in asymptomatic individuals who tested positive for SARS-CoV-2, but these antibodies were present at markedly lower levels than in symptomatic patients with COVID-19[74]. Also, initial studies have found that antibodies waned after infection[75–78]. Therefore, it is important to understand the degree and durability of the antibodies induced by SARS-CoV-2 infection or vaccination. Furthermore, to determine the importance of humoral immunity for protection and estimate the duration of protection, it is crucial to associate immune markers (e.g. antibodies) with specific levels of

protection, i.e. to determine ‘immune correlates of protection’. Correlates of protection are measurements of immune parameters that allow prediction of the degree of protection an individual has against a certain infectious disease and can be used as surrogate endpoints for vaccine efficacy. Given different variants have emerged with multiple mutations in the spike protein that can result in decreased neutralisation, it is also important to examine whether and to what extent protection from both prior infection and vaccination is reduced by variants.

2.7 Thesis aims and outline

The goal of this thesis is to examine the nature and duration of antibody responses after SARS-CoV-2 natural infection and vaccination, analyse the association between antibody levels and protection (i.e. correlates of protection), and estimate the duration of protection following infection or vaccination. Thus, these results could be used to inform global vaccination strategies and pandemic management. The UK’s COVID-19 Infection Survey (CIS) provided a huge amount of serological and outcome data which has enabled me to carry out the large-scale analyses presented (details described in Chapter 3).

Chapter 3 introduces the survey and data involved in my studies, and the general statistical methods applied. Chapters 4, 5, and 6 are the three main results chapters. Given different emerging variants and different stages of the vaccination programme, I present my study results in chronological order.

Chapter 4 focuses on antibody responses after natural SARS-CoV-2 infection using data up to June 2021, therefore results described in this chapter were mainly based on the wild-type and Alpha variant infections and so are less able to inform on further variants such as Delta and Omicron. I studied antibody trajectories following infection, and specifically

examined the proportion of non-responders and predictors of non-response. I then examined antibody peak levels and waning half-lives after infection and used the results to estimate the duration of protection.

In Chapter 5 I examine antibody responses after the first and second vaccination using data up to January 2022. Correlates of protection were explored for the Delta variant, therefore the results are less applicable to Omicron variant infections. I examined antibody trajectories after the first and second vaccination by vaccine type, and the impact of age, sex, ethnicity, long-term health conditions, dosing interval, and most importantly, prior infection, on the peak levels and half-lives after the first or second vaccination. Combining these results, I compared the duration of protection after the first and second vaccination.

In Chapter 6, with the initiation of the booster campaign in the UK and the circulation of the Omicron variant, I compare antibody responses after the third/booster vaccination versus breakthrough infection post second vaccination. I update the correlates of protection analyses for the Omicron variant, and determine the duration of protection following booster vaccination versus breakthrough infection.

Finally, in Chapter 7, I summarise and discuss the methodology, results, and limitations in my studies on SARS-CoV-2 antibody responses and protection, and conclude with implications for policymakers and directions for future work.

Chapter 3 Thesis Methods

3.1 UK's COVID-19 infection survey

The studies described in this thesis used data from the UK's Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) (ISRCTN registry number: ISRCTN21086382). The study was launched in April 2020 and is being delivered with the University of Oxford, IQVIA, and other partners. The data collection was officially paused in mid-March 2023. This is the largest household survey of COVID-19 infections and antibodies in the UK with more than 11 million swab tests and 3 million blood tests collected from households (described below)[79].

The survey randomly selects private households on a continuous basis from address lists and previous surveys conducted by the ONS or the Northern Ireland Statistics and Research Agency to provide a representative sample across the four countries comprising the UK (England, Wales, Northern Ireland, Scotland). The initial response rate was 51%, 41%, and 44% for those sampled from previous ONS studies in England, Wales, and Northern Ireland, respectively, and was 13%, 14%, and 13% for those invited from a randomly sampled list of addresses in England, Wales, and Scotland, respectively[80]. However, sampling bias may exist, for example, marginalised populations may have a lower response rate and thus underrepresented. Therefore, various statistical adjustments were conducted, including standard inverse probability weighting methods or post-stratification to weight observed data back to a representative UK population, based on response to the initial ONS letter in adults originally targeted, their characteristics compared with those to whom the original invite to participate in the ONS surveys, and those characteristics versus the general UK population[80].

Following verbal agreement to participate, a study worker visited each household to take written informed consent. This consent was obtained from parents/carers for those 2-15 years, while those 10-15 years also provided written assent. Children aged <2 years were not eligible for the study. Three options were available to the participants: 1) just have one visit; 2) have a visit every week for a month; 3) have a visit every week for a month and then continue to have visits every month until the end of the study (entirely voluntary).

Individuals were surveyed on their socio-demographic characteristics, behaviours, and vaccination status. Combined nose and throat swabs were taken from all consenting household members for SARS-CoV-2 PCR testing. For a random 10-20% of households, individuals ≥16 years were invited to provide blood samples monthly for serological testing. Household members of participants who tested positive were also invited to provide blood monthly for follow-up visits. Details on the sampling design are provided elsewhere[81]. From April 2021, additional participants were invited to provide blood samples monthly to assess vaccine responses, based on a combination of random selection and prioritisation of those in the study for the longest period (independent of test results). For the first two years of the pandemic, swab tests were done at a home visit undertaken by a trained individual, and blood samples were taken from a vein by a trained nurse, phlebotomist, or healthcare assistant. After two years, to make the survey more affordable, swab samples were taken by self-swabbing, and blood samples were taken by a fingerprick by the participant themselves. All samples were posted to ONS.

At each visit, participants were asked whether they present 12 specific symptoms, including cough, fever, myalgia, fatigue, sore throat, shortness of breath, headache, nausea, abdominal pain, diarrhoea, loss of taste, and loss of smell, and whether they thought they had unspecified symptoms compatible with COVID-19. From 26 April through 22 July

2020, questions referred to current symptoms, and from 23 July 2020 from the day of the visit to the preceding 7 days. Symptomatic infection was defined as having a positive response to any symptom question at a swab-positive visit.

The study protocol is available at https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-survey/protocol-and-information-sheets. The study received ethical approval from the South Central Berkshire B Research Ethics Committee (20/SC/0195).

3.2 Vaccination data

Vaccination information was obtained from participants at visits by self-report, including vaccination type, number of doses, and vaccination dates. Participants from England were also linked to the National Immunisation Management Service (NIMS), which contains all individuals' vaccination data in the English National Health Service COVID-19 vaccination programme. There was good agreement between self-reported and administrative vaccination data (98% on type and 95% on date[82]). I used vaccination data from NIMS where available for participants from England, and otherwise data from the survey.

3.3 Laboratory data

3.3.1 Antibody data

SARS-CoV-2 antibody levels were measured using an ELISA detecting anti-trimeric spike IgG developed by the University of Oxford[83]. All testing was performed at the University of Oxford. Normalised results are reported in ng/ml of mAb45 monoclonal antibody equivalents. Up to 26 February 2021, the assay was performed using a fluorescence detection mechanism as previously described, using a threshold of 8 million units to

identify positive samples[83]. Subsequent testing was performed with a CE-marked version of the assay, the Thermo Fisher OmniPATH 384 Combi SARS-CoV-2 IgG ELISA, which uses the same antigen, with a colorimetric detection system. mAb45 is the manufacturer-provided monoclonal antibody calibrant for this quantitative assay. To allow the conversion of fluorometrically determined values in arbitrary units, 3840 samples were run in parallel on both systems and compared. A piece-wise linear regression was used to generate the following conversion formula:

$\log_{10}$(mAb45 units) = 0.221738 + 1.751889e-07*fluorescence_units +
5.416675e-07*(fluorescence_units>9190310)*(fluorescence_units-9190310)

Each batch of 320 samples (diluted 1:50) was run with a negative control sample (Sigma human serum H6194, diluted 1:50) included in duplicate and a dilution series of 3 monoclonal antibodies[84–86] included in duplicate used for assay calibration and quality control (CR3022 – 4 dilutions [1000, 300, 100, 30 ng/ml], mAb45 – 5 dilutions [400, 300, 100, 30, 10 ng/ml], mAb269 – 4 dilutions [300, 100, 30, 10 ng/ml]). Values obtained for each control and calibration sample were compared to established historic control values and plates subjected to acceptance criteria that required all 28 controls and calibrants to fall within historic limits, namely, no more than 5 control samples >2 standard deviations different, no more than 2 samples >3 standard deviations different and no more than 1 sample >4 standard deviations different. The first two limits were based on rejecting batches where the probability of the observed variation excluded that expected 99% of the time. The latter rule allowed for one-off robotic error.

Results were calibrated of the Thermo Fisher OmniPATH assay into WHO international units (binding antibody unit, BAU/mL) using serial dilutions of National Institute for Biological Standards and Control (NIBSC) Working Standard 21/234. The NIBSC 21/234

Working Standard has been previously calibrated against the WHO International Standard for anti-SARS-CoV-2 immunoglobulin (NIBSC code 20/136), with anti-spike IgG potency of 832 BAU/mL (95%CI 746-929). 2-fold dilutions of 21/234 between 1:400 and 1:8000 from three separate batches on three separate days were generated. Results from a total of 63 diluted samples were merged and a linear regression model fitted constrained to have an intercept of zero to convert mAB45 units in ng/ml for samples diluted at 1:50 to BAU/mL (**Figure *3.1***):

BAU/mL = 0.559 * [mAb45 concentration in ng/mL at 1:50]

I used ≥ 42 ng/ml (or 23 BAU/mL) as the threshold for determining anti-spike IgG positivity (corresponding to the 8 million units with fluorescence detection). This threshold was determined prior to the development of COVID-19 vaccinations and is CE marked by the Medicines and Healthcare products Regulatory Agency, providing 99% sensitivity and specificity in identifying unvaccinated people who have had a COVID-19 infection before ("natural immunity") from unvaccinated people who have not[80].

Antibodies were diluted at 1:50 for samples at the start of the survey. The dilution was changed to 1:400 from 28th January 2022 and further changed to 1:1600 from 29th April 2022 due to booster vaccinations and widespread Omicron infections causing saturated results. The 1:1600 dilution brought most test results within the dynamic range of the assay. The upper limit of quantification for the assay at 1:50, 1:400, and 1:1600 dilutions were 450 BAU/mL, 1,360 BAU/mL, and 8,000 BAU/mL, respectively.

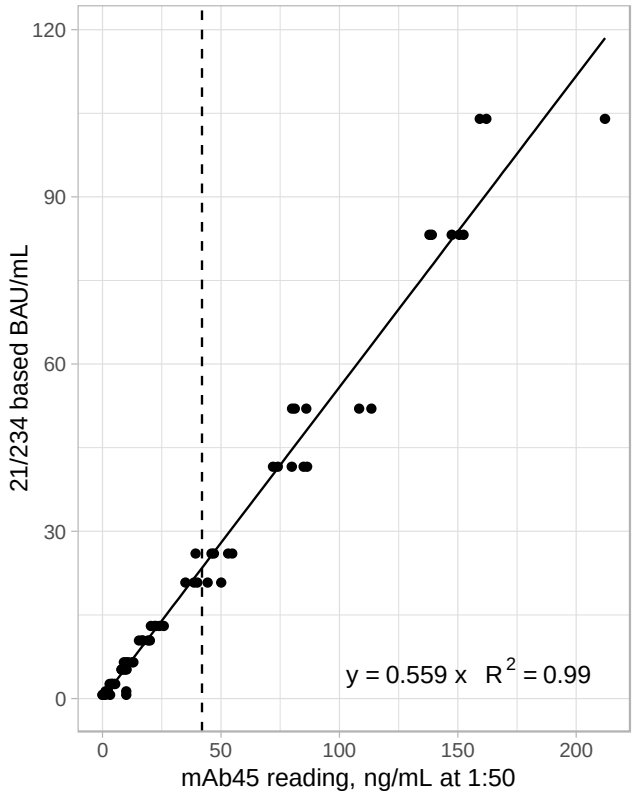


Figure 3.1 Correlation between anti-spike IgG levels in mAb45 units (ng/ml) and WHO international units (BAU/mL) in 63 dilution replicates.

Antibodies are usually evaluated by two quantities: peak level and half-life. Peak level is the maximum level reached after an immunogenic event (infection or vaccination). Half-life is the time required for antibody levels to decrease by half, assuming it decreases exponentially with time.

Between March 2021 and January 2022, samples for IgG immunoglobulins against the nucleocapsid (N) protein were also tested to distinguish between those with immunity due to natural infection (who would be anti-S and anti-N positive) and vaccination (anti-S positive, but anti-N negative). Anti-nucleocapsid IgG was measured using OmniPATH SARS-CoV-2 N-protein IgG ELISA. I used ≥ 30 ng/ml as the threshold for determining anti-N IgG positivity.

3.3.2 PCR data

PCR of combined nose and throat swabs was undertaken using the Thermo Fisher TaqPath SARS-CoV-2 assay at high-throughput national "Lighthouse" laboratories in Glasgow and Milton Keynes (to 8 February 2021). PCR outputs were analysed using UgenTec FastFinder 3.300.5, with an assay-specific algorithm and decision mechanism that allows conversion of amplification assay raw data into test results with minimal manual intervention. Samples are called positive if at least a single N-gene and/or ORF1ab are detected (although S-gene cycle threshold (Ct) values are determined, S-gene detection alone is not considered sufficient to call a sample positive[81]) and PCR traces exhibit an appropriate morphology. For contingency due to capacity issues, a small number of swabs were tested using endpoint PCR at the Rosalind Franklin laboratory (109,874 (12%) between 29 April and 12 September 2022, the period used in analyses of correlates of protection).

3.3 SARS-CoV-2 definition

3.3.1 Infection episode

Positive tests were grouped into episodes because PCR-positive results might be observed at multiple visits after infection. Whole-genome sequencing (WGS) was available on only a subset of positives, and only a subsample had antibody results, so episodes were defined using PCR results. WHO defined reinfection as positives occurring at least 90 days after the first infection[87], but this definition led to higher than plausible risk of a new episode between 90 and 120 days especially for infections with lower viral load, so I defined the start of a new infection episode or 'positive case' as the date of (1) the first PCR-positive test in the study (not preceded by any study PCR-positive test); (2) a PCR-positive test after four or more consecutive negative visits; (3) a PCR-positive test at least 120 days

after the start of a previous episode with one or more negative tests immediately preceding this. However since the emergence of Omicron variant in December 2021, more and more shorter reinfections occurred, so I modified the definition as (1) the first PCR-positive test in the study (not preceded by any study PCR-positive test); (2) a PCR-positive test after four or more consecutive negative visits; (3) a PCR-positive test at least 120 days after the start of a previous episode with one or more negative tests immediately preceding this; (4) a PCR-positive test at least 90 days after the start of a previous episode with the last 2 consecutive tests being negative or the last test being negative and from 20 December 2021 onwards; or (5) a PCR-positive test at least 60 days after the start of a previous episode with the last 3 consecutive tests being negative.

3.3.2 Infection variant

S-gene positivity pattern was examined for all swab samples by PCR, and the percentage of samples with a positive S-gene among all samples was calculated for each surveillance week (**Figure *3.2***). The variant associated with each infection episode was determined by WGS where this was available, otherwise as Pre-Alpha/Delta/Omicron BA.2-compatible if the S-gene was detected (with N/ORF1ab/both), or as Alpha-/Omicron BA.1/Omicron BA.4/5-compatible if positive at least once for ORF1ab+N (but not for the S gene, S-gene target failure, SGTF), using the dates when these variants were dominant in the UK, i.e. accounted for >50% of infections. For those without sequencing data and with only one gene-positive (N-only/ORF1ab-only) or no gene positivity/Ct data available, and samples from the England National testing programme, I assigned them to each variant type based on the dominant circulating variant nationally (>50% of infections) in each surveillance week. By this definition, an infection was designated as Pre-Alpha variant up to 6 December 2020, as Alpha variant from 7 December 2020 to 16 May 2021, as Delta variant

from 17 May 2021 to 12 December 2021, as Omicron BA.1 variant from 13 December 2021 to 20 February 2022, as Omicron BA.2 variant from 21 February to 5 June, and as Omicron BA.4/5 variant from 6 June 2022 to 6 November 2022.

After 7 November 2022, sequencing data in the UK showed a mixture of different sub-lineages including the BA.5 variant, BQ.1 variant (a sub-lineage of BA.5), CH.1.1 variant (a sub-lineage of BA.2.75), and XBB variant (a recombinant lineage derived from two BA.2 sub-lineages). This thesis did not include analyses on infections that occurred during this period.

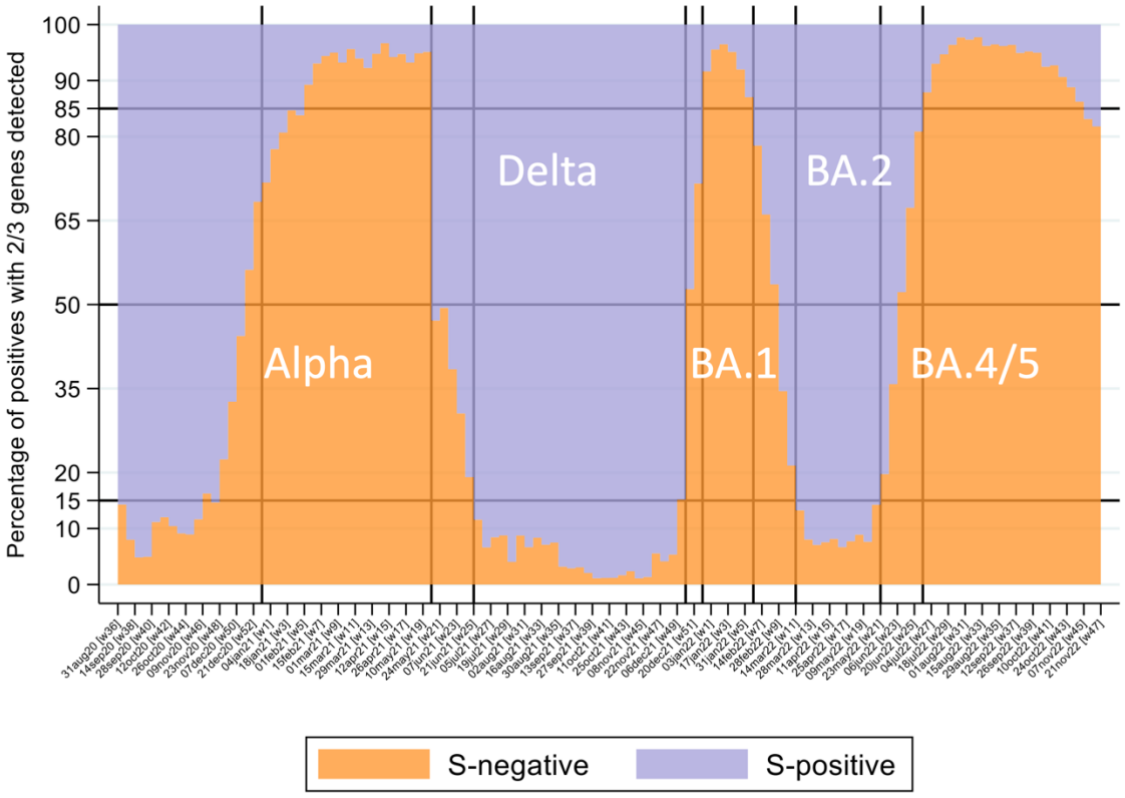


Figure 3.2 Gene positivity pattern by calendar time to determine Alpha, Delta, BA.1, BA.2, and BA.4/5 epochs.

3.4 Overall statistical methods

3.4.1 Generalised additive models

Commonly used generalised linear models (GLMs) assume a linear function of the response variable Y and a set of covariates X_1, X_2, …, X_p, in the form of $Y = \sum_1^p \beta_j X_j$. The parameters of the linear function are estimated by maximum likelihood. Typical models include linear regression models, logistic regression models, and Cox's proportional hazards model.

Generalised additive models (GAMs) are smooth extensions of generalised linear models[88]. They replace the linear form $\sum \beta_j X_j$ with a sum of flexible non-linear functions of the predictor variables $\sum f_j(X_j)$. The general formula is:

$$g\big(E(Y)\big) = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_j\big(x_j\big)$$

where g is the link function relating the expected value of Y to the predictor variables.

The functions f_j can be a parametric function such as polynomial and splines which are also commonly used in GLMs, or non-parametric or semi-parametric as 'smooth functions'. Each smoothing function can be used for one covariate (univariate smooth) or multiple covariates (multivariate smooth). GAMs are highly flexible models that are allowed to learn non-linear features with relaxed assumptions on the actual relationship between response and predictor, which can better fit the data than purely parametric model, but arguably lose some interpretability. However, it still has some of the explainability of linear regression through visualisation of non-linear relationships with the outcome.

Each smoothing function is represented by a sum of k simpler, fixed basis function multiplied by coefficients: $\beta_{j,k} b_{j,k}(x_j)$ ($b_{j,k}(x_j)$ are basis functions, and $\beta_{j,k}$ are

coefficients to be estimated). k is the 'basis size' or 'basis complexity' which controls the complexity of the smoothing function. To avoid large k leading to overfitting, a smoothing penalty is added to control the complexity and avoid excess wiggliness by a smoothing parameter λ[89,90]. The penalized log-likelihood for fitting the model is:

$$L - \lambda\beta^T S\beta$$

L is the model likelihood, S is the penalty matrix, β is the parameter vector.

The effective degrees of freedom (EDF) measure the complexity of penalized smooth terms, which at a maximum is the number of coefficients to be estimated in the model minus any constraints. A larger EDF represents more wiggly terms, and a smoother cannot have more than k EDF. Values around 1 means the term is closer to a linear term. The larger the basis dimension k, the more wiggliness one can allow for, so k should be set large enough to allow for potential variation in the smoother. However, at the meantime, the basis dimension k should not be set too large to retain computational efficiency and reduce computation time.

Smooths terms are represented using penalized regression splines with smoothing parameters (i.e. the parameters that control the smoothness of the predictive functions, λ) selected by generalised cross-validation statistic (GCV)/ Akaike Information Criteria (AIC)/restricted maximum likelihood (REML), or by regression splines with fixed degrees of freedom (pre-selecting λ). REML is the form of maximum likelihood estimation that can produce unbiased estimates of variance and covariance parameters. The REML method is recommended because it converges faster and is more robust to undersmoothing than GCV or AIC methods[91].

In R, GAMs can be fitted using the '*gam*' function in the '*mgcv*' package[92]. Each smooth term is fitted using *s(x)* in a *gam* formula. The default and most used smoother are penalised thin plate regression splines (TPRS). This is specified in *gam* as *s(x, bs="tp")*. TPRS are low rank isotropic smoothers of any number of covariates. 'Isotropic' means that rotation of the covariate coordinate system will not change the result of smoothing, and 'low rank' means that they have far fewer coefficients than there are data to smooth. They are the optimal smoother of any given basis dimension, and they do not require pre-specifying the 'knot' number and location as they are found during the procedure itself. TPRS tend to give the best mean squared error (MSE) performance but are slower to set up than the other smoothers[93]. Other commonly used smoother options in *mgcv* include cubic regression splines (*bs="cr"*), which needs to choose the knot location but have advantages in parameter interpretation.

For interaction between continuous covariates, isotropic multi-dimensional smooths specified via *s(x,y)* or tensor product smooths specified via *te, ti,* or *t2* terms can be used. A tensor product allows the creation of a new set of basis functions that allow for each marginal function to have its own marginal smoothness penalty, and different basis can be used in each marginal smooth. Tensor product smooths often have better performance than isotropic smooths when the covariates are not on the same scale, e.g. time and distance. Tensor product smooths are robust to different scales and results will remain the same with covariates measured in different units[94].

te smooths have one penalty per marginal basis that penalizes the average wiggliness in that term. *ti* smooths exclude the basis functions associated with the 'main effects' of the marginal smooths and interactions other than the highest order specified. It is more statistically stable than *te* when specifying models by decomposing main effects and

interactions, e.g. *ti(x)+ti(y)+ti(x,y)*. *t2* uses an alternative tensor product construction that leads to more penalties. It requires more smooth terms than *te*, but can be fit using other mixed effects software such as *lme4 and gamm4*.

After model fitting, it is important to perform model diagnostics using function *mgcv::gam.check* to visualise residuals, check model's convergence, and determine if the smoothing basis dimension *k* has been set too low.

To fit very large datasets, the '*bam*' function in the *mgcv* package can improve the performance. It uses less memory and can perform parallel computing, so it runs much faster than '*gam*'[95]. Its default uses a fast REML computation method, specified by '*method=fREML'*, and it can discretize covariate data into a smaller number of values to reduce the computation, specified by '*discrete=TRUE'*. It also uses a random subset of data to calculate the basis functions to save on memory. Therefore, '*bam*' function is preferred when fitting large datasets.

3.4.2 Linear mixed models

The linear mixed models (LMMs) are used to analyse non-independent, longitudinal, or hierarchical data. They are helpful where subjects have repeated measurements and intermittent missing at random (MAR) data, without requiring multiple imputations[96,97]. MAR means there might be systematic differences between the missing and observed values, but they can be explained by other observed variables. LMMs can handle MAR data through maximum likelihood estimation.

LMMs are an extension of the simple linear models to allow both fixed and random effects. Fixed effects are parameters that do not vary across groups, representing population-level (i.e. average) effects that should persist across participants, while random

effects are parameters that are themselves random variables, representing group-level effects that vary across levels of the grouping factor. Fixed effects should not have many levels and exhaust all possibilities, while random effects are usually multi-level categorical variables where the levels are a sample of all possibilities. There are two types of variability, random effect intercept, and random effect slope. Random effect intercepts allow the response to be higher or lower across different groups, and random effect slopes allow each group to have a different slope, thus the relationship between the response and the explanatory variable is different for each group. LMMs can also include the correlation between random effect intercept and slope using a multivariate Gaussian distribution. If the correlation is not included, it assumes that random effects are independently sampled from their respective distributions and can lead to a worse model fit. However, sometimes the data is not adequate for the reliable estimation of all correlation parameters among random effects, reducing the model complexity by forcing the correlation parameter as 0 can be helpful.

The general formula for LMMs is

$$Y = X\beta + Zu + \varepsilon$$

Y is an $N \times 1$ vector of the outcome variable, X is an $N \times p$ design matrix of the p predictor variables; β is a $p \times 1$ column vector of the fixed-effects coefficients; Z is the $N \times qJ$ design matrix for the q random effects and J groups; u is a $qJ \times 1$ vector of the random effects for J groups; ε is an $N \times 1$ column vector of the residuals.

I assume u follows a multivariate Gaussian distribution with $E(u)=0$ and variance G: $u \sim N(0, G)$. G is the variance-covariance matrix of the random effects. For example, if there are a random intercept and a random slope, then

$$G = \begin{bmatrix} \sigma^2_{int} & \sigma^2_{int,slope} \\ \sigma^2_{int,slope} & \sigma^2_{slope} \end{bmatrix}$$

I assume ε follows a normal distribution with E(ε)=0 and variance var(ε)=σ^2_{error}.

LMMs use maximum likelihood (ML) estimation method by maximising the multivariate Gaussian distribution function with respect to parameters β, σ^2_{int}, σ^2_{slope}, and σ^2_{error}. However, ML estimator is biased especially for high-dimensional data and smaller sample sizes due to using an unknown mean estimator. Restricted Maximum Likelihood (REML) can solve this problem by removing all the information about the mean estimator prior to minimizing the log-likelihood function, thus produce unbiased estimations. However, when comparing two models using likelihood ratio tests, REML requires exactly the same fixed effects specification, so ML should be used for model comparison if the fixed effects are different, and REML for estimating a single model.

In R, linear mixed models can be fitted using *lmer* function from the *lme4* package. The random effects are represented by vertical bars (|) separating expressions for design matrices from grouping factors. Two vertical bars (||) can be used to specify multiple uncorrelated random effects for the same grouping variable.

For example, 1 + (1 | g) represents 'random intercept with fixed mean'. 1 + x + (1 + x | g) represents 'Correlated random intercept and slope'. 1 + x + (1 + x || g) represents 'uncorrelated random intercept and slope'[98].

3.4.3 Latent class mixed models

Advantages of linear mixed models over linear models including their abilities to handle population heterogeneity by adding subject-to-subject variation as random effects. On top

of this, another method called latent class mixed model (LCMM) can incorporate additional heterogeneity which cannot be fully captured by the normally distributed random effects, and the overall population can be divided into K subpopulations which have their own trajectories[99].

When fitting the LCMM, latent class membership is defined by a discrete random variable c_i if the subject i belongs to latent class k (k=1, …, K). The variable c_i is latent. The probability of c_i is described by a multinomial logistic regression model according to covariates v_i.

$$P(c_i = k | X_{c_i}) = \frac{\exp(v_i \beta_k)}{\sum_{j=1}^{K} \exp(v_i \beta_j)}$$

β_k are the class-specific coefficients associated with covariates v_i that predict the latent class membership. β_1=0 for identifiability.

In contrast to standard linear mixed models, both fixed effects and the distribution of the random effects are class-specific. Each subpopulation has its own model describing the longitudinal data, which is represented by a linear mixed model with class effects:

$$Y_i|_{c_i=k} = X_{1i}\beta + X_{2i}\vartheta_k + Z_i\mu_{ik} + \varepsilon$$

The fixed effects are split into X_{1i} with common fixed effects β over classes and X_{2i} with class-specific fixed effects ν_k. The vector Zi is associated with individual random effects μ_{ik} whose distributions are class-specific[99,100].

The common trajectories within a subpopulation are captured through latent classes, while the variability among subjects in the same class is captured through random effects.

In R, the package *lcmm* could be used to estimate latent class mixed models for Gaussian longitudinal outcomes (*hlme* function) and curvilinear and ordinal univariate longitudinal outcomes (*lcmm* function)[99].

Argument 'fixed' defines the two-sided formula for the linear regression at the population level with the dependent variable (Y) on the left-hand side and the combination of covariates (X1 and X2) with fixed effects on the right-hand side. Argument 'random' defines the one-sided formula with the covariates having a random effect. 'mixture' indicates a one-sided formula with the subset of covariates having a class-specific fixed effect (X2), and 'classmb' provides the optional covariates explaining the latent class membership in the multinomial logistic model (Xc). 'ng' specifies the number of latent classes. Argument 'B' specifies the vector of initial values. One way of initializing is to use the parameters obtained from a baseline mixed model with ng=1[99]. Usually, models with different number of latent classes (2-7) are fitted and the optimal number of latent classes is selected by minimizing Bayesian information criteria (BIC).

3.4.4 Bayesian statistics vs frequentist statistics

There are two main approaches to statistical inferences, which are frequentist and Bayesian. Bayesian statistics are based on the Bayes' theorem, which describes the probability of an event based on prior knowledge of conditions that might be related to the event. The formula is[101]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A) is the prior probability of the outcome A, which is usually from the literature, historical data, or expert beliefs. P(A|B) is the posterior probability, which is calculated using the prior information and the new data B.

The main differences between Bayesian and frequentist statistics include[102]:

1) Bayesian statistics assumes the parameters of a model are random while the data is fixed, in contrast to the frequentist approach assumes parameters are fixed but the data is random. In other words, in Bayesian statistics the data are observed from real sample but parameters are unknown random variables and described probabilistically, while in frequentist statistics sampling is infinite and data are a repeatable random sample with a frequency, but underlying parameters remain constant during the repeatable sampling process.
2) In hypothesis testing, frequentist statistics estimate the desired confidence percentage in which some parameter is placed and then accept or reject the null hypothesis, while Bayesian statistics answer what is the probability of the hypothesis given the measured data and estimate the ratio of the probabilities of different hypotheses.
3) Bayesian statistics use prior information to infer the posterior estimates, while frequentist analysis is only driven by the data once a model has been decided.
4) The experimental design is not important in Bayesian statistics as it is based on prior and posterior distributions, while must be specified in advance for frequentist statistics and different null hypothesis could lead to different interpretation of the same outcome.
5) The interpretation of the results is more intuitive with a Bayesian analysis which can be interpreted as the probability that a hypothesis is true; the frequentist

approach calculates the probability of obtaining a test statistic that is the same or more extreme in another dataset than this dataset at hand given the hypothesis is true, which is denoted as a p-value.

6) The uncertainty around the estimates is called the 'confidence interval' for frequentist statistics and the 'credible interval' for Bayesian statistics. A 95% confidence interval means if to repeat the study many times, 95% of the calculated confidence intervals would capture the true mean, and therefore the chance for the true mean to be contained with the 95% CI of a particular study is 50% - it either contains it or not. In contrast, a 95% credible interval means directly there is a 95% probability of the true mean falling within the interval.

3.4.5 Bayesian linear mixed models

Compared with frequentist linear mixed models, two main advantages of the Bayesian linear mixed models are: 1) the prior knowledge about a parameter can be incorporated in the model and the sample size requirements are therefore less stringent than frequentist linear mixed models[103]; 2) complex models with a large number of random variance components, censored or missing observations, multilevel correlation structures can be fit[104]. In my studies, I chose to use Bayesian linear mixed models because although the number of patients was quite large, the number of antibody measurements per patients was small, thus the frequentist models would struggle to fit trajectories over time. Also, a certain proportion of antibody measurements were censored at the upper limit of quantification (see Section 3.3.1), so models allowing for censored outcome are needed. The drawbacks of the Bayesian model are: 1) it can be more computationally expensive than frequentist models; 2) as there is no single well-defined method for choosing a prior,

it can generate biased estimates if wrong priors are chosen, given the priors force the parameter estimates into a certain direction.

Frequentist LMMs are based on the maximum likelihood estimation (MLE), which is a method of estimating the model parameters by maximising the likelihood *L(y)* of observing the data *y,* while Bayesian LMMs are based on maximum a posterior probability (MAP) estimation, which assumes the data to be distributed with some likelihood L(y), and use prior probability to maximise posterior probability.

In R, I used the *brms* package to implement Bayesian LMMs, which uses the probabilistic programming language *Stan*. The model specification is similar to *lmer*, with random effects represented by vertical bars (|). A censored variable can be included in the model using the formula: y | cens(cen1) ~ 1, with 'cen1' a variable indicating whether the value is censored[105].

The prior distributions of the parameters can be explicitly specified using '*set_prior'* function, for example, a normal(0,10) represents a normal distribution with mean 0 and standard deviation 10, or a cauchy(1,2) represents a cauchy distribution with location 1 and scale 2. Usually, weakly informative priors are used because they can rule out unreasonable parameter values but are not too strong to rule out reasonable values. If there is a reasonably large amount of data in the model, the likelihood will dominate, and the prior will not be important. An example of a weakly informative prior is normal (0,1) for a parameter on unit scale, so there is no prior information on its direction while still incorporating some regularisation with a standard deviation of 1. A common strategy to check the plausibility of weakly informative priors is through the use of a prior predictive check, in which data are generated according to the prior predictive distribution to assess whether a prior is appropriate.

The correlations between random effects are estimated with Lewandowski-Kurowicka-Joe (LKJ)-Correlation prior[106] with parameter ζ. The LKJ distribution is a continuous probability distribution for a symmetric matrix and a commonly used prior for correlation or covariance matrices. If $\zeta = 1$ (the default) the density is uniform over correlation matrices, suggesting no prior information on whether there is a correlation or not. If $\zeta > 1$, non-zero correlations become less likely with values centred around 0, whereas $0 < \zeta < 1$ results in higher probabilities for non-zero correlations with both positive and negative correlations.

Parameters are estimated by Monte-Carlo Markov Chain (MCMC). MCMC is a sampling method for estimating model parameters that can handle more complex models. It is a stochastic procedure that repeatedly generates random samples (by Markov chain) that characterize the distribution of parameters (by Monte Carlo integration)[107]. A Markov chain draws random samples from more likely regions of the parameter space, which is proportionally to the probability (which is determined by the prior information). The stationary state of the chains is called convergence. Then, Monte Carlo integration is used to summarise statistics from the samples to obtain the distribution of the parameters of interest. The number of chains and iterations can be specified using '*n.chains'* and '*n.iter'* in brms[108]. A burn-in period is specified using '*warmup*'. Iterations during the burn-in period are those Markov chains wandering to the probable region of the sample space and should be discarded.

After fitting the model, MCMC diagnostics are needed for model checking and ensure the model is converged. MCMC trace plots and autocorrelation function (ACF) diagrams are traditional ways of diagnostics, using '*mcmc_trace'* and '*mcmc_acf*' functions in R.

1) Trace plots: The x axis is each iteration, and the y-axis is the drawn samples. If the different Markov chains in one model have good mixing and converge to the steady state distribution, the samples drawn should be random samples from the same range of the y-axis.
2) Autocorrelation plots: The x-axis is the number of lagged iterations, and the y-axis is the correlation of samples at each step of the chain with previous estimates of that variable. If it dies out after lag 1, samples from the stationary distribution are random and not dependent on the initial values.
3) Density plots: a posterior distribution of the parameters using kernel density estimation. It should be smooth and have only one peak of the density which is the mode of the distribution. Unexpected peaks or strange shapes may indicate poor convergence (**Figure *3.3***).

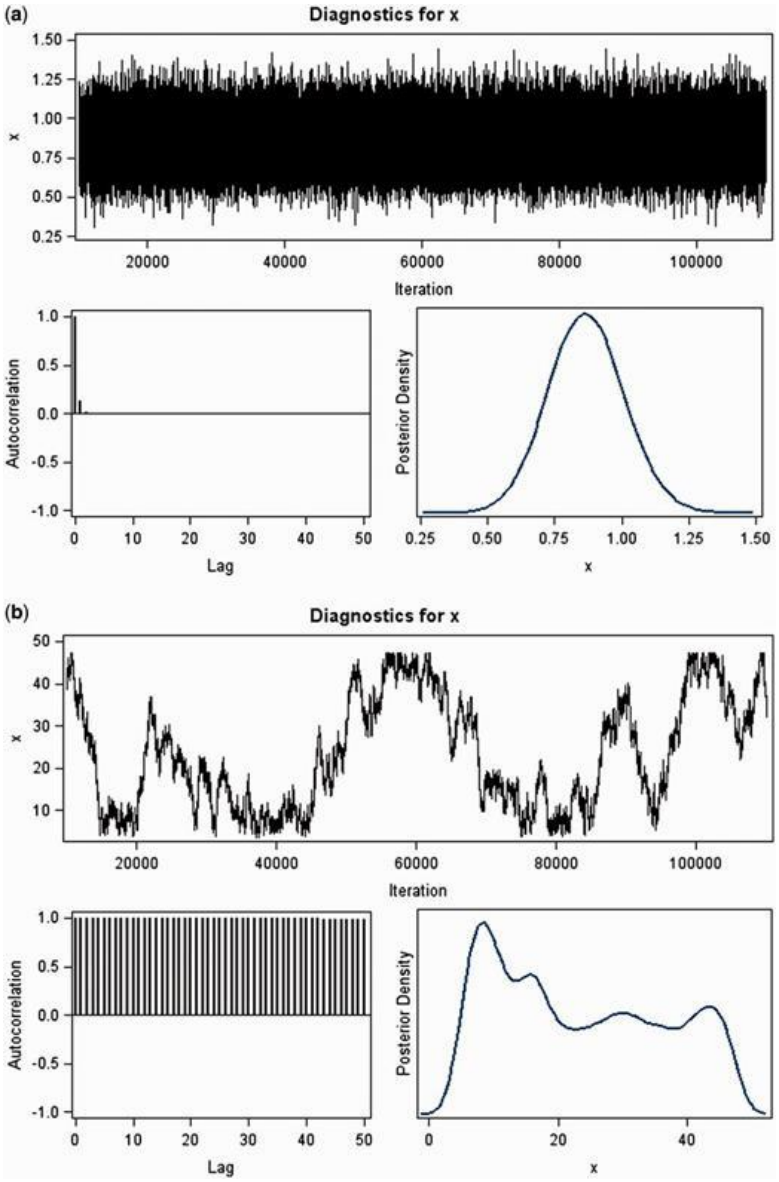


Figure 3.3 Diagnostic plots for MCMC convergence.

Note: Panel (a) shows an example of good convergence, while panel (b) shows an example of non-convergence. Taken from[107].

In addition, two statistics from the brms output can be used for model diagnostics[109]:

1) Gelman–Rubin statistic (rhat): the ratio of the (between-chain variance and within-chain variance)/within-chain variance. If difference chains converge to the stationary state, all chains should be similar and between-chain variation should be

close to 0, and the ratio should be close to 1. Usually, if the Gelman–Rubin statistic is <1.05, model is considered converged.

2) Effective sample size (ESS): the number of independent samples with the same estimation power as N autocorrelated samples. If there is a high correlation between samples, samples are dependent, and ESS will be lower. ESS should be as large as possible but >1000 is usually sufficient for estimation[105].

Chapter 4 Antibody responses to SARS-CoV-2 natural infection

This chapter is based on a previous publication: *Wei, J., Matthews, P.C., Stoesser, N. et al. Anti-spike antibody response to natural SARS-CoV-2 infection in the general population. Nat Commun 12, 6250 (2021).*

All data analysed and literature discussed in this chapter were up to June 2021, so the infections in this study were mainly Pre-Alpha and Alpha variants.

4.1 Introduction

To June 2021, over 170 million severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections and over 3 million associated deaths have been reported globally[10]. However, in the months following infection, reinfection is uncommon at that time and anti-spike SARS-CoV-2 antibodies are associated with protection[110–112]. The duration of post-infection immunity has important implications for the future of the pandemic and vaccination policy[113].

Seroconversion to viral spike and nucleocapsid antigens usually happens within 1-3 weeks after SARS-CoV-2 infection[114–116], with peak antibody levels achieved in 4-5 weeks[117,118]. However, 5-22% of individuals remain seronegative following infection[119–121]. Absence of seroconversion is more common following mild vs. severe disease (e.g. 22.2% vs. 2.6%, n=236[120]) and in asymptomatic vs. symptomatic individuals (11.0% vs 5.6% respectively, n=2,547[121]). However, the contribution of other factors, including viral load, has not been comprehensively assessed.

Among those who do seroconvert, data on the trajectory and duration of antibody responses to different SARS-CoV-2 antigens vary, partly reflecting assay-dependent

differences even where similar viral antigens are studied[122,123], as well as differences in the populations and disease groups investigated. Estimates for the half-life of anti-spike IgG antibodies (associated with neutralising activity[124]) vary from 36 to 244 days[123,125–131]. Similarly anti-nucleocapsid IgG half-lives have been estimated between 35 and 85 days[123,127,128,130].

Most studies have had small to moderate sample sizes or specific sub-populations; large-scale representative population studies are limited. In this chapter, I used data from the Office for National Statistics (ONS) COVID-19 Infection Survey (CIS), a large community-based survey representative of United Kingdom's general population, to investigate predictors of seroconversion following SARS-CoV-2 infection up to June 2021, identify anti-spike IgG antibody trajectories, and examine the peak and duration of IgG antibody responses, particularly considering the impact of demographic factors, PCR cycle threshold (Ct) values (inversely related to viral load), and self-reported symptoms on post-infection antibody responses.

4.2 Methods

4.2.1 Study population and data

Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and laboratory tests were already described in Chapter 3.

In this analysis, I included participants aged 16 years and over who had SARS-CoV-2 infection (defined by a positive PCR test) from 26 April 2020 to 14 June 2021. Since multiple positive swab tests could be obtained at follow-up visits, positive PCR tests were grouped into 'episodes'. I used the first episode (starting with the first positive PCR, or index positive) for each participant in the main analysis. Second episodes, defined by a

repeat PCR positive >120 days after the start of the first infection (associated with risk reductions for new positive episodes of similar magnitude to vaccination) were excluded.

Study visits occur on a fixed schedule, meaning that infection episodes could be identified up to 30 days or more after onset (as well as 'early' in some pre-symptomatic cases). As participants were told to obtain a test from the national testing programme if symptomatic, to improve my estimate of the start of each infection episode, I used study data linked by ONS to data on swab positivity from the English national testing programme (national testing programme data were not available for Scotland, Wales, and Northern Ireland). The national testing programme was intended for individuals with symptoms and contacts of those with symptoms (a substantial proportion report no symptoms), and so not all PCR-positive episodes in English study participants also have a positive test from the national testing programme. For this analysis I used the date of the first positive PCR test in the study or the national testing programme as the start of the episode, whichever came first. Ct values and gene positivity patterns were not available from the national testing programme when this study was conducted, and so these factors were obtained from PCR-positive samples in the ONS survey only.

I included all antibody measurements from 90 days before each participant's first swab positive date (index positive) through to 180 days after (approximate 95th percentile), to avoid undue influence from outliers at late time points. I also excluded all antibody measurements taken from 3 days after the first vaccination.

I used the Ct value as the proxy of viral burden, defined as the minimum from all positive swab tests in the infection episode and categorising at <30 to indicate moderate to higher viral burden. This threshold is used in the UK in algorithms for review of low-level positives at the laboratories where the PCR tests were performed and as a threshold for

attempting whole-genome sequencing[132]. Gene positivity pattern during the episode was classified as three groups: 1) a single *ORF1ab* gene or a single *N* gene positive; 2) Alpha (B.1.1.7) SARS-CoV-2 variant compatible (at least once positive for *ORF1ab+N* across the episode and never *S* positive), and 3) *S*-positive (*ORF1ab+N+S* or *ORF1ab+S* or *N+S* at least once across the episode). Participants with missing information on Ct values or gene positivity patterns or symptoms at all tests within the episode were excluded from analysis (n=133). Self-reported symptoms were those reported at any visit within 35 days after the index positive date or reported to the national testing programme. Fever, cough, loss of smell, and loss of taste were considered as 'classic symptoms'.

4.2.2 Statistical analysis

I first used latent class mixed models (LCMM) to identify distinct patterns of antibody response after natural infection, counting the date of the index positive in the survey as time 0. Restricted natural cubic splines (internal knots at -10, 30, 60 days, and boundary knots at -60 and 140 days) were used to model time since the index positive as the fixed effect. A random-effect intercept and random-effect slope on all time spline variables were added to account for variability between individuals. Age as a natural cubic spline (internal knots at 50 years, and boundary knots at 20, 80 years), presence of self-reported long-term health conditions, Ct value, and self-reported symptoms were included as covariates for class membership[133]. The number of classes, up to a maximum of 4, was determined by examining and comparing the shape of the class trajectories and measures of model fit using BIC.

I used Bayesian linear mixed interval censored models to estimate the decay in antibody responses from their peak level, excluding those who did not seroconvert, and any participant with a positive or equivocal antibody result strictly before their index positive

date (≥13 BAU/mL) (n=6) or a negative antibody measurement within 42 days of their first index positive (N=13) **(Figure *4.1*)**. Time zero (peak level) for this analysis was determined from the estimated trajectories for each class from the LCMM (see Results). I initially assumed an exponential fall in antibody levels over time, i.e., a linear decline on a $\log_2$ scale. This exponential decline model had been widely used in studies on antibody kinetics[125,134,135]. Population-level fixed effects, individual-level random effects for intercept and slope, and covariance between random effects were included in the model. The outcome was right-censored at 450 reflecting truncation of IgG values at 450 BAU/mL by the laboratory assay. I excluded a very small number of measurements (N=24) below 13 BAU/mL (likely reflecting mislabelled samples) to reduce the influence of outliers. I examined the association between peak levels and antibody half-lives with age, sex, ethnicity, reporting having long-term health conditions, Ct values, and self-reported symptoms. As sensitivity analysis, I also included a time-dependent variable to reflect pre-Alpha (before 16 Nov 2020) and Alpha (17 Nov 2020-16 May 2021) periods. Due to the short follow-up time when the Delta variant was dominant in this study (17 May 2021 onwards), there were no Delta period infections in the model. I tested for evidence of non-linearity in antibody declines on the log scale using two alterative models, a piecewise linear regression allowing for a biphasic exponential decline with knot at 28 days post-peak, and a model using natural cubic splines to allow for a more flexible non-linear fit. Model fits were compared using the leave-one-out cross-validation information criterion (LOOIC). Although the spline-based model provided a slightly better fit, extrapolating cubic polynomial models can lead to unstable estimates, so I used the biphasic model to predict times above different thresholds**.**

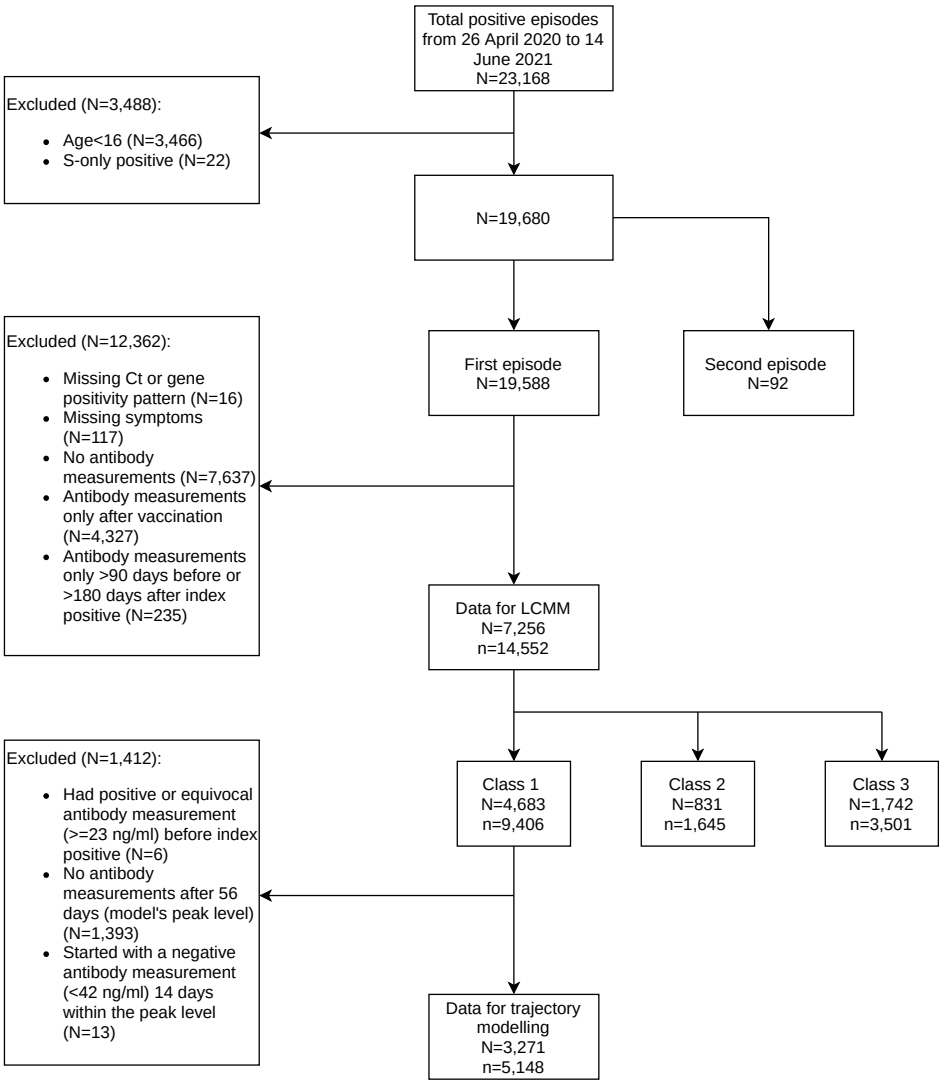


Figure 4.1 Flowchart of the study cohort.

Note: N represents the number of positive episodes/participants, and n represents the number of antibody measurements. LCMM=latent class mixed model. Class 1=’seroconverted in response to infection’, Class 2=’possible late/reinfection’, Class 3=’seronegative non-responders’, see ***Figure* 4.2**.

For each Bayesian linear mixed interval censored model, weakly informative priors were used **(Table *4.1*).** 4 chains were run per model with 4,000 iterations and a warm-up period of 2,000 iterations to ensure convergence, which was confirmed visually and by ensuring the Gelman-Rubin statistic was <1.05. 95% credibility intervals were calculated using highest posterior density intervals.

Table 4.1 Priors used in the Bayesian linear mixed interval censored models.

Model term	Priors
Intercept	normal (7.4, 2.8)
Slope	normal (0, 0.1)
Coefficient for change in intercept (age)	normal (0, 2)
Coefficient for change in slope (age)	normal (0, 0.1)
Coefficient for change in intercept (sex)	normal (0, 6)
Coefficient for change in slope (sex)	normal (0, 0.1)
Coefficient for change in intercept (ethnicity)	normal (0, 10)
Coefficient for change in slope (ethnicity)	normal (0, 0.2)
Coefficient for change in intercept (long-term health condition)	normal (0, 7)
Coefficient for change in slope (long-term health condition)	normal (0, 0.2)
Coefficient for change in intercept (Ct)	normal (0, 1)
Coefficient for change in slope (Ct)	normal (0, 0.1)
Coefficient for change in intercept (symptom)	normal (0, 6)
Coefficient for change in slope (symptom)	normal (0, 0.1)
Random effect SD: intercept	normal (0, 1)
Random effect SD: slope	normal (0, 0.01)
Random effect intercept & slope covariance	lkj_corr_cholesky (2)

As sensitivity analyses, I additionally used 225 and 280 as the censoring threshold for IgG levels and chose different starting points to examine robustness.

Data preparation was conducted using Stata MP16. All analyses were performed in R 3.6 using the following packages: tidyverse (version 1.3.0), brms (version 2.14.0), rstanarm (version 2.21.1), splines (version 3.6.1), lcmm (version 1.9.2), nnet (version 7.3-14), ggeffects (version 0.14.3), arsenal (version 3.4.0), cowplot (version 1.1.0), bayesplot (version 1.7.2).

4.3 Results

4.3.1 Study population

From 26 April 2020 to 14 June 2021, 467,450 participants had one or more throat and nose swab study results (median 10, IQR 8-12) during a median (IQR) 221 (141-251) days of follow-up. 19,588 (4.2%) participants ≥16 years were ever PCR-positive, 92 (0.5%) with a

second episode >120 days after their first PCR-positive result (median 149, IQR 134-174 days later). Analysis included the 7,256/19,588 (37%) participants with at least one anti-spike IgG antibody measurement within [-90,+180] days of the start of their first infection episode, who contributed 14,552 antibody measurements (median 2, IQR 1-3, range 1-10; excluding measurements from 3 days after first vaccination) **(Figure *4.1*).**

The median age of these 7,256 participants was 47 (IQR 34-59) years, and 3,874 (53.4%) were female **(Table *4.2*)**. 6,577 (90.6%) reported white ethnicity, 127 (1.8%) working in patient-facing healthcare, and 1,592 (21.9%) having a long-term health condition. Considering the minimum Ct value across all positive tests in the first infection episode, i.e. the maximum viral load, the median was 27 (IQR 19-32), with 4,420 (60.9%) having Ct<30. Three SARS-CoV-2 PCR target genes (ORF1ab, nucleocapsid protein (N), and spike protein (S)) were tested for in all participants: 1,505 (20.7%) were only positive on a single gene (ORF1ab or N); 2,822 (38.9%) were Alpha (B.1.1.7)-compatible (i.e. showed S gene target failure). 4,190 (57.7%) reported having any symptoms, with 2,773 (38.2%) reporting classic symptoms (fever, cough, loss of smell, or loss of taste). 5,169 (71%) participants only contributed antibody measurements after their index positive date.

Table 4.2 Characteristics of participants in classes identified from latent class mixed models for 7,256 participants infected with SARS-CoV-2.

Note: Continuous variables are presented graphically in **Figure 4.3**. Continuous variables were compared using Kruskal-Wallis tests, and categorical variables were compared using Chi-squared tests. Participants included in this study were not subject to influence from vaccination campaign as antibodies measured after vaccination were excluded from the analyses, and all included antibody measurements were after infection. 'Non-white' includes all ethnic groups except white ethnic groups, including black, Asian, and other minority ethnic groups. The term was used for conciseness reason throughout this thesis, but caution should be taken when using this term, and more appropriate phrase (e.g. 'all other ethnic groups combined') may be used in a broader context.

	Total (N=7256)	Class 1 'classical seroconversion' (N=4683)	Class 2 'possible late/reinfection' (N=831)	Class 3 'seronegative non responders' (N=1742)	p value (all)	P value (2 vs 1)	P value (3 vs 1)
Percentage	100%	64.5%	11.5%	24.0%			
Age					< 0.001	0.006	<0.001
Median	47	46	44	51			
IQR	34, 59	33, 58	32, 56	36, 65			
Sex					0.6	0.9	0.3
Female	3874 (53.4%)	2485 (53.1%)	440 (52.9%)	949 (54.5%)			
Male	3382 (46.6%)	2198 (46.9%)	391 (47.1%)	793 (45.5%)			
Ethnicity					0.001	0.2	0.002
White	6577 (90.6%)	4224 (90.2%)	738 (88.8%)	1615 (92.7%)			
Non-white	679 (9.4%)	459 (9.8%)	93 (11.2%)	127 (7.3%)			
Report working in patient facing healthcare					0.002	0.4	< 0.001
No	7129 (98.2%)	4590 (98.0%)	811 (97.6%)	1728 (99.2%)			
Yes	127 (1.8%)	93 (2.0%)	20 (2.4%)	14 (0.8%)			
Report having long-term health condition					< 0.001	0.002	0.004
No	5664 (78.1%)	3668 (78.3%)	690 (83.0%)	1306 (75.0%)			
Yes	1592 (21.9%)	1015 (21.7%)	141 (17.0%)	436 (25.0%)			
Minimum Ct value across the infection episode					< 0.001	< 0.001	< 0.001
Median	27	22	32	33			
IQR	19, 32	17, 28	30, 34	31, 34			
Ct positivity pattern (*S*-gene positivity)					< 0.001	< 0.001	< 0.001
OR+S/N+S/OR+N+S	2929 (40.4%)	2361 (50.4%)	221 (26.6%)	347 (19.9%)			
OR+N	2822 (38.9%)	2079 (44.4%)	321 (38.6%)	422 (24.2%)			
OR or N only	1505 (20.7%)	243 (5.2%)	289 (34.8%)	973 (55.9%)			
Self-reported symptoms					< 0.001	< 0.001	< 0.001
No	3066 (42.3%)	1044 (22.3%)	650 (78.2%)	1372 (78.8%)			
Yes	4190 (57.7%)	3639 (77.7%)	181 (21.8%)	370 (21.2%)			
Self-reported classic symptoms					< 0.001	< 0.001	< 0.001
No	4483 (61.8%)	2115 (45.2%)	742 (89.3%)	1626 (93.3%)			

	Total (N=7256)	Class 1 'classical seroconversion' (N=4683)	Class 2 'possible late/reinfection' (N=831)	Class 3 'seronegative non responders' (N=1742)	p value (all)	P value (2 vs 1)	P value (3 vs 1)
Yes	2773 (38.2%)	2568 (54.8%)	89 (10.7%)	116 (6.7%)			
Posterior class-membership probability (%)					< 0.001	< 0.001	< 0.001
Median	95	93	72	100			
IQR	80, 100	80, 98	60, 99	100, 100			
Symptoms & Ct value					< 0.001	< 0.001	< 0.001
No self-reported symptoms & Ct>=30	1779 (24.5%)	132 (2.8%)	528 (63.5%)	1119 (64.2%)			
No self-reported symptom & Ct<30	1287 (17.7%)	912 (19.5%)	122 (14.7%)	253 (14.5%)			
Self-reported symptoms & Ct>=30	1057 (14.6%)	641 (13.7%)	128 (15.4%)	288 (16.5%)			
Self-reported symptom & Ct<30	3133 (43.2%)	2998 (64.0%)	53 (6.4%)	82 (4.7%)			
Have >=2 positive swabs in the infection episode					< 0.001	< 0.001	< 0.001
No	4285 (59.1%)	1975 (42.2%)	628 (75.6%)	1682 (96.6%)			
Yes	2971 (40.9%)	2708 (57.8%)	203 (24.4%)	60 (3.4%)			
Days between first and last positive swab if >=2 positive swabs					< 0.001	0.002	< 0.001
Median	14	14	18	20			
IQR	7, 26	7, 26	10, 28	10, 29			
Have negative study swab before index positive					< 0.001	< 0.001	< 0.001
Yes	6203 (85.5%)	4032 (86.1%)	603 (72.6%)	1568 (90.0%)			
No	1053 (14.5%)	651 (13.9%)	228 (27.4%)	174 (10.0%)			
Days between previous negative and index positive					< 0.001	< 0.001	< 0.001
Median	26	25	29	25			
IQR	12, 35	13, 34	25, 60	7, 29			
Have national testing programme positive included within infection episode					< 0.001	< 0.001	< 0.001
No	5675 (78.2%)	3210 (68.5%)	756 (91.0%)	1709 (98.1%)			
Yes	1581 (21.8%)	1473 (31.5%)	75 (9.0%)	33 (1.9%)			
National testing programme positive before study index positive					< 0.001	< 0.001	< 0.001
No	6053 (83.4%)	3570 (76.2%)	764 (91.9%)	1719 (98.7%)			
Yes	1203 (16.6%)	1113 (23.8%)	67 (8.1%)	23 (1.3%)			
Days between national testing programme positive and study index positive					< 0.001	< 0.001	< 0.001
Median	7	6	15	16			
IQR	3, 14	3, 13	10, 24	10, 23			

4.3.2 Antibody trajectories following infection

A latent class analysis identified three classes of post-infection anti-spike IgG antibody responses: Class 1, ‘classical seroconversion’, Class 2 ‘possible late detection/reinfection’, and Class 3 ‘seronegative non-responders’. Class-membership probabilities were high, suggesting that participants’ responses could be reliably assigned to one of the three classes **(Table *4.2*, Figure *4.2*, Figure *4.3*, Figure *4.4*)**. Participants who showed classical seroconversion responses after infection comprised Class1 (n=4,683, 64.5%). These participants had risen in antibody levels over the 4-5 weeks following their first PCR-positive sample, followed by subsequent waning. Class 1 had lower Ct values (median [IQR] 22 [17-28] vs Class 2, 32 [30-34], Class 3, 33 [31-34]), and a higher percentage of reported symptoms (77.7% vs Class 2, 21.8%, Class 3, 21.2%) and classic symptoms (54.8% vs Class 2, 10.7%, Class 3, 6.7%). Class 1 also had a lower percentage of single gene positives (5.2% vs Class 2, 34.8%, Class 3, 55.9%). 57.8% had more than one positive swab test in their first infection episode and 23.8% had a positive test in national testing programme prior to their first study positive test, a significantly higher percentage than other classes (Class 2, 24.4%, 8.1%; Class 3, 3.4%, 1.3%) **(Table *4.2*).**

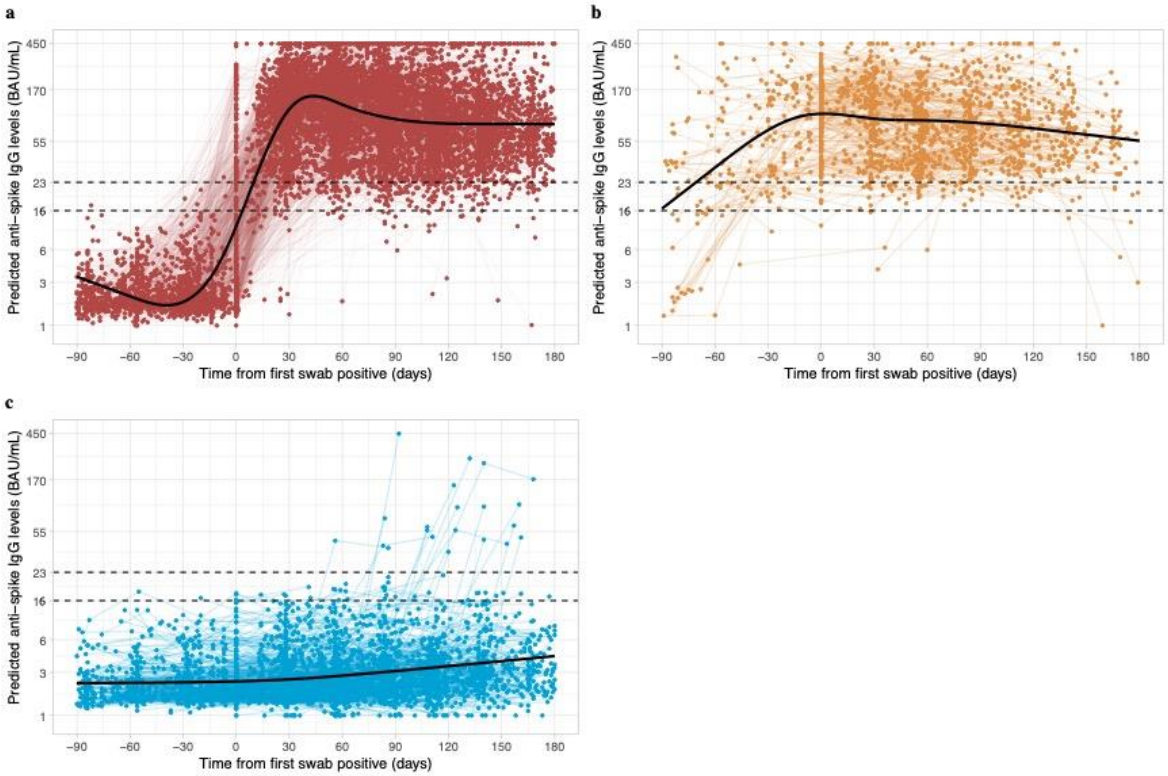


Figure 4.2 Individual trajectories for 7,256 participants infected with SARS-CoV-2 by class identified from latent class mixed models.

Note: **a,** Class 1, 'classical seroconversion in response to infection' (N=4,683, 64.5%). **b,** Class 2, 'possibly late/reinfection' (N=831, 11.5%). **c,** Class 3, 'seronegative non-responders' (N=1,742, 24.0%). Black dashed lines indicate the assay threshold for IgG positivity (23 BAU/mL; most strongly associated with prior infection[83]) and the level associated with 50% protection against reinfection[112] (16 BAU/mL). Restricted natural cubic splines (internal knots at -10,30,60 days, and boundary knots at -60 and 140 days) were used to model time (see methods). Distribution of the factors by class membership is shown in **Table 4.2.**

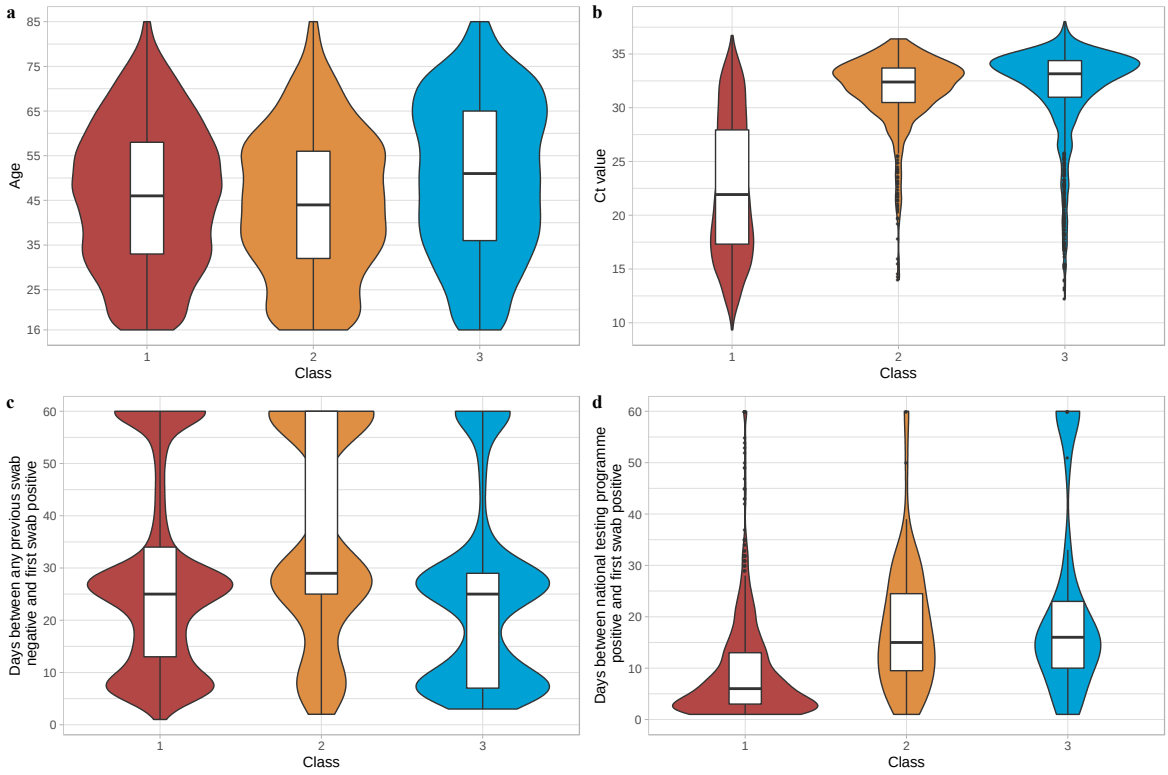


Figure 4.3 Distributions of continuous variables by classes identified from latent class mixed models.

Note: **a**, Participants' age. Participants in Class 3 were older than Class 1 and 2. **b**, Participants' minimum Ct value across the infection episode. Class 1 had a lower Ct value than Class 2 and 3. **c**, Days between any previous study swab negative test and the first swab positive test in the infection episode. More participants in Class 2 had a longer duration between previous negative swab and the index positive, supporting late detection. **d,** Days between any preceding national testing programme positive test and the first swab positive test (y-axis truncated at 60 days for visualisation). Class 1 had a shorter duration between the two tests than other two classes. Class 1='classical seroconversion in response to infection' (N=4,683, 64.5%), Class 2='possible late/reinfection' (N=831, 11.5%), Class 3='seronegative non-responders' (N=1,742, 24.0%). For the box and whisker inserts: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers.

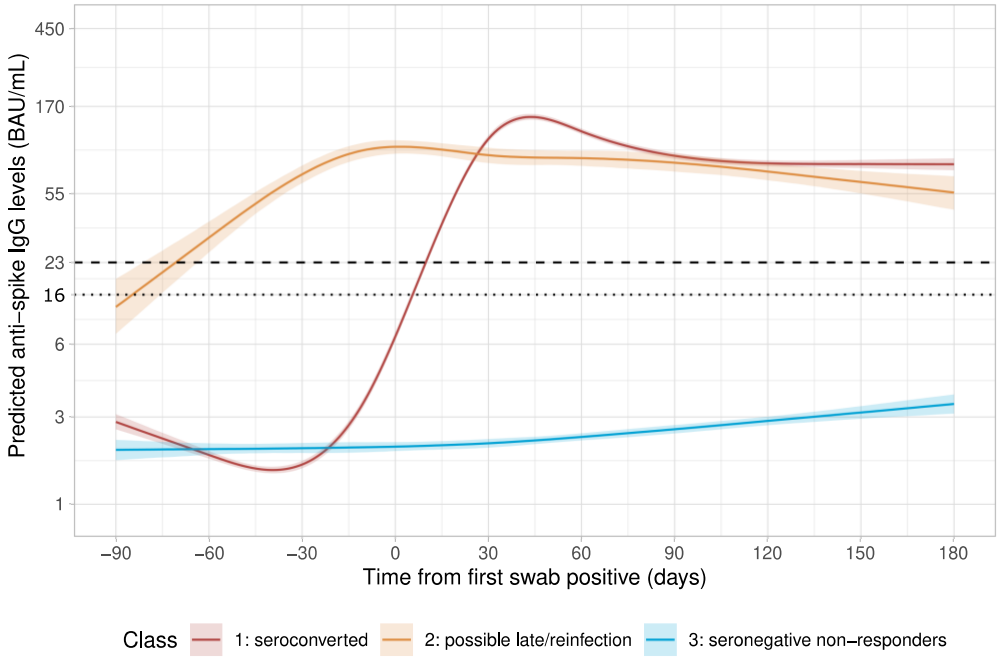


Figure 4.4 Predicted mean anti-spike IgG trajectories in participants with natural SARS-CoV-2 infection from latent class mixed models in 7,256 participants (with standard error of the prediction).

Note: Black dashed line indicates the assay threshold for IgG positivity (23 BAU/mL) and the dotted line indicates level associated with 50% protection against reinfection (16 BAU/mL). The 95% confidence interval are calculated by a Monte Carlo approximation of the posterior distribution of the predicted values. Class 1= 'classical seroconversion in response to infection' (64.5%, n=4683), Class 2= 'possible late/reinfection' (11.5%,n=831), Class 3='seronegative non-responders' (24.0%, n=1742).

Class 2 (n=831 (11.5%), 'possible late detection/reinfection') also had rises in anti-spike IgG levels but these started earlier, before the index positive PCR test. Their antibody levels reached a peak around the time of the index positive and then waned. This class likely partly reflects the study design, as study PCR testing was conducted at regular, usually monthly, intervals, irrespective of symptoms, with a proportion of missed visits (see Methods). Therefore, this group could represent those where infection was detected late, rather than reflecting any underlying biological difference. However, a subset may also represent reinfection with an undetected first infection. Supporting these possibilities, Ct values were higher (median [IQR] 32 [30-34]) than Class 1, self-reported symptoms

were less common (21.8%), as were multiple positive PCR tests (24.4%) **(Table *4.2*)**. For more participants, the index positive PCR was their first test in the study (27.4%); in the remainder, the median days since last negative was 29 days, higher than other classes and with considerable skew, with 369 (44.4%) being >31 days and 256 (30.8%) being >59 days, supporting late detection contributing to this group.

Lastly, 1,742 (24.0%) participants were assigned to Class 3 ('seronegative; non-responders'). Their IgG levels barely increased and were below the positivity threshold throughout (excepting 17 outlier individuals who appeared to mount a response >30 days after their index positive PCR test). Compared with Class 1, Class 3 had higher Ct values (median [IQR] 33 [31-34]), a lower percentage self-reporting symptoms (21.2%) or classic symptoms (6.7%). Very few had more than one positive swab in their first infection episode (3.4%) or an accompanying positive test in the national testing programme (1.9%) **(Table *4.2*)**. Whilst this class would be expected to be enriched for false-positives, of 1,742 participants in this class, 595 (34%) still had strong evidence for a true-positive PCR result (Ct ≤32 and ≥2 genes detected). Of the 5230 participants with Ct ≤32 and ≥2 genes detected, 595 (11%) were non-responders, 373 (7%) in Class 2, and 4259 (81%) in Class 1. Class 3 were also older **(Figure *4.3*)**, with fewer patient-facing healthcare workers (0.8%) and more participants with long-term health conditions (25.0%).

4.3.3 Predictors of non-responders

In the multinomial logistic regression model, independent predictors of remaining seronegative (Class 3) vs seroconverting (Class 1) were higher minimum Ct (i.e. lower viral load), not self-reporting symptoms, older age and not working in patient-facing healthcare **(Figure *4.5*, Table *4.3*)**, with no evidence of independent effects of sex, ethnicity, or long-term health conditions. For example, at the median age of 47 years (not

working in patient-facing healthcare), the Ct threshold at which seroconversion rates reached >90% were 26, 23 and 17 for those reporting classic symptoms, other symptoms or no symptoms (**Figure *4.5*b**). Excluding Ct from the model, there was still no evidence of independent effects of long-term health conditions, but non-white ethnicity was associated with lower odds of being in Class 3 (OR=0.70, 95%CI 0.55-0.90, p=0.005) than Class 1.

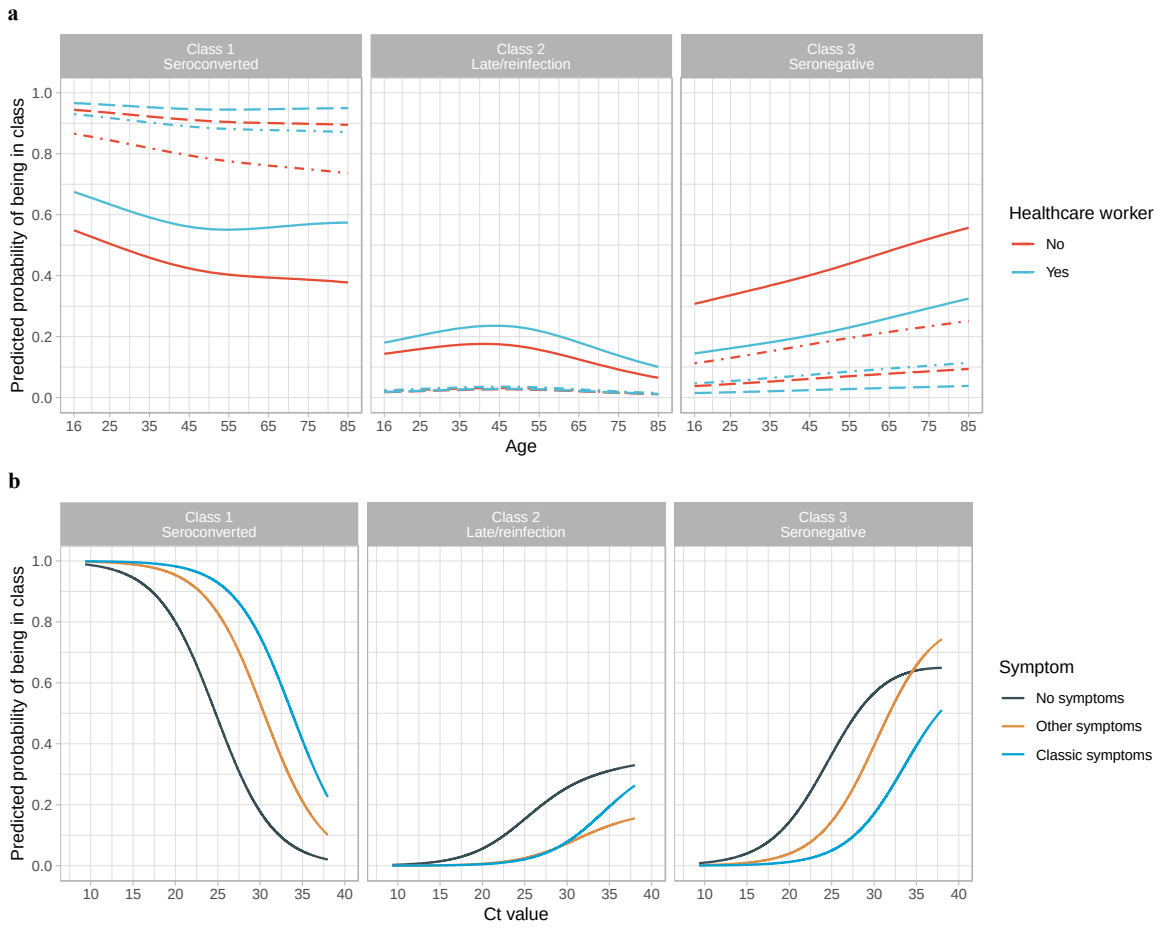


Figure 4.5 Predicted probability of being in Class 1 (classical seroconversion in response to infection), 2 (possible late/reinfection), and 3 (seronegative non-responders).

Note: **a,** By age and working in patient-facing healthcare, plotted at the reference category for other variables (female, white ethnicity, no long-term health condition, Ct=26, have only one positive swab test during the infection episode) and no symptoms (solid line), other symptoms (dash-dotted line), classic symptoms (dashed line). **b,** By Ct value and self-reported symptoms, plotted at the reference category for other variables (47-year-old, female, white ethnicity, no long-term health condition, not working in patient-facing healthcare, have only one positive swab test during the episode). Age was fitted using natural cubic spline with one internal knot placed at 50 years and two boundary knots at 20, 80 years. Full model results are shown in **Table 4.3**.

Table 4.3 Odds ratio with 95% confidence intervals from multinomial logistic regression, using Class 1 as the reference.

Note: See **Figure 4.5** for effect of age and other factors on the probability scale. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction; two-sided Wald p values are shown. Bold numbers indicate significant p values <0.05.

	Class 1 Classical aeroconversion	**Class 2 Possible late/reinfection**			**Class 3 Seronegative**		
		OR	95%CI	p-value	OR	95%CI	p-value
Age	reference	See **Fig. 2**					
Sex							
Female		1(ref)			1(ref)		
Male		0.98	0.81-1.18	0.8	0.88	0.74-1.04	0.1
Ethnicity							
White		1(ref)			1(ref)		
Non-white		1.03	0.75-1.40	0.9	0.80	0.60-1.08	0.2
Report having long-term health conditions							
No		1(ref)			1(ref)		
Yes		0.82	0.64-1.04	0.1	1.10	0.90-1.35	0.4
Report working in patient facing healthcare							
No		1(ref)			1(ref)		
Yes		1.02	0.52-1.96	1	0.39	0.18-0.83	**0.01**
Ct value (one unit higher)		1.35	1.32-1.39	**<0.001**	1.33	1.31-1.36	**<0.001**
Report having symptoms							
No symptoms		1(ref)			1(ref)		
Other symptoms		0.09	0.07-0.13	**<0.001**	0.23	0.19-0.29	**<0.001**
Classic symptoms		0.07	0.05-0.09	**<0.001**	0.07	0.06-0.09	**<0.001**
Have >=2 positive swabs in the infection episode							
No		1(ref)			1(ref)		
Yes		1.13	0.80-1.60	0.5	0.08	0.05-0.13	**<0.001**
Days between first and last positive (1 day longer)		0.99	0.98-1.00	0.1	1.00	0.98-1.02	0.8

To investigate associations with specific symptoms, I fitted a logistic regression model comparing only seroconversion (Class 1) vs non-response (Class 3), and omitting Ct and other test characteristics as these may mediate effects of symptoms. I found cough, loss of smell, fever, loss of taste, fatigue, headache, and sore throat were associated with lower odds of non-response, with cough (OR=0.20, 95%CI 0.15-0.25, p<0.001) and loss of smell (OR=0.21, 95%CI 0.13-0.33, p<0.001) most strongly associated. Results remained similar restricting seronegatives to those with stronger evidence of a true PCR-positive result (Ct ≤32 and ≥2 genes detected) **(Figure *4.6*, Table *4.4*,** with non-linear effect of age in **Figure *4.7*).** I additionally examined the association with specific comorbidities by incorporating them into the model but found no strong evidence of major impact **(Table *4.5*).**

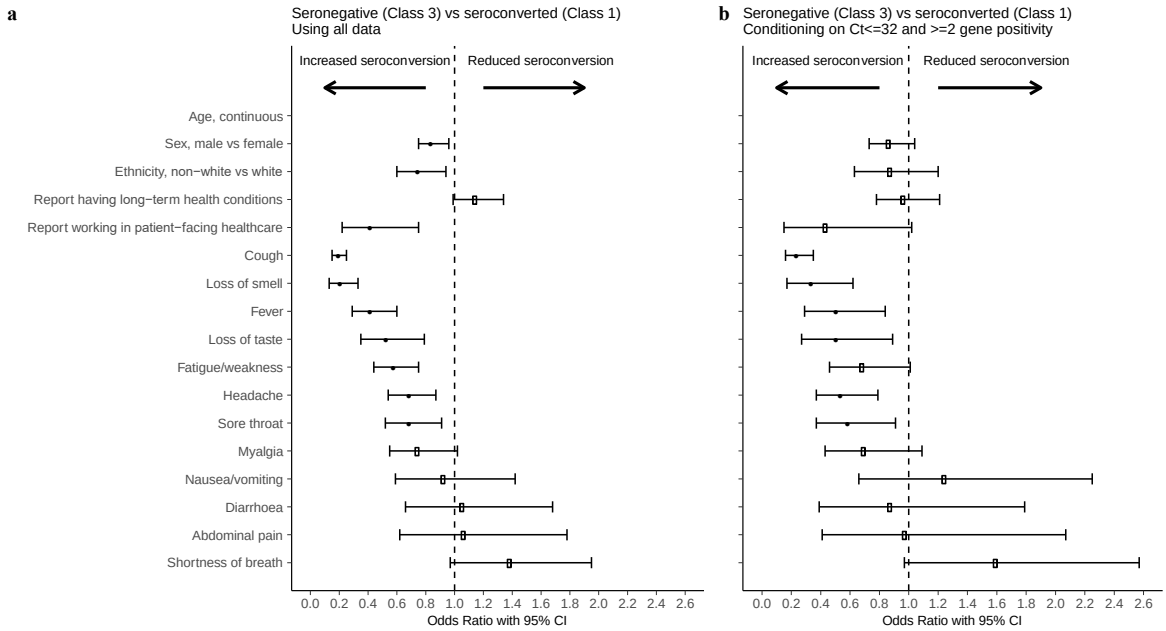


Figure 4.6 Odds ratio with 95% confidence intervals from logistic regression comparing seronegative vs seroconverting (Class 3 vs Class 1) using demographic factors and individual symptoms that would be available without a positive test result.

Note: **a**, Using all data from Class 3 (N=1,742) vs Class 1 (N=4,683). **b,** Restricting Class 3 to those with Ct value ≤32 and ≥2 genes detected (N=595) to decrease the impact of potential false positive swab tests. Age was fitted using natural cubic spline with one internal knot placed at 50 years and two boundary knots at 20, 80 years. Effect of age is presented in **Figure 4.7**. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction; solid dots indicate estimates with p-values<0.05, while hollow dots indicate those with p-values ≥0.05. Numbers of odds ratio, 95%CI, and p-values are presented in **Table 4.4**.

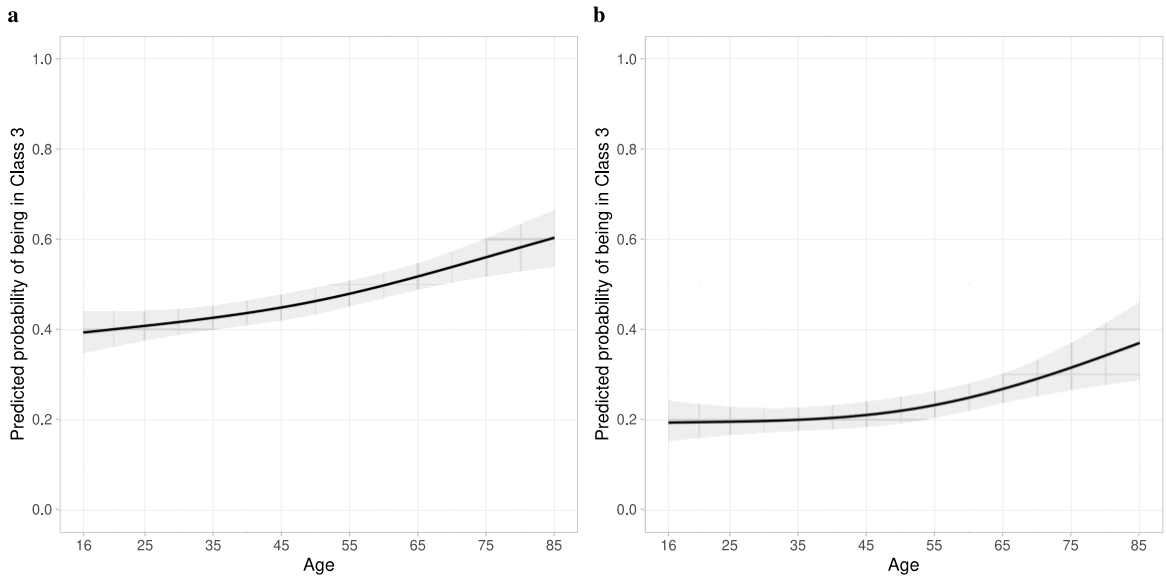


Figure 4.7 Predicted probability of being in Class 3 (seronegative) compared with Class 1 (seroconverted) by age.

Note: Plotted at the reference category for other variables (female, white ethnicity, no long-term health condition, not working in patient-facing healthcare, no self-reported symptoms of any kind)**. a,** Using all data from Class 3 (N=1,742). **b**, Conditioning on Ct value ≤32 and ≥2 genes detected (N=595) to decrease the impact of potential false positive swab tests. Age was fitted using natural cubic spline with one internal knot placed at 50 years and two boundary knots at 20, 80 years. Full model's results are presented in **Table 4.4, Figure 4.6**. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction.

Table 4.4 Odds ratio with 95% confidence intervals from logistic regression comparing seronegative vs seroconverting (Class 3 vs Class 1) using demographic factors and individual symptoms that would be available without a positive test result.

Note: **(A)** Using all data from Class 3 (N=1,742) vs Class 1 (N=4,683). **(B)** Restricting Class 3 to those with Ct value ≤32 and ≥2 genes detected (N=595) to decrease the impact of potential false positive swab tests. Age was fitted using natural cubic spline with one internal knot placed at 50 years and two boundary knots at 20, 80 years. Effect of age is presented in **Figure 4.7**. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction; two-sided Wald p values are shown. Bold numbers indicate significant p values <0.05. Odds ratios and 95% confidence intervals are visually presented in **Figure 4.6.**

	Class 1 'classical seroconversion'	**Class 3 (A: Using all data): 'seronegative non responders'**			**Class 3 (B: Conditioning on Ct ≤32 and ≥2 gene positivity)**		
		OR	95%CI	p-value	OR	95%CI	p-value
Age				**<0.001**			**<0.001**
Sex (Male vs Female)		0.84	0.75-0.96	**0.008**	0.87	0.73-1.04	0.1
Ethnicity (Non-White vs White)		0.75	0.60-0.94	**0.01**	0.88	0.63-1.20	0.4
Report having long-term health conditions (Yes vs No)		1.15	0.99-1.34	0.06	0.97	0.78-1.21	0.8
Report working in patient facing healthcare (Yes vs No)		0.42	0.22-0.75	**0.005**	0.44	0.15-1.02	0.09
Cough (Yes vs No)		0.20	0.15-0.25	**<0.001**	0.24	0.16-0.35	**<0.001**
Loss of smell (Yes vs No)		0.21	0.13-0.33	**<0.001**	0.34	0.17-0.62	**<0.001**
Fever (Yes vs No)		0.42	0.29-0.60	**<0.001**	0.51	0.29-0.84	**0.01**
Loss of taste (Yes vs No)	reference	0.53	0.35-0.79	**0.002**	0.51	0.27-0.89	**0.03**
Fatigue/weakness (Yes vs No)		0.58	0.44-0.75	**<0.001**	0.69	0.46-1.01	0.06
Headache (Yes vs No)		0.69	0.54-0.87	**0.003**	0.54	0.37-0.79	**0.002**
Sore throat (Yes vs No)		0.69	0.52-0.91	**0.01**	0.59	0.37-0.91	**0.02**
Myalgia (Yes vs No)		0.75	0.55-1.02	0.07	0.70	0.43-1.09	0.1
Nausea/Vomiting (Yes vs No)		0.93	0.59-1.42	0.7	1.25	0.66-2.25	0.5
Diarrhoea (Yes vs No)		1.06	0.66-1.68	0.8	0.88	0.39-1.79	0.7
Abdominal pain (Yes vs No)		1.07	0.62-1.78	0.8	0.98	0.41-2.07	1
Shortness of breath (Yes vs No)		1.39	0.97-1.95	0.07	1.60	0.97-2.57	0.06

Table 4.5 Odds ratio with 95% confidence intervals from a logistic regression model comparing seronegative vs seroconverting participants (Class 3 vs Class 1) using demographic factors, individual symptoms, and comorbidities.

Note: **(A)** Using all data from Class 3 (N=1,383) vs Class 1 (N=4,032). **(B)** Restricting Class 3 to those with Ct value ≤32 and ≥2 genes detected (N=487) to decrease the impact of potential false positive swab tests. Age was fitted using natural cubic spline with one internal knot placed at 50 years and two boundary knots at 20, 80 years. Comorbidities were obtained by linkage to the General Practice Extraction Service (GPES) Data for Pandemic Planning and Research and Hospital Episode Statistics (HES) via the NHS number by the ONS. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction; two-sided Wald p values are shown. Bold numbers indicate significant p values <0.05.

	Class 1 'classical seroconversion'	**Class 3 (A: Using all data): 'seronegative non responders'**			**Class 3 (B: Conditioning on Ct ≤32 and ≥2 gene positivity)**		
		OR	95%CI	p-value	OR	95%CI	p-value
Age	reference			**<0.001**			**<0.001**
Sex (Male vs Female)		0.84	0.73-0.97	**0.02**	0.85	0.69-1.04	0.1
Ethnicity (Non-White vs White)		1.26	0.98-1.62	0.07	1.17	0.82-1.69	0.4
Report having long-term health conditions (Yes vs No)		1.19	1.00-1.43	0.05	1.06	0.81-1.38	0.7
Report working in patient facing healthcare (Yes vs No)		0.42	0.21-0.84	**0.01**	0.35	0.11-1.13	0.08
Cough (Yes vs No)		0.18	0.13-0.24	**<0.001**	0.20	0.13-0.33	**<0.001**
Loss of smell (Yes vs No)		0.21	0.12-0.36	**<0.001**	0.31	0.15-0.66	**0.002**
Fever (Yes vs No)		0.38	0.24-0.58	**<0.001**	0.40	0.21-0.77	**0.006**
Loss of taste (Yes vs No)		0.45	0.28-0.72	**<0.001**	0.47	0.23-0.95	**0.03**
Fatigue/weakness (Yes vs No)		0.58	0.43-0.80	**<0.001**	0.70	0.44-1.11	0.1
Headache (Yes vs No)		0.64	0.48-0.86	**0.003**	0.60	0.39-0.94	**0.03**
Sore throat (Yes vs No)		0.63	0.46-0.88	**0.006**	0.55	0.32-0.92	**0.02**
Myalgia (Yes vs No)		0.83	0.59-1.18	0.3	0.66	0.38-1.15	0.1
Nausea/Vomiting (Yes vs No)		0.84	0.50-1.43	0.5	1.11	0.52-2.35	0.8
Diarrhoea (Yes vs No)		0.95	0.53-1.72	0.9	0.74	0.27-2.01	0.6
Abdominal pain (Yes vs No)		0.96	0.50-1.83	0.9	0.96	0.35-2.63	0.9
Shortness of breath (Yes vs No)		1.39	0.93-2.10	0.1	1.39	0.76-2.53	0.3
Angina (Yes vs No)		0.96	0.43-2.14	0.9	1.04	0.38-2.82	0.9
Asthma (Yes vs No)		0.84	0.66-1.06	0.1	0.93	0.67-1.30	0.7
Atrial fibrillation (Yes vs No)		1.07	0.59-1.94	0.8	1.04	0.46-2.35	0.9
Cancer (Yes vs No)		1.07	0.80-1.43	0.6	1.03	0.69-1.56	0.9
Chronic kidney disease (Yes vs No)		1.13	0.70-1.84	0.6	0.94	0.46-1.94	0.9
Chronic liver disease (Yes vs No)		0.88	0.39-1.99	0.8	0.92	0.29-2.94	0.9

	Class 1 'classical seroconversion'	Class 3 (A: Using all data): 'seronegative non responders'			Class 3 (B: Conditioning on Ct ≤32 and ≥2 gene positivity)		
COPD (Yes vs No)		0.69	0.40-1.17	0.2	1.37	0.72-2.60	0.3
Coronary heart disease (Yes vs No)		0.99	0.40-2.43	1	1.19	0.39-3.66	0.8
Frailty (Yes vs No)		0.64	0.41-1.01	0.05	0.48	0.24-0.98	**0.04**
Heart failure (Yes vs No)		0.97	0.37-2.60	1	0.93	0.26-3.35	0.9
Hypertension (Yes vs No)		0.87	0.64-1.17	0.3	0.81	0.52-1.24	0.3
Myocardial infarction (Yes vs No)		1.45	0.61-3.48	0.4	1.21	0.37-3.98	0.8
Osteoporosis (Yes vs No)		2.00	1.07-3.73	**0.03**	1.33	0.52-3.42	0.6
Peripheral arterial disease (Yes vs No)		0.66	0.32-1.37	0.3	0.63	0.22-1.81	0.4
Rheumatoid arthritis (Yes vs No)		0.82	0.32-2.12	0.7	0.89	0.23-3.53	0.9
Stroke (Yes vs No)		1.11	0.47-2.62	0.8	0.53	0.11-2.59	0.4
Transient ischaemic attack (Yes vs No)		0.77	0.32-1.84	0.6	1.50	0.50-4.50	0.5
Type 1 diabetes (Yes vs No)		3.06	0.95-9.85	0.06	3.10	0.52-18.55	0.2
Type 2 diabetes (Yes vs No)		1.29	0.76-2.18	0.3	0.94	0.43-2.05	0.9
Overweight (BMI 25 to <30 kg/m^2) (Yes vs No)		0.73	0.61-0.88	**<0.001**	0.80	0.61-1.05	0.1
Obese (BMI ≥30 kg/m^2) (Yes vs No)		0.84	0.68-1.03	0.1	0.88	0.65-1.19	0.4
Antihypertensive medication (Yes vs No)		1.00	0.76-1.33	1	0.95	0.64-1.42	0.8
Diabetes medication (Yes vs No)		0.47	0.26-0.86	**0.01**	0.45	0.17-1.15	0.09
Corticosteroids (Yes vs No)		1.53	0.87-2.71	0.1	0.83	0.34-2.04	0.7
Immunosuppressants (Yes vs No)		0.79	0.31-2.05	0.6	0.96	0.24-3.81	1

4.3.4 Determinants of the peak and half-life of antibody responses

In those who showed a classical antibody response, i.e. Class 1, I estimated anti-spike IgG peak antibody levels and half-life post-infection. Those in Class 2 where the timing of first infection was unclear and those who remained seronegative in Class 3 were not included because their antibody trajectories followed different patterns **(Figure *4.2*).** I estimated trajectories from 56 days after the first positive in the infection episode, when the IgG levels were close to the maximum level with high data completeness **(Figure *4.8*)**. 3,271 participants were included in this analysis, contributing 5,148 antibody measurements (interval censored at an assay upper limit of 450 BAU/mL), median (IQR) [range] 1 (1-2) [1-5] per participant. Using a Bayesian linear mixed model, assuming antibody levels fell exponentially (i.e. linearly on the log scale, in line with previous studies [125,134,135]) and accounting for variation in individuals' peak levels and half-lives using correlated random effects, the estimated mean anti-spike IgG half-life was 184 days (95% credibility interval, Crl 163-210), and peak level was 113 BAU/mL (95%Crl 106-117) (**Figure *4.9***). Estimated peak levels varied substantially between participants, ranging from 23 to 777 BAU/mL **(Figure *4.10*a)**. Longer half-lives were correlated with lower peak levels **(Figure *4.10*b)** (Spearman's rank coefficient=-0.50, $p<0.0001$; correlation between random intercept and slope -0.26). Results were similar in sensitivity analyses starting modelling from different times and using different interval censoring thresholds (280, 225 BAU/mL) (**Table *4.6***).

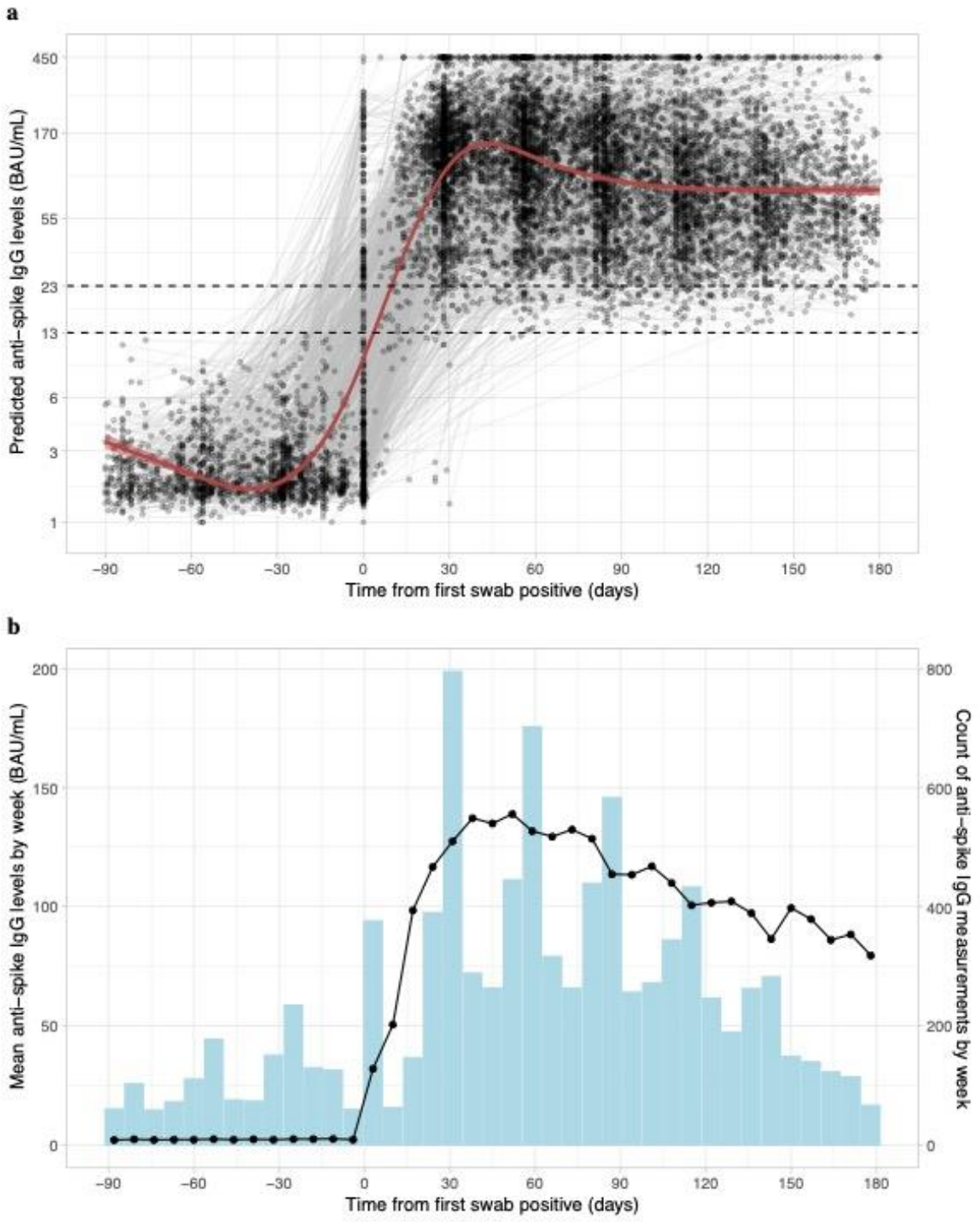


Figure 4.8 Determination of the 'peak level' for modelling anti-spike IgG trajectories.

Note: **a,** Class trajectory and individual trajectories for participants included in the model from Class 1. **b,** Mean anti-spike IgG levels and count of IgG measurements by week after index positive test. Around 56 days (8 weeks), the number of IgG measurements is high, and the mean IgG levels are close to the highest level.

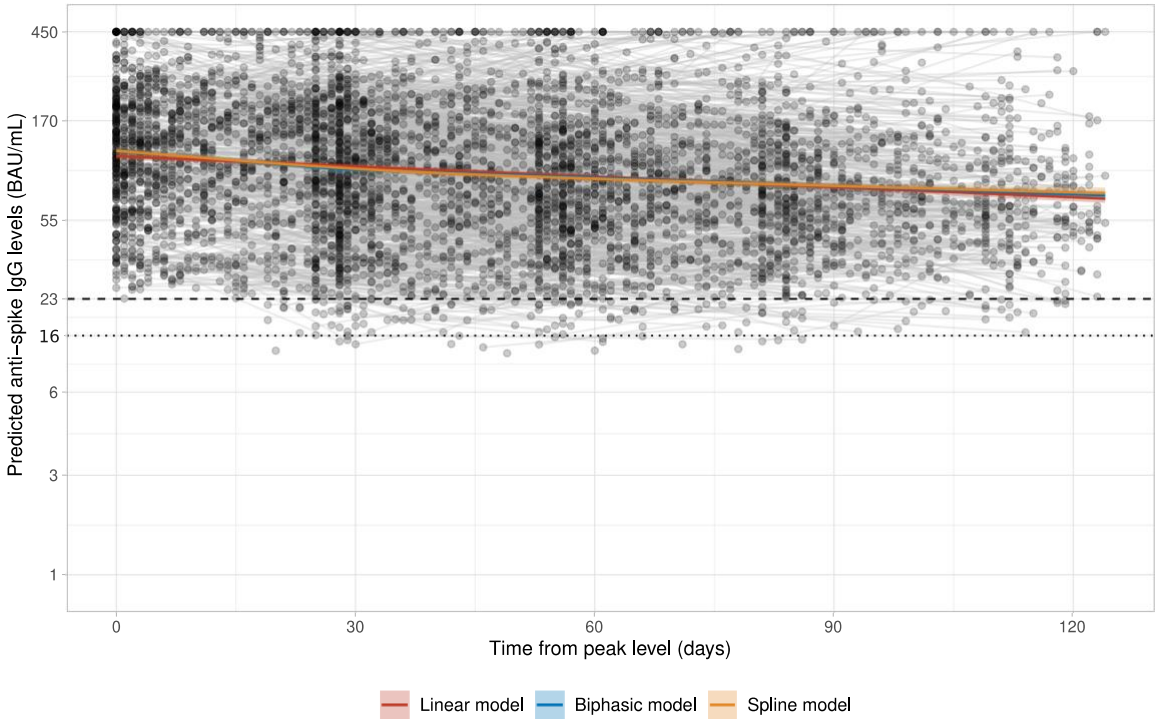


Figure 4.9 Estimated mean trajectory of anti-spike IgG antibody levels and individual trajectories in 3,271 participants in Class 1.

Note: The timing of the peak level 56 days after the first positive swab is determined from the latent class mixed model. Estimated trajectories from three models are presented: the model assuming a linear decline in log2 scale (red line), the biphasic exponential model (blue line), and the model using splines (orange line). For the biphasic model, knot is placed at 28 days. For the spline model, time is fitted using natural cubic splines with internal knots at 30, 70 and boundary knots at 5, 110. The posterior mean and 95% credibility interval are shown. Black dashed line indicates the assay threshold for IgG positivity (23 BAU/mL) and the dotted line the level associated with 50% protection against reinfection (16 BAU/mL).

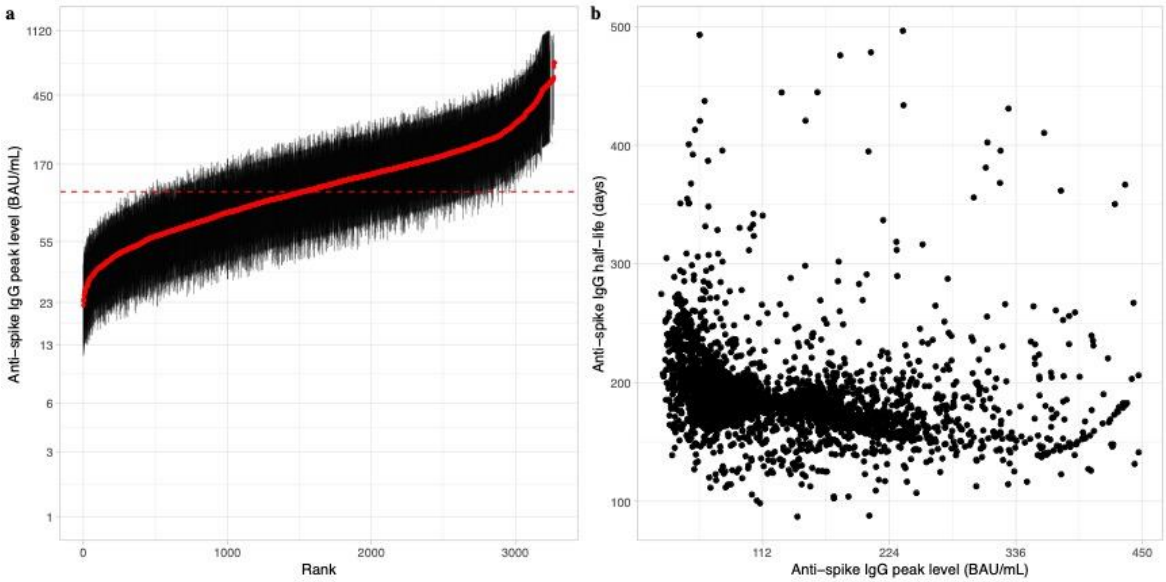


Figure 4.10 Posterior estimations for individual anti-spike IgG half-lives and peak level.

Note: **a,** Estimated peak anti-spike IgG level with 95% CrI for all participants (N=3,271) by rank. The red dot shows the mean level in each individual. The red dashed line shows the overall mean level (113 BAU/mL). **b,** Relationship between estimated IgG peak levels with half-lives. X and Y axes are truncated at 450 and 500 for visualization (Two-sided Spearman correlation test. Spearman's rank coefficient=-0.5, p<0.0001).

Table 4.6 Sensitivity analysis using different starting point (28, 84 days) and different censoring threshold (225, 280 BAU/mL) for anti-spike IgG trajectory modelling (peak and half-life estimation).

		Peak level (BAU/mL)			**Half-life (days)**		
		Posterior mean	95% CrI		Posterior mean	95% CrI	
Model from 56 days after index positive	Censored at 450 (base case)	113	106	117	184	163	210
	Censored at 280	108	106	112	187	168	209
	Censored at 225	106	102	108	192	172	215
Model from 28 days after index positive	Censored at 450	126	119	130	189	171	211
	Censored at 280	121	117	124	184	168	202
	Censored at 225	117	114	120	187	172	204
Model from 84 days after index positive	Censored at 450	99	92	103	225	182	287
	Censored at 280	94	91	98	233	192	290
	Censored at 225	92	88	95	242	199	305

In the multivariable linear mixed model, age, ethnicity, and Ct values were independently associated with IgG peak levels (model intercept), while sex and ethnicity were independently associated with IgG half-life (model slope) **(Table *4.7*, Figure *4.11*;**

posterior checks and MCMC diagnostics in **Table *4.8*, Figure *4.12*, Figure *4.13*)**. Conditional on having seroconverted (which occurred at lower rates in older individuals), older age was associated with higher IgG peak levels (adjusted 10 BAU/mL higher (95%Crl 7-16) per 10 years older). Males had a shorter half-life than females (adjusted 77 days shorter, 95%Crl 23-178). Non-white participants had higher IgG peak levels (adjusted 46 BAU/mL higher (95%Crl 31-63) than white participants, but a shorter half-life (adjusted 75 days shorter, 95% Crl 1-181). Higher Ct values (i.e. lower viral burden) were associated with a slightly higher peak level (adjusted 1 BAU/mL higher (95% Crl 0-1) per 1 unit higher). Conditional on inclusion in the analysis, i.e. seroconversion, I did not find any evidence of effect of reported long-term health conditions or self-reported symptoms on either IgG peak levels or half-life. In a sensitivity analysis, I did not find effects of time period (pre-Alpha vs Alpha) on either IgG peak levels or half-life (**Table *4.9***).

Table 4.7 Posterior mean and 95% credibility intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the univariable and multivariable models in 3,271 participants in Class 1.

Note: The reference category in the multivariable model is: 43-year-old, female, white ethnicity, no long-term health conditions, Ct value=22, and no self-reported symptoms. * indicates where the multivariable 95% CrI excludes 0 (no effect).

		Univariable model			**Multivariable model**		
		Posterior mean	95% CrI		Posterior mean	95% CrI	
Baseline	Peak level (Intercept) (BAU/mL)	113	106	117	103	88	112
	IgG half-life (slope) (days)	184	163	210	233	161	364
Age	Peak level: 43 years (median)	112	105	115			
	IgG half-life: 43 years (median)	204	177	237			
	Change in peak level: per 10-year older	10	7	12	10*	7	13
	Change in half-life: per 10-year older	-2	-20	21	-8	-36	23
Sex	Peak level: Female	111	102	117			
	IgG half-life: Female	232	189	291			
	Change in peak level: Male	4	-3	12	3	-4	11
	Change in half-life: Male	-79	-141	-30	-77*	-178	-23
Ethnicity	Peak level: White	110	103	114			
	IgG half-life: White	190	166	219			
	Change in peak level: Non-white	39	22	55	46*	31	63
	Change in half-life: Non-white	-46	-92	13	-75*	-181	-1
Long term health conditions	Peak level: No	111	102	115			
	IgG half-life: No	186	163	213			
	Change in peak level: Yes	13	3	24	7	-1	17
	Change in half-life: Yes	11	-50	100	10	-93	162
Ct value	Peak level: 22 (median)	113	106	117			
	IgG half-life: 22 (median)	184	163	209			
	Change in peak level: per 1 unit higher	1	0	1	1*	0	1
	Change in half-life: per 1 unit higher	0	-4	4	0	-7	7
Symptom	Peak level: No	115	101	124			
	IgG half-life: No	154	124	197			
	Change in peak level: Yes	-2	-11	8	1	-8	9
	Change in half-life: Yes	41	-8	86	57	-64	172

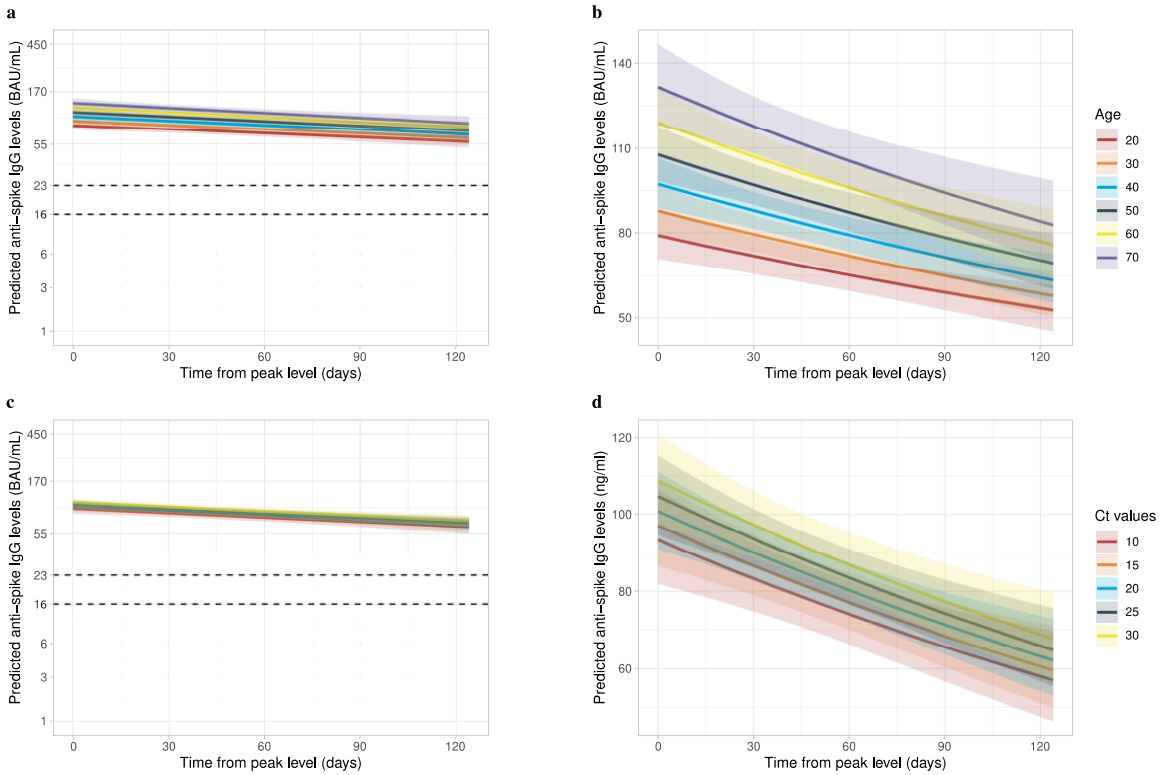


Figure 4.11 Posterior mean trajectories with 95% credibility intervals of anti-spike IgG after the peak level by age (a, b) and Ct values (c, d) in the multivariable model in 3,271 participants.

Note: Panel a and c show the relationship on the log10 scale (used for modelling). Black dashed lines indicate the assay threshold for IgG positivity (23 BAU/mL) and the level associated with 50 % protection against reinfection (16 BAU/mL). Panel b and d show the relationship on the original absolute scale.

Table 4.8 Model coefficients and MCMC diagnostics for the univariable and multivariable models.

Note: The reference category in the multivariable model is: 43-year-old, female, white ethnicity, no long-term health conditions, Ct value=22, and no self-reported symptoms.

		Estimate	95%Crl		Rhat
Baseline	Intercept	7.664	7.615	7.712	1.00
	time	-0.005	-0.006	-0.005	1.00
	sigma	0.517	0.494	0.540	1.00
	sd(Intercept)	1.041	0.993	1.090	1.00
	sd(time)	0.004	0.002	0.006	1.01
	cor(Intercept,time)	-0.258	-0.417	-0.047	1.00
Age	Intercept	7.641	7.593	7.688	1.00
	time	-0.005	-0.006	-0.004	1.00
	age	0.119	0.087	0.151	1.00
	time:age	0.000	-0.001	0.000	1.00
	sigma	0.516	0.494	0.540	1.02

		Estimate	95%Crl		Rhat
	sd(Intercept)	1.022	0.976	1.070	1.01
	sd(time)	0.004	0.002	0.006	1.04
	cor(Intercept,time)	-0.241	-0.401	-0.034	1.01
Sex	Intercept	7.638	7.570	7.704	1.00
	time	-0.004	-0.005	-0.003	1.00
	sex	0.053	-0.043	0.149	1.00
	time:sex	-0.002	-0.004	-0.001	1.00
	sigma	0.518	0.495	0.542	1.02
	sd(Intercept)	1.039	0.991	1.088	1.01
	sd(time)	0.004	0.002	0.006	1.04
	cor(Intercept,time)	-0.262	-0.433	-0.032	1.01
Ethnicity	Intercept	7.622	7.571	7.672	1.00
	time	-0.005	-0.006	-0.005	1.00
	ethnicity	0.434	0.266	0.593	1.00
	time:ethnicity	-0.002	-0.004	0.000	1.00
	sigma	0.516	0.494	0.539	1.01
	sd(Intercept)	1.034	0.988	1.083	1.00
	sd(time)	0.004	0.002	0.006	1.01
	cor(Intercept,time)	-0.251	-0.407	-0.043	1.00
LTHC	Intercept	7.629	7.575	7.683	1.00
	time	-0.005	-0.006	-0.005	1.00
	LTHC	0.165	0.042	0.288	1.00
	time:LTHC	0.000	-0.002	0.002	1.00
	sigma	0.519	0.496	0.543	1.03
	sd(Intercept)	1.035	0.987	1.083	1.01
	sd(time)	0.004	0.001	0.006	1.04
	cor(Intercept,time)	-0.241	-0.412	0.005	1.01
Ct value	Intercept	7.661	7.613	7.710	1.00
	time	-0.006	-0.006	-0.005	1.00
	ct	0.006	-0.002	0.013	1.00
	time:ct	0.000	0.000	0.000	1.00
	sigma	0.517	0.495	0.541	1.01
	sd(Intercept)	1.039	0.991	1.087	1.00
	sd(time)	0.004	0.002	0.006	1.02
	cor(Intercept,time)	-0.254	-0.416	-0.041	1.01
Symptom	Intercept	7.680	7.574	7.788	1.00
	time	-0.007	-0.008	-0.005	1.00
	symptom	-0.020	-0.140	0.100	1.00
	time:symptom	0.001	0.000	0.003	1.00
	sigma	0.516	0.492	0.541	1.02
	sd(Intercept)	1.041	0.990	1.091	1.01
	sd(time)	0.004	0.002	0.006	1.03
	cor(Intercept,time)	-0.257	-0.416	-0.041	1.01
Multivariable	Intercept	7.531	7.412	7.648	1.00
	time	-0.005	-0.006	-0.003	1.00
	age	0.134	0.101	0.167	1.00

		Estimate	95%CrI		Rhat
	sex	0.046	-0.051	0.142	1.00
	ethnicity	0.530	0.372	0.693	1.00
	LTHC	0.099	-0.018	0.218	1.00
	ct	0.009	0.001	0.017	1.00
	symptom	0.009	-0.107	0.127	1.00
	time:age	0.000	-0.001	0.000	1.00
	time:sex	-0.002	-0.003	-0.001	1.00
	time:ethnicity	-0.002	-0.004	0.000	1.00
	time:LTHC	0.000	-0.002	0.002	1.00
	time:ct	0.000	0.000	0.000	1.00
	time:symptom	0.001	-0.001	0.003	1.00
	sigma	0.519	0.495	0.543	1.00
	sd(Intercept)	1.006	0.958	1.056	1.00
	sd(time)	0.004	0.001	0.006	1.02
	cor(Intercept,time)	-0.216	-0.404	0.093	1.00

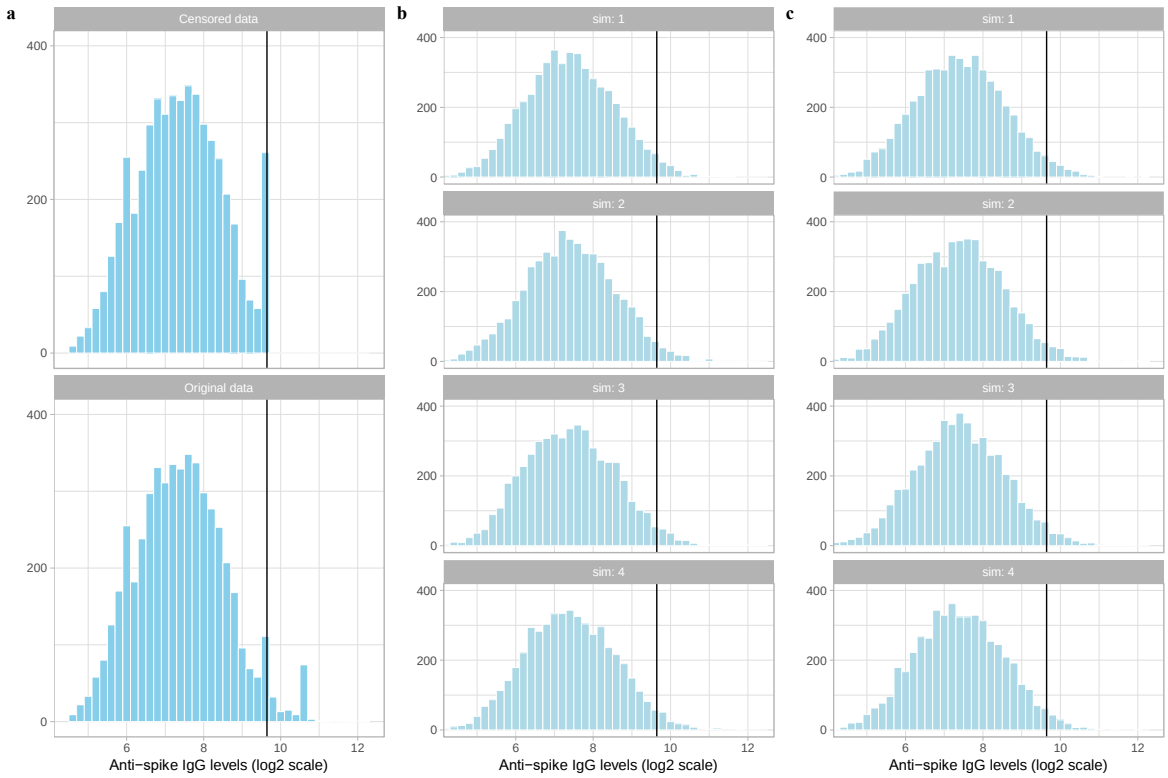


Figure 4.12 Posterior predictive check for the baseline model and the multivariable model.

Note: **a,** Distribution of the observed anti-spike IgG levels on log2 scale. **b,** Distribution of anti-spike IgG levels from four posterior simulated datasets for the baseline model. **c,** Distribution of anti-spike IgG levels from four posterior simulated datasets for the multivariable model.

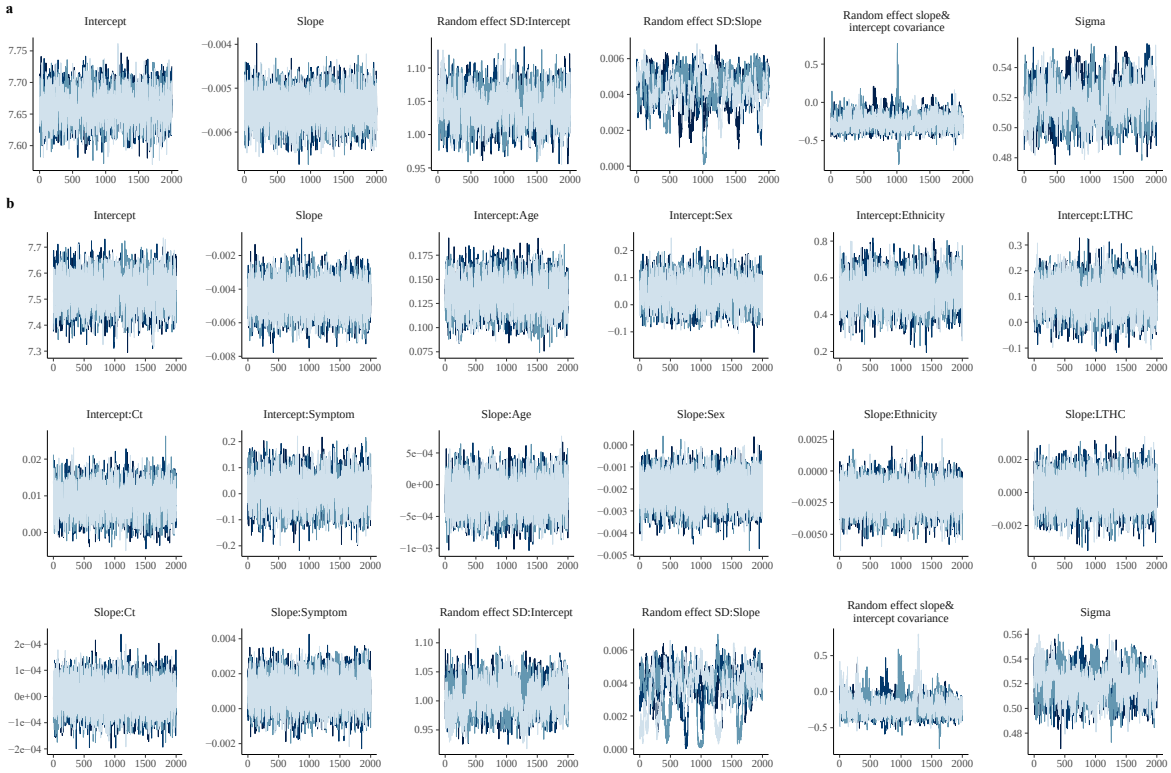


Figure 4.13 MCMC trace plots for assessing convergence of chains.

Note**: a,** Trace plots for the baseline model. **b,** Trace plots for the multinomial model.

Table 4.9 Posterior mean and 95% credibility intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in multivariable model including the time period factor (pre-Alpha vs Alpha) in 3,271 participants in Class 1.

Note: Pre-Alpha period: before 16 Nov 2020; Alpha period: 17 Nov 2020-16 May 2021. Due to the short follow-up time in Delta period (after 17 May 2021), no Delta period infections are included. The reference category is: 43-year-old, female, white ethnicity, no long-term health conditions, Ct value=22, no self-reported symptoms, and in pre-Alpha period. * indicates where the multivariable 95% CrI excludes 0 (no effect).

		Multivariable model		
		Posterior mean	95% CrI	
Baseline	Peak level (Intercept)	100	84	110
	IgG half-life (slope)	256	168	439
Age	Peak level: 43 years (median)			
	IgG half-life: 43 years (median)			
	Change in peak level: per 10-year older	10*	8	13
	Change in half-life: per 10-year older	-12	-54	22
Sex	Peak level: Female			
	IgG half-life: Female			
	Change in peak level: Male	3	-3	10
	Change in half-life: Male	-92*	-236	-25

		Multivariable model		
Ethnicity	Peak level: White			
	IgG half-life: White			
	Change in peak level: Non-white	42*	28	60
	Change in half-life: Non-white	-85	-223	5
LTHC	Peak level: No			
	IgG half-life: No			
	Change in peak level: Yes	7	-1	17
	Change in half-life: Yes	15	-115	227
Ct value	Peak level: 22 (median)			
	IgG half-life: 22 (median)			
	Change in peak level: per 1 unit higher	1*	0	1
	Change in half-life: per 1 unit higher	0	-9	9
Symptom	Peak level: No			
	IgG half-life: No			
	Change in peak level: Yes	1	-8	8
	Change in half-life: Yes	73	-76	236
Time period	Peak level: Pre-Alpha			
	IgG half-life: Pre-Alpha			
	Change in peak level: Alpha	7	-1	14
	Change in half-life: Alpha	-35	-153	46

4.3.5 Duration of antibody responses and immune protection

Multivariable linear mixed models allowing for a biphasic exponential decline, or a flexible non-linear decline, in antibody levels showed evidence of better model fit than the baseline model using exponential decline **(Figure *4.14*)**, although differences were modest. Over the 120 days following peak antibody levels, all 3 models were qualitatively similar (**Figure *4.9***); however the more flexible models showed the rate of decline in antibodies slowing over time. I therefore used a multivariable biphasic exponential mixed model to estimate that antibody responses were likely to remain positive, i.e. ≥23 BAU/mL, for 703 (95%Crl 371-2654), 490 (286-923), 561 (334-1456), 441 (285-809) days from the start of infection for white females, white males, non-white females, non-white males aged 60 years, respectively. From the start of infection to 16 BAU/mL, the antibody level associated with 50% protection against new infection in a study of those previously infected[112], the estimated time was 869 (482-3145), 600 (376-1123), 667 (407-1710), and 520 (343-962) days respectively. For a threshold of 3 BAU/mL, estimated to provide 50%

protection against severe infection (based on previous estimates that this was provided by neutralising antibody levels 3% of peak[135]), the estimated time was 1500 (871-5973), 1017 (685-1909), 1070 (669-2827), and 826 (571-1532) days respectively (**Figure *4.15*, Table *4.10***). To allow for emerging viral variants needing higher antibody concentrations to afford the same level of neutralising activity, a sensitivity analysis assumed 2- to 10-fold greater antibody concentrations were required. For example, if 5-fold higher concentrations were required, for an example 60-year-old white male, the estimated duration of response was 71 (57-247), 162 (57-391), and 581 (361-1093) days for levels associated with a positive result (5*23 BAU/mL), 50% protection from infection (5*16 BAU/mL), and 50% protection against severe infection (5*3 BAU/mL) respectively. I also estimated the duration of protection using the linear exponential model (**Figure *4.16*, Table *4.10***), which yielded shorter estimates of durations for each population group.

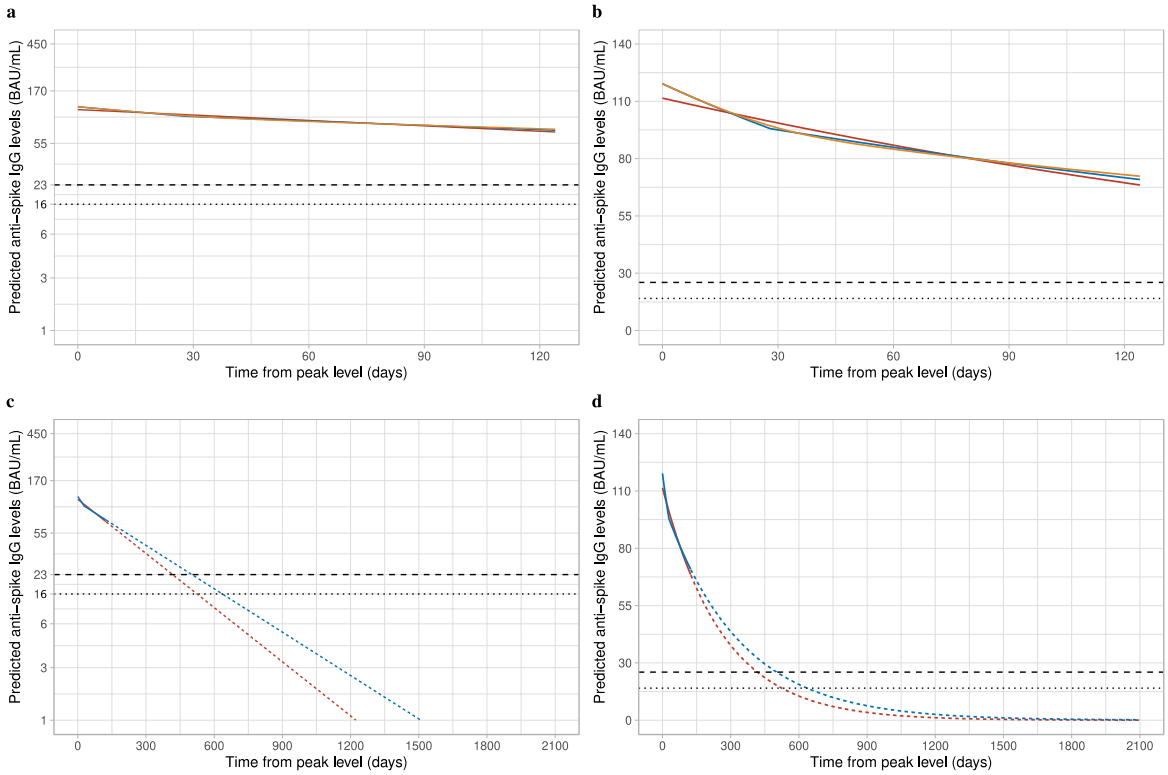


Figure 4.14 Comparison of anti-spike IgG decline in 3,271 participants among the model assuming a linear decline in log2 scale (red line), the biphasic exponential model (blue line), and the model using splines (orange line).

Note: For the biphasic model, knot is placed at 28 days. For the spline model, time is fitted using natural cubic splines with internal knots at 30, 70 and boundary knots at 5, 110. Panel a, b show the antibody trajectories from the three models within observed period (124 days). Panel c,d show the predicted trajectories from the linear and biphasic models over 2100 days (dotted line indicates the values are extrapolated). Panel a,c shows the trajectory on the log10 scale. Panel b,d shows the trajectory in its original scale. Black dashed line indicates the assay threshold for IgG positivity (23 BAU/mL) and the dotted line at 16 BAU/mL (indicates level associated with 50 % protection against reinfection). The spline model had the lowest LOOIC (11402), followed by the biphasic model (11594), and the linear model (11819).

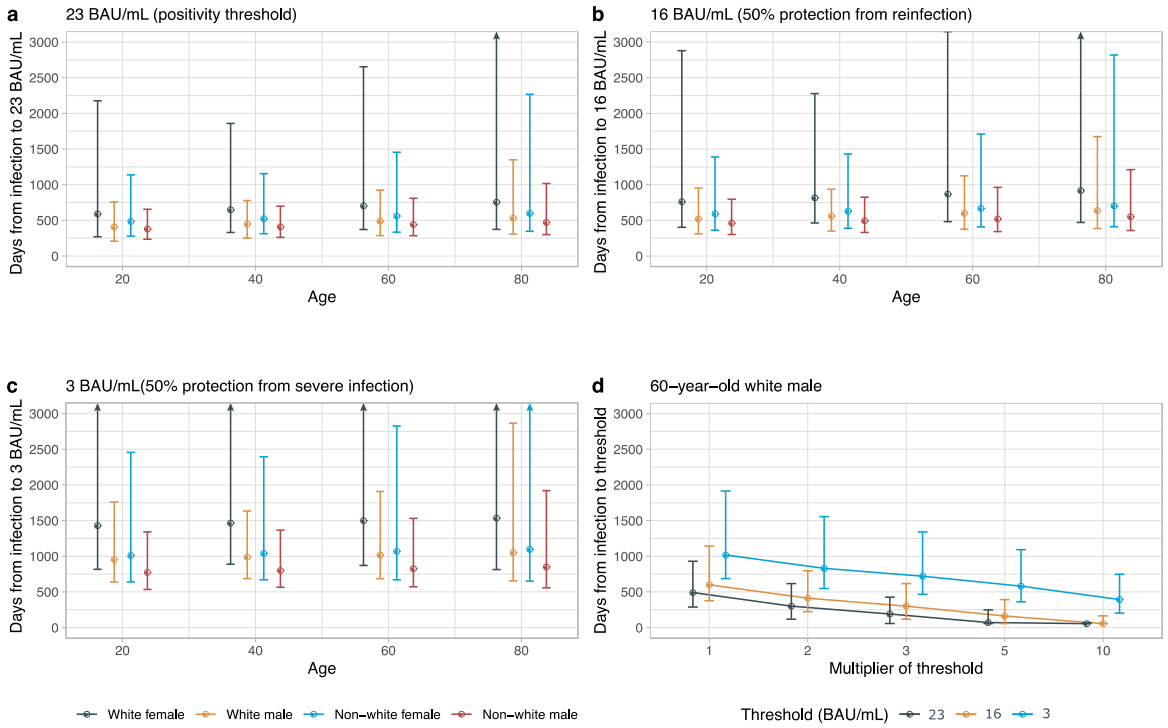


Figure 4.15 Posterior predicted time (95% credibility interval) from the start of infection to three anti-spike IgG thresholds (23, 16, and 3 BAU/mL) by age (20, 40, 60, 80 years), sex, and ethnicity from the multivariable biphasic exponential model in 3271 participants.

Note**: a**, Time from the start of infection to the positivity threshold of 23 BAU/mL. **b,** Time from the start of infection to the equivocal threshold of 16 BAU/mL, which corresponds to 50% protection against PCR-confirmed reinfection. **c**, Time from the start of infection to 3 BAU/mL, which corresponds to 50% protection against severe infection. **d**, Time from the start of infection to the above three thresholds multiplied by 2, 3, 5, and 10, in a 60-year-old white male as an example, to estimate the duration given the higher antibody level required for protection against variants of concern.

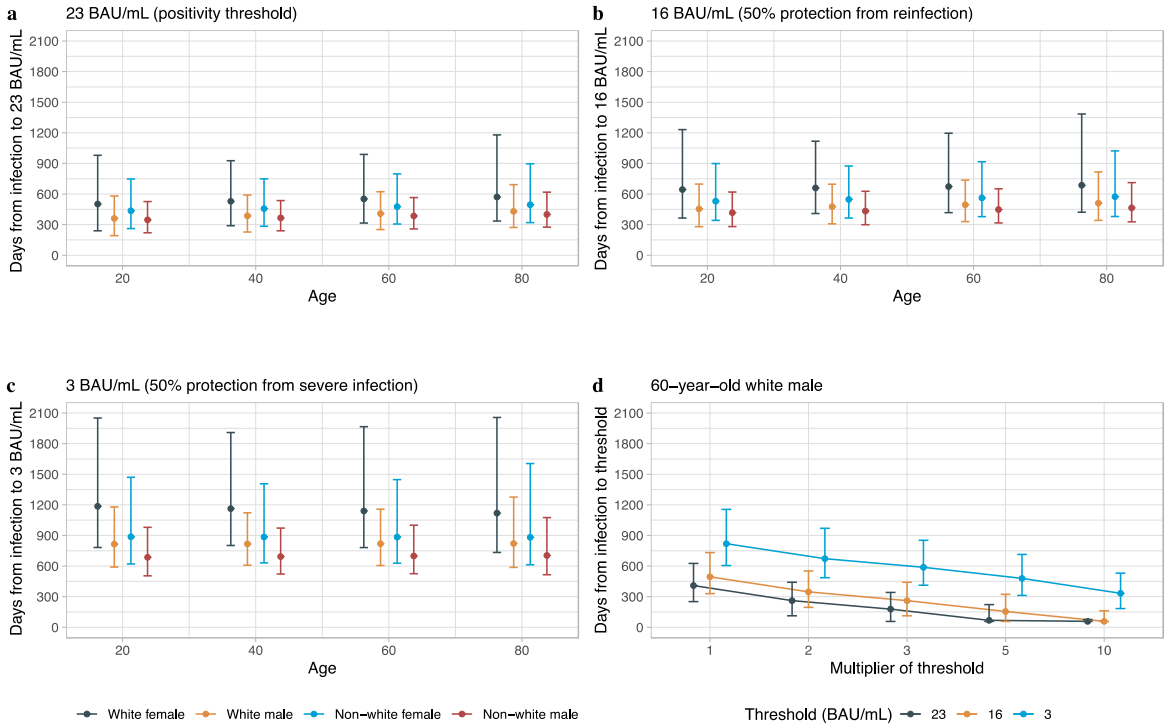


Figure 4.16 Posterior predicted time (95% credibility interval) the start of infection to three anti-spike IgG thresholds (23, 16, and 3 BAU/mL) by age (20, 40, 60, 80 years), sex, and ethnicity from the multivariable linear model in 3271 participants.

Note: **a**, Time from the start of infection to the positivity threshold of 23 BAU/mL. **b,** Time from the start of infection to the equivocal threshold of 16 BAU/mL, which corresponds to 50% protection against PCR-confirmed reinfection. **c**, Time from the start of infection to 3 BAU/mL, which corresponds to 50% protection against severe infection. **d**, Time from the start of infection to the above three thresholds multiplied by 2, 3, 5, and 10, in a 60-year-old white male as an example, to estimate the duration given the higher antibody level required for protection against variants of concern.

Table 4.10 Posterior predicted days of protection from the start of infection at different thresholds.

Note: A) Posterior predicted days with 95% credibility interval from the start of infection to three anti-spike IgG thresholds (23, 16, and 3 BAU/mL) by age (20, 40, 60, 80 years), sex, and ethnicity in the multivariable biphasic exponential model and linear exponential model. B) Posterior predicted days with 95% credibility interval from the start of infection to thresholds (23, 16, and 3 BAU/mL) multiplied by 2, 3, 5, 10 in 60-year-old white male and female population in the multivariable biphasic exponential model and linear exponential model.

		Biphasic exponential model						Linear exponential model					
A		**23 BAU/mL**		**16 BAU/mL**		**3 BAU/mL**		**23 BAU/mL**		**16 BAU/mL**		**3 BAU/mL**	
Age	**Sex&ethnicity**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**
20	White female	590	270-2176	762	402-2879	1430	817-5825	502	239-980	644	364-1232	1186	782-2051
	Non-white female	485	278-1138	593	362-1389	1012	640-2458	436	261-747	530	341-898	887	621-1470
	White male	407	207-758	520	309-952	955	639-1761	361	192-580	454	280-698	815	590-1180
	Non-white male	380	235-656	463	302-796	775	535-1341	347	220-526	417	281-619	686	504-980
40	White female	649	329-1859	817	463-2277	1466	888-4186	529	289-926	660	409-1117	1163	802-1909
	Non-white female	524	313-1155	630	389-1429	1040	670-2395	457	284-748	547	364-875	885	631-1406
	White male	449	251-778	562	350-936	988	687-1634	386	228-590	476	308-697	817	609-1122
	Non-white male	410	262-699	495	329-824	801	565-1365	366	240-536	433	299-627	693	522-972
60	White female	703	371-2654	869	482-3145	1500	871-5973	552	316-989	673	418-1196	1140	781-1966
	Non-white female	561	334-1456	667	407-1710	1070	669-2827	476	306-798	562	378-917	884	628-1448
	White male	490	286-923	600	376-1123	1017	685-1909	408	251-623	495	330-737	819	605-1158
	Non-white male	441	285-809	520	343-962	826	571-1532	385	258-565	448	317-651	699	525-1000
80	White female	755	374-7812	917	470-7812	1534	813-7812	571	335-1180	686	422-1385	1119	733-2056
	Non-white female	598	347-2267	703	411-2819	1097	650-4521	495	320-896	574	380-1023	882	613-1605
	White male	531	308-1348	638	385-1674	1050	655-2866	430	272-692	511	341-817	821	587-1276
	Non-white male	471	300-1017	551	357-1211	851	557-1920	400	275-618	464	327-712	704	515-1075

		Biphasic exponential model						Linear exponential model					
B		**23 BAU/mL**		**16 BAU/mL**		**3 BAU/mL**		**23 BAU/mL**		**16 BAU/mL**		**3 BAU/mL**	
Group	**Multiplier**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**	**Estimate**	**95%CrI**
60-year-old white male	1	490	286-923	600	376-1123	1017	685-1909	408	251-623	495	330-737	819	605-1158
	2	301	118-616	412	221-797	830	548-1557	261	112-441	347	195-552	673	486-970
	3	191	57-426	301	118-616	720	467-1341	178	57-341	261	112-441	588	412-853
	5	71	57-247	162	57-391	581	361-1093	68	57-222	155	57-323	479	312-714
	10	56	57-69	56	57-164	394	202-747	58	57-75	58	57-161	333	183-530
60-year-old white female	1	700	369-2616	868	489-3290	1498	873-5864	551	317-990	674	425-1184	1138	776-1955
	2	418	141-1498	586	283-2035	1218	701-4483	344	131-660	463	241-842	930	625-1596
	3	253	57-907	418	141-1498	1051	594-3942	218	57-486	344	131-660	809	525-1406
	5	70	57-390	210	57-775	841	457-3124	64	57-305	186	57-450	653	404-1154
	10	56	57-64	56	57-209	553	255-1999	58	57-69	58	57-196	444	224-826

4.4 Discussion

In this chapter, I used data from a representative national UK survey to determine predictors of seroconversion following a positive PCR test before widespread vaccination and investigate the duration of antibody responses and possible associated protection in those who do seroconvert. It should be noted again that the infections in this study were mainly Pre-Alpha and Alpha variants.

I found 24% of participants did not seroconvert after testing PCR-positive. However, restricting to participants with strong evidence for a true positive PCR result (Ct≤32, ≥2 genes detected), a lower proportion, 595/5230 (11%), did not seroconvert. Similar observations have been reported before, but with varying percentages of non-responders from 0%-25%[119–121,136–139]. Non-responders likely reflect a combination of genuine non-responders, false-positive PCR results, and false-negative antibody results. However, those not seroconverting typically had persistently low antibody levels, suggesting that any false-negative antibody results would not change if antibody positivity thresholds were adjusted within reasonable limits. Consistent with the first two possibilities, non-responders had fewer symptoms and higher Ct values (lower viral loads), but, more consistent with being genuine non-responders, they were also older. The sensitivity of the serological assay was previously reported as 99%[83], based on samples from predominantly symptomatic individuals, including those admitted to hospital, and it is possible that the greater proportion of those with asymptomatic or mild infection explains part of the difference seen here. I found no evidence of an independent effect of long-term health conditions on non-response, possibly reflecting the heterogeneity of this group including those with a range of cardiovascular and metabolic conditions not typically associated with impaired humoral immunity, as well as conditions more directly impacting antibody production. Other studies have reported that people taking immunosuppressive

medications or with impaired immunity have decreased antibody responses[140–144]. Although in some populations, antibodies are associated with protection from reinfection[111,112], the risk of reinfection and vaccine failure in PCR-positive seronegative individuals from specific immunocompromised groups needs further study.

Although the specificity of PCR testing in this cohort has been estimated as ≥99.995%[81,145], given the large number of tests performed in asymptomatic individuals, i.e. with a low pre-test probability of infection, assuming a sensitivity of 94%[146] and specificity 99.995%, the positive predictive value of PCR tests ranges between 95.0% and 99.7% for SARS-CoV-2 prevalence between 0.1% and 2%. Therefore, although some non-responders will have had false-positive PCR results, particularly since the majority (97%) of participants in Class 3 have only one positive swab test during the study, the relatively high PPV suggests most non-responders are likely to have had a true positive PCR result. It is also possible that some prior infections were missed by PCR testing, e.g. because participants were not tested, variation in self-performed swabbing technique, or the assay itself.

I found that apart from age, individual symptoms including cough, loss of smell/taste, fever, fatigue, headache, and sore throat were independently associated with generating antibodies following a positive PCR test. The strongest predictors were the four classic symptoms (cough, loss of smell/taste, fever).

I estimated the half-life of anti-spike IgG to be 184 days, indicating a sustained antibody response against infection, compared with previous reports between 36 and 244 days[123,125,127–131]. I found multiple factors associated with peak levels and decline. Variation in the literature may be explained by differences in study design, population (age, sex), and assay performance (different targets and assay types). Longer half-lives

were correlated with lower peak levels, suggesting some individuals, e.g. after mild disease[128,138] mount a lower antibody response that wanes more slowly, whilst others, produce higher antibody responses but that wane more quickly. This contrasts with a previous healthcare worker study which found a positive correlation between IgG half-life and peak levels[77], but agrees with another reporting a faster decay of IgG in hospitalized patients with high initial responses than individuals with asymptomatic or mild infections[131]. Since most SARS-CoV-2 infection is mild/asymptomatic, the duration of antibody responses in my study are likely to best generalise to the population at large.

As expected from previous studies of humoral immunity, older age was associated with lower seroconversion rates. However, among those that did seroconvert, peak IgG levels were higher in older individuals. Similar findings have been reported in healthcare workers, where older age (in those of working age) was associated with higher maximum anti-nucleocapsid IgG levels and longer half-lives[123]. Others have also reported associations between older age and higher immune responses, including IgG and memory B cells[147,]. In my study, selection bias may contribute, as my findings are conditional on participants seroconverting, and the subset of older participants who seroconvert may have more robust immune responses than younger participants overall, amongst whom more may seroconvert despite more heterogenous underlying immunity.

Females previously infected with SARS-CoV-2 have been found to have more robust T cell activation and develop stronger antibody responses than males[148,149]. I found that males were equally likely to seroconvert, but among those that did seroconvert, males had a shorter IgG half-life than females, despite no evidence of difference in peak IgG levels, consistent with a previous healthcare worker study[150]. Another study found no difference in IgG antibody between males and females in mild infection and recovering patients, but a

higher IgG in females than males in severe infections and early phases of infection[151]. I found non-white participants were more likely to seroconvert than white participants (in models not adjusting for Ct value) and to develop higher antibody levels that then waned more quickly. Higher antibody levels in individuals of non-white ethnicity have been reported in several healthcare worker populations[123,152]. The observed sex and ethnicity effects likely arise from a combination of genetic and societal factors, with further adjustment for confounding arising from social differences and structural inequalities required to estimate the relative contributions of each mechanism.

While lower Ct values were associated with seroconversion, I found that higher Ct values were associated with slightly higher peak IgG levels, which was counterintuitive as higher Ct values (lower viral burden) have been previously associated with lower antibody titres[116,128,153]. The most likely explanation is that as testing was conducted at regular intervals, rather than in response to symptoms, measured Ct values do not fully reflect peak viral load in my study. I found no evidence of association between self-reported symptoms and IgG peak levels or half-life, although symptoms were associated with seroconversion; previous findings suggest that symptomatic infections develop stronger antibody responses than asymptomatic infections[74]. This could be because my models conditioned on those who seroconverted, or because infections in this general population were generally mild.

Important findings from this study are the predictions about the duration of antibody responses associated with protection from infection, albeit that these related to thresholds previously associated with protection from reinfection or protection from severe infection in vaccine trials. Other immune responses may last for differing time periods, and also memory responses may mean that protection lasts longer than measurable antibody levels.

Furthermore, consistent with previous studies suggesting antibody waning following a biphasic exponential pattern[154,155], I observed that the rate of antibody decline slows over time and antibody levels can be sustained for longer than assuming antibody levels fall exponentially. I estimated the time from peak level to three thresholds, the positivity threshold 23 BAU/mL, 16 BAU/mL (50% protection from any symptomatic/asymptomatic infection[112]), and 3 BAU/mL (3% of my estimated peak level, providing 50% protection against severe infection according to[135]). Based on extrapolations from other studies correlating anti-spike IgG antibody titres with neutralising activity and early protection (i.e. within a year) from re-infection with currently circulating variants, I found that 50% protection against infection might be expected to last 1.5-2 years, with protection against severe infection potentially lasting several years, but with uncertainty in the precise estimates given the assumptions relating neutralising activity to antibody titres over time and in estimating levels associated with protection against severe infection[135]. However, given that variants may require higher antibody levels for the same level of neutralisation, the duration of protection might be substantially reduced. It may also be the case that the functional quality of antibodies changes over time[156]; this was not evaluated in this study. Overall, at least in the short-term, protection against re-infection appears high.

Study limitations include the fact that only anti-spike IgG was measured using a single assay; seronegative non-responders in Class 3 might have antibodies detected using other assays or other target antigens. Neutralizing antibodies or T cell responses were not measured; however, neutralizing antibody responses are strongly correlated (Spearman ρ=0.87) with anti-spike binding antibodies following infection as previously reported[157]. This community survey had visits scheduled independent of infection or symptom status, so I could not precisely identify the start of infection or symptom onset; I therefore also incorporated positives from the national testing programme (targeting symptomatic

infections) and used the first swab positive test and latent class models to indirectly estimate the start of infection. Similarly, I was not able to model antibody trajectories from each participant's maximum levels since antibody data were collected monthly. However, I chose a starting point that was close to but slightly after the peak IgG level; while this could slightly underestimate peak IgG levels, the half-life will be unbiasedly estimated if the assumption of exponential decline is correct. Re-infections were rare during the calendar period included in this study, with only 92 (0.5%) participants with antibody data having potential re-infections >120 days after their first infection episode**.** Most had only one antibody result, so it was impossible to investigate any boosting of antibody levels following re-infection.

In conclusion, in this representative study of infected individuals from the UK general population, around 1 in 4 people did not develop anti-spike IgG antibodies following a positive-PCR test in regular screening. Non-responders were more likely to be older and not report symptoms. Among participants who seroconverted, anti-spike IgG antibodies remained above the positivity threshold for an average of 380-590 days for 20 year olds, 410-649 days for 40 year olds, 441-703 days for 60 year olds, and 471-755 days for 80 year olds. These estimates of the durability of natural immunity may aid planning of the vaccination strategies. Further studies are required to determine the extent to which waning antibody levels impact immunity and protection following infection and vaccination, which will be discussed in my next chapter.

Chapter 5 Antibody responses to first and second COVID-19 vaccination and correlates of protection against SARS-CoV-2 Delta variant

This chapter is based on previous publications:

*Wei, J., Stoesser, N., Matthews, P.C. et al. Antibody responses to SARS-CoV-2 vaccines in 45,965 adults from the general population of the United Kingdom. Nat Microbiol **6**, 1140–1149 (2021).*

*Wei, J., Pouwels, K.B., Stoesser, N. et al. Antibody responses and correlates of protection in the general population after two doses of the ChAdOx1 or BNT162b2 vaccines. Nat Med **28**, 1072–1082 (2022).*

*Wei, J., Matthews, P.C., Stoesser, N. et al. SARS-CoV-2 antibody trajectories after a single COVID-19 vaccination with and without prior infection. Nat Commun **13**, 3748 (2022).*

All data analysed and literature discussed in this chapter were up to January 2022. The correlates of protection against Delta variant were used data from 17th May 2021 to 4th October 2021.

5.1 Introduction

I have explored the antibody responses after natural infection in previous chapter, and in this chapter, I will explore antibody responses after the first and second vaccination.

Multiple vaccines have been developed which offer protection against COVID-19 by generating immune responses against the spike antigen of SARS-CoV-2. On 8 December 2020, the UK started its national vaccination programme with the Pfizer-BioNTech BNT162b2 vaccine[61], followed by the approval of the Oxford-AstraZeneca ChAdOx1

nCoV-19 vaccine, first used outside a clinical trial on 4 January 2021[62]. Both vaccines have been widely used in the UK. Vaccines were initially administered to priority groups, including care home residents, people >80 years old, healthcare workers, and those clinically vulnerable (≥16 years), and then offered to the rest of the adult (≥18 years) population in decreasing age order[63].

The efficacy of the ChAdOx1 and BNT162b2 vaccines against symptomatic laboratory-confirmed SARS-CoV-2 infection has been reported in large randomized controlled clinical trials as 52% (95%CI 30-86%) after the first dose and 95% (90-98%) after the second dose of BNT162b2[55], and 70% (55-81%) after the second dose of ChAdOx1[57]. Several studies have examined the immunogenicity of vaccines in healthcare workers, typically the earliest groups to be vaccinated. A study of 3610 healthcare workers found 99.5% and 97.1% seroconverted after a single dose of BNT162b2 and ChAdOx1 respectively, and that higher quantitative IgG levels were achieved in previously infected individuals[158]. Other studies have also found single-dose BNT162b2 elicited higher antibody levels in previously seropositive individuals, comparable to the levels after two doses of vaccines in seronegative individuals[159–161]. Outside of trials, there were initially limited data on post-vaccine antibody responses in other groups, especially older adults who were underrepresented in the ChAdOx1 trial[57]. A study of 185 individuals aged >70 years showed high seropositivity after one or two BNT162b2 doses[162]. Another, of 100 individuals aged 80-100 years, showed almost universal high antibody responses 3 weeks after a single dose of BNT162b2, with spike-specific cellular responses in 63%[163]. However, the representativeness of these small cohorts is unclear.

With wide-spread Alpha variant transmission, in January 2021 the UK government extended the dosing interval from 3-4 to 12 weeks for all vaccines to maximize first dose

coverage, based on preliminary data showing high short-term efficacy from single BNT162b2 (90%) and ChAdOx1 (70%) doses[164]. This approach raises several questions. Although the ChOxAd1 trial found higher vaccine efficacy with dosing intervals ≥6 weeks[57], BNT162b2 trials did not compare different dosing intervals. Subsequent UK studies showed extended BNT162b2 dosing intervals generated higher antibody responses than the 3-week interval[165–167]. However, these studies were based on relatively small sample sizes (N<600) or specific population groups such as healthcare workers, potentially reducing generalisability, and antibody levels were only measured at specific times after second doses. With the rapid emergence of the Delta variant, and greater protection after second than first vaccine doses[168–170], from mid-2021 dosing intervals were reduced to 8 weeks[171] to achieve greater protection faster. However, large population-based investigations of how these different dosing intervals, or other factors, affect longer-term antibody changes after the second dose are limited, but essential to assess the duration of protection and the need for booster doses.

Anti-spike IgG and pseudovirus neutralisation titres following the ChAdOx1 vaccine are associated with protection against symptomatic SARS-CoV-2 infection[172]. Similarly for the mRNA-1273 vaccine, anti-spike antibody and neutralisation titres were inversely associated with SARS-CoV-2 infection, with 68.5% of vaccine efficacy up to 126 days post-second dose mediated by day 29 neutralisation titres[173]. However, how these measurements relate to antibody levels and durations of protection in populations over time, and following BNT162b vaccination, is not fully understood.

Real-world data provide information on populations who may not participate in clinical trials as well as assessing the effectiveness of interventions as deployed. In this chapter, I used data from the UK's national COVID-19 Infection Survey (CIS) to investigate

population-wide anti-trimeric spike IgG antibody responses after the first and second SARS-CoV-2 vaccination by time since vaccination. Specifically, I explored the following questions:

1. What is the impact of different factors, including demographics, vaccine type, dosing interval, and prior SARS-CoV-2 infection status on antibody peak levels and declines?
2. What is the association between anti-spike IgG levels and protection against SARS-CoV-2 infection (correlates of protection)?
3. Can correlates of protection analyses be combined with estimated antibody trajectories to predict the duration of protection following first vaccination, second vaccination, and natural infection?

5.2 Methods

5.2.1 Population and data

Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and laboratory tests were already described in Chapter 3.

This chapter is based on results from three different papers (hereinafter referred to as 'initial vaccination study', 'second dose study', 'first dose study') thus the cohorts and the methods are separated by papers. To improve readability, I used a diagram to summarise the three different cohorts, and analyses performed using data from each cohort (**Figure *5.1***).

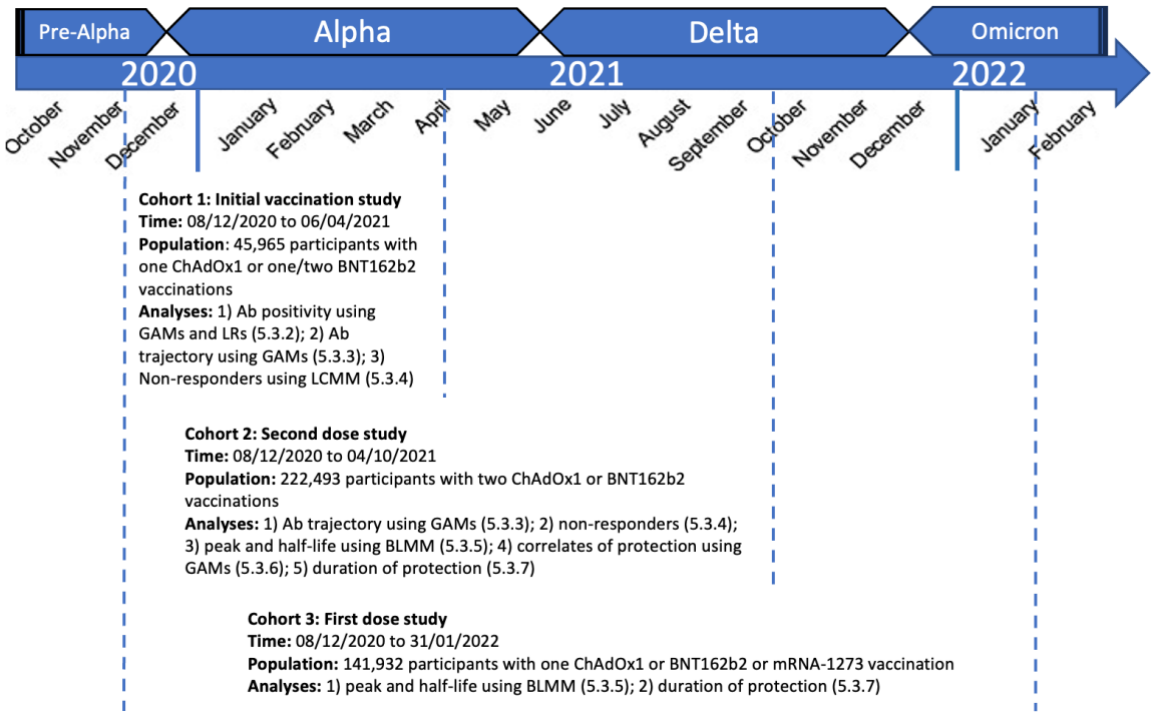


Figure 5.1 Diagram summarising the time periods, populations, and analyses from the three different papers in Chapter 5.

Note: The corresponding results section for each analysis is included in parenthesis.

5.2.1.1 Population and data for the initial vaccination study (cohort 1)

This study included participants ≥16 years who had received at least one dose of vaccine from 8 December 2020 onwards with one or more antibody measurements from 91 days before their first vaccination date through to 6 April 2021.

Participants with a SARS-CoV-2 PCR-positive nose/throat swab or a prior positive anti-spike IgG antibody result at any time prior to vaccination were considered to have been previously infected with SARS-CoV-2, irrespective of whether they had reported previous symptoms or not. Regular PCR results from the survey were included in this classification, but not self-reported PCR or lateral flow test results obtained outside the study.

Given the prior hypothesis that response to vaccination would vary differentially by age and time according to vaccine type and prior infection, I fitted separate models to populations defined by vaccine type, for those receiving one or two vaccinations, and by prior infection status. For participants receiving one vaccine dose, four models were fitted, for each vaccine and in those with and without evidence of prior infection. Two dose models were only fitted for those receiving BNT162b2 vaccine without evidence of prior infection, as sample sizes were small for other groups (315 participants with prior infection receiving two BNT162b2doses; 533 participants without prior infection receiving two ChAdOx1 doses; 66 participants with prior infection receiving two ChAdOx1 doses, **Figure *5.2***).

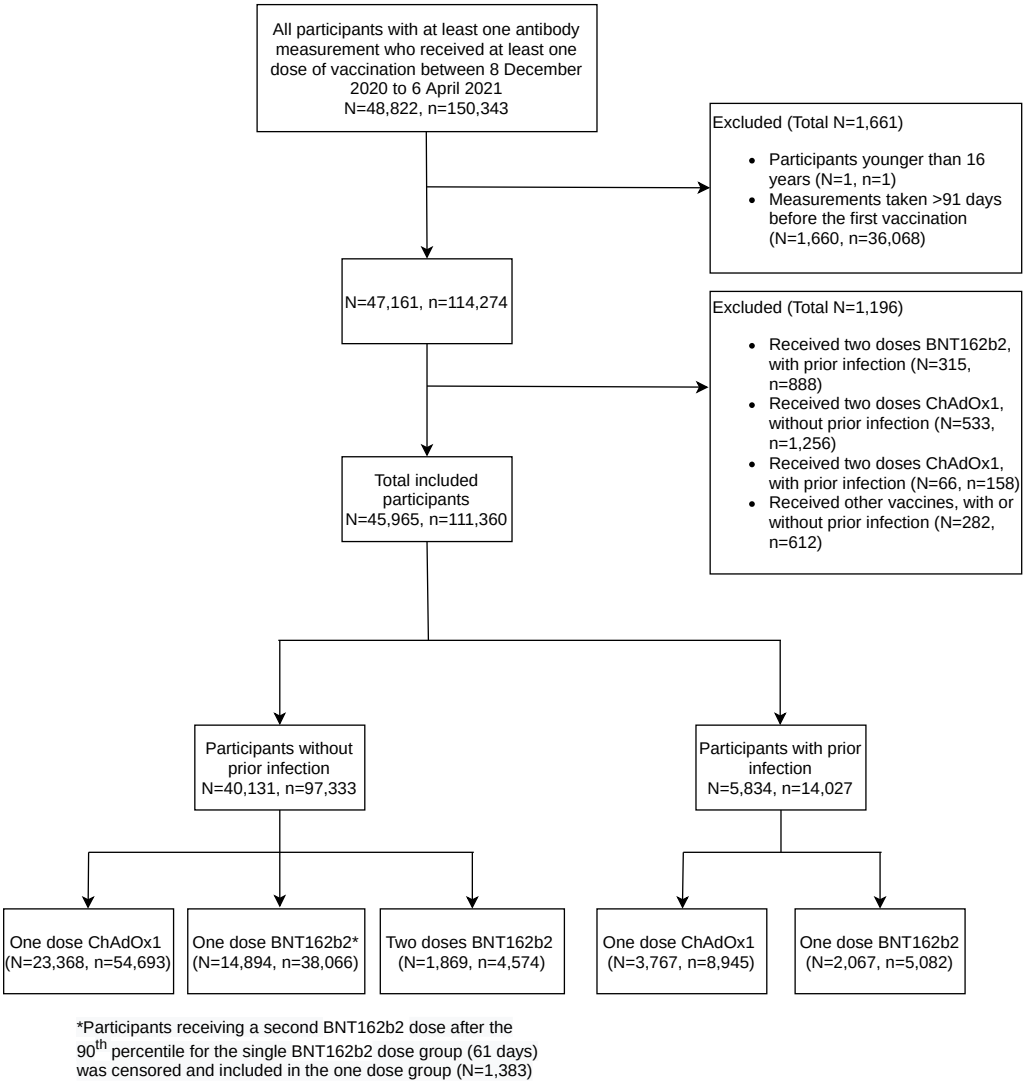


Figure 5.2 Flowchart of the study cohort.

Note: N represents the number of participants, and n represents the number of antibody measurements.

5.2.1.2 Population and data for the second dose study (cohort 2)

This study initially considered 402,348 participants ≥16 years who received two ChAdOx1 or BNT162b2 vaccinations from 8th December 2020 to 4th October 2021. The median (IQR) age was 56 (41-68) years, 216,301 (53.8%) were females, and 375,880 (93.4%) reported white ethnicity. The median (IQR) deprivation percentile was 62 (38-82), and dosing interval (time between first and second vaccinations) was 74 (63-78) days. 11,565

(2.9%) reported working in patient-facing healthcare, and 112,074 (27.9%) reported having a long-term health condition. Among these 402,348 participants, following the design described above with restricted blood sampling, 222,493 participants gave at least one blood sample for antibody testing from 91 days before the first vaccination onwards and were included in my analyses.

Analysis of antibody levels included participants aged ≥16 years who received two doses of ChAdOx1 or BNT162b2 vaccines with or without prior SARS-CoV-2 infection. Age was truncated at 85 years in all analyses to reduce the influence of outliers. Prior infection was defined as having a PCR-positive swab test recorded in the survey or the English national testing programme (national testing data were not available for Scotland, Wales, and Northern Ireland), or a prior positive anti-spike IgG result (≥23 BAU/mL) any time before the first vaccination. Where participants were known to have become infected after vaccination (based on positive PCR tests from the survey or national linked data), antibody measured after infection were excluded from the analyses. The dosing interval was calculated from the first and second vaccination dates. For the main analysis, I excluded a small number of participants who were considered as non-responders after the first or second dose, which was defined as all antibody measurements being <16 BAU/mL and having at least one antibody measurement 21 days after the first or second dose (N=5,098 excluded for ChAdOx1, N=1,649 excluded for BNT162b2) **(Figure *5.3*)**. I also excluded participants with recorded dosing interval <49 days or >91 days for ChAdOX1 (N=4,748 excluded), and 29-48 days or >91 days for BNT162b2 (N=6,374 excluded). 17-28 days were classified as a 3-week interval for BNT162b2.

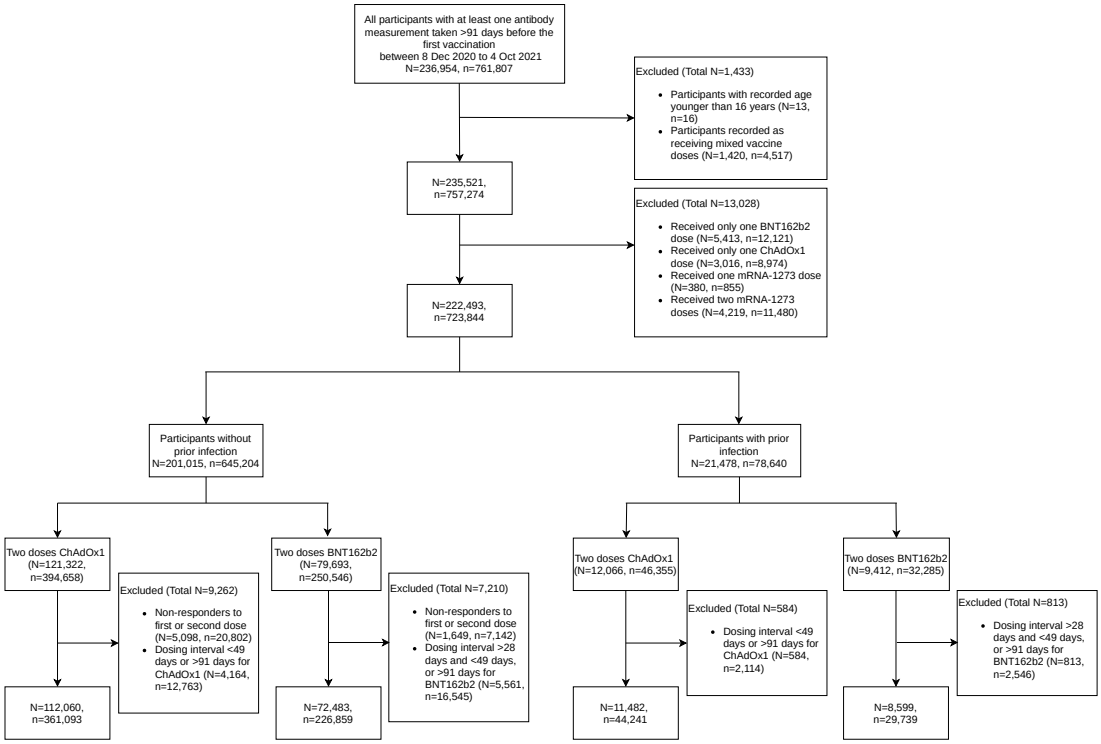


Figure 5.3 Flowchart of the study cohort.

Note: N represents the number of participants, and n represents the number of antibody measurements.

5.2.1.2 Population and data for the first dose study (cohort 3)

In this study, participants aged ≥16 years who received at least a single vaccination with ChAdOx1 or BNT162b2 or mRNA-1273 with antibody measurements from 8th December 2020 until 31st January 2022 were included. Participants with prior infection (before vaccination) were defined by 1) having a positive PCR swab test in the survey or the linked English national testing programme; 2) having a positive anti-spike IgG result (≥23 BAU/mL) before vaccination; 3) having two consecutive positive anti-nucleocapsid IgG results (≥30 ng/ml); or 4) self-reporting a positive swab test in the survey. The infection date was defined as the earliest recorded date from the above definitions. Age was truncated at 85 years to reduce the influence of outliers.

To estimate antibody waning, I excluded a small number of participants who were considered as non-responders after the first dose, defined as all antibody measurements being <16 BAU/mL and having at least one antibody measurement 21 days after the first dose (N=4,488 excluded for ChAdOx1, N=1,450 excluded for BNT162b2, N=17 excluded for mRNA-1273).

5.2.2 Statistical analysis

Data cleaning and preparation were performed in STATA MP 16 by Professor Sarah Walker. All analyses were performed in R 4.1 using the following packages: tidyverse (version 1.3.1), brms (version 2.15.0), mgcv (version 1.8-31), splines (version 3.6.1), lcmm (version 1.9.2), ggeffects (version 0.14.3), arsenal (version 3.4.0), emmeans (version 1.5.1), cowplot (version 1.1.0), gmodels (version 2.18.1), mgcViz (version 0.1.6), bayesplot (version 1.8.1), and tidybayes (version 3.0.1).

5.2.2.1 Statistical analysis for the vaccination study (cohort 1)

I used multivariable logistic and linear generalized additive models (GAMs) to investigate binary (positive/negative) and quantitative ($\log_{10}$(mAb45 units)) anti-spike IgG antibody measurements post-first vaccination. Models were adjusted for participant age using a tensor product of B-splines to allow for non-linearity and interaction between age and time since vaccination. The smoothing penalty was selected using fast restricted maximum likelihood (fREML) as implemented in the mcgv R package. I included a random intercept for each participant to account for repeated measurements using a random effect smoother with the number of basis functions equal to the number of participants. The 95% confidence intervals were calculated using the formula: prediction ± 1.96*standard error of prediction. The date of the first vaccination was set as t=0. For those with no prior evidence of infection, I truncated time at t=0 and t=-14 for logistic and linear models

respectively (t=-14 for linear models to estimate IgG baseline pre-vaccination). I excluded measurements taken after the 90th percentile of observed time points for all models to avoid undue influence from outliers at late time points. Any participant receiving a second BNT162b2 dose after the 90th percentile for the single BNT162b2 dose group (61 days) was censored at this timepoint and included in the one dose group (1,383 [3%] participants were censored in this way). Age was truncated at 85 years in all analyses to avoid outlier influence.

To investigate predictors of antibody response in those without prior evidence of infection, I considered the latest antibody measurement per participant between 14-60 days post first vaccine. I used multivariable logistic regression to examine the association between antibody positivity and vaccine type and doses received by this measurement time, demographic factors (age, sex, ethnicity), household size, deprivation ranking (index of multiple deprivation (IMD) in England and equivalent percentile ranking in Wales, Northern Ireland and Scotland), whether the participant reported working in patient-facing healthcare or social care, whether they reported working in a care home (any role), and whether they reported having a long-term health condition. Non-linearity in age was accounted for using restricted natural cubic splines with internal knots at the 20th, 40th, 60th, and 80th percentiles of unique values, and boundary knots at 5th and 95th percentiles. I tested for and added interactions between age and other variables if the interaction p-value was <0.05.

For those without evidence of prior infection who received a single dose of vaccine I also investigated whether I could identify distinct patterns of antibody responses, using latent class mixed models (LCMM) to identify subgroups with different antibody trajectories after first vaccination. Natural cubic splines (internal knots at the 20th, 40th, 60th, 80th

percentiles of unique values, and boundary knots at 5th and 95th percentiles) were used to model time since vaccination as a fixed effect and a random intercept was added to account for individual variability. Within-class between-individual heterogeneity may also be present in the trajectories; however, models accounting for random slopes failed to converge. Age with natural cubic splines (same as above), sex, reported long-term health conditions, and whether the participant was a healthcare worker were included as covariates for class membership. The 95% confidence interval of the estimation was calculated by a Monte Carlo approximation of the posterior distribution of the predicted values, using the 2.5% and 97.5% percentiles. The number of classes was determined by minimizing the Bayesian information criterion (BIC) for each vaccine, and then fitting the maximum number of classes (four) to both groups for comparability.

To compare prevalence of long-term health conditions across different subgroups identified, English participants were linked by the ONS to the General Practice Extraction Service Data for Pandemic Planning and Research via their NHS number (equivalent data not available for participants from the Devolved Administrations (Wales, Northern Ireland, Scotland)). A range of pre-existing conditions across organ systems were identified from diagnosis codes over the 10-year look-back period 1 January 2010 to 24 January 2020 (the date of the first known case of COVID-19 in the UK). Body mass index (BMI) was the most recently recorded measurement over the look-back period, without imputation. Participants were recorded as being on antihypertensive medication, diabetes medication, corticosteroids, or immunosuppressants if they were prescribed these treatments within 90 days of the end of the look-back period. All clinical variables were derived from primary care records only; hospital admissions data were not used.

5.2.2.2 Statistical analysis for the second dose study (cohort 2)

I used linear generalized additive models (GAMs) to model anti-spike IgG antibody measurements after the first and second dose to identify the most appropriate timepoints to model antibody declines post second vaccination in more detail. Antibody measurements truncated at 450 BAU/mL were counted as 450 BAU/mL. I built separate models by vaccine type and prior infection status given the hypothesis that antibody response would vary by these two factors. Each model was minimally adjusted for only age and dosing interval using a tensor product of B-splines to allow for non-linearity and interaction among age, dosing interval, and time since vaccination, setting the date of the second vaccination as t=0. The smoothing penalty was selected using fast restricted maximum likelihood as implemented in the mcgv R package. The 95% CIs were calculated using the following formula: prediction ± 1.96 × standard error of prediction. I only included antibody measurements from 14 days before the first dose (setting the most recent measurement prior to 14 days before the first dose as 14 days) for those with no evidence of prior infection, and excluded measurements taken after the 95th percentile of the observed t>0 time points to avoid the outlier influence.

I used Bayesian linear mixed interval-censored models to estimate changes in antibody levels after the second ChAdOx1 or BNT162b2 dose. I included measurements from 21 days post-second dose reflecting the peak level from the GAMs (except for 3-week BNT162b2, see below). Measurements taken after the 95th percentile of the observed time points from 21 days post-second dose were excluded to avoid outlier influence. I assumed an exponential fall in antibody levels over time, i.e., a linear decline on a log2 scale. To examine non-linearity in antibody declines, especially the assumption that the rate of antibody decline would flatten, I additionally fitted a model using 4-knot splines for time (knots placed at 10th, 40th, 60th, and 90th of observed time points) and compared the model

fit with the log-linear model using the leave-one-out cross-validation information criterion (LOOIC). I found that the spline model had a higher LOOIC (indicating a worse model fit) than the log-linear model for ChAdOx1 (382954 vs 378640) and BNT162b2 (252730 vs 240472). For both vaccines, the estimated trajectories were similar and there was no evidence of antibody decline flattening (**Figure *5.4***), so I used the log-linear model for the rest of the analysis.

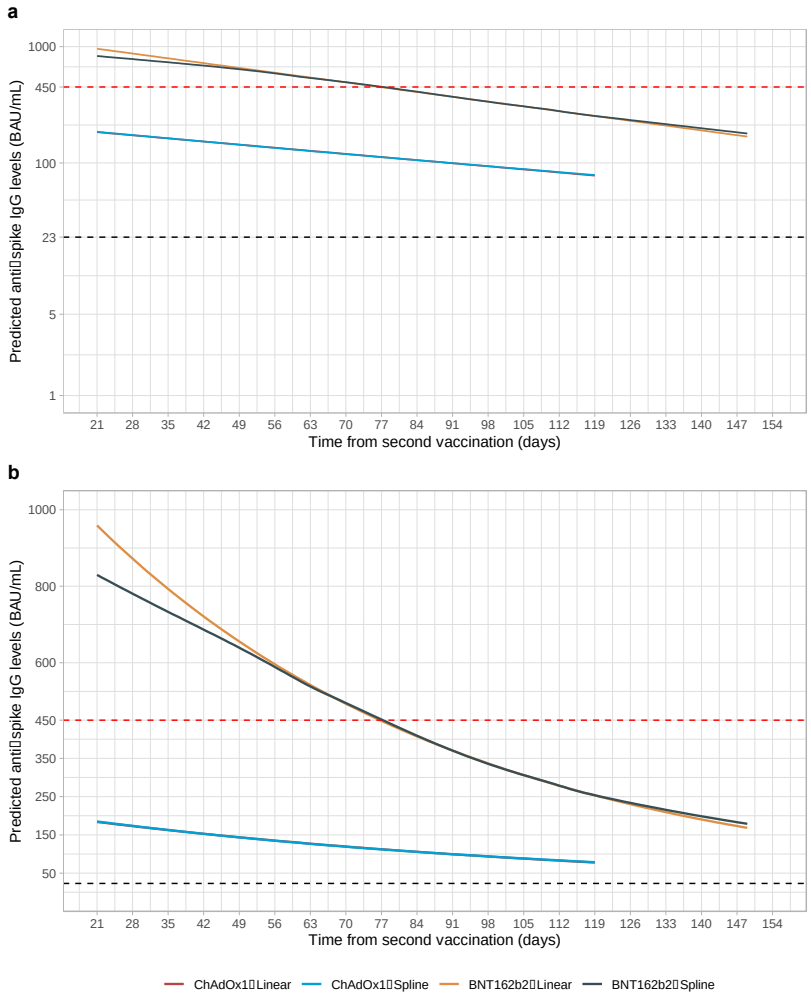


Figure 5.4 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline.

Note: **a**, Plotted in log10 scale. **b**, Plotted in original scale. The estimated trajectory from the spline model (with 4 knots placed at 10th, 40th, 60th, and 90th of observed time points) is similar to the log-linear model for both ChAdOx1 and BNT162b2, indicating that there was no evidence of antibody decline flattening. Log-linear models provided better fit. Black dotted line shows the positivity threshold of 23 BAU/mL. Red dotted line shows the upper quantification limit of 450 BAU/mL.

Population-level fixed effects, individual-level random effects for intercept and slope, and correlation between random effects were included in both models. The outcome was right-censored at 450 BAU/mL reflecting truncation of IgG values at the upper limit of quantification (i.e. all measurements truncated to 450 BAU/mL were considered to be >450 BAU/mL in analyses). I built a multivariable model to examine the association between peak levels and antibody half-lives with continuous age (16-85 years), sex, ethnicity, report having a long-term health condition, report working in patient-facing healthcare, deprivation percentile, continuous dosing interval (7-13 weeks), and prior infection status for both vaccines. For BNT162b2, I additionally examined the impact of a 3-week dosing interval (17-28 days) by creating a binary variable and excluding antibody measurements ≤42 days post-second dose for the 3-week group (identified from the GAM as they peaked at around 42 days post-second dose).

For each Bayesian linear mixed interval-censored model, weakly informative priors were used **(Table *5.1*).** Four chains were run per model with 4,000 iterations and a warm-up period of 2,000 iterations to ensure convergence, which was confirmed visually and by ensuring the Gelman-Rubin statistic was <1.05. 95% credible intervals were calculated using highest posterior density intervals.

Table 5.1 Priors used in the Bayesian linear mixed interval-censored models.

Note: Age (16-85 years) and deprivation percentile (0-100) were scaled by dividing by 10; dosing interval (42-91 days) was scaled by dividing by 7.

Model term	Priors for ChAdOx1	Priors for BNT162b2
Intercept	normal (7.5,1)	normal (9.5, 1)
Slope	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (age)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (age)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (sex)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (sex)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (ethnicity)	normal (0, 1)	normal (0, 1)

Model term	Priors for ChAdOx1	Priors for BNT162b2
Coefficient for change in slope (ethnicity)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (long-term health condition)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (long-term health condition)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (healthcare worker)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (healthcare worker)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (deprivation)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (deprivation)	normal (0, 0.01)	normal (0, 0.01)
Coefficient for change in intercept (3-week interval)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (3-week interval)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (dosing interval)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (dosing interval)	normal (0, 0.1)	normal (0, 0.1)
Coefficient for change in intercept (prior infection)	normal (0, 1)	normal (0, 1)
Coefficient for change in slope (prior infection)	normal (0, 0.1)	normal (0, 0.1)
Random effect SD: intercept	cauchy(0, 0.5)	cauchy(0, 0.5)
Random effect SD: slope	cauchy(0, 0.01)	cauchy(0, 0.01)
Random effect intercept & slope covariance	lkj_corr_cholesky(2)	lkj_corr_cholesky(2)

For the analysis of correlates of protection, I used data from study visits from 17th May 2021 to 4th October 2021. These visits were from 35-298 days after the first vaccination, which captured the time period over which real-world vaccine effectiveness had been observed to wane (**Table *5.2***, **Figure *5.5***). I grouped positive tests into episodes because PCR-positive results might be observed at multiple visits after the same infection. Following previous work, I defined the start of a new episode or 'positive case' as the date of (1) the first PCR-positive test in the study (not preceded by any study PCR-positive test); (2) a PCR-positive test after four or more consecutive negative visits; or (3) a PCR-positive test at least 120 days after the start of a previous episode with one or more negative tests immediately preceding this[170]. Analyses were based on visits, dropping any visits where participants were not at risk due to a recent new positive PCR test, with new PCR-positive episodes as the outcome. I used separate logistic generalised additive models for three outcomes: any positive PCR episode, a positive PCR episode with a moderate to high viral load (Ct value <30), a positive PCR episode with self-reported symptoms. Two

exposure specifications were investigated. Firstly, considering the effect of the most recent antibody measurement obtained 21-59 days prior to the current visit, excluding more recent measurements to avoid changes in antibody levels arising from recent infection that might only be detected at the routine study visit despite occurring before this. Alternatively, I considered the maximum antibody measurement obtained ≥21 days prior to the visit. The relationship between antibody levels and the outcome was modelled using thin plate splines. I adjusted for the following confounders in all models: geographic area (12 regions in England, or Wales, Scotland or Northern Ireland) and age in years, rural-urban home address, sex, ethnicity (white vs. non-white), household size, multigenerational household, deprivation, presence of long-term health conditions, working in a care-home, having a patient-facing role in health or social care, direct or indirect contact with a hospital or care-home, smoking status, and visit frequency. Calendar time and age were included using a tensor spline which was allowed to vary by region/country[170].

Three groups were investigated, unvaccinated participants with or without evidence of prior infection (to assess the impact of prior infection), participants vaccinated with ChAdOx1, and participants vaccinated with BNT162b2. Vaccinated participants with evidence of prior infection were excluded, as there were insufficient data to model these groups separately and the relationship between antibody levels and protection may differ in these groups (see Chapter 6). Visits occurring in the 21 days prior to vaccination were excluded, as previously reported infection rates change in the run-up to vaccination. Visits in vaccinated participants were included from ≥21 days after first vaccination. In secondary analyses I grouped vaccinated individuals into one category.

Table 5.2 Frequency of visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection.

Note: Study visits were included from 17th May 2021 to 4th October 2021, i.e. while the Delta variant accounted for nearly all cases. The median (IQR) [range] days from first vaccination was 149 (115-182) [36-273] days for ChAdOx1 and 167 (123-204) [35-298] days for BNT162b2. Note that UK started its booster programme from 16th September 2021, so a small number of visits occurred in the first two weeks after initiating the booster campaign.

Days from first vaccination	ChAdOx1	BNT162b2
35-49	1,364	1,875
49-63	4,766	4,327
63-77	9,681	5,027
77-91	14,081	5,830
91-105	18,268	7,318
105-119	22,644	9,288
119-133	26,749	10,725
133-147	27,934	11,090
147-161	28,251	12,327
161-175	27,636	13,392
175-189	26,770	14,111
189-203	21,210	14,195
203-217	14,786	13,397
217-231	10,285	11,202
231-245	4,640	7,638
245-259	993	3,966
259-273	65	1,673
273-287	1	647
287-298	0	103

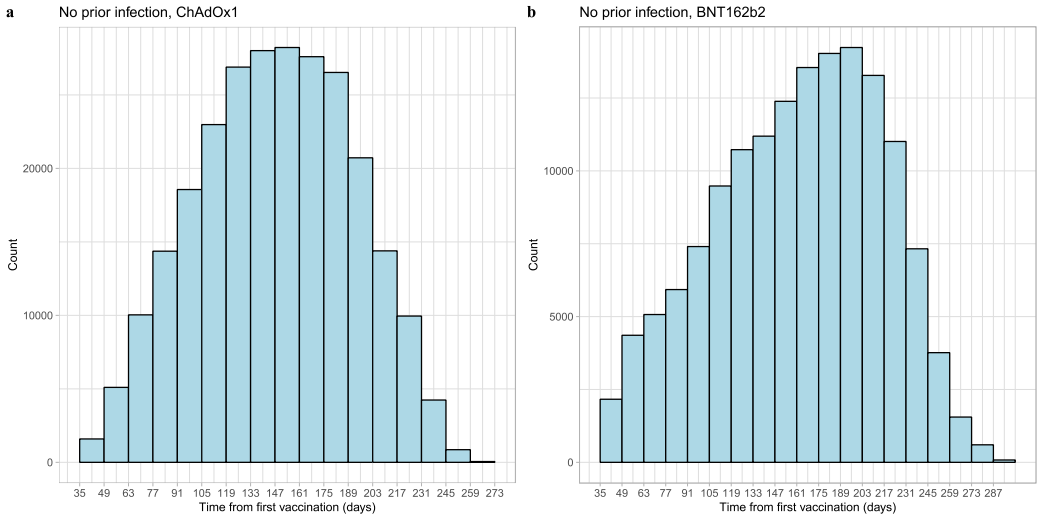


Figure 5.5 Distribution of study visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection.

Note: Study visits were included from 17th May 2021 to 4th October 2021, i.e. while the Delta variant accounted for nearly all cases. The median (IQR) [range] days from first vaccination was 149 (115-182) [36-273] days for ChAdOx1 and 167 (123-204) [35-298] days for BNT162b2.

5.2.2.3 Statistical analysis for the first dose study (cohort 3)

Again I used Bayesian linear mixed interval-censored models to estimate changes in antibody levels after the first vaccination with ChAdOX1, BNT162b2, or mRNA-1273. Antibody measurements taken after the second vaccination or after infection that happened post-first dose were excluded. I included antibody measurements from 28 days post-first dose to estimate changes from the peak level. Measurements taken after the 90th percentile of the observed time points were excluded to avoid outlier influence (76, 77, and 74 days post-first dose for ChAdOx1, BNT162b2, and mRNA-1273).

I used a multivariable model to examine the association between peak levels and antibody half-lives with continuous age (16-85 years), sex, ethnicity (white vs non-white), reporting having a long-term health condition, reporting working in patient-facing healthcare, deprivation percentile, and prior infection status. Non-white ethnicities were combined because of small sample sizes. I assumed an exponential fall in antibody levels over time, i.e., a linear decline on a log2 scale[174]. To examine non-linearity in antibody declines, especially the potential for the rate of antibody decline to flatten, I additionally fitted a model using 4-knot splines for time (knots placed at 10th, 33rd, 67th, and 90th of observed time points) and compared the model fit with the log-linear model using the leave-one-out cross-validation information criterion (LOOIC). I found that the spline model had a higher LOOIC (indicating a worse model fit) than the log-linear model for ChAdOx1 (155,331 vs 138,883) and mRNA-1273 (7,586 vs 7,570). The model fit for BNT162b2 was slightly better in the spline model than the log-linear model (107,858 vs 108,863); however, for all three models the estimated trajectories were similar and there was no evidence of antibody decline flattening up to 77 days post-first vaccination (**Figure *5.6***), so I used the log-linear model for the analysis.

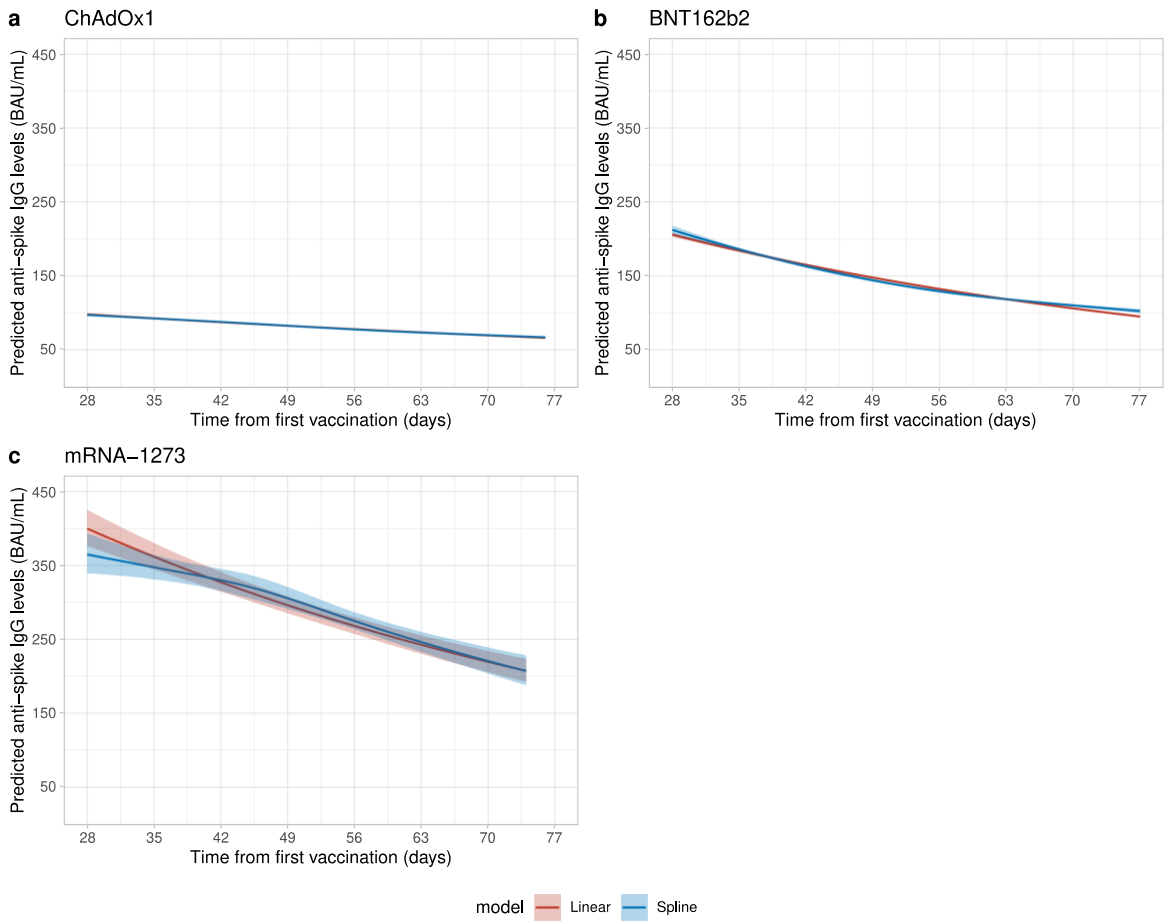


Figure 5.6 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline.

Note: The estimated trajectory from the spline model (with 4 knots placed at 10th, 33rd, 67th, and 90th of observed time points) is similar with the linear exponential model for all three vaccines, indicating that there was no evidence of antibody decline flattening

Population-level fixed effects, individual-level random effects for intercept and slope, and correlation between random effects were included in the models. The outcome was right-censored at 450 BAU/mL reflecting truncation of IgG values at the upper limit of quantification (i.e. all measurements truncated to 450 BAU/mL were considered to be >450 BAU/mL in analyses). For each model, weakly informative priors were used**.** Four chains were run per model with 4,000 iterations and a warm-up period of 2,000 iterations to ensure convergence, which was confirmed visually and by ensuring the Gelman-Rubin

statistic was <1.05**.** 95% credible intervals were calculated using highest posterior density intervals.

To evaluate whether the censored Bayesian linear model could recover the true intercept and slope, and correlation between them, in the presence of similar percentage of observations being above the upper limit of quantification as observed in my study, I performed a simulation study. The correlation between intercept and slope were varied across scenarios between -0.5 and 0.5, and slopes and intercepts were chosen to be similar to those observed in the real-world data. Within each scenario, I chose two parameter sets to mimic the different proportion of censored measurements (the proportion of censored data is 11% for ChAdOx1, 19% for BNT162b2, and 38% for mRNA-1273). I used 1000 as the sample size and performed 100 repetitions of each simulation, to reduce computational burden, providing conservative estimates of model performance in terms of bias and coverage.

5.3 Results

Figure *5.1* summarised the population and analyses from three different cohorts, which are also clearly indicated below after each subheading.

5.3.1 Study population

5.3.1.1 Study population for the initial vaccination study (cohort 1)

45,965 participants aged ≥16 years from the general population who were first vaccinated between 8th December 2020 and 6th April 2021 contributed a total of 111,360 SARS-CoV-2 anti-spike IgG measurements taken at any point between 91 days before their first vaccination through to 6 April 2021 (**Figure *5.2***). The median (IQR) age was 64 (54-71) years and 25,330 (55.1%) were female. 2,745 (6.0%) were healthcare workers, and 15,334

(33.4%) had a long-term health condition (**Table *5.3***). 5,834 (12.7%) participants with a SARS-CoV-2 PCR-positive study nose/throat swab or anti-spike IgG positive study antibody result at any time prior to vaccination were considered to have been previously infected with SARS-CoV-2, irrespective of whether they had reported previous symptoms or not. Using this definition, 3,767 (8.2%) and 2,067 (4.5%) previously infected participants then received one dose of ChAdOx1 or BNT162b2, respectively. 23,368 (50.8%), 14,894 (32.4%), and 1,869 (4.1%) participants without evidence of prior infection received one dose of ChAdOx1, one dose of BNT162b2, and two doses of BNT162b2, respectively. Among 1,869 (4.1%) participants without evidence of prior infection who received two doses of BNT162b2, the median (IQR) duration between two doses was 31 (21-47) days, with 1,020 (54.6%) ≤31 days (**Table *5.3***). Participant characteristics varied across the different vaccination groups, generally reflecting vaccine prioritisation, with proportionately more healthcare workers and the oldest individuals having received two doses of BNT162b2.

Table 5.3 Participants' characteristics by cohort in the initial vaccination study.

Note: Continuous variables were compared using ANOVA, and categorical variables were compared using Chi-squared tests. Note: higher deprivation percentile means living in a less deprived area.

	No evidence of prior infection			Evidence of prior infection		Total	p
	One dose ChAdOx1 (N=23368)	One dose BNT162b2 (N=14894)	Two doses BNT162b2 (N=1869)	One dose ChAdOx1 (N=3767)	One dose BNT162b2 (N=2067)	(N=45965)	value
Duration between two doses [Median (IQR)]			31 (21-47)				
Age							< 0.001
Median	63	66	70	61	62	64	
IQR	55, 70	54, 73	53, 80	52, 68	50, 70	54, 71	
Sex							< 0.001
Female	12510 (53.5%)	8497 (57.0%)	1153 (61.7%)	2034 (54.0%)	1136 (55.0%)	25330 (55.1%)	
Male	10858 (46.5%)	6397 (43.0%)	716 (38.3%)	1733 (46.0%)	931 (45.0%)	20635 (44.9%)	
Ethnicity							< 0.001
Non-white	1110 (4.8%)	817 (5.5%)	118 (6.3%)	241 (6.4%)	148 (7.2%)	2434 (5.3%)	
White	22258 (95.2%)	14077 (94.5%)	1751 (93.7%)	3526 (93.6%)	1919 (92.8%)	43531 (94.7%)	

	No evidence of prior infection			Evidence of prior infection		Total	p
	One dose ChAdOx1 (N=23368)	One dose BNT162b2 (N=14894)	Two doses BNT162b2 (N=1869)	One dose ChAdOx1 (N=3767)	One dose BNT162b2 (N=2067)	(N=45965)	value
Household size							< 0.001
1	5019 (21.5%)	3302 (22.2%)	501 (26.8%)	721 (19.1%)	395 (19.1%)	9938 (21.6%)	
2	12990 (55.6%)	8260 (55.5%)	985 (52.7%)	1916 (50.9%)	1072 (51.9%)	25223 (54.9%)	
3	2863 (12.3%)	1688 (11.3%)	168 (9.0%)	545 (14.5%)	312 (15.1%)	5576 (12.1%)	
4	1797 (7.7%)	1130 (7.6%)	145 (7.8%)	403 (10.7%)	202 (9.8%)	3677 (8.0%)	
5+	699 (3.0%)	514 (3.5%)	70 (3.7%)	182 (4.8%)	86 (4.2%)	1551 (3.4%)	
Deprivation percentile							< 0.001
Median	63	64	64	60	57	63	
IQR	40, 82	40, 83	38, 83	35, 81	32, 79	39, 82	
Report working in patient facing healthcare							< 0.001
No	22860 (97.8%)	13430 (90.2%)	1466 (78.4%)	3636 (96.5%)	1828 (88.4%)	43220 (94.0%)	
Yes	508 (2.2%)	1464 (9.8%)	403 (21.6%)	131 (3.5%)	239 (11.6%)	2745 (6.0%)	
Report working in person facing social care							< 0.001
No	23149 (99.1%)	14548 (97.7%)	1800 (96.3%)	3717 (98.7%)	2014 (97.4%)	45228 (98.4%)	
Yes	219 (0.9%)	346 (2.3%)	69 (3.7%)	50 (1.3%)	53 (2.6%)	737 (1.6%)	
Report working in a care home (any role)							< 0.001
No	23195 (99.3%)	14551 (97.7%)	1778 (95.1%)	3701 (98.2%)	2011 (97.3%)	45236 (98.4%)	
Yes	173 (0.7%)	343 (2.3%)	91 (4.9%)	66 (1.8%)	56 (2.7%)	729 (1.6%)	
Report having long-term health condition							< 0.001
No	15774 (67.5%)	9597 (64.4%)	1217 (65.1%)	2645 (70.2%)	1398 (67.6%)	30631 (66.6%)	
Yes	7594 (32.5%)	5297 (35.6%)	652 (34.9%)	1122 (29.8%)	669 (32.4%)	15334 (33.4%)	

5.3.1.2 Study population for the second dose study (cohort 2)

From 8 December 2020 to 4th October 2021, 222,493 participants received two ChAdOx1 or two BNT162b2 vaccinations and had at least one antibody measurement from 91 days before the first vaccination onwards. The median (interquartile range, IQR) age was 57 (43-68) years, 120,866 (54.3%) were female, and 209,898 (94.3%) reported white ethnicity. 7,071 (3.2%) reported working in patient-facing healthcare, and 62,814 (28.2%) having a long-term health condition. 121,322 (54.5%) and 79,693 (35.8%) participants without evidence of prior infection (see Methods) received two doses of ChAdOx1 or

BNT162b2, as did 12,066 (5.4%) and 9,412 (4.2%) with evidence of prior infection, respectively. These four cohorts contributed 723,844 anti-spike IgG measurements **(Figure *5.3*)**. The median (IQR) [range] dosing interval was 76 (68-78) [17-237] and 76 (66-78) [17-225] days for those receiving ChAdOx1 without or with prior infection, respectively, and 71 (58-77) [17-289] and 65 (56-76) [17-238] days respectively for BNT162b2 **(Table *5.4*).**

Table 5.4 Characteristics of participants with two vaccine doses and at least one antibody measurement from 91 days before the first vaccination through to 4th October 2021 in the second dose study.

Note: Continuous variables were compared using Kruskal-Wallis tests, and categorical variables were compared using Chi-squared tests.

	ChAdOx1 two dose no prior infection (N=121,322)	BNT162b2 two dose no prior infection (N=79,693)	ChAdOx1 two dose prior infection (N=12,066)	BNT162b2 two dose prior infection (N=9,412)	Total (N=222,493)	p value
Percentage (%)	54.5	35.8	5.4	4.2	100	
Dosing interval (days)						
Median	76	71	76	65	74	
Q1, Q3	68, 78	58, 77	66, 78	56, 76	64, 78	
Age (years)						< 0.001
Median	58	56	53	37	57	
Q1, Q3	48, 67	36, 70	45, 62	29, 57	43, 68	
Sex						< 0.001
Female	64,470 (53.1%)	44,717 (56.1%)	6,477 (53.7%)	5,202 (55.3%)	120,866 (54.3%)	
Male	56,852 (46.9%)	34,976 (43.9%)	5,589 (46.3%)	4,210 (44.7%)	101,627 (45.7%)	
Ethnicity						< 0.001
Non-white	5,698 (4.7%)	4,858 (6.1%)	932 (7.7%)	1,107 (11.8%)	12,595 (5.7%)	
White	115,624 (95.3%)	74,835 (93.9%)	11,134 (92.3%)	8,305 (88.2%)	209,898 (94.3%)	
Household size						< 0.001
1	22,809 (18.8%)	15,069 (18.9%)	1,928 (16.0%)	1,291 (13.7%)	41,097 (18.5%)	
2	59,704 (49.2%)	39,965 (50.1%)	5,065 (42.0%)	3,890 (41.3%)	108,624 (48.8%)	
3	18,176 (15.0%)	11,691 (14.7%)	2,167 (18.0%)	1,859 (19.8%)	33,893 (15.2%)	
4	15,215 (12.5%)	9,143 (11.5%)	2,075 (17.2%)	1,593 (16.9%)	28,026 (12.6%)	
5+	5,418 (4.5%)	3,825 (4.8%)	831 (6.9%)	779 (8.3%)	10,853 (4.9%)	
Deprivation percentile						< 0.001
Median	64	62	60	56	62	

	ChAdOx1 two dose no prior infection (N=121,322)	BNT162b2 two dose no prior infection (N=79,693)	ChAdOx1 two dose prior infection (N=12,066)	BNT162b2 two dose prior infection (N=9,412)	Total (N=222,493)	p value
Q1, Q3	40, 83	38, 82	35, 81	32, 78	39, 82	
Report working in patient facing healthcare						< 0.001
No	119,937 (98.9%)	74,840 (93.9%)	11,805 (97.8%)	8,840 (93.9%)	215,422 (96.8%)	
Yes	1,385 (1.1%)	4,853 (6.1%)	261 (2.2%)	572 (6.1%)	7,071 (3.2%)	
Report having a long-term health condition						< 0.001
No	87,034 (71.7%)	56,147 (70.5%)	9,083 (75.3%)	7,415 (78.8%)	159,679 (71.8%)	
Yes	34,288 (28.3%)	23,546 (29.5%)	2,983 (24.7%)	1,997 (21.2%)	62,814 (28.2%)	

5.3.1.3 Study population for the first dose study (cohort 3)

From 8th December 2020 to 31st January 2022, 80,353 included participants received at least one dose of ChAdOx1 (10,207 [12.7%] with prior infection before the first vaccination), 57,181 at least one dose of BNT162b2 (10,116 [17.7%] with prior infection, reflecting many SARS-CoV-2 exposed healthcare workers receiving BNT162b2 early in the vaccination programme), and 4,398 at least one dose of mRNA-1273 (807 [18.3%] with prior infection, reflecting this vaccine being used later in the pandemic in younger age groups with higher infection rates) (**Table *5.5***). The median age was 50 years (interquartile range IQR: 37-63). 75,593 (53.3%) were female, 131,365 (92.6%) reported white ethnicity, 1,243 (0.9%) black ethnicity, 5,699 (4.0%) Asian ethnicity, 2,241 (1.6%) mixed ethnicity, and 1,384 (1.0%) another ethnicity. 2,577 (1.8%) were healthcare workers, and 33,599 (23.7%) reported having a long-term health condition.

Table 5.5 Characteristics of participants with a single ChAdOx1, BNT162b2, or mRNA-1273 vaccination in the first dose study.

Note: Participants had at least one antibody measurement from 91 days before the first vaccination up to the second vaccination (if received) or breakthrough infections post-first vaccination. Continuous variables were compared using Kruskal-Wallis tests, and categorical variables were compared using Chi-squared tests.

	Without prior infection			With prior infection				
	ChAdOx1 (N=70,146)	**BNT162b2 (N=47,065)**	**mRNA-1273 (N=3,591)**	**ChAdOx1 (N=10,207)**	**BNT162b2 (N=10,116)**	**mRNA-1273 (N=807)**	**Total (N=141,932)**	**p value**
Age								< 0.001
Median	55	39	35	52	34	32	50	
Q1, Q3	46, 64	30, 63	28, 41	44, 61	24, 48	26, 39	37, 63	
Sex								< 0.001
Female	37,109 (52.9%)	25,469 (54.1%)	1,751 (48.8%)	5,424 (53.1%)	5,466 (54.0%)	374 (46.3%)	75,593 (53.3%)	
Male	33,037 (47.1%)	21,596 (45.9%)	1,840 (51.2%)	4,783 (46.9%)	4,650 (46.0%)	433 (53.7%)	66,339 (46.7%)	
Ethnicity								< 0.001
White	65,757 (93.7%)	43,412 (92.2%)	3,315 (92.3%)	9,333 (91.5%)	8,857 (87.6%)	691 (85.6%)	131,365 (92.6%)	
Black	563 (0.8%)	365 (0.8%)	18 (0.5%)	124 (1.2%)	162 (1.6%)	11 (1.4%)	1,243 (0.9%)	
Asian	2,366 (3.4%)	1973 (4.2%)	135 (3.7%)	510 (5.0%)	660 (6.5%)	55 (6.8%)	5,699 (4.0%)	
Mixed	860 (1.2%)	845 (1.8%)	80 (2.2%)	130 (1.3%)	295 (2.9%)	31 (3.8%)	2,241 (1.6%)	
Others	600 (0.9%)	470 (1.0%)	43 (1.2%)	110 (1.1%)	142 (1.4%)	19 (2.4%)	1,384 (1.0%)	
Household size								< 0.001
1	11,496 (16.4%)	6,679 (14.2%)	406 (11.3%)	1,466 (14.4%)	1,064 (10.5%)	83 (10.3%)	21,194 (14.9%)	
2	31,449 (44.8%)	20,430 (43.4%)	1,297 (36.1%)	3,994 (39.1%)	3,586 (35.4%)	276 (34.2%)	61,032 (43.0%)	
3	11,660 (16.6%)	8,371 (17.8%)	759 (21.1%)	1,897 (18.6%)	2,064 (20.4%)	162 (20.1%)	24,913 (17.6%)	
4	11,297 (16.1%)	7,960 (16.9%)	835 (23.3%)	2,010 (19.7%)	2,193 (21.7%)	194 (24.0%)	24,489 (17.3%)	
5+	4,244 (6.1%)	3,625 (7.7%)	294 (8.2%)	840 (8.2%)	1,209 (12.0%)	92 (11.4%)	10,304 (7.3%)	
Deprivation percentile								< 0.001
Median	62	60	58	59	56	55	61	
Q1, Q3	38, 82	36, 81	35, 81	34, 81	31, 78	29, 79	37, 81	

	Without prior infection			With prior infection				
	ChAdOx1 (N=70,146)	BNT162b2 (N=47,065)	mRNA-1273 (N=3,591)	ChAdOx1 (N=10,207)	BNT162b2 (N=10,116)	mRNA-1273 (N=807)	Total (N=141,932)	p value
Region								< 0.001
North East	3,019 (4.3%)	1,790 (3.8%)	105 (2.9%)	450 (4.4%)	426 (4.2%)	33 (4.1%)	5,823 (4.1%)	
North West	7,288 (10.4%)	4,921 (10.5%)	326 (9.1%)	1,499 (14.7%)	1,380 (13.6%)	103 (12.8%)	15,517 (10.9%)	
Yorkshire and the Humber	6,060 (8.6%)	3,632 (7.7%)	417 (11.6%)	931 (9.1%)	847 (8.4%)	99 (12.3%)	11,986 (8.4%)	
East Midlands	4,822 (6.9%)	3,190 (6.8%)	120 (3.3%)	708 (6.9%)	646 (6.4%)	20 (2.5%)	9,506 (6.7%)	
West Midlands	5,822 (8.3%)	3,696 (7.9%)	191 (5.3%)	854 (8.4%)	787 (7.8%)	33 (4.1%)	11,383 (8.0%)	
East of England	6,671 (9.5%)	4,366 (9.3%)	345 (9.6%)	956 (9.4%)	862 (8.5%)	57 (7.1%)	13,257 (9.3%)	
London	8,546 (12.2%)	7,932 (16.9%)	765 (21.3%)	2,015 (19.7%)	2,473 (24.4%)	232 (28.7%)	21,963 (15.5%)	
South East	10,237 (14.6%)	5,926 (12.6%)	567 (15.8%)	1,217 (11.9%)	1,017 (10.1%)	110 (13.6%)	19,074 (13.4%)	
South West	6,154 (8.8%)	4,075 (8.7%)	313 (8.7%)	513 (5.0%)	542 (5.4%)	44 (5.5%)	11,641 (8.2%)	
Northern Ireland	1,859 (2.7%)	1,745 (3.7%)	21 (0.6%)	187 (1.8%)	278 (2.7%)	8 (1.0%)	4,098 (2.9%)	
Scotland	5,861 (8.4%)	3,835 (8.1%)	346 (9.6%)	514 (5.0%)	529 (5.2%)	60 (7.4%)	11,145 (7.9%)	
Wales	3,807 (5.4%)	1,957 (4.2%)	75 (2.1%)	363 (3.6%)	329 (3.3%)	8 (1.0%)	6,539 (4.6%)	
Report working in patient-facing healthcare								< 0.001
No	69,449 (99.0%)	45,763 (97.2%)	3,560 (99.1%)	10,032 (98.3%)	9,751 (96.4%)	800 (99.1%)	139,355 (98.2%)	
Yes	697 (1.0%)	1,302 (2.8%)	31 (0.9%)	175 (1.7%)	365 (3.6%)	7 (0.9%)	2,577 (1.8%)	
Report having a long-term health condition								< 0.001
No	52,348 (74.6%)	35,934 (76.3%)	3,131 (87.2%)	7,844 (76.8%)	8,349 (82.5%)	727 (90.1%)	108,333 (76.3%)	
Yes	17,798 (25.4%)	11,131 (23.7%)	460 (12.8%)	2,363 (23.2%)	1,767 (17.5%)	80 (9.9%)	33,599 (23.7%)	

5.3.2 Anti-spike IgG antibody positivity after vaccination and its predictors in the initial vaccination study (cohort 1)

In participants without evidence of prior infection, models of binary (positive vs. negative) post-vaccine antibody responses showed positive anti-spike IgG results increased over the 2-4 weeks post first vaccination and varied significantly by age (**Figure *5.7*, Figure *5.8*, Figure *5.9***; observed numbers/percentages in **Figure *5.10*, Figure *5.11*, Figure *5.12*, Figure *5.13*, Figure *5.14***). Fewer older participants were seropositive after receiving a single dose of ChAdOx1 or BNT162b2, e.g., the estimated percentage of seropositive 80-year-olds was 74% (95%CI 66-80%) and 85% (80-89%) respectively 28 days after first vaccination, respectively, compared with 79% (75-83%) and 91% (89-93%) respectively for 60-year-olds and 84% (76-89%) and 95% (92-97%) respectively for 40-year-olds (**Table *5.6***). In contrast, two doses of BNT162b2 vaccine achieved >90% seropositivity 28-72 days after the first vaccination regardless of age, although there was some evidence of waning in those only receiving a first BNT162b2 doses at older ages. There was no evidence of differences in seropositivity rates 14-42 days post-first vaccine in those of younger ages (e.g. 20, 40 years) receiving one dose or two doses of BNT162b2 but greater rates of seroconversion were seen in older individuals (e.g. 80 years) receiving two doses (**Figure *5.8*, Figure *5.9***). There was no evidence of seropositivity declines following first vaccine dose in older individuals receiving a single dose of ChAdOx1.

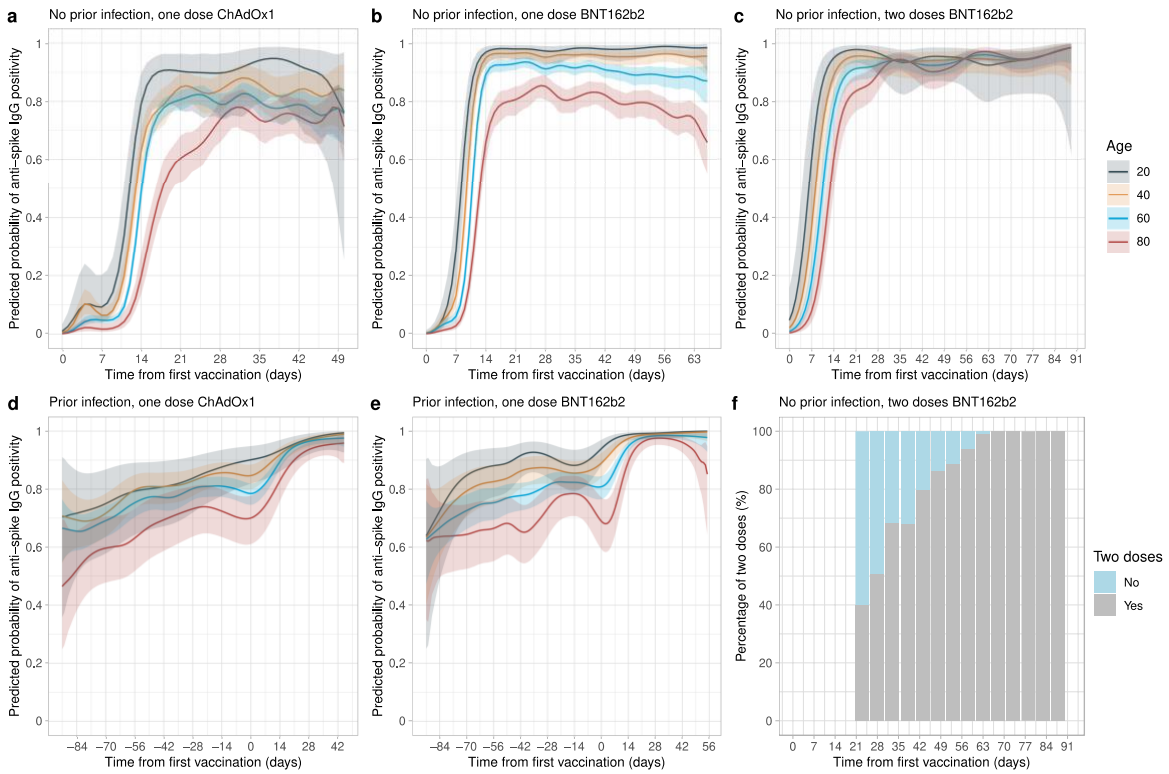


Figure 5.7 Predicted probability of anti-spike IgG positivity by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by vaccination and prior infection.

Note: **a,** No prior infection, received one dose of ChAdOx1. **b**, No prior infection, received one dose of BNT162b2. **c**, No prior infection, received two doses of BNT162b2. **d**, With prior infection, received one dose of ChAdOx1. **e**, With prior infection, received one dose of BNT162b2. **f,** For those who received two doses of BNT162b2 without prior infection, the percentage of participants having had two vaccine doses by each timepoint (Grey: had two doses; Blue: had only one dose). Different x-axis scales reflect different durations of follow-up post-vaccination in the different cohorts. Line colour indicates antibody response predicted for ages 20, 40, 60, and 80 years (see **Figure 5.9** for full model across all ages and comparisons of vaccine type by age). The 95% confidence intervals are calculated by prediction ± 1.96*standard error of prediction.

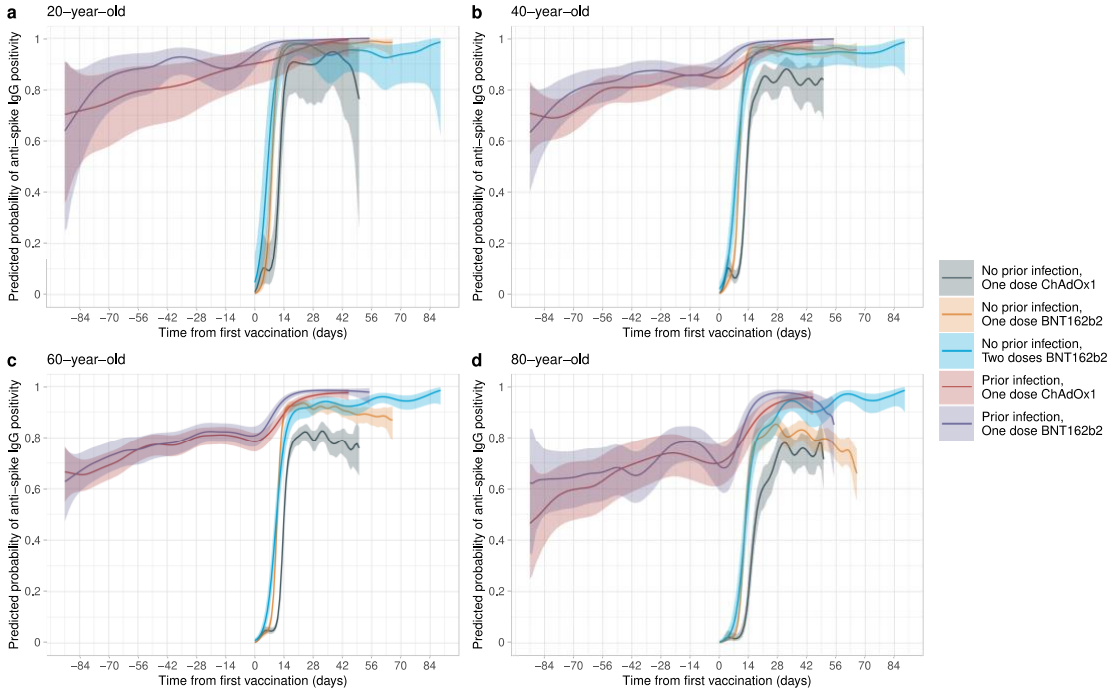


Figure 5.8 Predicted probability of anti-spike IgG positivity by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by age.

Note: **a**, 20-year-old. **b**, 40-year-old. **c**, 60-year-old. **d**, 80-year-old. Line colour indicates different vaccine type and prior infection status. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of prediction. Data identical to **Figure 5.7**, but **Figure 5.7** panels represent age rather than vaccine type.

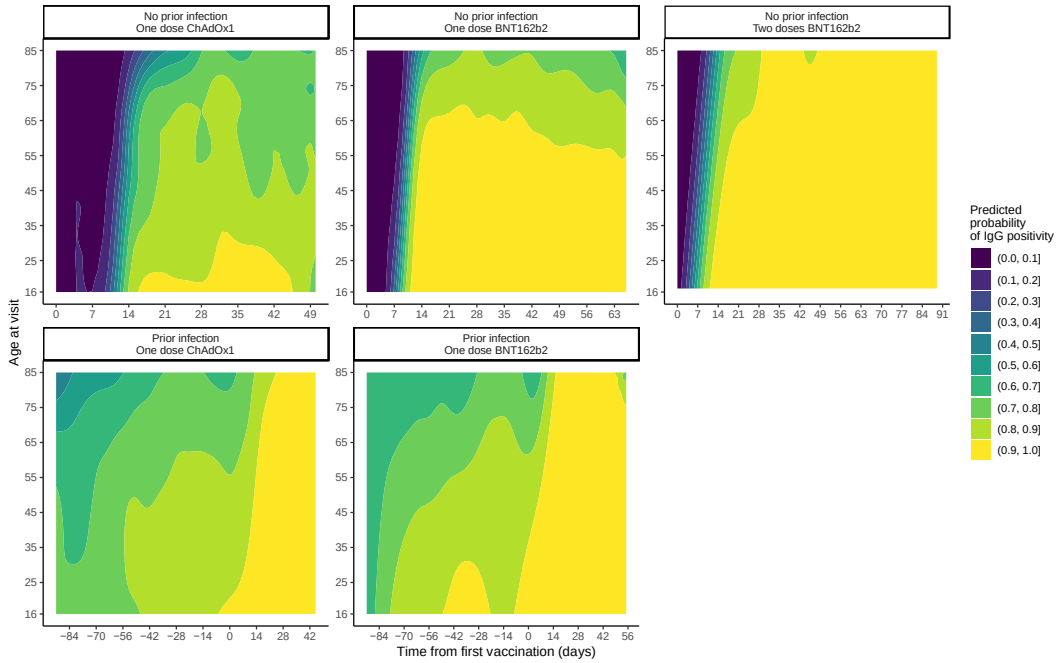


Figure 5.9 Predicted probability of anti-spike IgG positivity by time from first vaccination and age, according to vaccine type and prior infection status (full model).

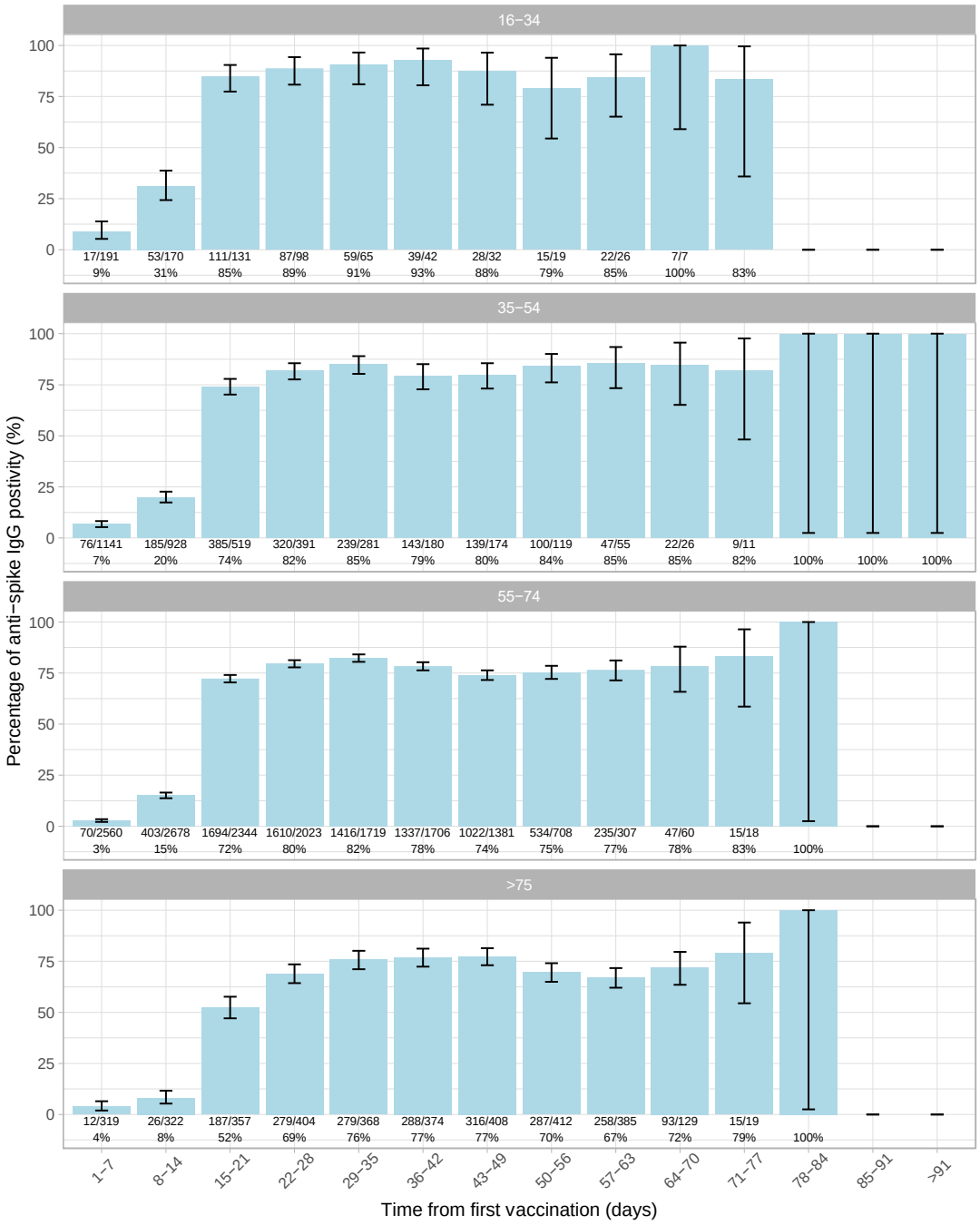


Figure 5.10 Unadjusted percentage (95% CI) anti-spike IgG positive by days after first vaccination in 23,368 participants receiving a single dose of ChAdOx1 vaccine and without evidence of prior infection.

Note: Results were divided into four age groups: 16-34, 35-54, 55-74, and >75 years. Exact binomial test was performed to obtain the 95% confidence intervals. Numbers above the x-axis represent positive number/total number, and the percentage of positivity (%).



Figure 5.11 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 3,767 participants receiving a single dose of ChAdOx1 vaccine and with evidence of prior infection.

Note: Results were divided into four age groups: 16-34, 35-54, 55-74, and >75 years. Exact binomial test was performed to obtain the 95% confidence intervals. Numbers above the x-axis represent positive number/total number, and the percentage of positivity (%).

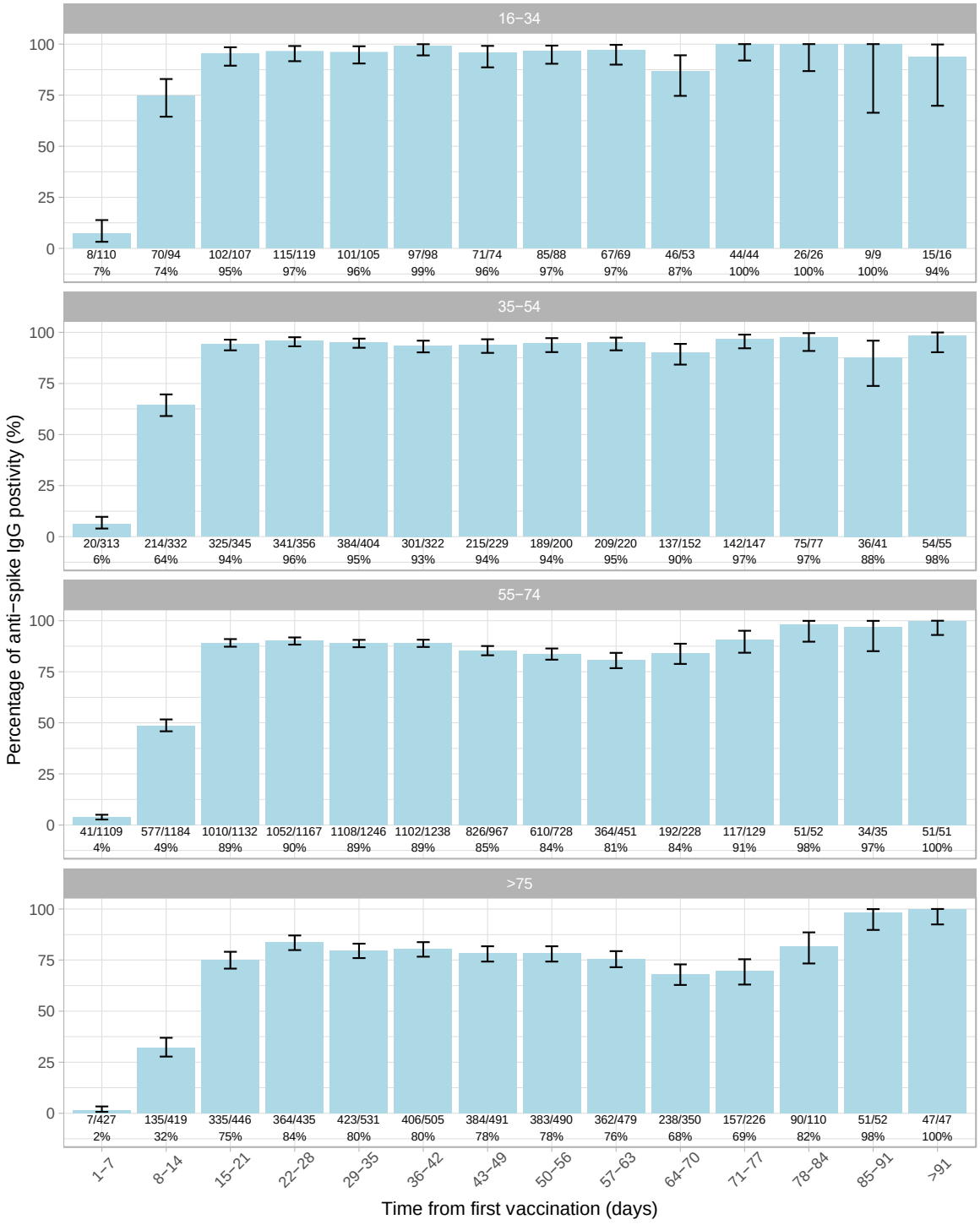


Figure 5.12 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 14,894 participants receiving a single dose of BNT162b2 vaccine and without evidence of prior infection.

Note: Results were divided into four age groups: 16-34, 35-54, 55-74, and >75 years. Exact binomial test was performed to obtain the 95% confidence intervals. Numbers above the x-axis represent positive number/total number, and the percentage of positivity (%).



Figure 5.13 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 2,067 participants receiving a single dose of BNT162b2 vaccine and with evidence of prior infection.

Note: Results were divided into four age groups: 16-34, 35-54, 55-74, and >75 years. Exact binomial test was performed to obtain the 95% confidence intervals. Numbers above the x-axis represent positive number/total number, and the percentage of positivity (%).

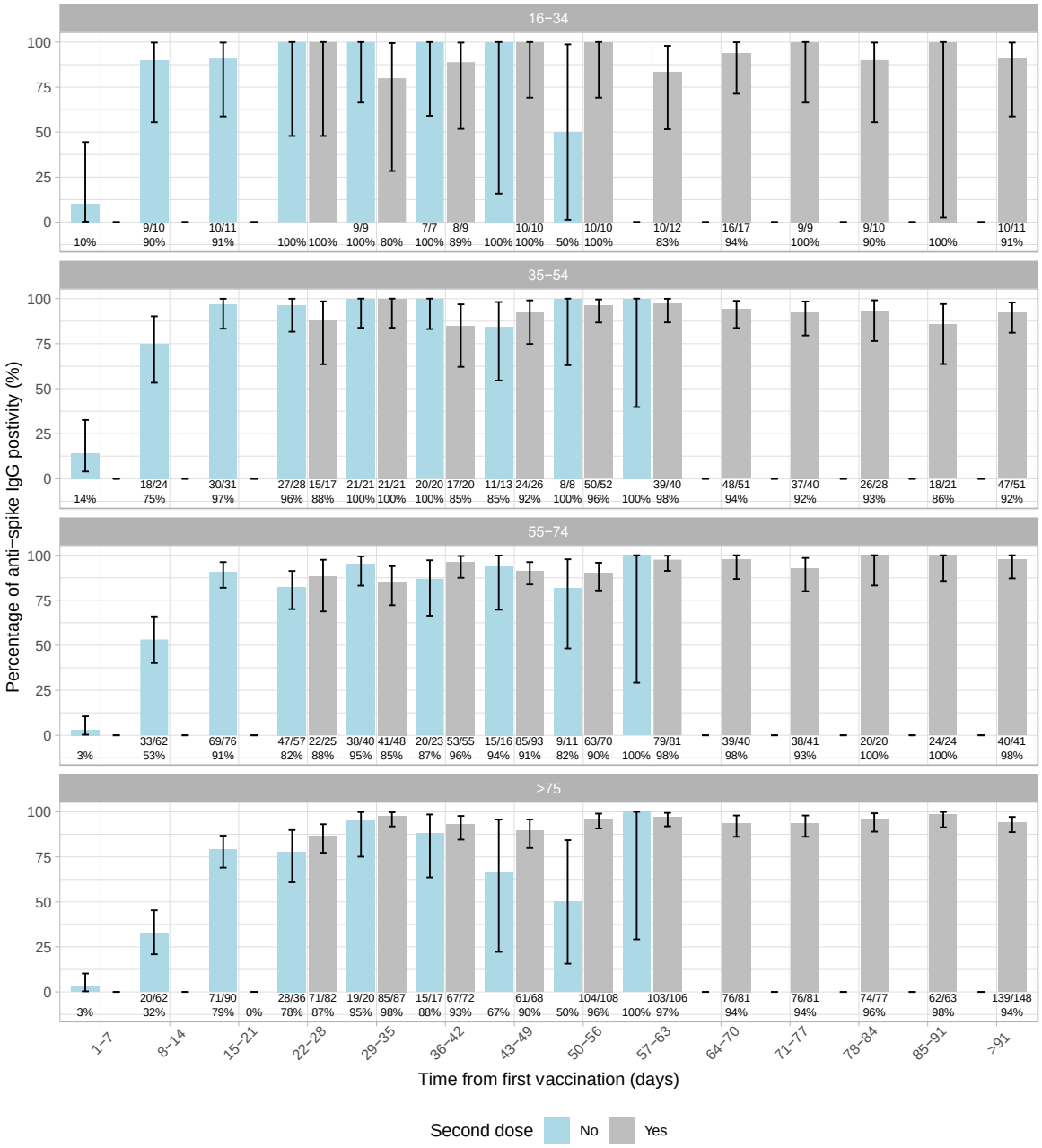


Figure 5.14 Unadjusted percentage (95% CI) anti-spike IgG positives by days after first vaccination in 1,869 participants receiving two doses of BNT162b2 vaccine and without evidence of prior infection.

Note: Results were divided into four age groups: 16-34, 35-54, 55-74, and >75 years. Exact binomial test was performed to obtain the 95% confidence intervals. Numbers above the x-axis represent positive number/total number, and the percentage of positivity (%).

Table 5.6 Predicted probability of anti-spike IgG seropositivity with 95% confidence interval (CI) according to age, vaccine type and dose, and prior infection status by weeks after vaccination.

Note: The 95% confidence intervals are calculated by prediction ± 1.96*standard error of prediction.

		No prior infection			Prior infection	
Time (days)	**Age**	**ChAdOx1 one dose**	**BNT162b2 one dose**	**BNT162b2 two doses**	**ChAdOx1 one dose**	**BNT162b2 one dose**
Logistic model		**Predicted probability of anti-spike IgG positivity (95%CI)**				
0	20	1 (0-3)	0 (0-1)	5 (1-16)	90 (82-95)	94 (88-97)
0	40	1 (0-1)	0 (0-1)	2 (1-5)	85 (80-88)	89 (84-92)
0	60	0 (0-0)	0 (0-0)	1 (1-2)	78 (75-82)	81 (77-84)
0	80	0 (0-0)	0 (0-0)	0 (0-1)	70 (61-78)	69 (59-77)
7	20	9 (4-20)	29 (17-45)	60 (34-81)	91 (83-96)	97 (94-99)
7	40	6 (4-10)	13 (9-18)	36 (23-52)	88 (83-91)	94 (90-96)
7	60	5 (4-6)	6 (4-8)	18 (12-25)	82 (78-85)	87 (83-90)
7	80	2 (1-4)	3 (2-5)	8 (4-14)	74 (66-81)	73 (64-81)
14	20	79 (62-90)	97 (92-99)	95 (87-98)	93 (86-97)	99 (97-100)
14	40	63 (55-71)	95 (92-97)	89 (80-94)	92 (89-94)	97 (95-99)
14	60	49 (44-53)	89 (86-91)	78 (70-83)	89 (87-92)	95 (93-96)
14	80	20 (14-28)	66 (59-72)	60 (49-70)	83 (76-88)	90 (84-93)
21	20	90 (77-96)	98 (94-99)	98 (93-99)	96 (89-98)	99 (97-100)
21	40	85 (78-90)	96 (94-98)	96 (91-98)	95 (92-97)	99 (97-99)
21	60	81 (77-84)	93 (91-95)	91 (87-94)	95 (93-96)	98 (97-98)
21	80	60 (52-68)	81 (75-85)	84 (77-89)	90 (85-93)	96 (93-98)
28	20	90 (74-96)	97 (93-99)	97 (89-99)	97 (92-99)	99 (97-100)
28	40	84 (76-89)	95 (92-97)	95 (90-98)	97 (94-98)	99 (98-100)
28	60	79 (75-83)	91 (89-93)	93 (89-95)	96 (95-97)	98 (97-99)
28	80	74 (66-80)	85 (80-89)	89 (84-93)	94 (90-96)	98 (95-99)
35	20	94 (79-99)	98 (95-99)	93 (83-98)	99 (93-100)	99 (97-100)
35	40	86 (78-91)	96 (93-97)	94 (88-97)	98 (95-99)	99 (98-100)
35	60	80 (75-84)	91 (89-93)	94 (91-96)	97 (96-98)	98 (97-99)
35	80	74 (67-80)	81 (76-85)	94 (90-97)	95 (92-97)	97 (95-99)
42	20	92 (72-98)	98 (94-99)	95 (86-98)	99 (93-100)	100 (98-100)
42	40	84 (75-90)	96 (93-97)	94 (88-97)	99 (95-100)	99 (98-100)
42	60	79 (73-84)	91 (88-93)	93 (89-95)	97 (95-99)	98 (97-99)
42	80	75 (67-81)	82 (77-86)	91 (85-95)	96 (91-98)	96 (93-98)
49	20	80 (39-96)	98 (94-99)	95 (87-98)	-	100 (98-100)
49	40	84 (71-92)	95 (92-97)	94 (89-97)	-	100 (97-100)
49	60	78 (67-86)	89 (85-92)	93 (89-95)	-	98 (96-99)
49	80	78 (67-86)	79 (74-84)	91 (85-94)	-	92 (86-96)

In participants with prior evidence of infection, before vaccination younger participants were more likely to be seropositive, e.g. on the day of vaccination 90% (95%CI 82-95%) for 20-year-olds, 85% (80-88%) for 40-year-olds, 78% (75-82%) for 60-year-olds, and 70% (61-78%) for 80-year-olds receiving ChAdOx1 (**Table *5.6*;** same trend for BNT162b2). A high percentage of participants achieved positive antibody responses 28 days after vaccination (≥94%) regardless of age and vaccine given, similar to the positivity rate in participants without prior infection who received two doses of BNT162b2 (**Figure *5.8*, Figure *5.9***).

28,144 participants had an anti-spike IgG measurement 14-60 days after their first ChAdOx1 or BNT162b2 vaccination, of whom 24,977 (88.7%) had no evidence of prior infection and were included in a separate logistic regression analysis to investigate associations with antibody positivity. 20,505 (82.1%) had a positive post-vaccine anti-spike IgG result. Age, sex, vaccine type, ethnicity, social deprivation, healthcare roles, and long-term health conditions were associated with seropositivity after vaccination (**Table *5.7***). Consistent with **Figure *5.7***, anti-spike IgG positivity decreased with older age; the association was non-linear, with seropositivity dropping faster in those >75 years (**Figure *5.15***). There was evidence of effect modification between age and sex: at younger ages (30-55y) similar rates of seroconversion were seen in males and females (e.g. in 40 years, adjusted odds ratio, aOR=0.91 [95%CI 0.68-1.22]), but at older ages (>60y) men were less likely to seroconvert (e.g. aOR=0.65 [0.59-0.72] for 70 year-olds) (**Figure *5.15***, interaction p=0.02). Seroconversion by 60 days was less common following ChAdOx1 than after BNT162b2 (aOR=0.47 [95%CI 0.44-0.51]), while receiving two doses of BNT162b2 increased seroconversion compared to one BNT162b2 dose (aOR=2.11 [1.69-2.66]). Patient-facing healthcare workers were more likely to be anti-spike IgG positive by 60 days post-vaccination (aOR=1.63 [1.29-2.08]) and participants who had long-term

health conditions were less likely (aOR=0.64 [0.60-0.69]). There was evidence of greater seropositivity post-vaccination in participants from non-white ethnic groups (aOR=1.54 [1.27-1.90]). A 10-unit increase in deprivation percentile (i.e. decrease in deprivation) resulted in higher seropositivity post-vaccination (aOR=1.28 [1.13-1.46]). There was no evidence of independent associations between antibody positivity and household size or working in social care or long-term care facilities.

Table 5.7 Predictors of antibody positivity 14-60 days post first vaccination in participants without evidence of prior infection from univariable and multivariable logistic regression models.

Note: The 95% confidence intervals are calculated by estimate ± 1.96*standard error of the estimate; Wald p values are shown.

	14-60 days post vaccination			Univariable			Multivariable		
	Negative (N=4472)	**Positive (N=20505)**	**p value**	**OR**	**95%CI**	**p value**	**OR**	**95%CI**	**p value**
Age			< 0.001			<0.001	See **Figure *5.15*** for		
Median	69	67							
IQR	62, 74	58, 73					combined effect of		
Sex			< 0.001						
Female	2085 (46.6%)	11726 (57.2%)		1(ref)					
Male	2387 (53.4%)	8779 (42.8%)		0.65	0.61-0.70	<0.001	age and sex		
Vaccine type			< 0.001						
One dose BNT162b2	1323 (29.6%)	9141 (44.6%)		1(ref)			1(ref)		
One dose ChAdOx1	3058 (68.4%)	10231 (49.9%)		0.48	0.45-0.52	<0.001	0.47	0.44-0.51	<0.001
Two doses BNT162b2	91 (2.0%)	1133 (5.5%)		1.80	1.45-2.26	<0.001	2.11	1.69-2.66	<0.001
Ethnicity			< 0.001						
White	4356 (97.4%)	19529 (95.2%)		1(ref)			1(ref)		
Non-white	116 (2.6%)	976 (4.8%)		1.88	1.55-2.29	<0.001	1.54	1.27-1.90	<0.001
Household size			< 0.001						
1	1160 (25.9%)	4704 (22.9%)		1(ref)			1(ref)		
2	2621 (58.6%)	11900 (58.0%)		1.12	1.04-1.21	0.004	1.10	1.02-1.19	0.02
3	405 (9.1%)	2140 (10.4%)		1.30	1.15-1.48	<0.001	1.03	0.90-1.17	0.7
4	202 (4.5%)	1217 (5.9%)		1.49	1.27-1.75	<0.001	0.96	0.80-1.14	0.6
5+	84 (1.9%)	544 (2.7%)		1.60	1.27-2.04	<0.001	1.01	0.79-1.31	0.9
Deprivation percentile			0.001						
Median	63	64		(10 unit increase)			(10 unit increase)		
IQR	37, 82	40, 83		1.22	1.08-1.38	0.001	1.28	1.13-1.46	<0.001
Report working in patient facing healthcare			< 0.001						
No	4386 (98.1%)	19249 (93.9%)		1(ref)			1(ref)		
Yes	86 (1.9%)	1256 (6.1%)		3.33	2.69-4.18	<0.001	1.63	1.29-2.08	<0.001
Report working in person facing social care			<0.001						
No	4427 (99.0%)	20154 (98.3%)		1(ref)			1(ref)		
Yes	45 (1.0%)	351 (1.7%)		1.71	1.27-2.37	<0.001	1.00	0.72-1.40	1
Report working in care home (any role)			< 0.001						
No	4445 (99.4%)	20177 (98.4%)		1(ref)			1(ref)		
Yes	27 (0.6%)	328 (1.6%)		2.68	1.84-4.06	<0.001	1.24	0.82-1.93	0.3
Report having long-term health condition			< 0.001						
No	2404 (53.8%)	13427 (65.5%)		1(ref)			1(ref)		
Yes	2068 (46.2%)	7078 (34.5%)		0.61	0.57-0.65	<0.001	0.64	0.60-0.69	<0.001

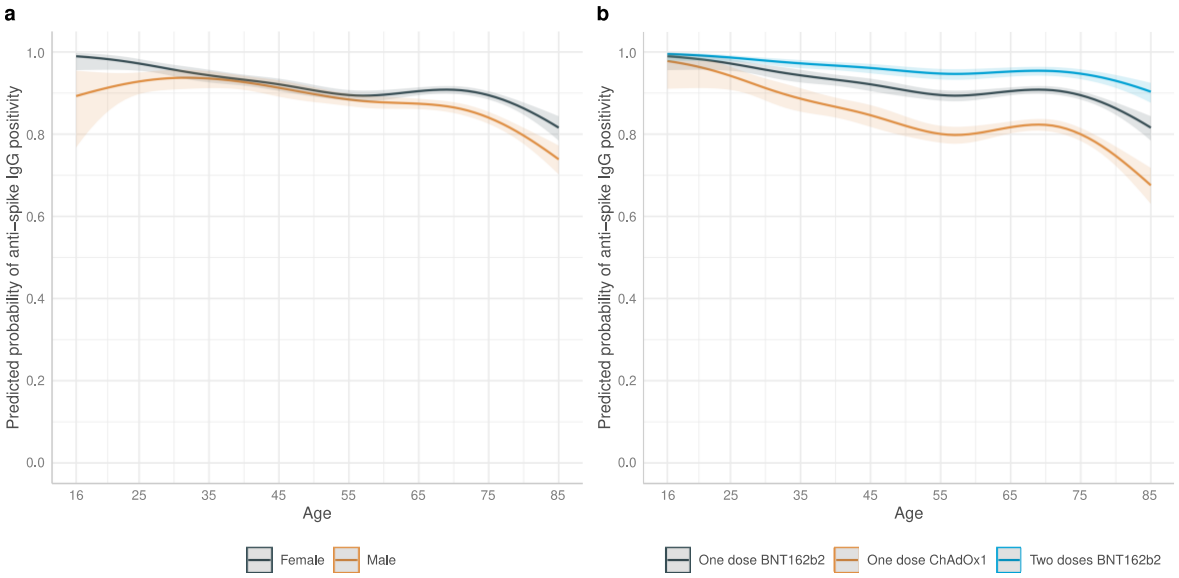


Figure 5.15 Predicted probability with 95% confidence intervals of spike IgG positivity 14-60 days after first vaccination in 24,977 participants without evidence of prior infection from a multivariable logistic regression model.

Note**: a,** Predicted probability of anti-spike IgG positivity by age and sex. **b**, Predicted probability of anti-spike IgG positivity by age and vaccine type. Age was fitted using natural cubic spline with four internal knots placed at 20th, 40th, 60th, and 80th percentile (30, 44, 57, 71 years) and two boundary knots at 5th and 95th percentile (19, 82 years). The 95% confidence intervals were calculated by prediction ± 1.96*standard error of the prediction and are shown as the shaded area. The interaction between sex and age was tested using a likelihood ratio test (p=0.02). Values are plotted at the reference category for other variables (BNT162b2 one dose (Panel A)/female (Panel B), white ethnicity, IMD=55, household size=1, did not work in patient-facing healthcare or social care, did not work in a care home, no long-term health condition).

5.3.3 Quantitative anti-spike IgG antibody responses after the first and second vaccination in the initial vaccination and second dose study (cohort 1 and 2)

5.3.3.1 After the first vaccination (cohort 1)

In participants without evidence of prior infection, quantitative anti-spike IgG levels followed similar patterns to binary IgG positivity post-vaccination (**Figure *5.16***, **Figure *5.17***, **Figure *5.18*)**. Following a single dose of BNT162b2 or ChAdOx1, older participants reached lower peak levels and levels rose more slowly than in younger participants. Anti-spike IgG levels were initially lower after a single dose of ChAdOx1 than after BNT162b2, e.g. by 28 days post-vaccine, the IgG levels were 41 (95%CI 36-45) and 63 (58-69) for 80-year-olds, 53 (49-56) and 91 (86-98) for 60-year-olds, 63 (56-72) and 132

(120-146) for 40-year-olds, and 71 (53-96) and 187 (149-235) BAU/mL for 20-year-olds, respectively. As context, in a prior study, protection from re-infection began to rise at antibody titres of ~10 BAU/mL, and increased as antibody titres rose to ~140 BAU/mL, with 50% protection against any PCR-positive result (symptomatic or asymptomatic) achieved at titres of 16 BAU/mL[112]. The rate of increase in antibody levels was also slightly slower following the ChAdOx1 vaccine, e.g. the estimated mean time to reaching the threshold for antibody positivity post-first vaccine in 40-year-olds was 10 days after receiving BNT162b2 but 14 days after receiving ChAdOx1 (**Figure *5.16***). However, antibody levels gradually decreased from ~35 days post-vaccination in participants receiving a single dose of BNT162b2 (**Figure *5.16***), while there was no evidence of decrease in those receiving a single ChAdOx1 dose up to 49 days post-vaccination. Hence, differences in mean antibody levels between single doses of the two vaccines attenuated over time, particularly at older ages.

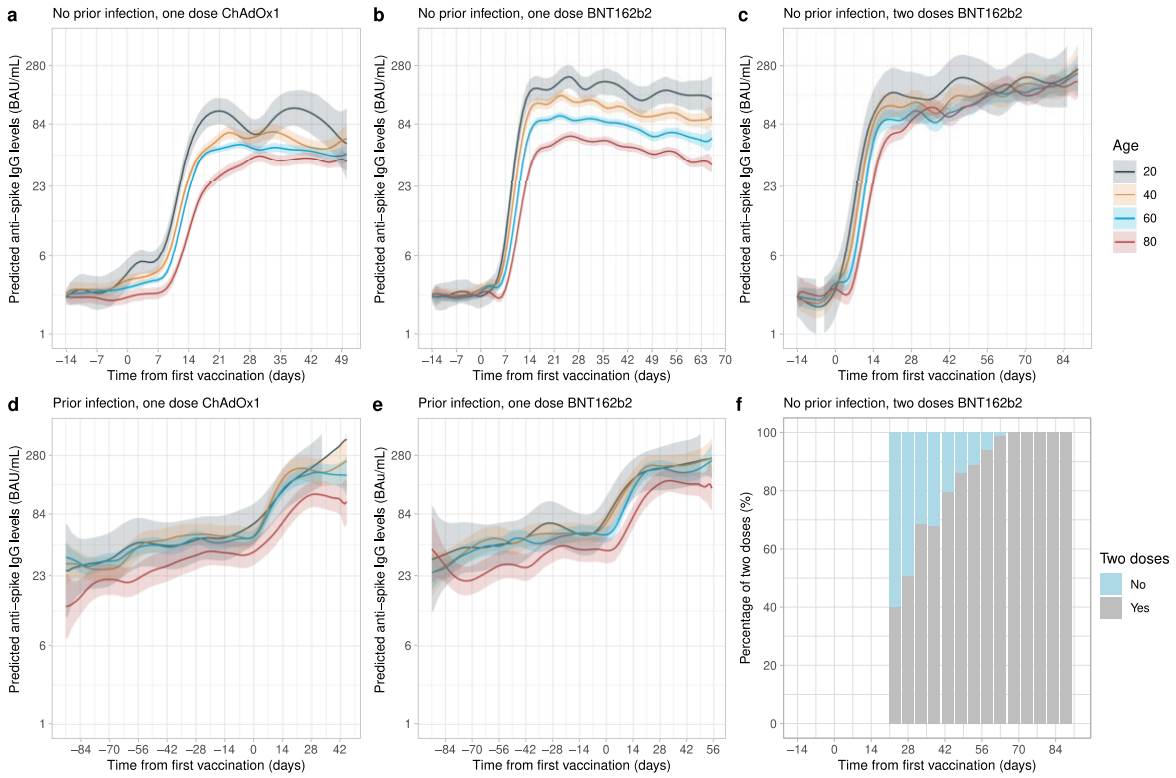


Figure 5.16 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by vaccination type and prior infection.

Note**: a,** No prior infection, received one dose of ChAdOx1. **b**, No prior infection, received one dose of BNT162b2. **c**, No prior infection, received two doses of BNT162b2. **d**, With prior infection, received one dose of ChAdOx1. **f,** For those who received two doses of BNT162b2 without prior infection, the percentage of participants having had two vaccine doses by each timepoint (Grey: had two doses; Blue: had only one dose). Different x-axis scales reflect different durations of follow-up post-vaccination in the different cohorts. Predicted levels are plotted on a log scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Line colour indicates response predicted for ages 20, 40, 60, and 80 years (see **Figure 5.17** for comparisons of vaccine type by age and **Figure 5.18** for all ages). The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction.

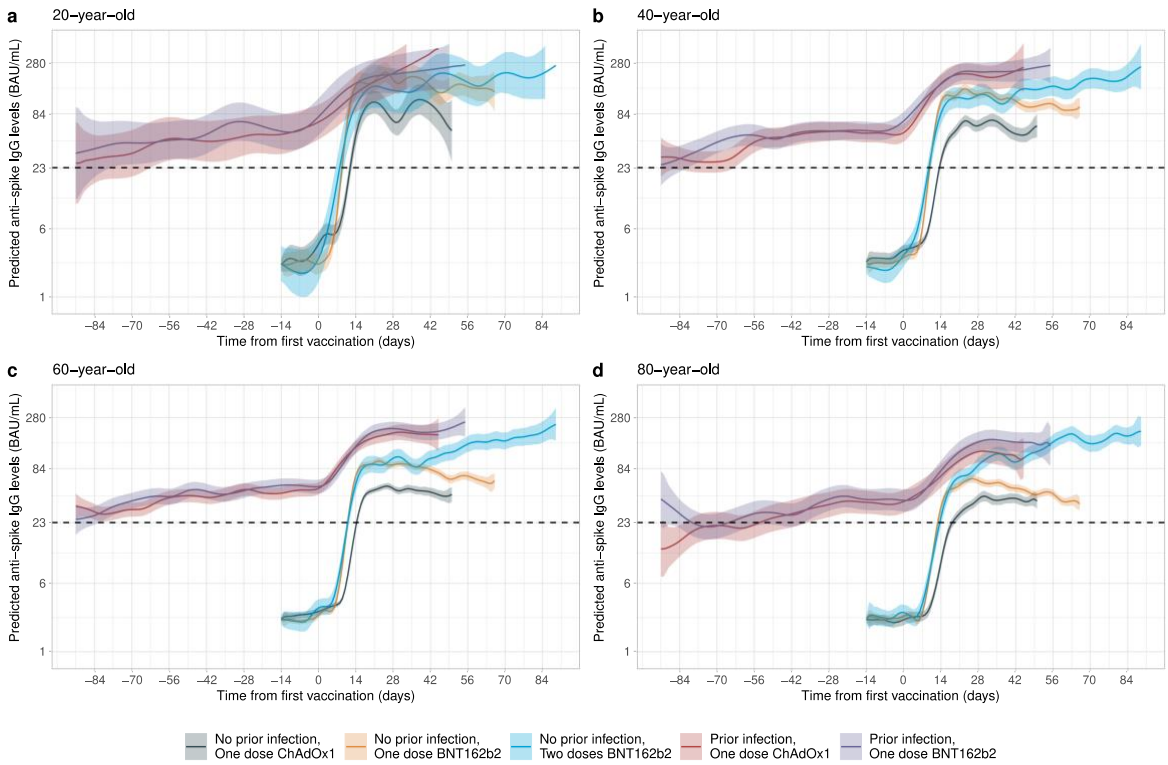


Figure 5.17 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, faceted by age.

Note: **a,** 20-year-old. **b,** 40-year-old. **c,** 60-year-old. **d,** 80-year-old. Line colour indicates predicted response for different vaccine type and prior infection status (full models shown in **Figure 5.18**, plotted by vaccine in **Figure 5.16**). The 95% confidence intervals are calculated by prediction ± 1.96*standard error of the prediction.

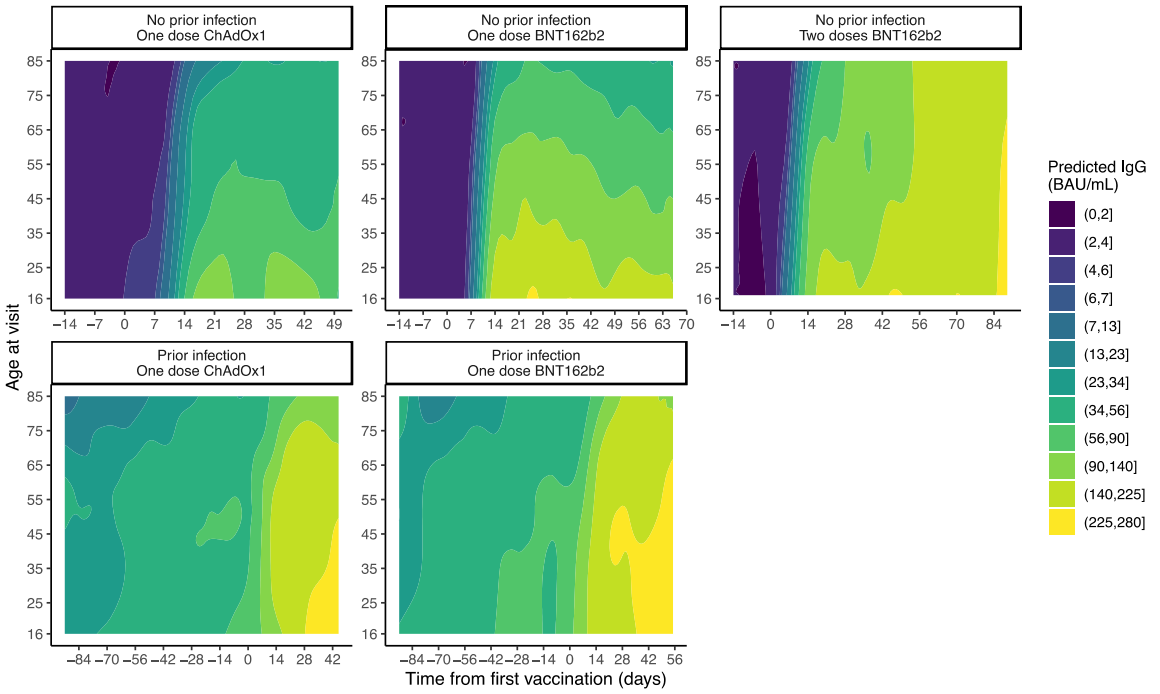


Figure 5.18 Predicted anti-spike IgG levels (BAU/mL) by time from first vaccination and age based on data from 40,131 participants without prior infection and 5,834 participants with prior infection, according to vaccine type and prior infection status (full model).

In participants with evidence of prior infection, whilst vaccination increased antibody levels at all ages, the absolute increases were more modest. Older previously-infected participants had lower IgG levels compared with younger ages before and after vaccination (**Figure *5.16***). There was no evidence of a difference in post-vaccine response after prior infection between those receiving BNT162b2 or ChAdOx1 (**Figure *5.16***, **Figure *5.17***). At intermediate ages, antibody levels were significantly higher with a single dose following natural infection than two BNT162b2 doses, whereas two doses achieved similar antibody levels to one dose following natural infection at younger and older ages.

5.3.3.2 After the second vaccination (cohort 2)

In participants receiving two vaccinations without prior infection, generalised additive models (GAMs) adjusting only for age and dosing interval showed generally similar

antibody trajectories for both vaccines, but higher antibody levels were achieved with BNT162b2 vs. ChAdOx1 **(Figure *5.19*a, c, Table *5.8*a; observed data: Figure *5.20*)**. Anti-spike IgG levels increased after the first dose, peaked ~21 days later, then gradually declined until the second dose, after which they reached even higher peak levels ~21 days later, then gradually declined again. Post-first dose, peak levels were lower in older participants, but age differences were attenuated after the second dose **(Figure *5.21*a,c,Table *5.8*b)**. In these minimally adjusted analyses (adjusted for age and dosing interval), there was no evidence of differences in antibody levels and declines after the second dose across 8-week to 12-week dosing intervals for both ChAdOx1 and BNT162b2 vaccines. However, the antibody trajectories differed for the 3-week BNT162b2 dosing interval, with antibody levels gradually increasing from the start of the first dose until around 42 days after the second dose, after which antibody levels were generally similar to those with 8–12-week dosing intervals, although slightly lower in 80-year-olds **(Figure *5.19*c)**.

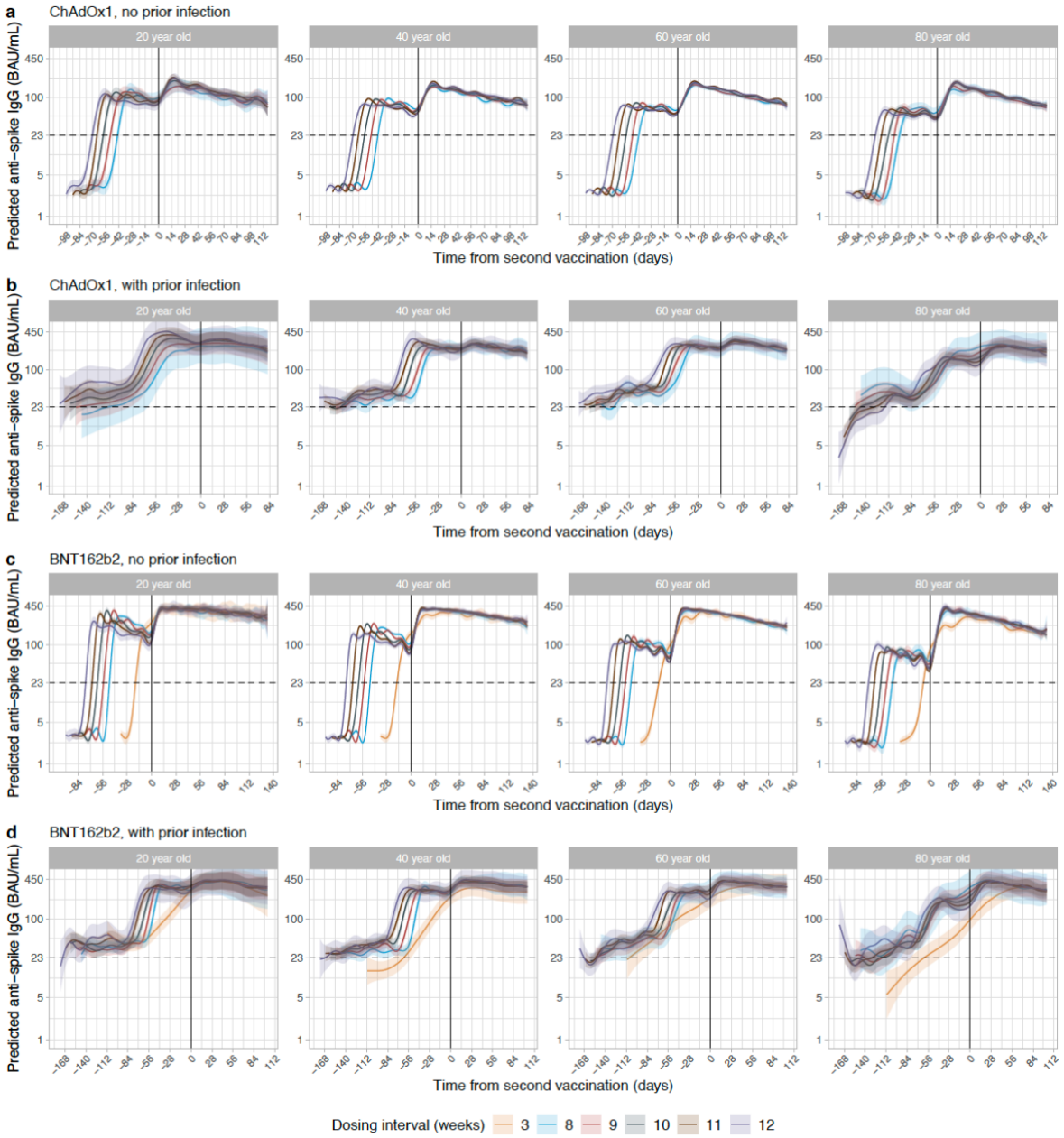


Figure 5.19 Mean anti-spike IgG levels by time from second vaccination in 20-, 40-, 60- and 80-year-olds according to dosing interval using generalised additive models adjusted for age and dosing interval.

Note: **a**, Participants who received two doses of ChAdOx1 without prior infection. **b**, Participants who received two doses of ChAdOx1 with prior infection. **c,** Participants who received two doses of BNT162b2 without prior infection. **d**, Participants who received two doses of BNT162b2 with prior infection. Different x axis scales reflect different durations of follow-up post-vaccination in the different cohorts. Predicted levels are plotted on a log scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Black solid line indicates the second vaccination date. Line colour indicates response predicted for 3 weeks, 8-12 weeks dosing interval. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction.

Table 5.8 Predicted anti-spike IgG levels (BAU/mL) with 95% confidence interval (CI) from generalised additive models.

Note: a, Comparison of 8-, 10-, 12-, and 3-week dosing interval 21, 42, 63 days after the second dose in 60-year-olds. **b,** Comparison of 20-, 40-, 60-, and 80-year-olds 21 days after the first and second dose in those with 10-week dosing interval. The 95% confidence intervals are calculated by prediction ± 1.96*standard error of prediction. Values truncated at 450 BAU/mL counted as =450 BAU/mL.

(a)

			Without prior infection		With prior infection	
Time from second dose (days)	**Dosing interval (days)**	**Age (years)**	**ChAdOx1**	**BNT162b2**	**ChAdOx1**	**BNT162b2**
21	21	60	NA	276 (236-324)	NA	305 (212-440)
21	56	60	159 (147-172)	371 (340-405)	312 (219-443)	418 (284-614)
21	70	60	175 (168-182)	400 (386-415)	325 (278-380)	425 (355-509)
21	84	60	169 (161-178)	392 (367-418)	280 (218-360)	426 (308-588)
42	21	60	NA	328 (284-378)	NA	347 (243-495)
42	56	60	132 (123-141)	357 (324-394)	277 (197-390)	392 (261-587)
42	70	60	137 (132-142)	375 (359-391)	294 (251-344)	406 (342-481)
42	84	60	143 (137-150)	356 (330-383)	274 (215-347)	395 (286-547)
63	21	60	NA	287 (249-331)	NA	371 (260-530)
63	56	60	109 (101-116)	321 (295-350)	230 (163-326)	372 (241-572)
63	70	60	111 (107-115)	321 (310-333)	254 (217-298)	379 (318-453)
63	84	60	121 (116-127)	323 (304-344)	266 (209-339)	395 (288-542)

(b)

			Without prior infection		With prior infection	
Time from second dose (days)	**Dosing interval (days)**	**Age (years)**	**ChAdOx1**	**BNT162b2**	**ChAdOx1**	**BNT162b2**
-49	70	20	108 (90-130)	295 (208-417)	351 (317-388)	347 (273-440)
21	70	20	183 (154-217)	317 (225-447)	404 (362-450)	420 (308-574)
-49	70	40	81 (76-87)	247 (203-300)	219 (208-231)	293 (243-353)
21	70	40	156 (148-165)	289 (241-348)	394 (377-413)	410 (342-492)
-49	70	60	65 (62-68)	255 (216-301)	142 (135-150)	234 (195-282)
21	70	60	175 (168-182)	325 (278-380)	400 (386-415)	425 (355-509)
-49	70	80	37 (33-42)	112 (85-148)	91 (83-100)	141 (107-185)
21	70	80	175 (160-193)	254 (193-334)	404 (377-432)	422 (324-548)

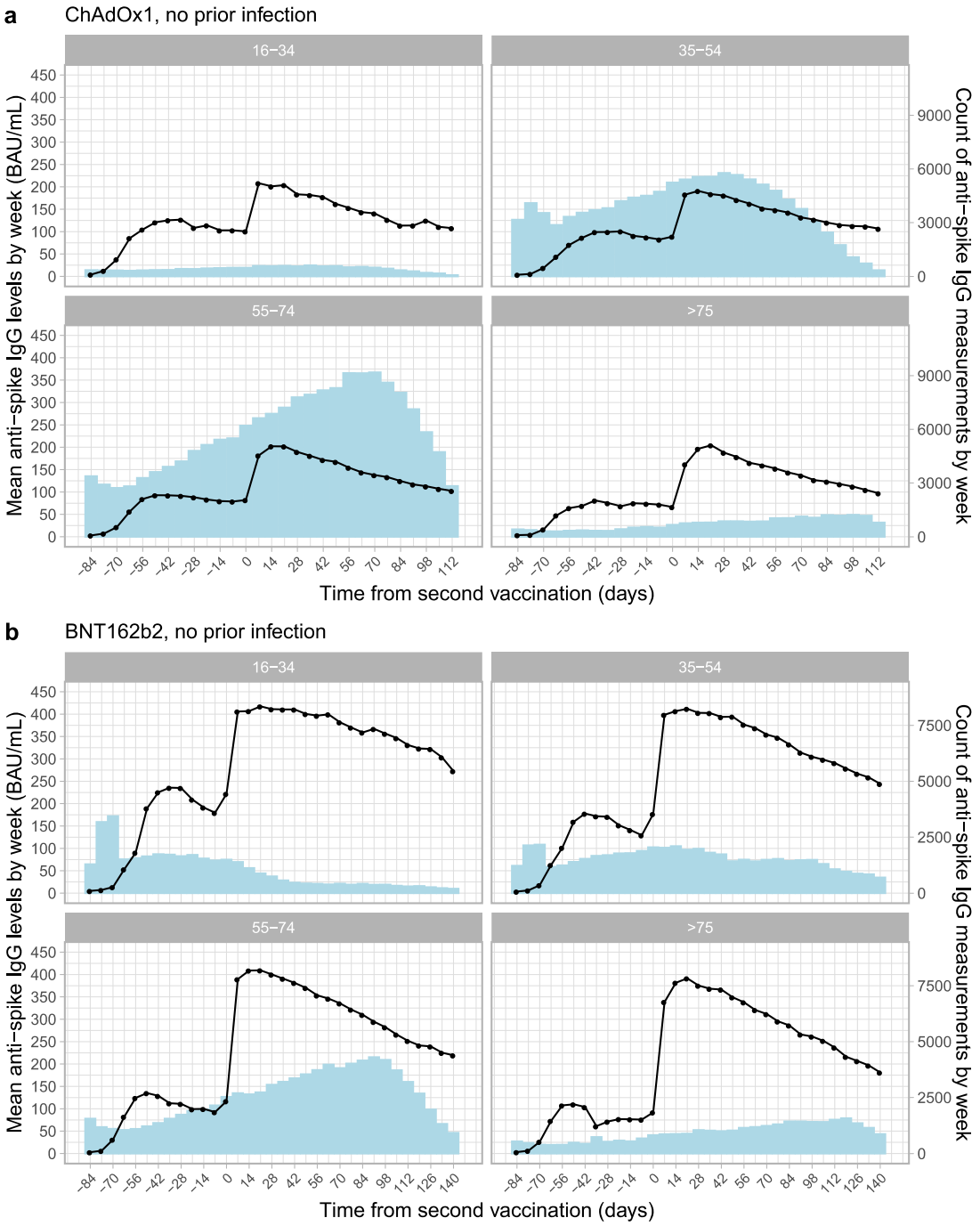


Figure 5.20 Unadjusted mean anti-spike IgG levels (BAU/mL) and count of IgG measurements by week from second vaccination in those without prior infection.

Note: Plots are separated by age group (16-34, 35-54, 55-74, and ≥75 years) and are not separated by dosing intervals. **a**, in those who received two ChAdOx1 vaccinations. **b**, in those who received two BNT162b2 vaccinations. Values truncated at 450 BAU/mL counted as =450 BAU/mL in calculation of unadjusted means.

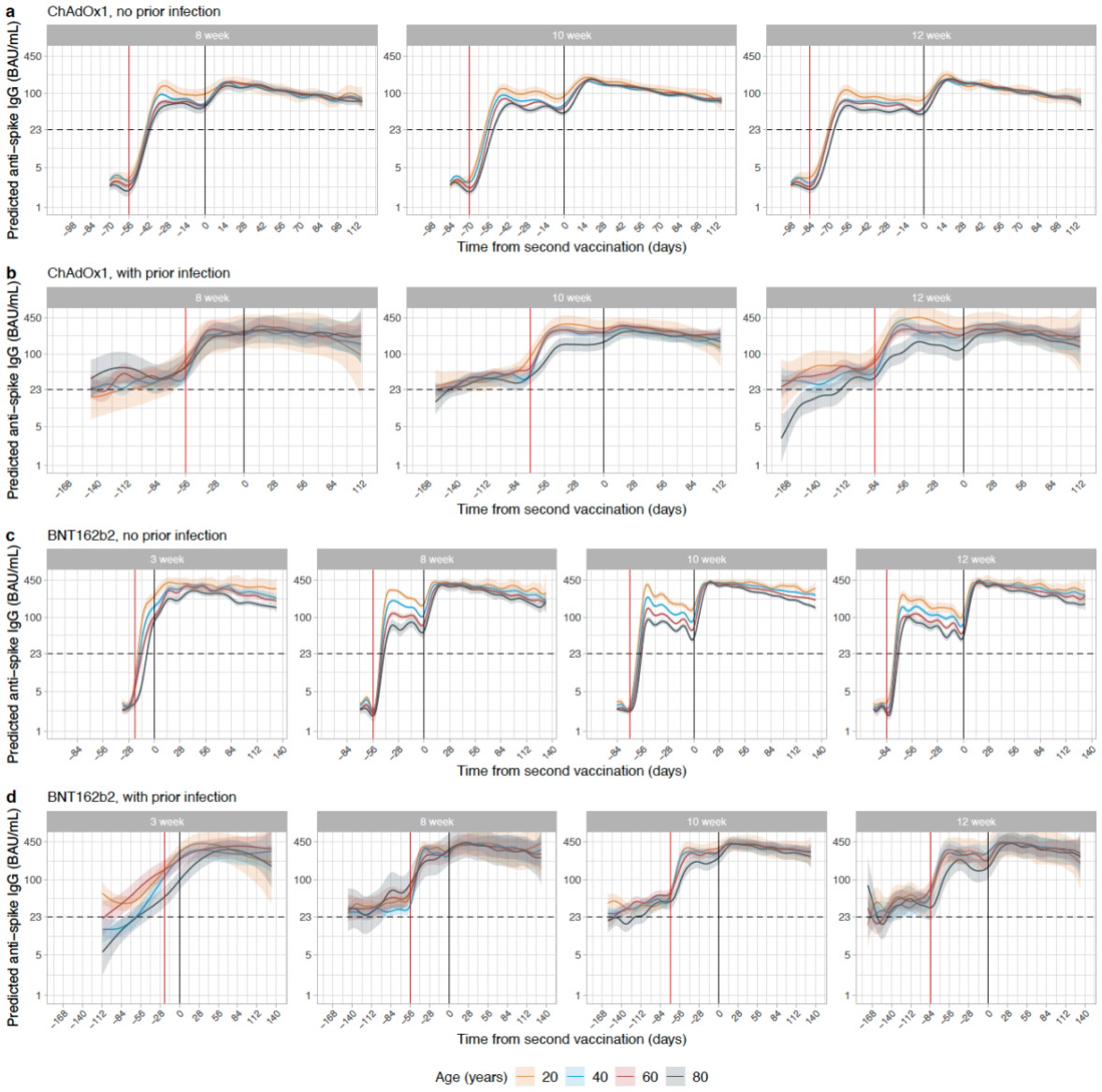


Figure 5.21 Mean anti-spike IgG levels by time from second vaccination in 3-, 8-, 10-, and 12-week dosing interval groups according to age using generalised additive models adjusted for age and dosing interval.

Note: **a**, Participants who received two doses of ChAdOx1 without prior infection. **b**, Participants who received two doses of ChAdOx1 with prior infection**. c,** Participants who received two doses of BNT162b2 without prior infection. **d,** Participants who received two doses of BNT162b2 with prior infection. Different x axis scales reflect different durations of follow-up post-vaccination in the different cohorts. Predicted levels are plotted on a log scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Red solid line indicates the first vaccination and black solid line indicates the second vaccination. Line colour indicates response predicted for age 20, 40, 60, and 80 years. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction.

In participants with evidence of prior infection, antibody levels started from levels above

23 BAU/mL (the positivity threshold, see Methods) and gradually increased for both

vaccines; there was no evidence of differences in antibody levels and declines after the second dose for dosing intervals from 8 to 12 weeks **(Figure *5.19*b, d; observed data shown in Figure *5.22*).** Following prior infection, antibody levels rose to lower levels in older vs. younger participants after the first dose, but the difference was attenuated after the second dose **(Figure *5.21*b, d)**. In those with prior infection, there was less boosting seen from the second dose given higher levels after one dose, compared to those not previously infected, but there was a second dose boosting effect for 80-year-olds **(Figure *5.23*)**. For ChAdOx1, participants without prior infection had lower antibody levels post-second dose than those with prior infection post-first dose; but for BNT162b2, two vaccinations without prior infection led to higher antibody levels than previously infected participants having only one dose, especially for 80-year-olds **(Figure *5.23***; trajectories for other dosing intervals in **Figure *5.24*).**

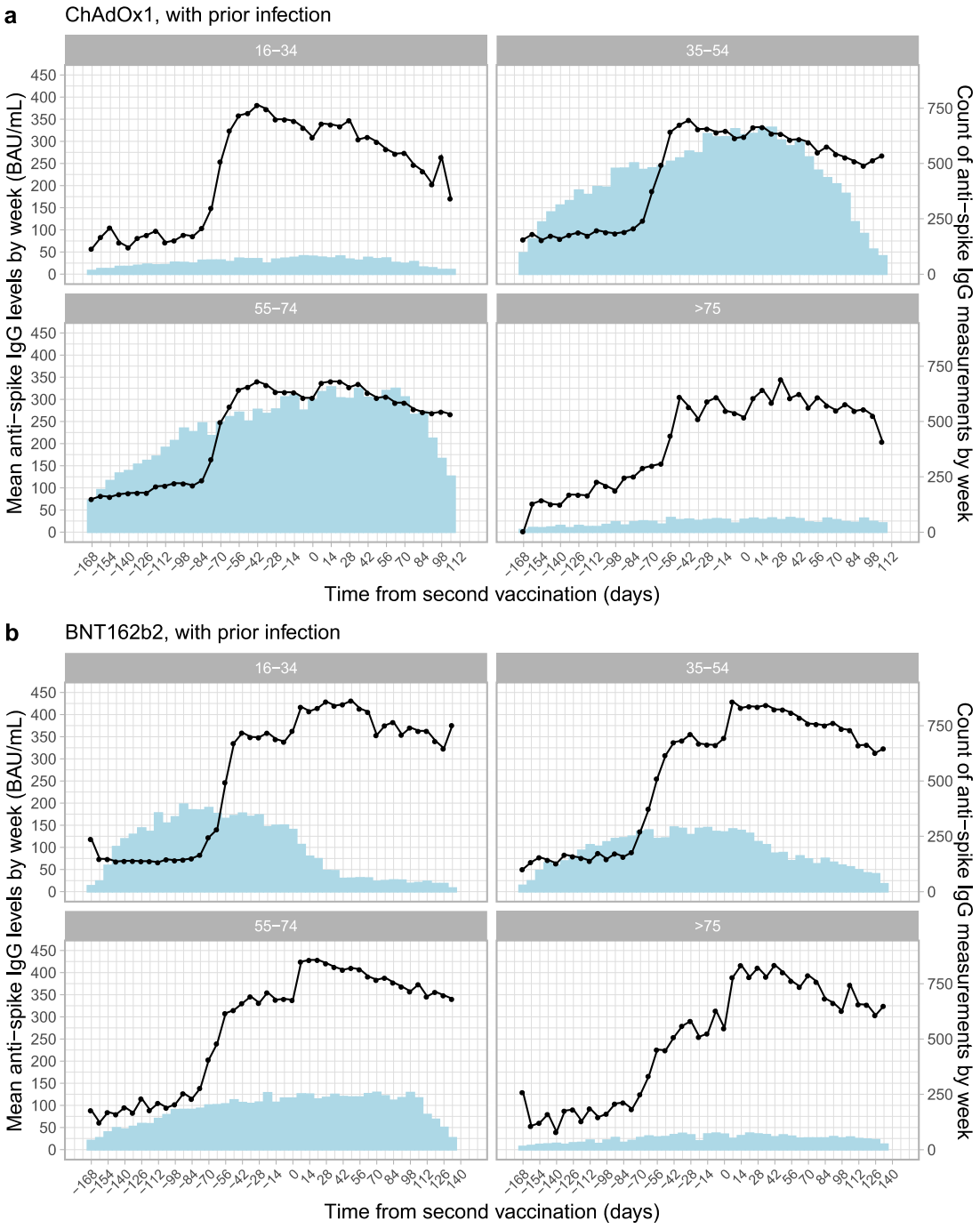


Figure 5.22 Unadjusted mean anti-spike IgG levels (BAU/mL) and count of IgG measurements by week from second vaccination in those with prior infection.

Note: Plots are separated by age group (16-34, 35-54, 55-74, and ≥75 years) and are not separated by dosing intervals. **a**, in those who received two ChAdOx1 vaccinations. **b**, in those who received two BNT162b2 vaccinations. Values truncated at 450 BAU/mL counted as =450 BAU/mL in calculation of unadjusted means.

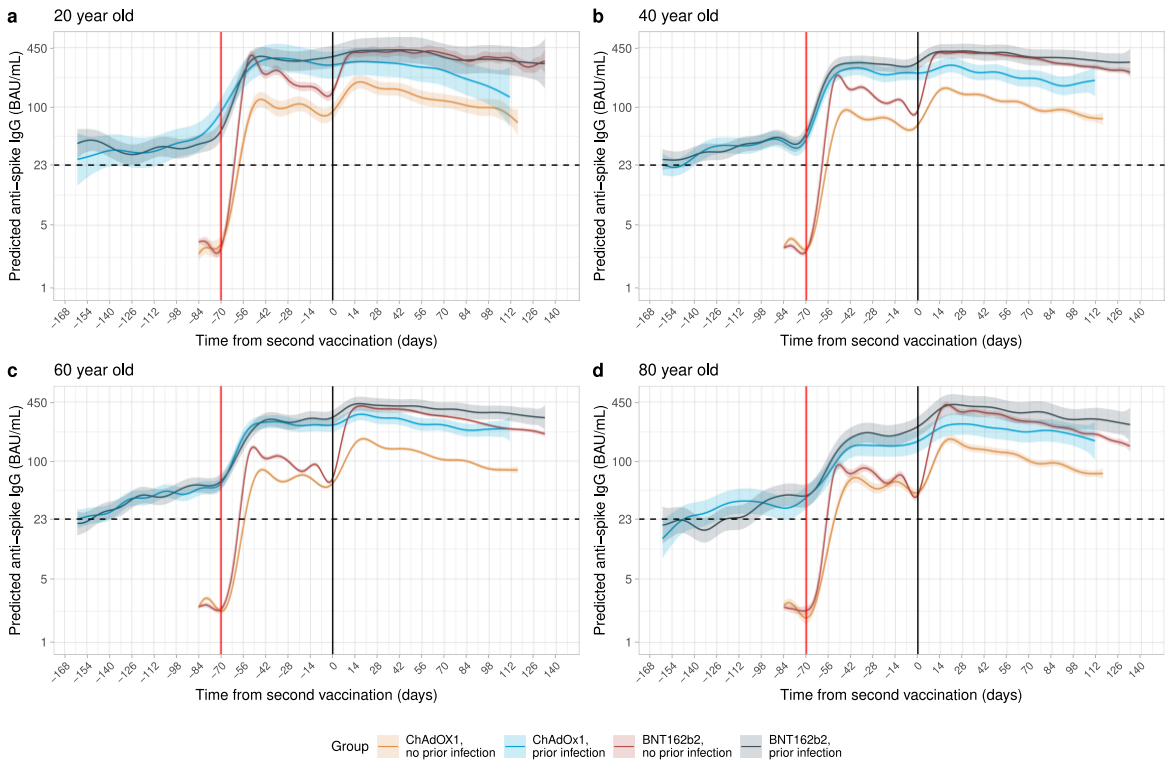


Figure 5.23 Predicted mean anti-spike IgG levels in participants with 10-week dosing interval by time from second vaccination according to vaccine type, age and prior infection status using generalised additive models adjusted for age and dosing interval.

Note: a, 20-year-old. b, 40-year-old. c, 60-year-old. d, 80-year-old. Predicted levels are plotted on a log scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Red solid line indicates the first vaccination and black solid line indicates the second vaccination. Line colour indicates response predicted for ChAdOx1 and BNT162b2, with or without prior infection. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction.

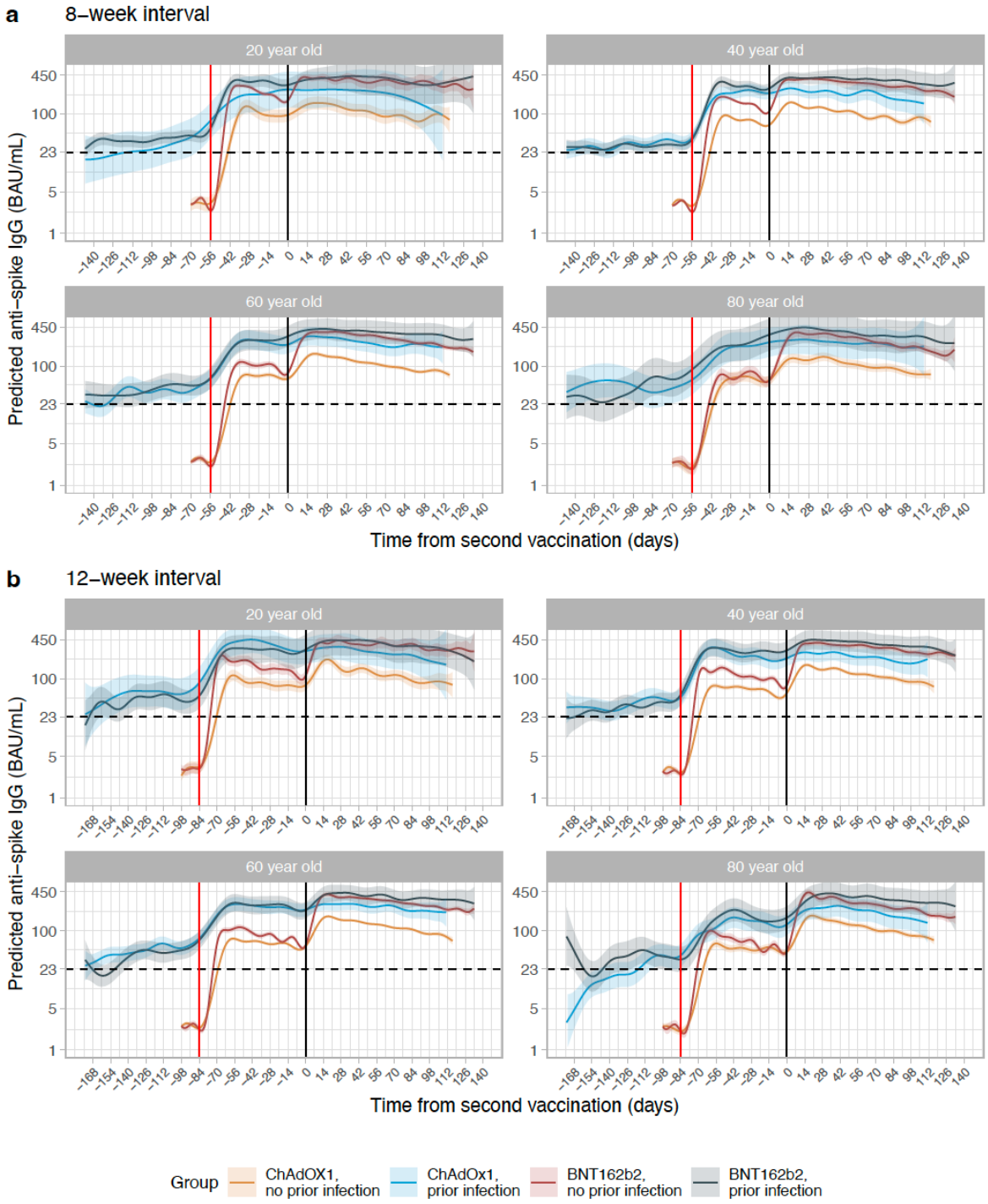


Figure 5.24 Mean anti-spike IgG levels by time from second vaccination according to vaccine type, prior infection status, and age using generalised additive models adjusted for age and dosing interval.

Note: **a**, Participants with a 8-week dosing interval**. b,** Participants with a 12-week dosing interval. Predicted levels are plotted on a log scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Red solid line indicates the first vaccination and black solid line indicates the second vaccination. Line colour indicates response predicted for ChAdOx1 and BNT162b2, with or without prior infection. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction.

5.3.4 Vaccine non-responders after the first and second vaccination and their predictors (cohort 1 and 2)

5.3.4.1 After the first vaccination (cohort 1)

I used descriptive latent class mixed models to identify evidence for different subgroups of responses after the first vaccination. The best-fitting models identified four classes of antibody responses for both vaccines (**Table *5.9*, Figure *5.25*, Figure *5.26***). In a 'plausibly previously infected' group (class 1, grey line, comprising 3.9% of those receiving single-dose ChAdOx1 or BNT162b2), anti-spike IgG levels started higher pre-vaccination (but below the threshold for positivity) and rose rapidly. In a 'high response' group (class 2, orange line, 31.6% and 63.5% of ChAdOx1 and BNT162b2 recipients respectively) IgG levels increased rapidly and to a higher level before plateauing. In a medium response' group (class 3, blue line, 58.7% and 27.5% of ChAdOx1 and BNT162b2 recipients), mean antibody levels rose to slightly below the 'high' response group but still above the positivity threshold. Lastly, participants in a 'low response' group (class 4, red line) had mean IgG levels below the positivity threshold throughout, peaking at ~10 ng/ml, and their response was delayed. A similar percentage, 5.8% and 5.1% of participants receiving the ChAdOx1 and BNT162b2 vaccines, respectively, fell in this group. 'Low response' participants were older and 'high response' participants younger for both vaccines (**Figure *5.27***). 'Low responders' also had a higher proportion of males for ChAdOx1 vaccine and people with long-term health conditions for both vaccines ($p<0.001$). Many health conditions were more common in 'low responders' (**Table *5.10***), with taking immunosuppressants (aOR for class 4 vs. class 2 or 3 responses, 3.91 [95%CI 2.64-5.78]), rheumatoid arthritis (2.50 [1.66-3.76]), chronic liver disease (2.34 [1.06-5.19]), cancer (1.62 [1.31-1.99]), taking corticosteroids (1.59 [1.21-2.10]), type 2 diabetes (1.44 [1.07-1.93]), obesity (BMI ≥30 kg/m^2, 1.25 [1.05-1.48]) and asthma (1.25 [1.03-1.52]) independently-associated with 'low responses'.

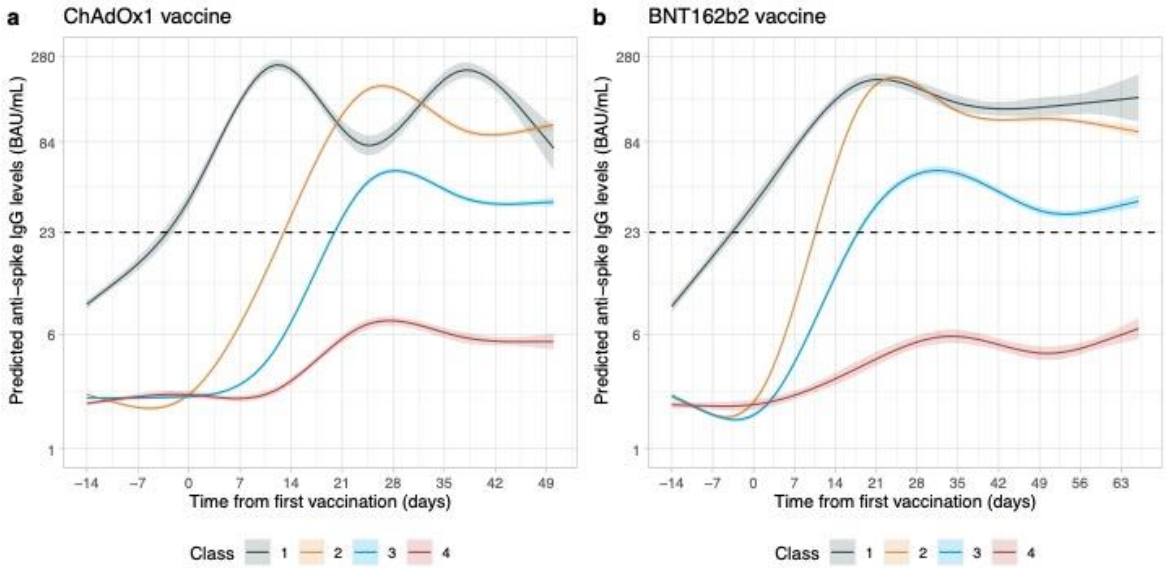


Figure 5.25 Predicted anti-spike IgG trajectory in 36,518 participants without prior infection by class identified from latent class mixed models, using data from 14 days before vaccination to the 90th percentile of the observed time points after vaccination.

Note**: a,** One dose ChAdOx1 vaccine, no evidence of prior infection, N=22,424 participants. **b**, One dose BNT162b2 vaccine, no evidence of prior infection, N=14,094 participants. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). The 95% confidence interval are calculated by a Monte Carlo approximation of the posterior distribution of the predicted values. Distribution of factors by class membership shown in Table **5.9**. Class 1='plausibly previously infected' group (3.9% ChAdOx1, 3.9% BNT162b2), 2='high response' group (31.6% ChAdOx1, 63.5% BNT162b2), 3='medium response' group (58.7% ChAdOx1, 27.5% BNT162b2), 4='low response group' (5.8% ChAdOx1, 5.1% BNT162b2).

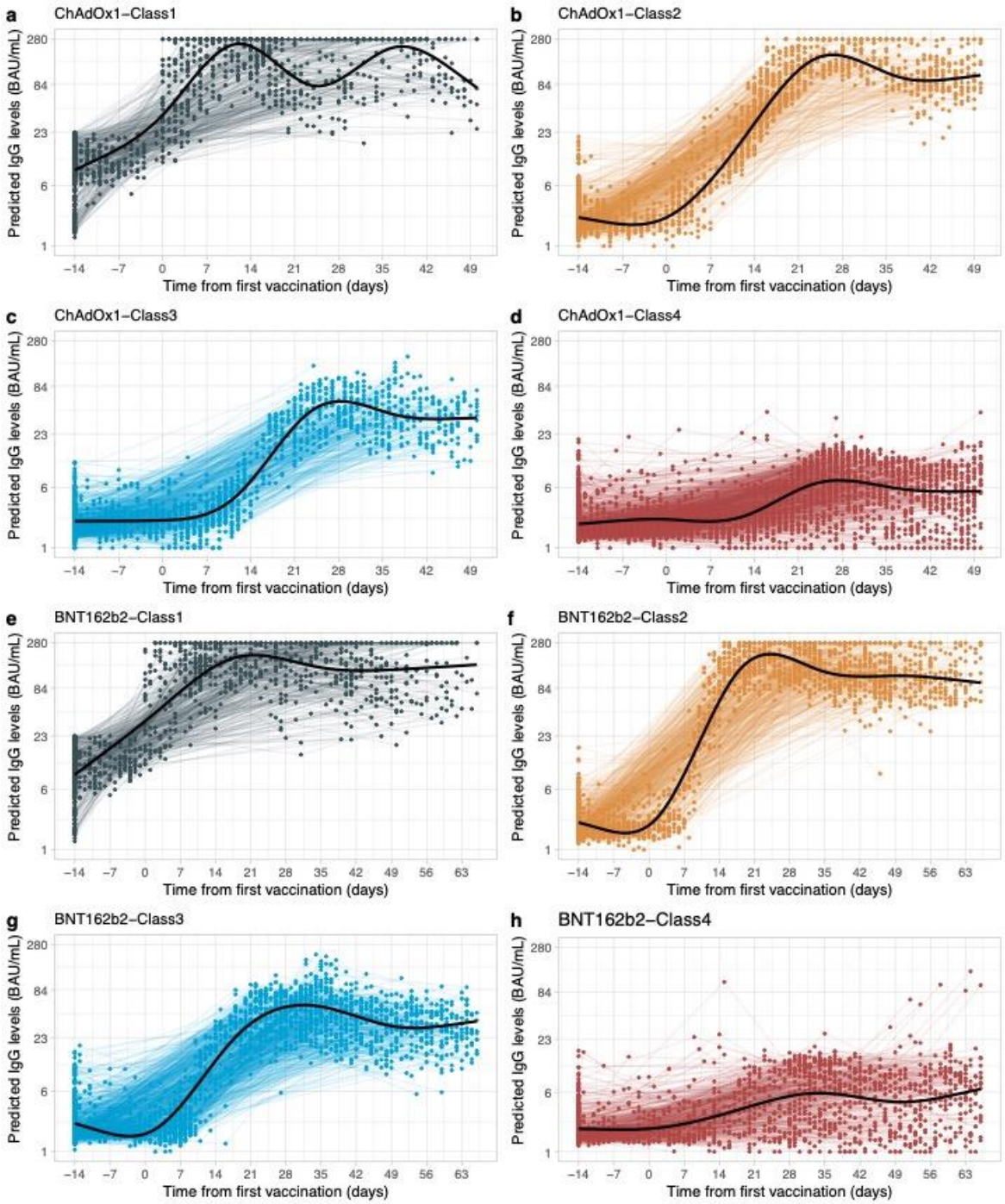


Figure 5.26 Individual trajectories by classes identified from latent class mixed models for single dose ChAdOx1 and BNT162b2 vaccine.

Note: In Class 1 and Class 4, all participants were plotted (ChAdOx1-Class 1: N = 867, Class 4: N = 1297. BNT162b2- Class 1: 547, Class 4: 720). In Class 2 and 3, 1,000 randomly selected participants were plotted for visualisation due to the large number. Colours: navy blue, Class 1; orange, Class 2; mid blue, Class 3; red, Class 4.

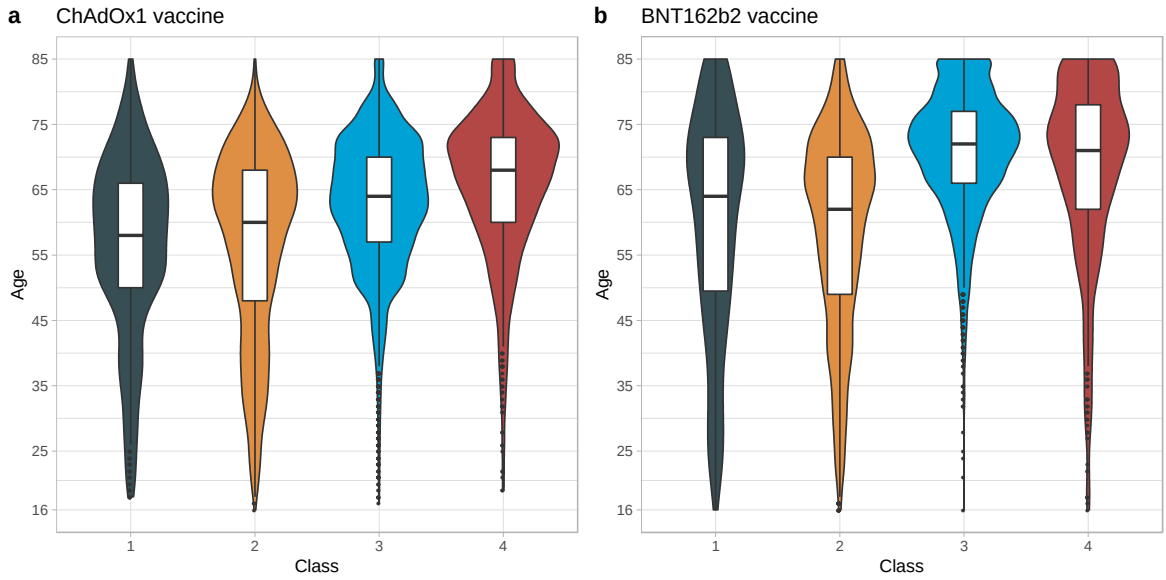


Figure 5.27 Age distribution by classes identified from latent class mixed models for single dose ChAdOx1 (N = 22,424) and BNT162b2 (N = 14,094) vaccines.

Note: 1 = 'plausibly previously infected' group, 2 = 'high response' group, 3 = 'medium response' group, 4 = 'low response group'. Same area for each violin, see **Table 5.9** for numbers in each group. For the box and whisker inserts: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. Colours: navy blue, Class 1; orange, Class 2; mid blue, Class 3; red, Class 4.

Table 5.9 Characteristics of classes identified from latent class mixed models for single dose ChAdOx1 and BNT162b2 vaccines in those without evidence of prior infection.

Note: Class 1=’plausibly previously infected’ group, 2=’high response’ group, 3=’medium response’ group, 4=’low response group’. Continuous variables were compared using Kruskal-Wallis, and categorical variables were compared using Chi-squared tests.

	ChAdOx1					BNT162b2				
	Class 1 (N=867)	**Class 2 (N=7097)**	**Class 3 (N=13163)**	**Class 4 (N=1297)**	**p value**	**Class 1 (N=547)**	**Class 2 (N=8951)**	**Class 3 (N=3876)**	**Class 4 (N=720)**	**p value**
Percentage	3.9%	31.6%	58.7%	5.8%		3.9%	63.5%	27.5%	5.1%	
Age					< 0.001					< 0.001
Median	58	60	64	68		64	62	72	71	
IQR	50, 66	48, 68	57, 70	60, 73		50, 73	49, 70	66, 77	62, 78	
Age group					< 0.001					< 0.001
16-34	73 (8.4%)	709 (10.0%)	167 (1.3%)	22 (1.7%)		63 (11.5%)	760 (8.5%)	12 (0.3%)	26 (3.6%)	
35-54	274 (31.6%)	1863 (26.3%)	2412 (18.3%)	165 (12.7%)		119 (21.8%)	2240 (25.0%)	231 (6.0%)	76 (10.6%)	
55-74	461 (53.2%)	4119 (58.0%)	9185 (69.8%)	862 (66.5%)		250 (45.7%)	4863 (54.3%)	2209 (57.0%)	354 (49.2%)	
>75	59 (6.8%)	406 (5.7%)	1399 (10.6%)	248 (19.1%)		115 (21.0%)	1088 (12.2%)	1424 (36.7%)	264 (36.7%)	
Sex					< 0.001					< 0.001
Female	453 (52.2%)	4467 (62.9%)	6505 (49.4%)	600 (46.3%)		305 (55.8%)	5577 (62.3%)	1721 (44.4%)	364 (50.6%)	
Male	414 (47.8%)	2630 (37.1%)	6658 (50.6%)	697 (53.7%)		242 (44.2%)	3374 (37.7%)	2155 (55.6%)	356 (49.4%)	
Report working in patient facing healthcare					< 0.001					< 0.001
No	834 (96.2%)	6902 (97.3%)	12966 (98.5%)	1276 (98.4%)		475 (86.8%)	7976 (89.1%)	3791 (97.8%)	692 (96.1%)	
Yes	33 (3.8%)	195 (2.7%)	197 (1.5%)	21 (1.6%)		72 (13.2%)	975 (10.9%)	85 (2.2%)	28 (3.9%)	
Report having long-term health condition					< 0.001					< 0.001
No	592 (68.3%)	5104 (71.9%)	8859 (67.3%)	636 (49.0%)		355 (64.9%)	6219 (69.5%)	2096 (54.1%)	344 (47.8%)	
Yes	275 (31.7%)	1993 (28.1%)	4304 (32.7%)	661 (51.0%)		192 (35.1%)	2732 (30.5%)	1780 (45.9%)	376 (52.2%)	
Posterior class-membership probability (%)					< 0.001					< 0.001
Median	98	82	71	99		94	89	86	100	
IQR	78, 100	61, 97	54, 92	84, 100		75, 100	74, 98	64, 97	91, 100	

Table 5.10 Previous health conditions identified from linked primary care records in 29,575 English participants without evidence of prior infection**.**

Note: Class 1=’plausibly previously infected’ group, 2=’high response’ group, 3=’medium response’ group, 4=’low response group’. ‘-’ indicates value not shown due to insufficient sample size. Overall linkage (by ONS) to the General Practice Extraction Service Data for Pandemic Planning and Research via the NHS number was achieved for 29,575 (93%) of 31,796 English participants (not available for participants from Wales, Scotland, or Northern Ireland). Univariable ORs were estimated from all 28,454 participants in Classes 2-4. Multivariable ORs were estimated from a complete-case analysis of 18,288 participants in Classes 2-4 with known BMI (18,449) and no missing values for any other characteristics. Multivariable ORs are adjusted for: age (using a third-degree polynomial with knots at the 25th, 50th and 75th percentiles) interacted with sex; region (nine former Government Office Regions); ethnic group (white or non-white); Index of Multiple Deprivation (IMD) quintile group; working in health or social care; and household size. The 95% confidence intervals are calculated by estimates ± 1.96*standard error of estimates; Wald p values are shown. *Weights were available in 18,449 participants, with denominators by response Class 1-4 of 723, 8029, 8654, and 1043 respectively.

Condition	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Univariable OR (95% CI) Class 4 vs 2+3 combined**	**Univariable p value**	**Multivariable OR (95% CI) Class 4 vs 2+3 combined**	**Multivariable p value**
Total with primary care data	**1121**	**13,057**	**13,808**	**1589**				
Angina	38 (3.4%)	336 (2.6%)	577 (4.2%)	95 (6.0%)	1.81 (1.45-2.25)	<0.0001	0.88 (0.49-1.58)	0.67
Asthma	**130 (11.6%)**	**1336 (10.2%)**	**1319 (9.6%)**	**198 (12.5%)**	**1.30 (1.11-1.51)**	**0.001**	**1.25 (1.03-1.52)**	**0.03**
Atrial fibrillation	42 (3.7%)	389 (3.0%)	578 (4.2%)	115 (7.2%)	2.09 (1.71-2.55)	<0.0001	1.16 (0.88-1.52)	0.28
Cancer	**54 (4.8%)**	**650 (5.0%)**	**970 (7.0%)**	**194 (12.2%)**	**2.17 (1.85-2.54)**	**<0.0001**	**1.62 (1.31-1.99)**	**<0.0001**
Chronic kidney disease	47 (4.2%)	491 (3.8%)	800 (5.8%)	178 (11.2%)	2.50 (2.12-2.95)	<0.0001	1.23 (0.98-1.54)	0.07
Chronic liver disease	**-**	**29 (0.2%)**	**30 (0.2%)**	**13 (0.8%)**	**3.75 (2.05-6.85)**	**<0.0001**	**2.34 (1.06-5.19)**	**0.04**
COPD	32 (2.9%)	289 (2.2%)	553 (4.0%)	99 (6.2%)	2.05 (1.66-2.55)	<0.0001	1.19 (0.90-1.58)	0.23
Coronary heart disease	50 (4.5%)	465 (3.6%)	762 (5.5%)	130 (8.2%)	1.86 (1.54-2.25)	<0.0001	1.03 (0.56-1.88)	0.92
Frailty	77 (6.9%)	670 (5.1%)	1109 (8.0%)	223 (14.0%)	2.30 (1.98-2.67)	<0.0001	1.18 (0.97-1.44)	0.11
Heart failure	10 (0.9%)	76 (0.6%)	185 (1.3%)	41 (2.6%)	2.70 (1.93-3.77)	<0.0001	1.16 (0.72-1.87)	0.54
Hypertension	211 (18.8%)	2436 (18.7%)	3351 (24.3%)	494 (31.1%)	1.64 (1.47-1.83)	<0.0001	1.03 (0.86-1.23)	0.77
Myocardial infarction	23 (2.1%)	187 (1.4%)	294 (2.1%)	50 (3.1%)	1.78 (1.33-2.40)	<0.0001	1.11 (0.66-1.87)	0.71
Osteoporosis	29 (2.6%)	317 (2.4%)	387 (2.8%)	73 (4.6%)	1.79 (1.40-2.29)	<0.0001	1.39 (1.01-1.91)	0.05
Peripheral arterial disease	-	77 (0.6%)	129 (0.9%)	30 (1.9%)	2.49 (1.69-2.66)	<0.0001	1.52 (0.95-2.43)	0.08

Condition	Class 1	Class 2	Class 3	Class 4	Univariable OR (95% CI) Class 4 vs 2+3 combined	Univariable p value	Multivariable OR (95% CI) Class 4 vs 2+3 combined	Multivariable p value
Total with primary care data	**1121**	**13,057**	**13,808**	**1589**				
Rheumatoid arthritis	**-**	**88 (0.7%)**	**161 (1.2%)**	**85 (5.3%)**	**6.04 (4.70-7.77)**	**<0.0001**	**2.50 (1.66-3.76)**	**<0.0001**
Stroke	18 (1.6%)	128 (1.0%)	206 (1.5%)	31 (2.0%)	1.58 (1.09-2.29)	0.02	0.81 (0.49-1.35)	0.42
Transient ischaemic attack	14 (1.2%)	118 (0.9%)	197 (1.4%)	29 (1.8%)	1.57 (1.07-2.30)	0.02	0.91 (0.55-1.49)	0.70
Type 1 diabetes	14 (1.2%)	101 (0.8%)	67 (0.5%)	10 (0.6%)	1.01 (0.53-1.91)	0.98	-	-
Type 2 diabetes	**72 (6.4%)**	**848 (6.5%)**	**1227 (8.9%)**	**246 (15.5%)**	**2.19 (1.90-2.53)**	**<0.0001**	**1.44 (1.07-1.93)**	**0.02**
Overweight (BMI 25 to <30 kg/m^2)*	301 (41.6%)	2871 (35.7%)	3306 (38.2%)	355 (34.0%)	0.88 (0.77-1.00)	0.05	0.95 (0.81-1.12)	0.55
Obese (BMI ≥30 kg/m^2)*	**177 (24.5%)**	**2264 (28.2%)**	**2580 (29.8%)**	**385 (36.9%)**	**1.43 (1.26-1.63)**	**<0.0001**	**1.25 (1.05-1.48)**	**0.01**
Antihypertensive medication	312 (27.8%)	3641 (27.9%)	5096 (36.9%)	763 (48.0%)	1.92 (1.73-2.12)	<0.0001	1.17 (0.97-1.41)	0.11
Diabetes medication	69 (6.2%)	812 (6.2%)	1108 (8.0%)	212 (13.3%)	2.00 (1.72-2.33)	<0.0001	1.10 (0.80-1.51)	0.56
Corticosteroids	**27 (2.4%)**	**325 (2.5%)**	**469 (3.4%)**	**120 (7.6%)**	**2.68 (2.20-3.27)**	**<0.0001**	**1.59 (1.21-2.10)**	**0.001**
Immunosuppressants	**10 (0.9%)**	**86 (0.7%)**	**178 (1.3%)**	**103 (6.5%)**	**6.98 (5.53-8.82)**	**<0.0001**	**3.91 (2.64-5.78)**	**<0.0001**

5.3.4.2 After the second vaccination (cohort 2)

In the initial vaccination study I used latent class mixed models to identify 5.8% and 5.1% of participants receiving one ChAdOx1 or BNT162b2 dose, respectively, as non-responders. However, the study population was large in the second vaccination study and latent class models would not fit with larger numbers, therefore I used a heuristic rule based on these previous observations to define non-response as all antibody measurements <16 BAU/ml (similar levels to the previous non-response class) and having at least one antibody measurement 21 days after the first or second dose. To examine robustness, I also restricted to those having at least two antibody measurements 21 days or later, and after both doses (rather than each separately). Across different assumptions **(Table *5.11*)**, I found that 5.8-7.7% and 3.5-6.0% of participants were classed as non-responders to the first ChAdOx1 or BNT162b2 dose, respectively, similar to previously. Non-responders were older, had a higher percentage of males and white ethnicity, were more deprived, less likely to be working in healthcare, and more likely to have long-term health conditions **(Table *5.12*)**. However, only 0.4-1.0% and 0.1-0.5% participants were non-responders to the second ChAdOx1 or BNT162b2 dose, respectively, and 0.3-0.5% and 0.1-0.2% were non-responders to both first and second doses of ChAdOx1 or BNT162b2, respectively.

Table 5.11 Percentages of non-responders to first or second dose of ChAdOx1 or BNT162b2 using a heuristic rule defined as all antibody measurements being <16 BAU/mL and having at least one antibody measurement 21 days after the first or second dose.

Note**: a,** Non-responders in those who had antibody measurements after first dose or second dose. **b,** Non-responders in those who had antibodies measured after both first and second dose by responses to each dose. To examine the robustness of this definition, this rule was further applied to those having at least two antibody measurements.

a)		**ChAdOx1**			**BNT162b2**		
		Total	**Non-responder**	**%**	**Total**	**Non-responder**	**%**
First dose	At least one antibody measurement >=21 days after first dose and before second dose	57,650	4,280	7.4	30,843	1,374	4.5
	Two or more antibody measurements, including at least one antibody measurement >=21 days after first dose and before second dose	29,277	1,706	5.8	15,525	546	3.5
Second dose	At least one antibody measurement >=21 days after second dose	100,776	1,010	1.0	57,313	312	0.5
	Two or more antibody measurements, including at least one antibody measurement >=21 days after second dose	77,947	496	0.6	44,311	146	0.3
b)		**ChAdOx1**			**BNT162b2**		
First and second dose	At least one Ab measurement >=21 days after first dose and before second dose, and after second dose	**Second dose: non-responder**	**Second dose: responder**	**Total**	**Second dose: non-responder**	**Second dose: responder**	**Total**
	First dose: non-responder	192 (0.5%)	3,243 (7.2%)	3,435 (7.7%)	37 (0.2%)	1,036 (5.8%)	1,073 (6.0%)
	First dose: responder	99 (0.2%)	41,212 (92.1%)	41,311 (92.3%)	15 (0.1%)	16,909 (93.9%)	16,924 (94.0%)
	Total	291 (0.7%)	44,455 (99.3%)	44,746 (100%)	52 (0.3%)	17,945 (99.7%)	17,997 (100%)
	Two or more antibody measurements, including at least one antibody measurement >=21 days after first dose and before second dose, and after second dose	**Second dose: non-responder**	**Second dose: responder**	**Total**	**Second dose: non-responder**	**Second dose: responder**	**Total**
	First dose: non-responder	59 (0.3%)	1,092 (5.7%)	1,151 (6.0%)	8 (0.1%)	364 (5.1%)	372 (5.2%)
	First dose: responder	10 (0.1%)	17,965 (93.9%)	17,975 (94.0%)	0 (0%)	6,731 (94.8%)	6,731 (94.8%)
	Total	69 (0.4%)	19,057 (99.6%)	19,126 (100%)	8 (0.1%)	7,095 (99.9%)	7,103 (100%)

Table 5.12 Comparison of the characteristics of responders with non-responders to first or second dose of ChAdOx1 or BNT162b2.

Note: Non-responders were identified using a heuristic rule as all antibody measurements being <15 BAU/mL and having at least one antibody measurement 21 days after the first or second dose. Continuous variables were compared using Kruskal-Wallis tests, and categorical variables were compared using Chi-squared tests.

	Responders (N=215,746)	Non-responders (N=6,747)	Total (N=222,493)	p value
Dosing interval (days)				< 0.001
Median	74	77	74	
Q1, Q3	64, 78	70, 79	64, 78	
Age (years)				< 0.001
Median	56	63	57	
Q1, Q3	43, 67	53, 72	43, 68	
Sex				< 0.001
Female	11,7859 (54.6%)	3,007 (44.6%)	120,866 (54.3%)	
Male	97,887 (45.4%)	3,740 (55.4%)	101,627 (45.7%)	
Ethnicity				< 0.001
Non-white	12,383 (5.7%)	212 (3.1%)	12,595 (5.7%)	
White	203,363 (94.3%)	6,535 (96.9%)	209,898 (94.3%)	
Household size				< 0.001
1	39,556 (18.3%)	1,541 (22.8%)	41,097 (18.5%)	
2	105,134 (48.7%)	3,490 (51.7%)	108,624 (48.8%)	
3	33,046 (15.3%)	847 (12.6%)	33,893 (15.2%)	
4	27,405 (12.7%)	621 (9.2%)	28,026 (12.6%)	
5+	10,605 (4.9%)	248 (3.7%)	10,853 (4.9%)	
Deprivation percentile				< 0.001
Median	63	60	62	
Q1, Q3	39, 82	36, 81	39, 82	
Report working in patient facing healthcare				< 0.001
No	208,765 (96.8%)	6,657 (98.7%)	215,422 (96.8%)	
Yes	6,981 (3.2%)	90 (1.3%)	7,071 (3.2%)	
Report having a long-term health condition				< 0.001
No	156,009 (72.3%)	3,670 (54.4%)	159,679 (71.8%)	
Yes	59,737 (27.7%)	3,077 (45.6%)	62,814 (28.2%)	

5.3.5 Determinants of anti-spike IgG peak levels and half-life after the first and second vaccination (cohort 2 and 3)

5.3.5.2 After the second vaccination (cohort 2)

ChAdOx1

Of the 133,388 participants (with or without prior infection) who received two ChAdOx1 doses, 100,639 participants contributed 191,137 antibody measurements ≥21 days after the second dose, median (IQR) [range] 2 (1-2) [1-5] measurements per participant (**Table**

5.13*)**. Antibody measurements were taken a median (IQR) [range] 61 (41-83) [21-119] days post second vaccination (**Table *5.13*, Figure *5.28).

Table 5.13 Frequency of antibody measurements per participant, and frequency of antibody measurements relative to the time of the second vaccination, for the Bayesian linear mixed model.

Note: 100,639 ChAdOx1 participants contributed 191,137 antibody measurements ≥21 days after the second dose, median (IQR) [range] 2 (1-2) [1-5] measurements per participant. 55,053 BNT162b2 participants contributed 120,728 antibody measurements ≥21 days after the second dose, median (IQR) [range] 2 (1-3) [1-6] per participant. The antibody measurements were taken a median (IQR) [range] 61 (41-83) [21-119] days after the second vaccination for ChAdOx1 and 79 (51-106) [21-149] days for BNT162b2.

(a) Frequency of antibody measurements per participant		
	ChAdOx1	**BNT162b2**
1	37,478	17,863
2	39,268	16,444
3	20,450	13,839
4	3,442	6,074
5	1	832
6	0	1
(b) Frequency of antibody measurements relative to the second vaccination		
(days)	**ChAdOx1**	**BNT162b2**
21-28	15,693	6,964
28-35	16,403	7,161
35-42	16,535	6,934
42-49	16,427	7,061
49-56	16,200	6,973
56-63	16,788	7,310
63-70	16,284	7,507
70-77	15,750	7,471
77-84	14,258	7,709
84-91	12,962	7,925
91-98	11,070	8,064
98-105	8,873	7,934
105-112	7,219	7,215
112-119	5,886	6,436
119-126	789	5,837
126-133	0	4,721
133-140	0	3,732
140-147	0	2,831
147-149	0	945

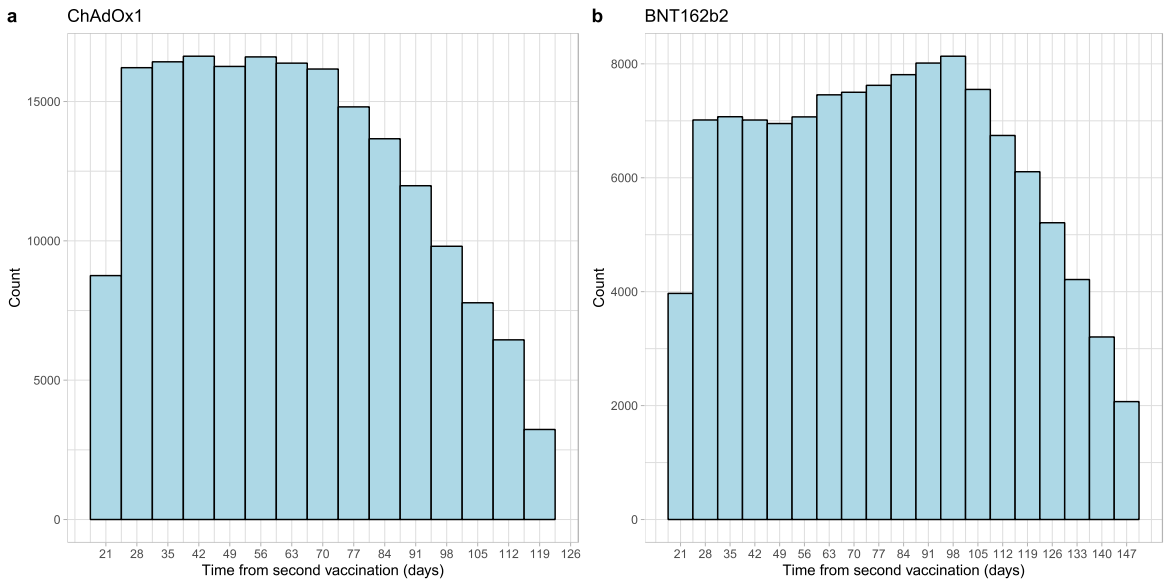


Figure 5.28 Distribution of timing of antibody measurements relative to the time of the second vaccination for the Bayesian linear mixed model.

Note: The median (IQR) [range] was 61 (41-83) [21-119] days for ChAdOx1 and 79 (51-106) [21-149] days for BNT162b2.

Assuming antibody levels declined exponentially, using Bayesian (log)linear mixed models I estimated a mean peak anti-spike IgG level of 184 BAU/mL (95% credible interval, Crl 183-185), and mean half-life of 79 days (78-80) **(Figure *5.30*)** following 2 ChAdOx1 doses, versus 113 BAU/mL (106-117) and 184 days (163-210) respectively following natural infection before vaccination[175] (Chapter 4). There was no evidence that rates of antibody decline flattened over time (up to 119 days post second vaccination, see Methods). In a multivariable model, all factors considered (age, sex, ethnicity, reporting a long-term health condition, healthcare work, deprivation, dosing interval, and prior infection status) were independently associated with anti-spike IgG peak levels 21 days post-second dose, but most effects were small **(Figure *5.30*, Table *5.14*)**. The largest effects were associated with prior infection, peak 219 BAU/mL higher (210-227) and ethnicity, peak 41 BAU/mL (36-47) higher in those reporting non-white ethnicity. Peak levels were slightly lower in males, those reporting a long-term health condition, those not

working in healthcare, with shorter dosing intervals, younger age, and less deprivation (higher deprivation percentile). Prior infection extended the half-life by 13 days (95%CrI 9-17). There were very small reductions in half-life at older ages, with non-white ethnicity, and having a long-term health condition. There was no evidence of associations between half-life and sex, being a healthcare worker, and dosing interval in participants who received ChAdOx1.

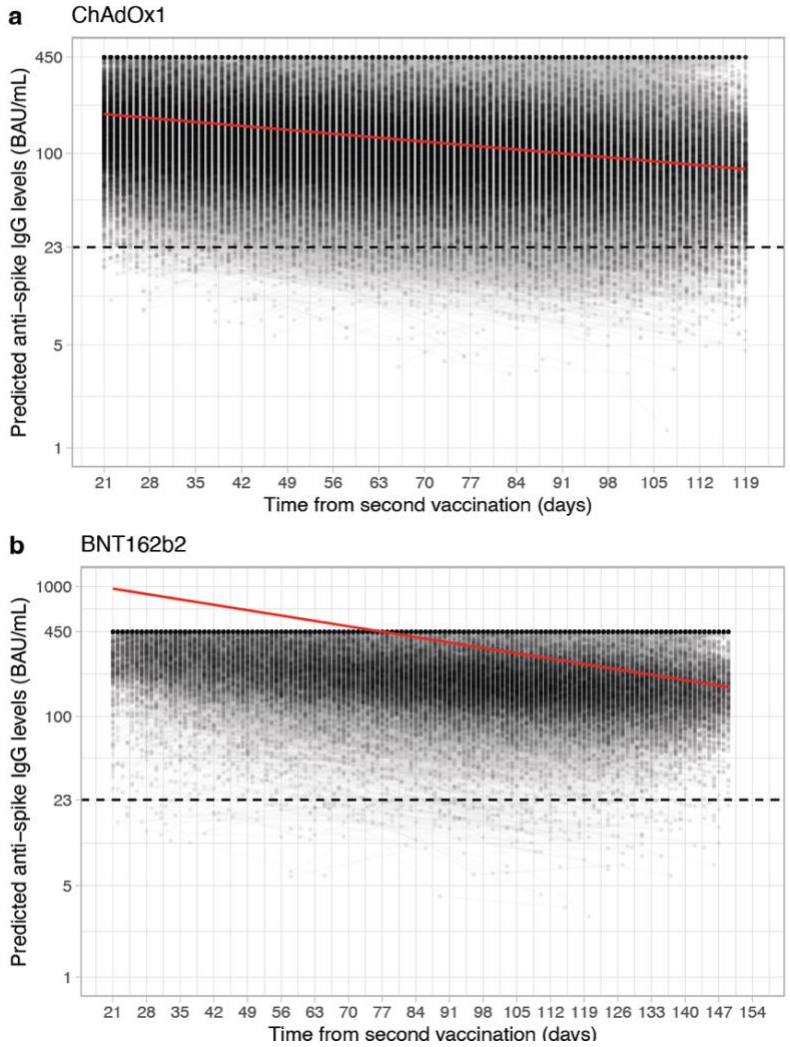


Figure 5.29 Estimated unadjusted mean trajectory of anti-spike IgG antibody levels and individual trajectories from 21 days after the second vaccination**.**

Note**: a,** In participants who received two doses of ChAdOx1. **b,** In participants who received two doses of BNT162b2. Black dashed line indicates the assay threshold for IgG positivity (23 BAU/mL). For BNT162b2, interval censored regression takes account of the fact that 52% measurements were above the upper quantification limit of 450 BAU/mL, and the timing of these measurements, so the estimated peak levels were higher than 450 BAU/mL.

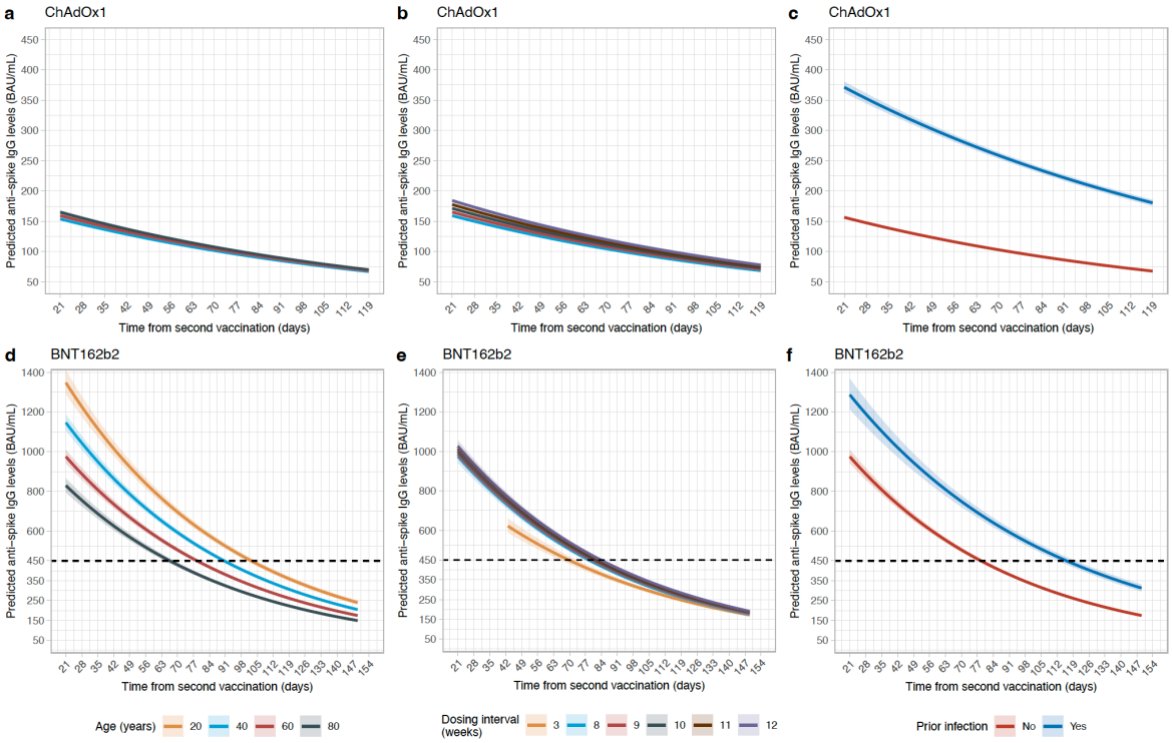


Figure 5.30 Posterior predicted mean trajectories (95% credible interval) of anti-spike IgG levels from 21 days post-second dose using Bayesian linear mixed interval censored models.

Note: Models are adjusted for age, sex, ethnicity, long-term health conditions, healthcare role, deprivation, dosing interval, and prior infection status. **Panel a,d** show the trajectories by age. **Panel b,e** show the trajectories by dosing interval. **Panel c,f** show the trajectories by prior infection status. **Panel a,b,c** show participants who received two doses of ChAdOx1. **Panel d,e,f** show participants who received two doses of BNT162b2. Black dotted line shows the upper quantification limit of 450 BAU/mL. For BNT162b2, as 52% measurements were above the upper limit, the estimated peak levels were higher than 450 BAU/mL (interval censoring accounted for this in analysis). Plotted at reference categories: 60 years, female, white ethnicity, not reporting a long-term health condition, not a healthcare worker, deprivation percentile=60, 8-week dosing interval, and no prior infection. In **panel a**, 20-year-old group is not plotted because the vast majority of those receiving ChAdOx1 were ≥40 years.

Table 5.14 Posterior median and 95% credible intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the unadjusted models and multivariable models in 100,639 participants with two ChAdOx1 doses and 55,053 participants with two BNT162b2 doses.

Note: The reference categories in the multivariable model are: 60-year-old, female, white ethnicity, not reporting a long-term health condition, not a healthcare worker, deprivation percentile=60, 8-week dosing interval, and no prior infection. Bold numbers indicate an effect not compatible with chance (95% credible interval excludes no effect).
*The predicted peak level in 3-week vs 8-week is extrapolated and for modelling purpose only, it should not be interpreted as an actual peak, since trajectories in the 3-week group were only modelled from 42 days post second vaccination due to late increases, see Methods.

		ChAdOx1			BNT162b2		
		Posterior mean	95% CrI		Posterior mean	95% CrI	
Unadjusted baseline	Peak level (Intercept)	184	183	185	959	944	974
	IgG half-life (slope)	79	78	80	51	50	52
Adjusted baseline	Peak level (Intercept)	160	155	162	976	903	1011
	IgG half-life (slope)	81	79	83	52	50	53
Age	Peak level: 60 years						
	IgG half-life: 60 years						
	Change in peak level: per 10-years older	**3**	**2**	**4**	**-76**	**-84**	**-68**
	Change in half-life: per 10-years older	**-1**	**-1**	**0**	0	0	0
Sex	Peak level: Female						
	IgG half-life: Female						
	Change in peak level: Male	**-5**	**-7**	**-3**	**-140**	**-164**	**-117**
	Change in half-life: Male	0	-1	2	-1	-2	1
Ethnicity	Peak level: White						
	IgG half-life: White						
	Change in peak level: Non-white	**41**	**36**	**47**	**141**	**78**	**208**
	Change in half-life: Non-white	**-8**	**-11**	**-5**	0	-3	2
Report having a long-term health condition	Peak level: No						
	IgG half-life: No						
	Change in peak level: Yes	**-5**	**-7**	**-3**	**-79**	**-104**	**-55**

		ChAdOx1			BNT162b2		
		Posterior mean	**95% CrI**		**Posterior mean**	**95% CrI**	
	Change in half-life: Yes	**-2**	**-4**	**-1**	**-1**	**-2**	**0**
Report working in patient-facing healthcare	Peak level: No						
	IgG half-life: No						
	Change in peak level: Yes	**10**	**2**	**19**	**287**	**221**	**358**
	Change in half-life: Yes	0	-6	6	**-5**	**-6**	**-3**
Deprivation	Peak level: 60 (median)						
	IgG half-life: 60 (median)						
	Change in peak level: per 10 percentile higher	**-1**	**-1**	**-1**	**5**	**1**	**10**
	Change in half-life: per 10 percentile higher	**1**	**0**	**1**	0	0	0
3-week interval	Peak level: No						
	IgG half-life: No						
	Change in 'peak level': Yes (vs 8 weeks) *				**-163**	**-227**	**-95**
	Change in half-life: Yes (vs 8 weeks)				**6**	**2**	**10**
Dosing interval	Peak level: 8 weeks (median)						
	IgG half-life: 8 weeks (median)						
	Change in peak level: per 1 week longer	**6**	**5**	**7**	**12**	**1**	**23**
	Change in half-life: per 1 week longer	0	-1	0	0	0	1
Prior infection	Peak level: No						
	IgG half-life: No						
	Change in peak level: Yes	**219**	**210**	**227**	**312**	**248**	**381**
	Change in half-life: Yes	**13**	**9**	**17**	**11**	**8**	**15**

BNT162b2

In 89,105 participants (with or without prior infection) who received two BNT162b2 doses, 55,053 participants contributed 120,728 antibody measurements ≥21 days after the second dose (≥42 days for those with 3-week dosing interval, see Methods, **Figure *5.24***), median (IQR) [range] 2 (1-3) [1-6] per participant (**Table *5.13***). Antibody measurements were taken a median (IQR) [range] 79 (51-106) [21-149] days post second vaccination (**Table *5.13*, Figure *5.28***). The estimated mean peak level was 959 BAU/mL (95%Crl 944-974), and the mean half-life was 51 days (50-52) (**Figure *5.29*).** There was again no evidence of antibody decline flattening on the log scale (up to 149 days post second vaccination, see Methods). Factors had greater effects on peak levels for BNT162b2 than ChAdOx1 **(Figure *5.30*, Table *5.14*)**. Peak levels were lower at older ages (76 BAU/mL lower per 10-years older, 95%Crl 68-84), in males (140 BAU/mL lower, 117-164), those reporting long-term health conditions (79 BAU/mL lower, 55-104), and were higher in those reporting non-white ethnicity (141 BAU/mL higher, 78-208), working in healthcare (287 BAU/mL higher, 221-358), and who lived in less deprived areas (5 BAU/mL higher per 10 percentiles higher, 1-10). However, these factors had little or no effect on half-life. For dosing intervals between 8 and 12 weeks, longer dosing intervals were associated with higher peak levels (12 BAU/mL higher per week longer, 1-23), but had no effect on half-life. Compared with an 8-week extended schedule, a 3-week dosing interval was associated with a lower peak level 42 days post-second dose but a slightly longer half-life (6 days longer, 2-10), leading to similar antibody levels at 149 days across different dosing groups. Similar to ChAdOX1, prior infection was associated with a much higher peak level (312 BAU/mL higher, 248-381) and a longer half-life (11 days longer, 8-15).

Comparing the effects of factors on response to the two vaccines, and with my previous findings on natural infection in **Chapter 4**, effects of some factors were relatively

consistent between ChAdOx1, BNT162b2 and/or natural infection, albeit with differing effect sizes (e.g sex, ethnicity, long-term health condition, working in healthcare, dosing interval on peak; prior infection on peak and half-life), whilst for others effects were in opposite directions (e.g age, deprivation on peak) or only associated with one vaccine (e.g. ethnicity, working in healthcare on half-life). In general, effects on peak levels were greater for BNT162b2 than ChAdOx1; and, other than prior infection, effects on half-life were limited for both vaccines.

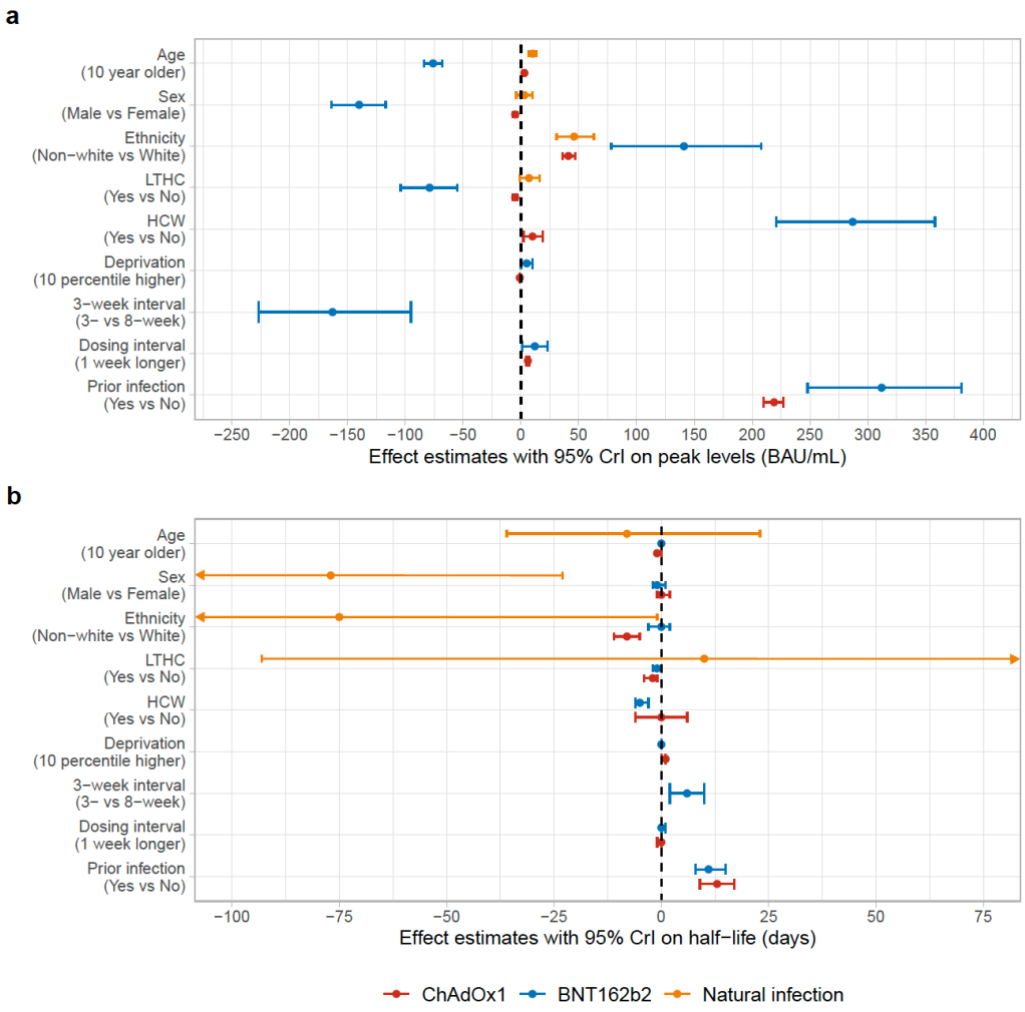


Figure 5.31 Comparison of effect of factors on antibody responses in participants who received two doses of ChAdOx1 or BNT162b2 or had natural SARS-CoV-2 infection.

Note**: a,** Effect on anti-spike IgG peak levels. **b,** Effect on anti-spike IgG half-lives. 100,639 participants received two doses of ChAdOx1, 55,053 participants received two doses of BNT162b2, and 3,271 had natural infection. Mean estimates with 95% credible intervals are presented. In panel b, 95% credible intervals are truncated at -100 and 75 days for visualisation.

5.3.5.2 After the first vaccination (cohort 3)

For first vaccination, I modelled antibody trajectories using measurements from 28 days post-first dose for all participants (i.e., from the time of approximate peak levels in observed and modelled data, **Figure *5.32*, Figure *5.33***), accounting for the assay upper limit of quantification using interval censored models and adjusting for age, sex, ethnicity, reporting a long-term health condition or working in healthcare, and deprivation in order to directly estimate peak levels and antibody declines by these characteristics. I excluded participants who did not mount an anti-S antibody response to first vaccination (defined as all antibody measurements <16 BAU/mL, including ≥1 measurement ≥21 days after the first dose): 4,488 (7.4%), 1,450 (4.2%), and 17 (0.7%) participants without prior infection, and 149 (1.9%), 52 (0.8%), 0 (0%) with prior infection receiving ChAdOx1, BNT162b2, and mRNA-1273, respectively.

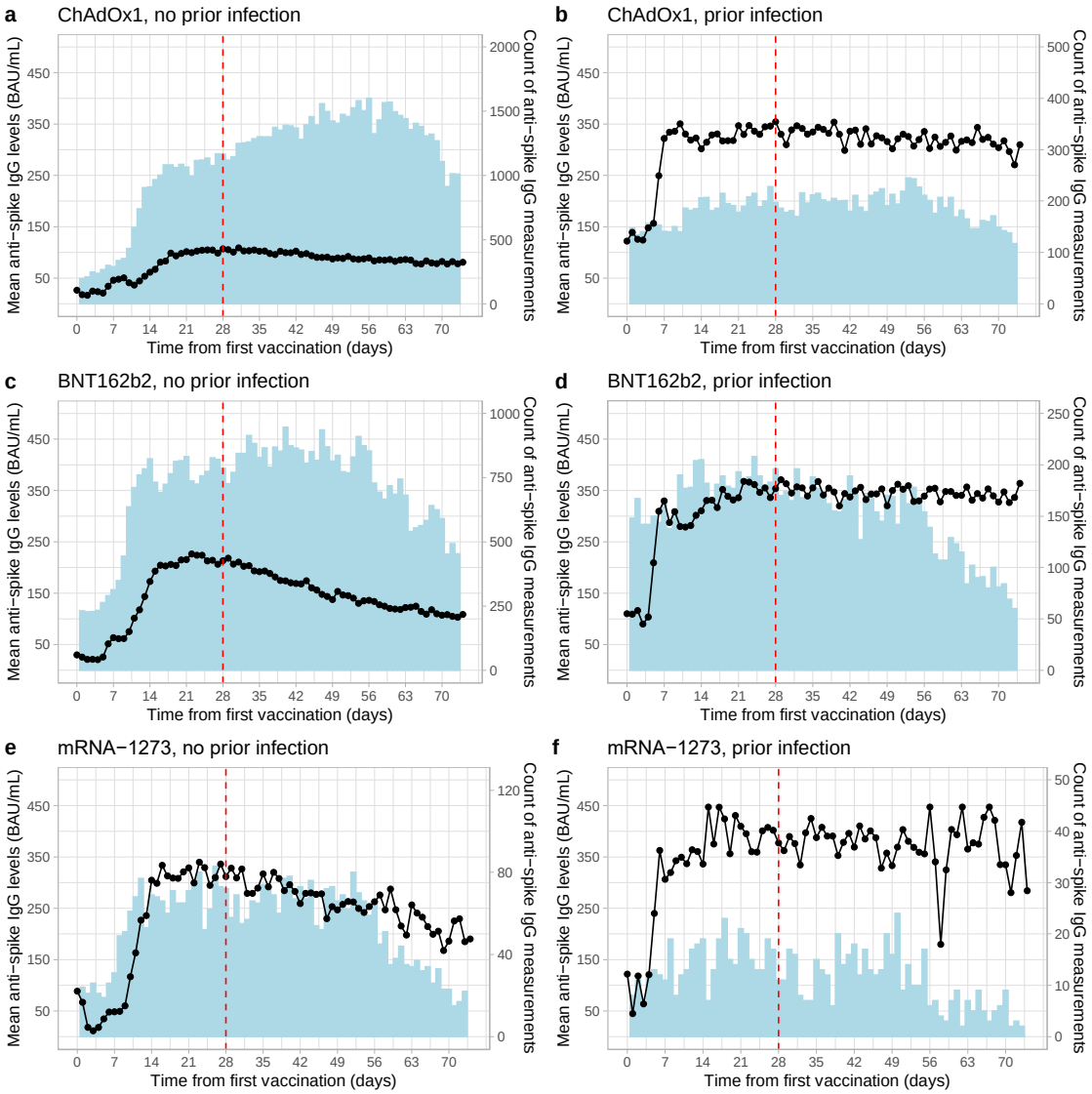


Figure 5.32 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day after the first vaccination.

Note: Panels are separated by vaccines (ChAdOx1, BNT162b2, mRNA-1273) and prior infection status. Red dotted lines represent the time of the 'peak level', which is 28 days. Values truncated at 450 BAU/mL counted as =450 BAU/mL for calculation of unadjusted means only.

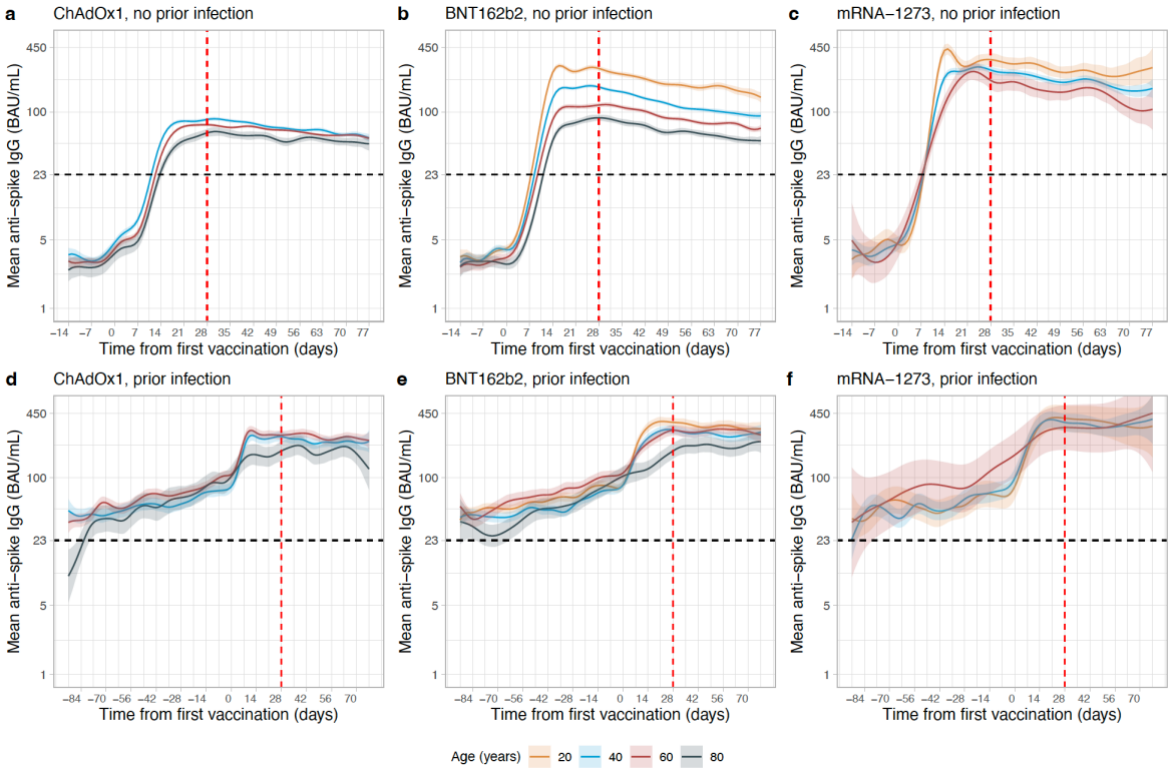


Figure 5.33 Mean anti-spike IgG levels (95% Cl) by time from first vaccination by age, vaccine type, and prior infection status.

Note: a, No prior infection and received one ChAdOx1 vaccination, N=70,146. b, No prior infection and received one BNT162b2 vaccination, N=47,065. c, No prior infection and received one mRNA-1273 vaccination, N=3,591. d, With prior infection and received one ChAdOx1 vaccination, N=10,207. e, With prior infection and received one BNT162b2 vaccination, N=10,116. f, With prior infection and received one mRNA-1273 vaccination, N=807. Predicted levels are plotted on a log 10 scale. Black dotted line indicates the threshold of IgG positivity (23 BAU/mL). Red dotted lines represent the time of the 'peak level', which is approximately 28 days. Line colour indicates response predicted for ages 20, 40, 60, and 80 years. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. Values truncated at 450 BAU/mL counted as =450 BAU/mL in this minimally adjusted descriptive analysis.

ChAdOx1

59,101 participants with a single ChAdOx1 vaccination (7,191 (12.2%) previously infected) contributed 73,963 antibody measurements ≥28 days post-first dose, median (IQR) [range] 2 (1-2) [1-3] measurements per participant. Assuming antibody levels declined exponentially, using multivariable Bayesian (log)linear mixed models (and accounting for censoring arising from the upper limit of quantification of the antibody assay) I estimated a median peak anti-spike IgG level of 447 BAU/mL (95% credible

interval, Crl 432-462), and a median half-life of 85 days (95%Crl 76-97) for those who had already been infected with SARS-CoV-2 before one vaccination in the reference category (50 years, female, white ethnicity, not reporting a long-term health condition, not a healthcare worker, and deprivation percentile=60). This peak was substantially higher than estimates after a second ChAdOx1 vaccination in participants without prior infection (160 BAU/mL [157-162] for the same reference category), and the half-life was similar (81 days [79-83]). Those receiving one ChAdOx1 vaccination without prior infection had significantly lower peak levels, 84 BAU/mL (81-85), but a similar antibody half-life to those with prior infection, 93 days (89-99) (**Figure *5.34*a, Table *5.15***). **Table *5.16*** compares peak levels and half-lives for different age groups, sex, and ethnicity. Previously infected participants reporting non-white ethnicity had higher post-first dose peak levels but slightly shorter half-lives regardless of prior infection (**Table *5.17*, Table *5.15***).

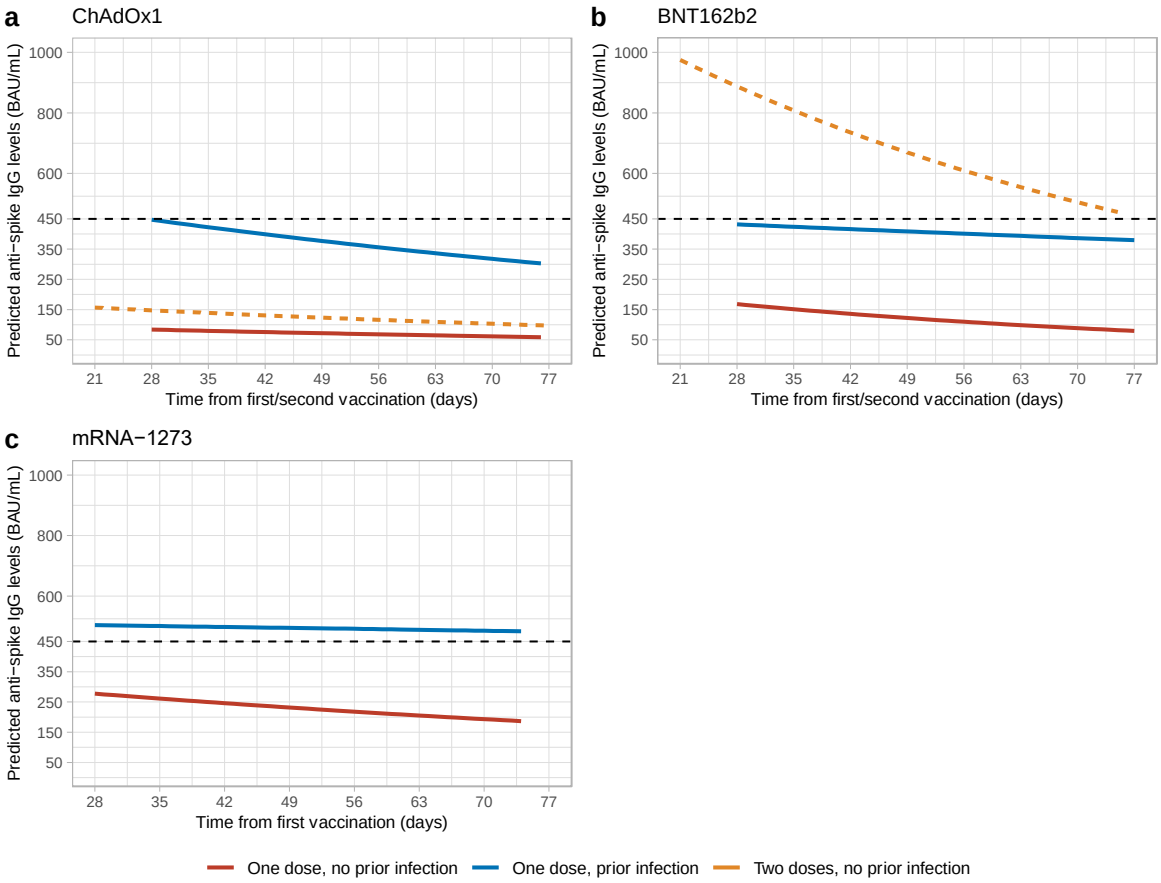


Figure 5.34 Posterior predicted trajectories (95% credible interval) of anti-spike IgG levels from 28 days post-first dose by prior infection status.

Note: **a,** 59,101 participants who received at least a single ChAdOx1 vaccination. **b,** 35,638 participants who received at least a single BNT162b2 vaccination. **c,** 2,597 participants who received at least a single mRNA-1273 vaccination. Plotted at reference categories: 50 years, female, white ethnicity, not reporting a long-term health condition, not a healthcare worker, and deprivation percentile=60. Black dotted line shows the upper quantification limit of 450 BAU/mL. Orange dotted lines in panel a and b were predicted trajectories starting from 21 days post-second dose for 92,584 and 51,034 participants who received two ChAdOx1 and BNT162b2 vaccination without prior infection, plotted at the same reference categories.

Table 5.15 Posterior median and 95% credible intervals for anti-spike IgG peak level (intercept) (BAU/mL) and half-life (slope) (days) in the multivariable models.

Note: 59,101 participants with first ChAdOx1 vaccination, 35,638 participants with first BNT162b2 vaccination, and 2,597 participants with first mRNA-1273 vaccination. The reference categories in the multivariable model are: no prior infection, 50-year-old, female, white ethnicity, not reporting a long-term health condition, not a healthcare worker, and deprivation percentile=60. Bold numbers indicate an effect not compatible with chance (95% credible interval excludes no effect). In mRNA-1273, some numbers are not estimable because antibody responses in certain groups were not estimated to decline.

		ChAdOx1			**BNT162b2**			**mRNA-1273**		
		Estimate	95% credible interval		Estimate	95% credible interval		Estimate	95% credible interval	
Adjusted baseline	Peak level (Intercept)	84	81	85	168	165	172	278	245	317
	IgG half-life (slope)	93	89	99	46	44	47	80	51	178
Prior infection	Peak level: No									
	IgG half-life: No									
	Change in peak level: Yes	**363**	**348**	**378**	**263**	**248**	**280**	**227**	**148**	**325**
	Change in half-life: Yes	-8	-18	3	**220**	**130**	**497**	**33**	**7**	**Not estimable**
Age	Peak level: 50 years									
	IgG half-life: 50 years									
	Change in peak level: per 10-years older	**-4**	**-5**	**-4**	**-28**	**-29**	**-27**	**-56**	**-67**	**-44**
	Change in half-life: per 10-years older	**6**	**3**	**9**	**2**	**1**	**3**	**32**	**10**	**Not estimable**
Sex	Peak level: Female									
	IgG half-life: Female									
	Change in peak level: Male	**-6**	**-7**	**-4**	**-13**	**-17**	**-9**	-32	-64	-3
	Change in half-life: Male	**-9**	**-15**	**-4**	0	-3	2	-21	-105	10
Ethnicity	Peak level: White									
	IgG half-life: White									
	Change in peak level: Non-white	**36**	**31**	**41**	**33**	**24**	**43**	-23	-78	40
	Change in half-life: Non-white	**-13**	**-22**	**-1**	-2	-6	3	20	-9	Not estimable
	Peak level: No									

		ChAdOx1			BNT162b2			mRNA-1273		
Report having a long-term health condition	IgG half-life: No									
	Change in peak level: Yes	**-5**	**-7**	**-3**	**-17**	**-22**	**-13**	7	-38	60
	Change in half-life: Yes	-3	-10	4	0	-3	3	-31	-119	8
Report working in patient-facing healthcare	Peak level: No									
	IgG half-life: No									
	Change in peak level: Yes	6	-1	14	-11	-22	2	-54	-161	153
	Change in half-life: Yes	-18	-33	7	**25**	**12**	**49**	Not estimable		
Deprivation percentile	Peak level: 60 (median)									
	IgG half-life: 60 (median)									
	Change in peak level: per 10 percentile higher	0	-1	0	0	-1	1	3	-3	9
	Change in half-life: per 10 percentile higher	1	0	2	0	0	1	0	-15	13

Table 5.16 Posterior predicted peak levels (BAU/mL), half-lives (days), and time from first/second dose to the positivity threshold (days) with 95% credible intervals.

Note: Results are shown in participants who received one vaccination with prior infection, two vaccinations without prior infection, and one vaccination without prior infection, by vaccine type. Results are separated by age (20, 40, 60, 80-year-old), sex (female vs male) and ethnicity (white vs non-white). Estimations for two vaccinations without prior infection were based on results from the previous section. NE: Not estimable, due to the antibody levels not declining in the posterior median or upper credible interval.

			Peak level (BAU/mL)			Half-life (days)			Time to positivity threshold 23 BAU/mL (days)		
			One dose with prior infection	Two doses without prior infection	One dose without prior infection	One dose with prior infection	Two doses without prior infection	One dose without prior infection	One dose with prior infection	Two doses without prior infection	One dose without prior infection
ChAdOx1	40	White female	471 (455-488)	157 (154-159)	89 (87-90)	81 (72-92)	81 (79-83)	88 (83-95)	381 (318-453)	246 (170-322)	200 (141-260)
		White male	439 (424-456)	152 (150-154)	83 (81-84)	74 (66-83)	82 (80-84)	80 (75-85)	342 (287-404)	244 (167-318)	175 (123-230)
		Non-white female	672 (640-707)	197 (192-203)	126 (121-131)	71 (62-83)	73 (70-77)	77 (68-88)	376 (314-448)	248 (181-318)	217 (165-275)
		Non-white male	627 (595-659)	191 (186-197)	118 (113-123)	66 (58-76)	74 (70-77)	70 (63-79)	342 (288-406)	246 (176-316)	194 (146-246)
	60	White female	424 (409-439)	162 (160-165)	80 (78-81)	90 (80-103)	80 (78-82)	99 (94-105)	407 (337-488)	246 (169-323)	206 (142-274)
		White male	395 (381-409)	157 (155-160)	74 (73-75)	81 (73-92)	80 (78-82)	89 (84-94)	362 (300-430)	244 (169-318)	178 (119-238)
		Non-white female	604 (574-636)	204 (198-211)	114 (109-118)	78 (67-94)	72 (69-76)	85 (75-99)	398 (329-484)	248 (182-317)	224 (165-288)
		Non-white male	563 (535-593)	198 (192-204)	106 (102-110)	72 (62-84)	73 (69-76)	77 (69-88)	359 (299-432)	245 (178-314)	198 (145-255)
	80	White female	381 (366-396)	168 (164-172)	72 (70-73)	102 (87-122)	79 (76-81)	113 (102-127)	440 (358-546)	247 (174-321)	214 (138-293)
		White male	355 (341-369)	163 (159-166)	67 (65-68)	91 (79-107)	79 (76-82)	100 (91-111)	386 (316-472)	243 (173-318)	181 (114-249)
		Non-white female	543 (513-574)	212 (204-219)	102 (97-107)	87 (73-109)	71 (68-75)	95 (81-116)	424 (341-534)	248 (181-315)	233 (166-311)
		Non-white male	506 (478-536)	205 (198-212)	95 (91-100)	79 (67-96)	71 (68-75)	86 (74-102)	379 (310-472)	247 (178-316)	204 (146-270)
BNT162b2	20	White female	741 (711-771)	1,243 (1,196-1,293)	289 (282-296)	150 (115-220)	51 (50-53)	40 (38-42)	780 (558-1,183)	323 (238-405)	175 (131-220)
		White male	682 (654-711)	1,066 (1,024-1,110)	266 (259-273)	147 (112-212)	51 (49-53)	40 (38-42)	745 (530-1,110)	308 (223-392)	169 (125-214)
		Non-white female	887 (836-942)	1,421 (1,331-1,520)	346 (329-364)	129 (94-207)	51 (48-54)	38 (35-43)	707 (489-1,141)	330 (244-416)	179 (136-224)
		Non-white male	817 (769-869)	1,218 (1,139-1,302)	319 (303-335)	126 (92-200)	50 (48-53)	38 (35-42)	679 (466-1,077)	315 (233-397)	173 (130-217)
	40	White female	517 (497-537)	1,057 (1,023-1,094)	202 (198-206)	212 (150-358)	52 (50-53)	44 (42-46)	979 (647-1,692)	312 (228-397)	164 (118-213)
		White male	476 (458-494)	906 (876-938)	186 (182-189)	206 (146-340)	51 (49-53)	43 (42-45)	924 (609-1,560)	297 (213-381)	159 (110-205)
		Non-white female	619 (583-657)	1,208 (1,134-1,290)	242 (230-254)	171 (114-340)	51 (48-54)	42 (38-46)	838 (528-1,672)	319 (237-405)	169 (122-217)

			Peak level (BAU/mL)			Half-life (days)			Time to positivity threshold 23 BAU/mL (days)		
			One dose with prior infection	Two doses without prior infection	One dose without prior infection	One dose with prior infection	Two doses without prior infection	One dose without prior infection	One dose with prior infection	Two doses without prior infection	One dose without prior infection
		Non-white male	570 (537-605)	1,035 (971-1,105)	222 (211-233)	167 (111-321)	51 (48-53)	41 (38-46)	798 (511-1,531)	305 (223-388)	163 (117-210)
	60	White female	361 (346-375)	898 (866-933)	141 (138-144)	356 (207-1205)	52 (50-53)	48 (45-50)	1,440 (780-4,837)	300 (216-384)	152 (100-206)
		White male	332 (319-346)	770 (742-798)	129 (127-132)	339 (199-1006)	51 (50-53)	47 (45-50)	1,338 (733-3,945)	285 (202-368)	146 (95-197)
		Non-white female	432 (406-459)	1,027 (960-1,099)	168 (160-177)	254 (143-1058)	51 (48-54)	45 (41-51)	1,102 (595-4,346)	307 (225-393)	158 (107-211)
		Non-white male	398 (374-423)	880 (822-941)	155 (147-163)	245 (139-897)	51 (48-54)	45 (40-50)	1,030 (572-3,682)	293 (211-377)	152 (102-203)
	80	White female	252 (240-264)	764 (728-801)	98 (95-101)	1120 (316-NE)	56 (52-60)	52 (49-56)	3,925 (1,064-NE)	289 (203-373)	138 (80-196)
		White male	231 (221-243)	655 (624-685)	90 (88-93)	952 (300-NE)	54 (50-59)	52 (49-56)	3,184 (948-NE)	274 (190-358)	131 (74-188)
		Non-white female	301 (282-322)	873 (809-942)	118 (111-124)	491 (190-NE)	49 (44-57)	49 (44-57)	1857 (696-NE)	297 (213-383)	144 (90-203)
		Non-white male	277 (260-296)	748 (694-805)	108 (102-114)	460 (183-NE)	48 (43-55)	49 (43-56)	1687 (664-NE)	283 (201-366)	138 (85-194)
mRNA-1273	20	White female	997 (812-1,232)		550 (475-637)	76 (39-1090)		41 (30-65)	437 (232-5,023)		216 (144-338)
		White male	881 (724-1,078)		486 (423-560)	57 (34-178)		35 (27-49)	325 (195-937)		180 (122-264)
		Non-white female	912 (686-1,223)		503 (392-649)	162 (42-NE)		58 (30-681)	894 (252-NE)		286 (155-2,764)
		Non-white male	806 (603-1,086)		445 (344-576)	94 (35-NE)		46 (27-174)	507 (206-NE)		224 (131-710)
	40	White female	634 (530-760)		349 (317-386)	191 (62-NE)		61 (46-91)	943 (313-NE)		267 (168-419)
		White male	560 (470-670)		309 (283-337)	102 (49-NE)		48 (39-62)	495 (237-NE)		207 (131-304)
		Non-white female	581 (443-762)		320 (255-401)	NE (68-NE)		106 (43-NE)	NE (361-NE)		427 (179-NE)
		Non-white male	513 (391-674)		282 (224-355)	363 (51-NE)		72 (36-NE)	1686 (262-NE)		287 (144-576)
	60	White female	403 (314-513)		222 (185-267)	NE (87-NE)		118 (55-NE)	NE (394-NE)		413 (180-NE)
		White male	356 (280-452)		196 (164-234)	514 (64-NE)		77 (45-297)	2105 (277-NE)		265 (135-928)
		Non-white female	368 (269-506)		203 (154-269)	NE (111-NE)		708 (57-NE)	NE (488-NE)		2252 (200-NE)
		Non-white male	326 (237-446)		179 (136-238)	NE (72-NE)		169 (45-NE)	NE (317-NE)		528 (145-NE)

Table 5.17 Posterior predicted median peak levels (BAU/mL) and half-lives (days) with 95% credible intervals in participants who received one vaccination with prior infection, two vaccinations without prior infection, and one vaccination without prior infection, by vaccine type.

Note: Estimates for two vaccinations without prior infection were based on results from the previous section. All estimates are at the reference age (50-year-old) and separated by sex (female vs male) and ethnicity (white vs non-white).

			Peak level (BAU/mL)			Half-life (days)		
Vaccine	Age	Group	One dose with prior infection	Two doses without prior infection	One dose without prior infection	One dose with prior infection	Two doses without prior infection	One dose without prior infection
ChAdOx1	50	White female	447 (432-462)	160 (157-162)	84 (83-85)	85 (76-97)	81 (79-83)	93 (89-99)
	50	White male	416 (402-431)	154 (152-157)	78 (77-80)	77 (70-87)	81 (79-83)	84 (80-89)
	50	Non-white females	637 (607-670)	201 (195-207)	120 (115-125)	75 (65-88)	73 (70-76)	81 (72-92)
	50	Non-white males	594 (565-625)	195 (189-201)	112 (107-116)	69 (60-80)	73 (70-77)	74 (66-83)
BNT162b2	50	White female	432 (416-449)	974 (942-1,009)	168 (165-172)	266 (175-544)	52 (50-53)	46 (44-47)
	50	White male	397 (382-413)	836 (807-864)	155 (152-158)	256 (168-500)	51 (50-53)	45 (43-47)
	50	Non-white females	517 (487-549)	1,114 (1,045-1,190)	202 (192-212)	204 (127-505)	51 (48-54)	43 (39-48)
	50	Non-white males	476 (448-506)	955 (894-1,019)	186 (176-195)	198 (124-470)	51 (48-54)	43 (39-48)
mRNA-1273	50	White female	505 (413-619)		278 (245-317)	768 (75-Not estimable)		80 (51-178)
	50	White male	447 (365-546)		246 (218-278)	169 (57-Not estimable)		59 (43-97)
	50	Non-white females	462 (348-615)		255 (200-325)	Not estimable (86-Not estimable)		184 (50-Not estimable)
	50	Non-white males	409 (306-544)		225 (176-288)	Not estimable (62-Not estimable)		100 (41-Not estimable)

BNT162b2

35,638 participants with a single BNT162b2 vaccination (5,523 [15.5%] previously infected) contributed 43,163 antibody measurements ≥28 days post-first dose, median (IQR) [range] 1 (1-2) [1-3] measurements per participant. For those who had already been infected with SARS-CoV-2 before their first vaccination, the estimated median peak antibody level was 432 BAU/mL (95%Crl 416-449), and the half-life was 266 days (175-544) at the reference category. The peak levels were lower than previously estimated following two BNT162b2 vaccinations without prior infection (974 BAU/mL [942-1009]), but half-life was substantially longer (51 days [50-53]). Those with one vaccination but no prior infection had lower peak levels (168 BAU/mL [165-172]) and shorter half-lives (46 days [44-47]) (**Figure *5.34*b, Table *5.15***). Similar to ChAdOx1, participants reporting non-white ethnicity had higher peak levels (**Table *5.17*, Table *5.15***).

mRNA-1273

2,597 participants with one mRNA-1273 vaccination (414 (15.9%) previously infected) contributed 2,975 antibody measurements ≥21 days post-first dose, median (IQR) [range] 1 (1-2) [1-2] measurements per participant. For those who had already been infected with SARS-CoV-2 before their first vaccination, the estimated peak level was 505 BAU/mL (413-619) at the reference category, and the half-life was 768 days (75-Not estimable). I was not able to estimate peak and half-life after two vaccinations in my previous study of initial vaccinations; in this study, participants without prior infection had lower peak levels after one vaccination (278 BAU/mL (245-317)) and half-life was estimated to be 80 days (51-178) (**Table *5.17***, **Figure *5.34*c, Table *5.15***). mRNA-1273 elicited higher antibody levels (especially in those without prior infection) and had a lower percentage of seronegative non-responders versus BNT162b2 (0.7% vs 4.2%), consistent with findings after two vaccinations among US healthcare workers[176].

5.3.6 Antibody correlates of protection against infection in the Delta period (cohort 2)

(Note: Dr. Koen Pouwels wrote the codes and performed the analyses in this section; I created the plots and summarised the results)

To interpret anti-spike IgG levels over time, I investigated whether recent antibody levels could be considered as correlates of protection from infection following prior infection or vaccination in the general population. I used data from 17th May 2021 to 4th October 2021, i.e. while the Delta variant accounted for nearly all cases[177], and fitted logistic generalised additive models for detected new infections, investigating the effect of the most recent antibody measurement obtained 21-59 days earlier on PCR test results at each study visit (distribution of visits relative to first vaccination shown in **Figure *5.35*** and **Table *5.18***). Three groups were investigated, unvaccinated participants with or without evidence of prior infection (6833 participants; 12,560 visits), participants vaccinated with ChAdOx1 ≥21 days previously (83,924 participants; 221,380 visits, 36-273 days post-first vaccination), and participants vaccinated with BNT162b2 ≥21 days previously (49,820 participants; 124,822 visits, 35-298 days post-first vaccination). Participants tested PCR-positive at 202 (1.6%), 1327 (0.6%), and 591 (0.5%) visits, respectively. Vaccinated participants with evidence of prior infection were excluded, as there were insufficient data to model these groups separately, and associations may differ. Adjustment was made for confounders, including age, geography, and calendar time (see Methods).

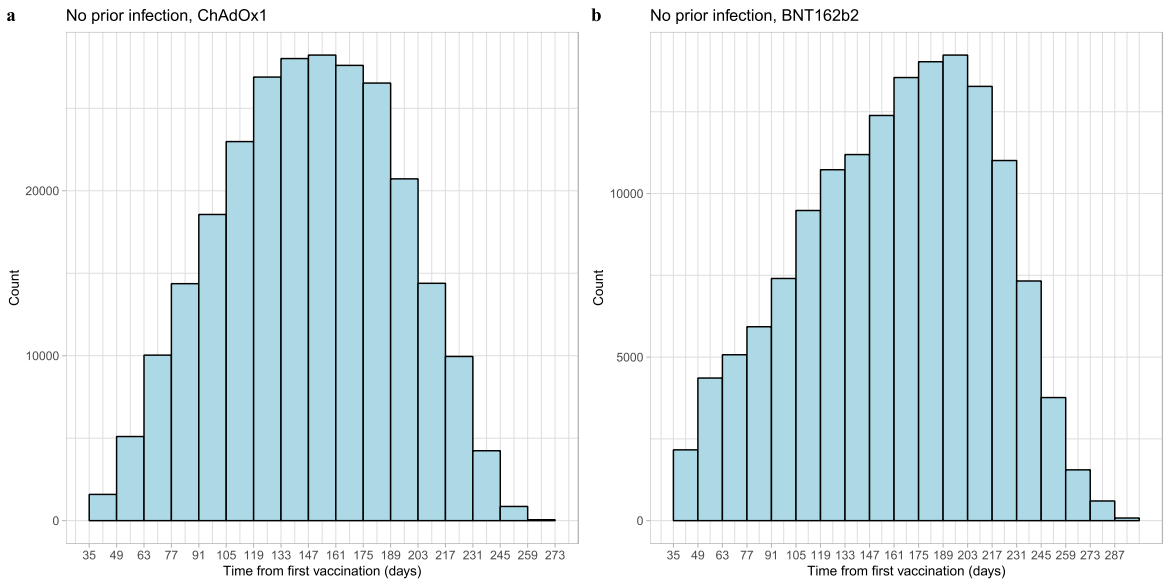


Figure 5.35 Distribution of study visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection.

Note: Study visits were included from 17th May 2021 to 4th October 2021, i.e. while the Delta variant accounted for nearly all cases. The median (IQR) [range] days from the first vaccination was 149 (115-182) [36-273] days for ChAdOx1 and 167 (123-204) [35-298] days for BNT162b2.

Table 5.18 Frequency of visits relative to the time of the first vaccination in the generalised additive models estimating correlates of protection.

Note: Study visits were included from 17th May 2021 to 4th October 2021, i.e. while the Delta variant accounted for nearly all cases. The median (IQR) [range] days from the first vaccination was 149 (115-182) [36-273] days for ChAdOx1 and 167 (123-204) [35-298] days for BNT162b2. Note that UK started its booster programme from 16th September 2021, so a small number of visits occurred in the first two weeks after initiating the booster campaign.

Days from first vaccination	ChAdOx1	BNT162b2
35-49	1,364	1,875
49-63	4,766	4,327
63-77	9,681	5,027
77-91	14,081	5,830
91-105	18,268	7,318
105-119	22,644	9,288
119-133	26,749	10,725
133-147	27,934	11,090
147-161	28,251	12,327
161-175	27,636	13,392
175-189	26,770	14,111
189-203	21,210	14,195
203-217	14,786	13,397
217-231	10,285	11,202

Days from first vaccination	ChAdOx1	BNT162b2
231-245	4,640	7,638
245-259	993	3,966
259-273	65	1,673
273-287	1	647
287-298	0	103

Compared with unvaccinated individuals with a most recent anti-spike IgG measurement of 1 BAU/mL, protection against infection increased steeply as antibody levels rose in unvaccinated participants, consistent with protection from prior infection, with 50% protection at 20 BAU/mL (below the 23 BAU/mL positivity threshold) and 67% protection at 33 BAU/mL. Higher antibody levels were needed to achieve the same level of protection post vaccination, with no evidence of differences in protection between ChAdOx1 and BNT162b2 at any given antibody level (**Figure *5.36*a**). For example, considering antibody levels where two thirds (67%) of individuals were protected against infection, participants vaccinated with ChAdOx1 or BNT162b2 required estimated levels of 107 BAU/mL and 94 BAU/mL (100 BAU/mL in models pooling both vaccines), respectively. However, as described above, antibody levels rose to higher levels following BNT162b2 than ChAdOx1, explaining higher vaccine effectiveness after BNT162b vs. ChAdOx1, with natural infection resulting in the lowest antibody levels but the greatest protection at a given antibody level (**Figure *5.36*d,e,f**). Protection against infection with moderate to high viral loads (Ct values <30) (**Figure *5.36*b**) and symptomatic infection (**Figure *5.36*c**) was similar. Findings were also similar when considering the maximum prior antibody measurement (which had a worse model fit), in part because of the limited time for antibody waning to occur between maximum and most recent antibody measurements (**Figure *5.37***).

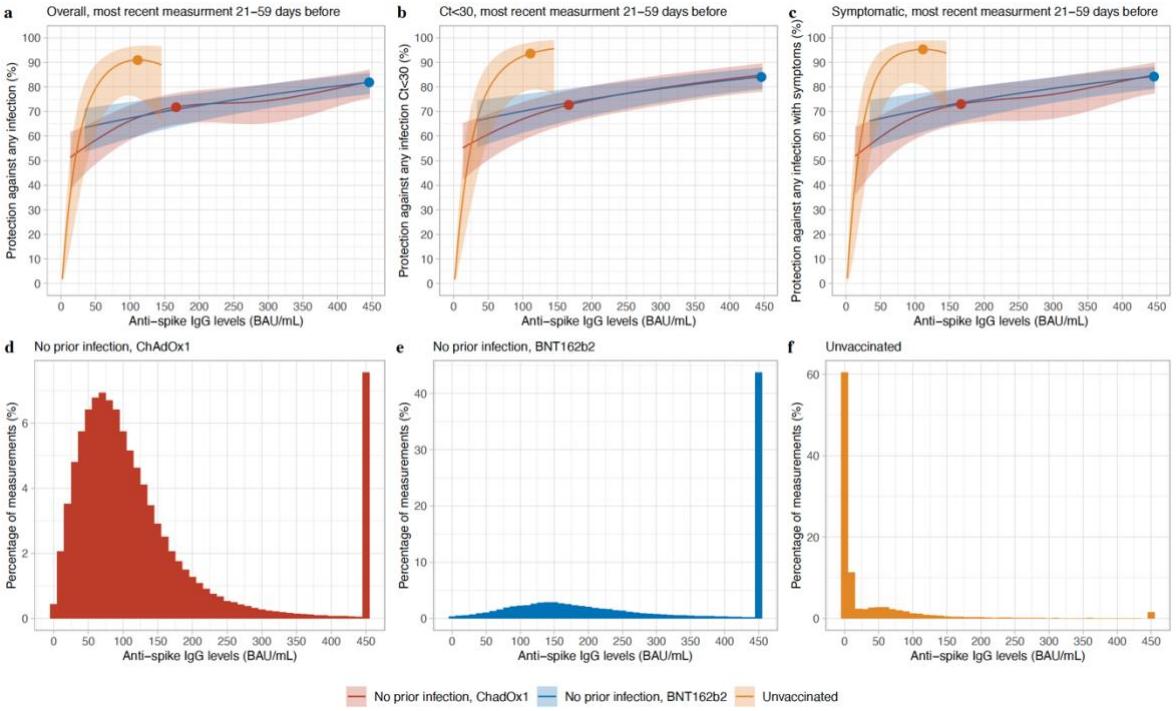


Figure 5.36 Association between anti-spike IgG levels and protection from SARS-CoV-2 infection using the most recent antibody measurement obtained 21-59 days prior to the current visit.

Note: **a,** protection against any infection; **b**, protection against infection with a moderate to high viral load (Ct value <30); **c**, protection against infection with self-reported symptoms. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. Three groups are investigated, unvaccinated participants with or without evidence of prior infection, participants vaccinated with ChAdOx1 without evidence of prior infection, and participants vaccinated with BNT162b2 without evidence of prior infection. Dots represent the median predicted individual peak levels from the Bayesian linear mixed models: 1026 BAU/mL for BNT162b2 (plotted at the upper quantification limit of 450 BAU/mL), 167 BAU/mL for ChAdOx1, and 111 BAU/mL for unvaccinated participants. Distribution of the most recent anti-spike IgG measurements for the three population groups are shown in panel **d, e, f**.

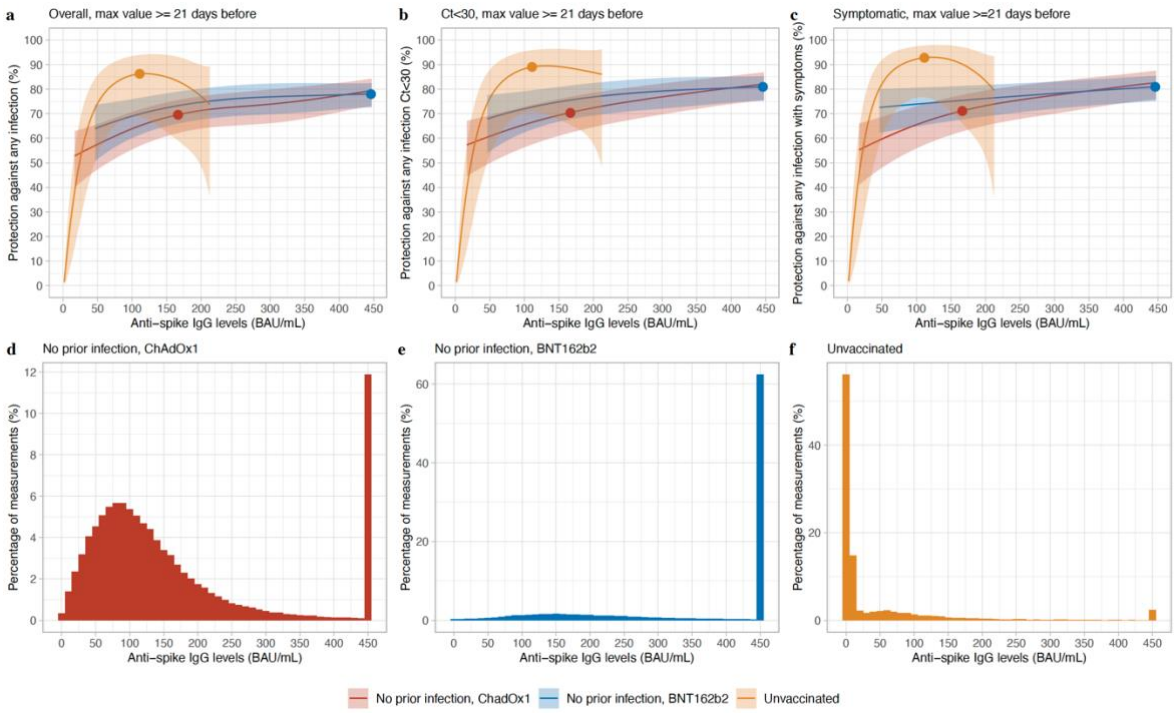


Figure 5.37 Association between anti-spike IgG levels and protection from SARS-CoV-2 infection using the maximum antibody measurement obtained ≥21 days prior to the visit.

Note: **a,** protection against any infection; **b,** protection against infection with a moderate to high viral load (Ct value <30); **c,** protection against infection with self-reported symptoms. The 95% Cis are calculated by prediction ± 1.96 × standard error of the prediction. Three groups are investigated, unvaccinated participants with or without evidence of prior infection, participants vaccinated with ChAdOx1 without evidence of prior infection, and participants vaccinated with BNT162b2 without evidence of prior infection. Dots represent the median predicted individual peak levels from the Bayesian linear mixed models: 1026 BAU/mL for BNT162b2 (plotted at the upper quantification limit of 450 BAU/mL), 167 BAU/mL for ChAdOx1, and 111 BAU/mL for unvaccinated participants. Distribution of the maximum anti-spike IgG measurements for the three population groups are shown in panel **d, e, f.**

5.3.7 Comparison of antibody trajectories and duration of protection (cohort 2 and 3)

5.3.7.1 After the second vaccination (cohort 2)

Models of correlates of protection allowed post-vaccination antibody levels corresponding to 67% protection from Delta variant infection to be estimated (data were insufficient to estimate lower protection percentages for BNT162b). The estimated mean time from second vaccination to antibody measurements reaching levels associated with 67% protection was 55-86 days for ChAdOx1 vaccinated participants without prior infection, with limited variation across age, sex, dosing interval, or long-term health conditions

(**Figure *5.38*a**). For BNT162b2, the estimated mean durations were 161-227 days, and were longer at younger ages, in females, those without long-term health conditions, but were similar across different dosing intervals (**Figure *5.38*b**). For unvaccinated participants infected previously, using the model of antibody declines after natural infection from Chapter 4, the duration from infection to the level associated with 67% protection was estimated to be 1-2 years (**Figure *5.38*c**). Older people had a longer duration of protection after natural infection, but these results were conditional on seroconverting, and older individuals had lower seroconversion rates (**Figure *5.38*d**). Although data were insufficient to estimate antibody levels correlated with protection for vaccinated participants with prior infection, conservatively assuming threshold levels were similar to those vaccinated without prior infection and given their half-lives were longer, the duration of protection could last for >1 year. Times for antibody levels to fall to the threshold for positivity, i.e. 23 BAU/mL, were longer, but followed the same patterns (**Figure *5.39***).

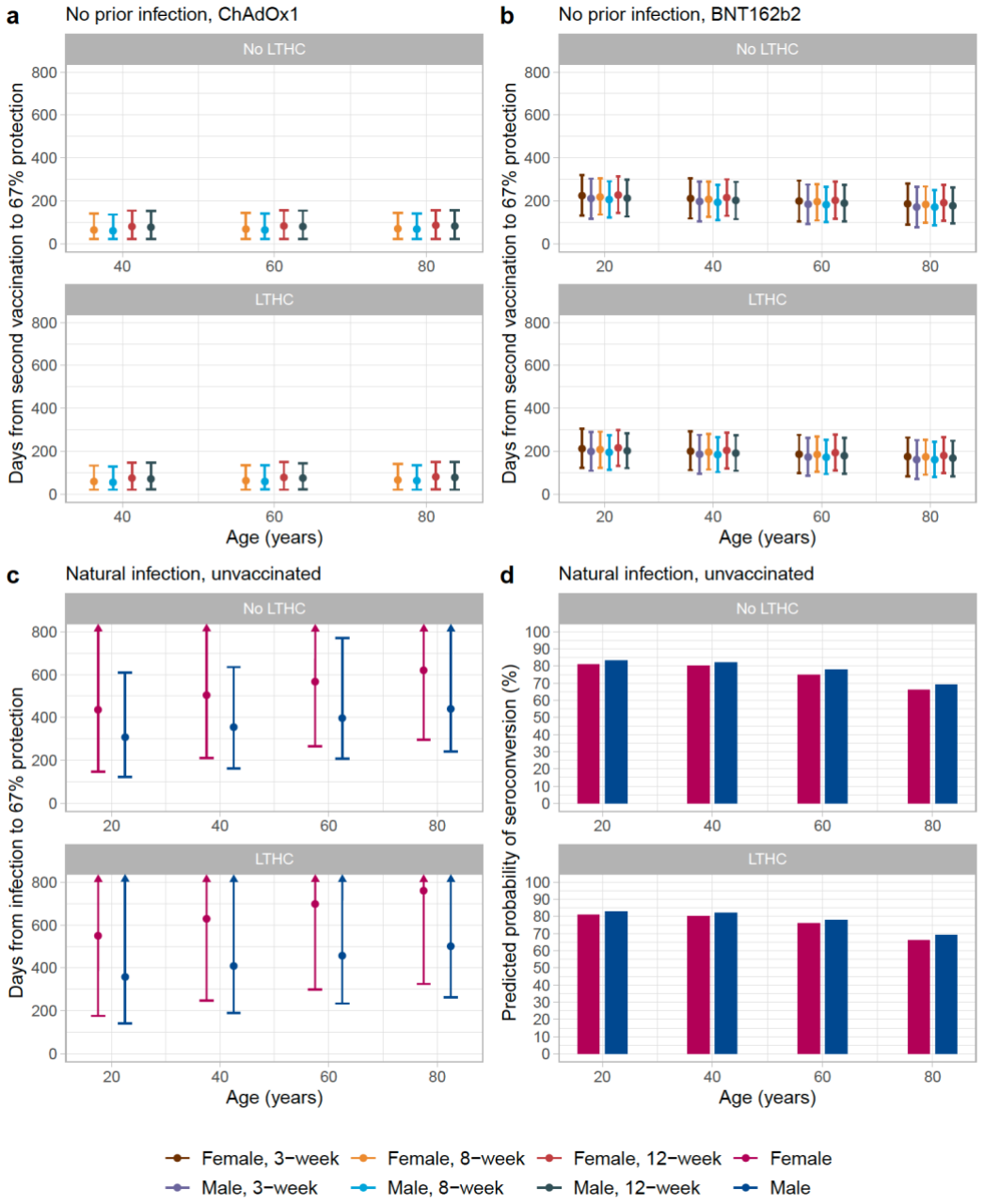


Figure 5.38 Posterior predicted mean days (95% credible interval) from the second vaccination or first infection to the threshold level associated with 67% protection.

Note: The threshold level is 107 BAU/mL for ChAdOx1, 94 BAU/mL for BNT162b2, and 33 BAU/mL for unvaccinated participants. **a,** In participants without prior infection and vaccinated with ChAdOX1 (N=92,584). **b,** In participants without prior infection and vaccinated with BNT162b2 (N=51,034). **c,** In unvaccinated participants who had natural infection (N=3,271). **d,** Predicted probability of seroconverting in unvaccinated participants who had natural infection, based on Chapter 4**.** Estimates were separated by age, sex, dosing interval, long-term health condition (LTHC), and vaccine type for vaccinated people, and by age, sex, and LTHC for unvaccinated people. y-axis is truncated at 800 days for visualisation. For ChAdOx1, 20-year-old group is not plotted because the vast majority of those receiving ChAdOx1 were ≥40 years. Panel a,b,c are conditional on participants having antibody response/seroconverting (see section of "Vaccine non-responders" for ChAdOx1 and BNT162b2, and discussion in Chapter 4 for unvaccinated individuals).

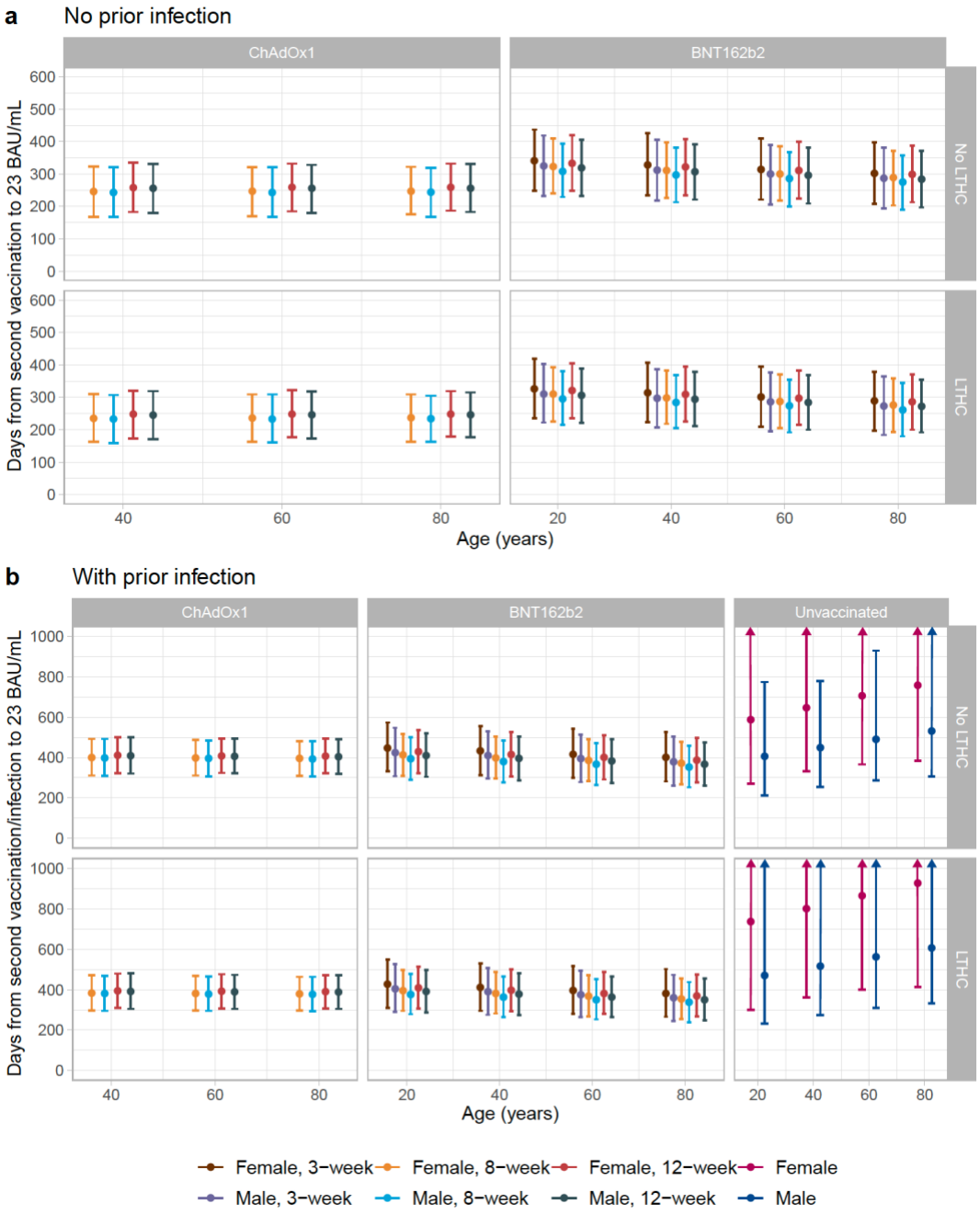


Figure 5.39 Posterior predicted mean days (95% credible interval) from the second vaccination or first infection to the positivity threshold of 23 BAU/mL.

Note**: a**, In those without evidence of prior infection, 92,584 participants vaccinated with ChAdOx1, 51,034 participants vaccinated with BNT162b2. **b**, In those with evidence of prior infection, 8,055 vaccinated with ChAdOx1, 4,019 vaccinated with BNT162b2, 3,271 had natural infection and unvaccinated. Estimates were separated by age, sex, dosing interval, long-term health condition (LTHC), and vaccine type for vaccinated people, and by age, sex, and LTHC for unvaccinated people. y-axis is truncated at 1000 days (panel b) for visualisation. For ChAdOx1, 20-year-old group is not plotted because the vast majority of those receiving ChAdOx1 were ≥40 years. Estimates for the unvaccinated group in panel b are based on a previously reported biphasic model assuming the rate of antibody decline slowing over time, which may explain the longer point estimates vs ChAdOx1 or BNT162b2, however credible intervals also overlap.

I estimated the proportion of participants with 67% protection at 90, 180, 270, and 360 days from second dose or natural infection based on individual-level predictions, conditional on seroconverting. At 180 days, 10% of ChAdOx1 participants without prior infection remained above the level associated with 67% protection, with little variation across different factors. For BNT162b2, the proportion varied between 40-80% for those without prior infection and was higher in younger ages, females, and those without long-term health conditions. Over 90% of unvaccinated participants with natural infection were above the level required for 67% protection 180 days post infection. At 270 days, almost everyone receiving ChAdOx1 or BNT162b2 fell below the level of 67% protection, while over 80% of unvaccinated participants were still above this threshold level (**Figure *5.40***).

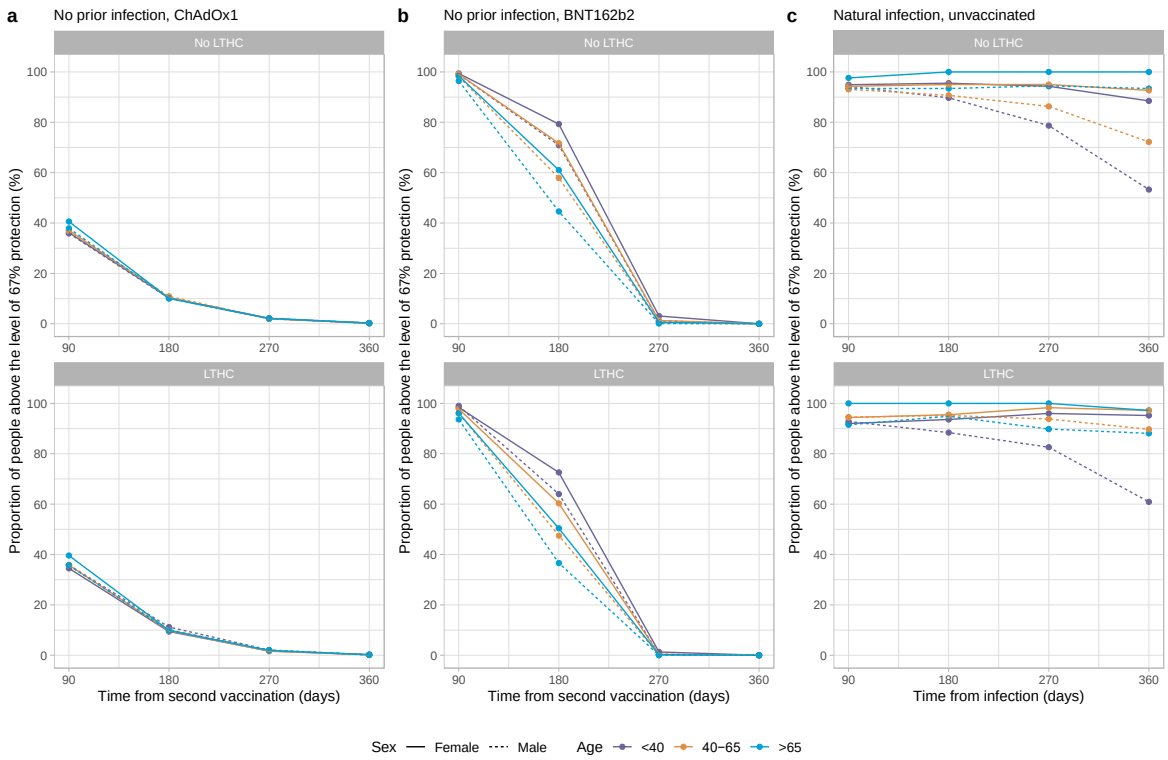


Figure 5.40 Percentage of participants above the threshold level associated with 67% protection by time from second vaccination/infection.

Note: The threshold level is 107 BAU/mL for ChAdOx1, 94 BAU/mL for BNT162b2, and 33 BAU/mL for unvaccinated participants. a, In participants without prior infection and vaccinated with ChAdOX1. b, In participants without prior infection and vaccinated with BNT162b2. c, In participants who had natural infection and were unvaccinated. Estimates were separated by age, sex, and long-term health condition (LTHC). Numbers of participants in each panel were (in the order of no LTHC, LTHC. Numbers in bracket represents <40, 40-65, >65: a, N=72,121 [4,973, 46,160, 20,988], 28,518 [2,166, 14,315, 12,037]; b, N=36,662 [6,650, 14,108, 15,904], 18,391 [1,800, 6,696, 9,895]; c, N=2,625 [1,207, 1,243, 175], 646 [194, 322, 130]. All panels are conditional on participants having antibody response/seroconverting (see section of "Vaccine non-responders" for ChAdOx1 and BNT162b2, and Chapter 4 for unvaccinated individuals).

Emerging viral variants may need higher antibody levels for the same level of neutralising activity[178]; in a sensitivity analysis assuming 2- to 8-fold increases above the threshold associated with 67% protection against the Delta variant, protection from two vaccine doses was short-lived. For example, in a 40-year-old female without any long-term health conditions and without prior infection, if 3-fold higher antibody levels were required, BNT162b2 would provide around 120 days of protection, while ChAdOx1 never reached

the required antibody level; further, if 8-fold higher antibody levels were required, the duration of protection from BNT162b2 would be reduced to only 40 days (**Figure *5.41***).

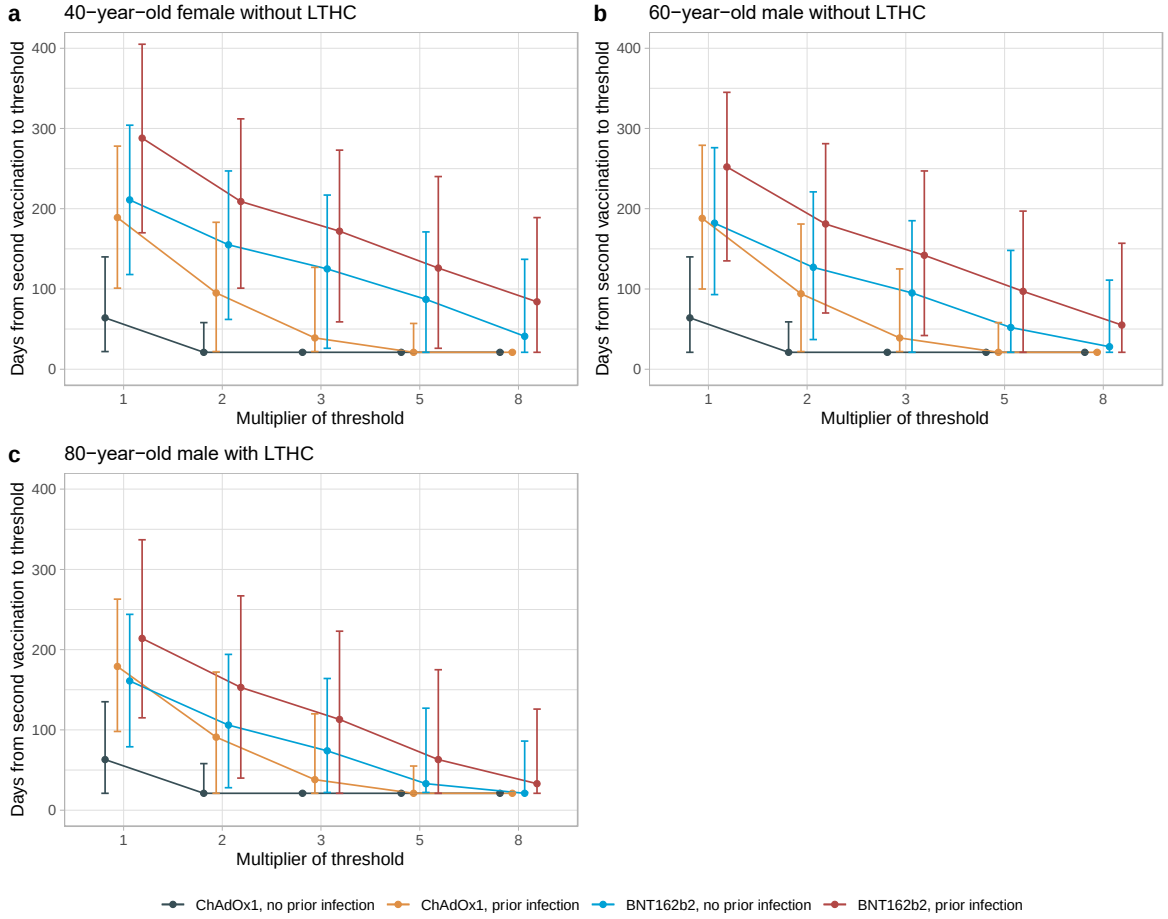


Figure 5.41 Posterior predicted days (95% credible interval) from the second vaccination to the threshold level associated with 67% protection (ChAdOx1: 107 BAU/mL, BNT162b2: 94 BAU/mL) multiplied by 2, 3, 5, 8, according to prior infection status and vaccine type.

Note: **a**, in a 40-year-old female without long-term health conditions. **b**, in a 60-year-old male without long-term health conditions. **c**, in an 80-year-old male with long-term health conditions. Estimates are based on 100,639 participants vaccinated with ChAdOx1 and 55,053 participants vaccinated with BNT162b2. All three panels were plotted at an 8-week dosing interval. LTHC: long-term health condition. Multipliers reflect the fact that higher antibody level may be required for protection against variants of concern.

5.3.7.2 After the first vaccination (cohort 3)

I estimated the time from first vaccination to levels falling to the antibody positivity threshold (23 BAU/mL, see Methods) in different subgroups. For those with prior

infection, for a 20-year-old following one vaccination, the estimated median durations were 340-440 days for those with ChAdOx1, >680 days for those with BNT162b2, and 300-900 days for those with mRNA-1273. For 40-60-year-olds with mRNA-1273, antibody levels in some groups were not estimated to decline and the upper credible intervals could not be estimated (**Figure *5.42***). Conditional on seroconverting after one dose as described above, older participants had longer durations of seropositivity than younger participants following one dose of any of the three vaccines due to longer half-lives despite lower peak levels (**Table *5.15***). Females and those without long-term health conditions also had longer estimated durations of seropositivity following one dose of any of the three vaccines. In analysis of responses post-second vaccination, the time from the second dose to antibody levels falling to 23 BAU/mL in those without prior infection was estimated to be around 250 and 270-330 days for those receiving two ChAdOx1 or BNT162b2 doses respectively, much shorter than with prior infection and one vaccination in this analysis, especially in the older age groups.

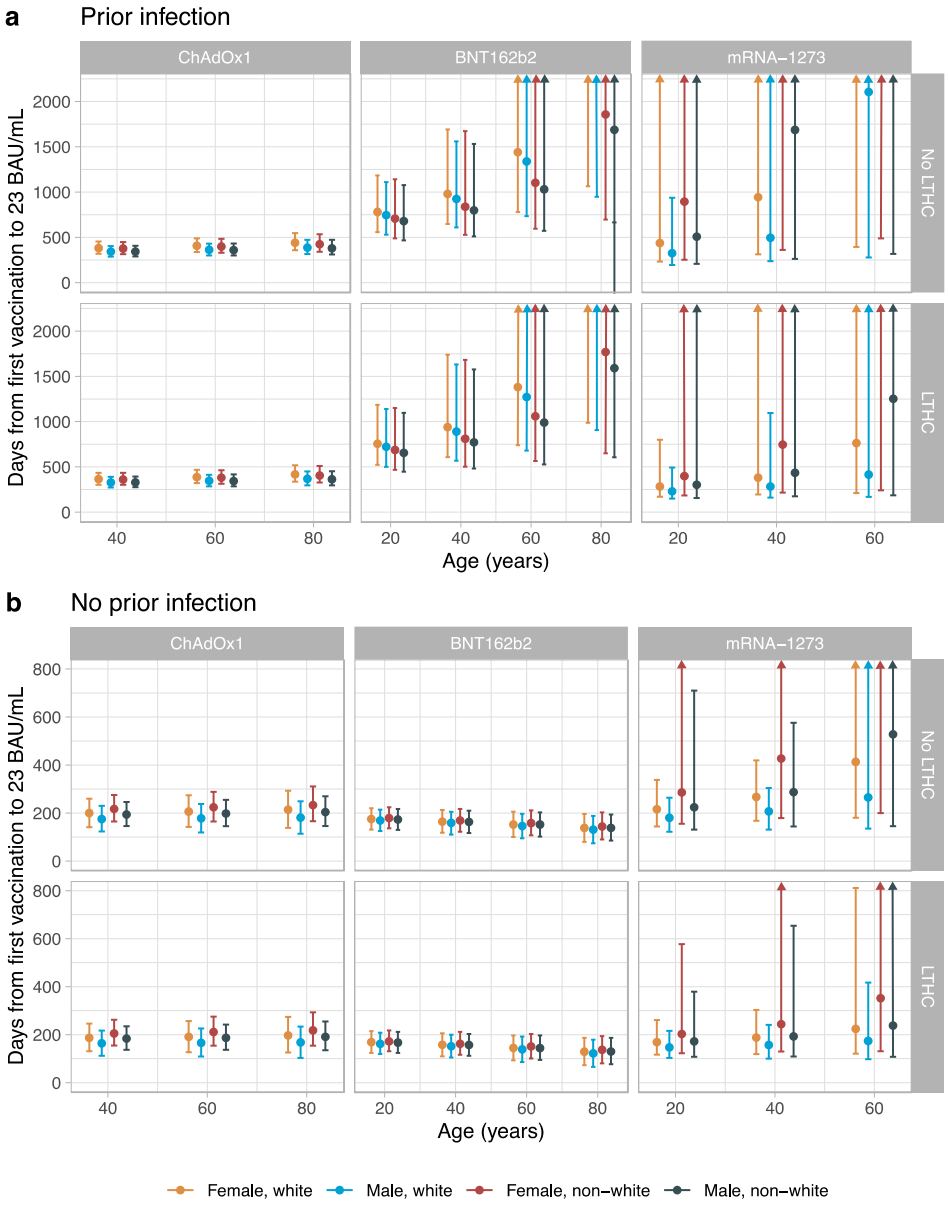


Figure 5.42 Posterior predicted mean days (95% credible interval) from the first vaccination to the positivity threshold of 23 BAU/mL.

Note: **a**, in those with evidence of prior infection (n=7,191, 5,523, 414 for ChAdOx1, BNT162b2, mRNA-1273, respectively). **b**, in those without evidence of prior infection (n=51,910, 30,115, 2,183 for ChAdOx1, BNT162b2, mRNA-1273, respectively). Estimates were separated by age (predicted at 20, 40, 60, and 80-year-old), sex, ethnicity, long-term health condition (LTHC), and vaccine type. The y-axis is truncated at 2,000 days (panel a) for visualisation. For ChAdOx1, the 20-year-old group is not plotted because the vast majority of those receiving ChAdOx1 were ≥40 years. For mRNA-1273, the 80-year-old is not plotted because the vast majority of those receiving mRNA-1273 were ≤60 years.

To provide context to the observed antibody levels in this study, in the previous section I showed that levels of 107 and 94 BAU/mL were associated with 67% protection against

new Delta infection in those without prior infection but vaccinated with ChAdOx1 and BNT162b2, respectively, compared with 33 BAU/mL in those unvaccinated with prior infection. Equivalent levels could not be estimated for those with prior infection and vaccinated due to insufficient data. However, since levels associated with the same degree of protection were lower for unvaccinated but previously infected individuals, if I conservatively assume the threshold levels are similar post any vaccination, in those with prior infection the duration providing >67% protection is estimated to be around 170-220 days for a single ChAdOX1 vaccination, and over a year for a single BNT162b2 vaccination. Since the duration providing >67% protection in unvaccinated individuals with natural infection was estimated to be 1-2 years, it is highly likely that the duration of protection is >1 year for those with prior infection receiving a single vaccination. Using similar assumptions for those with prior infection, two vaccinations would result in an estimated duration of 67% protection of 180-200 days for ChAdOX1 and 240-300 days for BNT162b2, respectively. Therefore, for people with prior infection, two vaccinations did not provide longer protection than a single vaccination. In those without prior infection, using these thresholds, a single ChAdOx1 vaccination would not reach the required antibody level, while a single BNT162b2 vaccination would provide 50-100 days of protection for people <60 years (**Figure *5.43***).

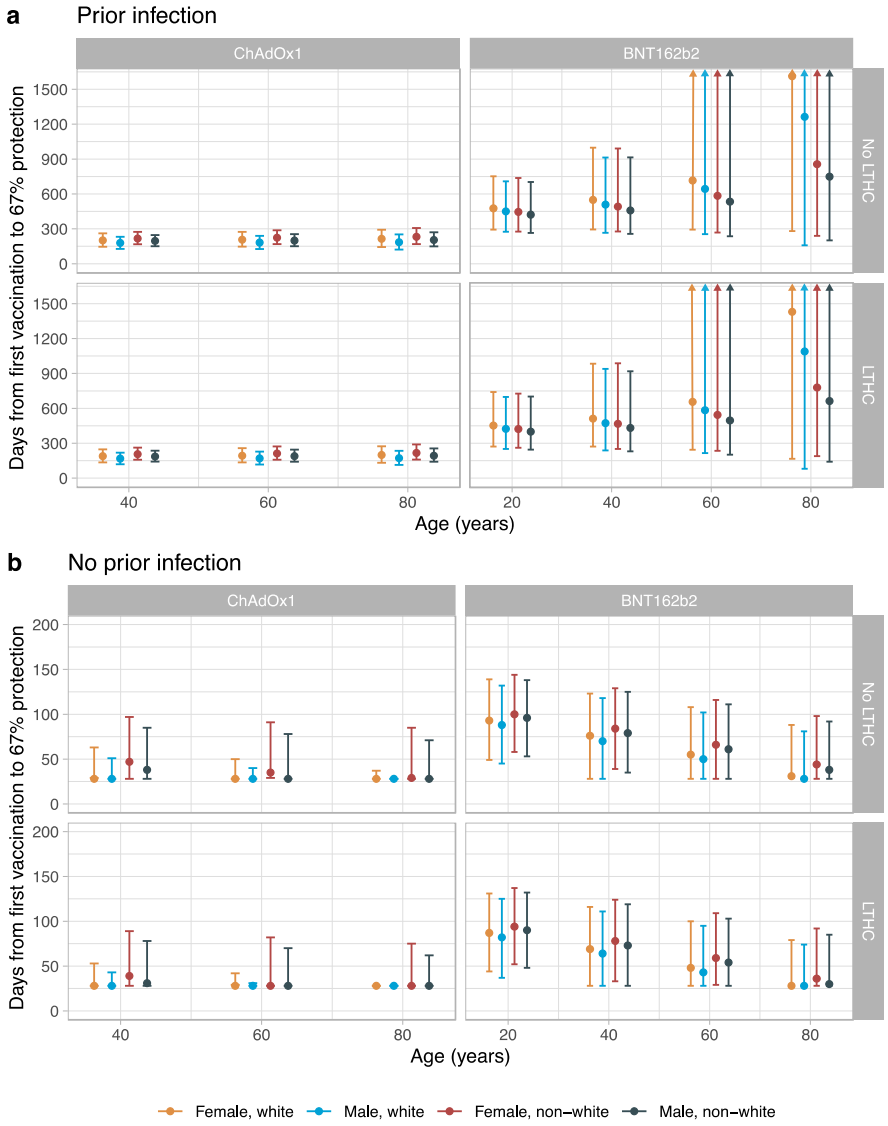


Figure 5.43 Posterior predicted mean days (95% credible interval) from the first vaccination to the threshold level associated with 67% protection (107 BAU/mL for ChAdOx1 and 94 BAU/mL for BNT162b2).

Note: **a,** in those with evidence of prior infection (n=7,191, 5,523 for ChAdOx1, BNT162b2, respectively). **b,** in those without evidence of prior infection (n=51,910, 30,115 for ChAdOx1, BNT162b2, respectively). Estimates were separated by age (predicted at 20, 40, 60, and 80-year-old), sex, ethnicity, long-term health condition (LTHC), and vaccine type. The correlates of protection were reported in my previous study on two vaccine doses[174]. For ChAdOx1, the 20-year-old group is not plotted because the vast majority of those receiving ChAdOx1 were ≥40 years. mRNA-1273 is not plotted because of insufficient data to estimate the equivalent level of protection.

5.4 Discussion

In this chapter, I examined the antibody responses and associated protection after first and second vaccinations against SARS-CoV-2, combining results from three papers. Based on these three studies using data from a large random sample of the UK population, post-vaccine anti-spike IgG responses vary by prior infection status, age, sex, vaccine type, and the number of doses received. In those who were previously infected, all age groups achieved high antibody response after the first vaccination. In those without evidence of prior infection, older participants had lower and slower responses post-first vaccine dose than younger participants. A single dose of ChAdOx1 vaccine resulted in lower absolute antibody levels and slower responses compared to a single dose of BNT162b2 vaccine. However, antibody levels after a single dose of BNT162b2 waned over time, whereas levels remained approximately constant after a single dose of ChAdOx1 over the short-term. Two vaccine doses achieved high responses across all age groups, and particularly increased the number of older people seroconverting to similar levels to those receiving one dose after prior infection. I found significant boosting of anti-spike IgG following second doses of both ChAdOx1 and BNT162b2 vaccines in all age groups and using different dosing intervals, including the three-week dosing interval for BNT162b2. Consistent with a single dose, those receiving second BNT162b2 doses had significantly higher peak anti-spike IgG responses than those receiving second ChAdOx1 doses; however, their antibody levels fell faster than following ChAdOx1. After the second vaccination, older age, male sex, and long-term health conditions were all associated with substantially lower peak levels in participants who received BNT162b2 but had smaller effects on peak levels with ChAdOx1. Participants with prior infection had significantly higher peak levels for both vaccines, and their antibody levels fell more slowly.

After the first vaccination, the relative differences in vaccine response by age and prior infection status are similar to those reported by the REACT-2 study of binary point-of-care lateral flow assay results after a single dose of BNT162b2 vaccine[179]. However, my results showed a much higher antibody response than reported in REACT-2, especially in older people, despite studying the same UK population. These differences likely reflect the lower sensitivity of the assay used in REACT-2, despite efforts to adjust for this[179]. In this study, mean quantitative responses were not far from the positivity threshold, particularly for older age groups, demonstrating the challenge of applying binary thresholds to what are essentially continuous data. My study provides additional comparative data on antibody responses following the ChAdOx1 vaccine. Studies in healthcare workers also support an inverse association between antibody response and age in those receiving a single dose of BNT162b2[180,181] or ChAdOx1 vaccine[158].

I found in those without evidence of prior infection, at older ages females had a higher probability of being IgG positive post-vaccination than males, and females were more likely to be in the 'high response' latent class. Sex differences in antibody levels have also been described following natural infection[150]. These findings are consistent with observations that females generate stronger humoral immunity and greater vaccine efficacy than males[182,183]. However, a UK study of 3,610 healthcare workers (median age 41 years) did not find any association between sex and post-single dose ChAdOx1 or BNT162b2 antibody responses,[158] possibly explained by my finding that sex differences in antibody responses become more marked above 60 years of age.

Antibody trajectories after vaccination differed substantially by prior infection status. Prior infection was associated with a significantly higher peak level and a longer half-life for both ChAdOX1 and BNT162b2. This is consistent with studies in healthcare workers

showing that prior SARS-CoV-2 infection leads to higher antibody levels and slower waning post second BNT162b2 vaccination[184–187]. Previous studies also reported that a single dose of BNT162b2 or Sputnik V, an adenovirus-based vaccine, elicited post-vaccination antibody levels that were similar to, or higher than, those without prior infection who received two doses[188–191]. From the generalised additive models, I found slightly lower IgG levels after a single BNT162b2 dose in previously infected participants vs. those with two BNT162b2 doses without prior infection, particularly at older ages. This suggests that a second BNT162b2 dose may still be helpful for previously infected individuals where supplies are sufficient, especially for older age groups. However, for ChAdOx1, the post-second dose IgG levels in those without prior infection were lower than in those previously infected with one ChAdOx1 dose.

In an adjusted Bayesian linear mixed model, longer dosing intervals between 8 and 12 weeks resulted in higher peak antibody levels for ChAdOx1 (6 BAU/mL per week) and BNT162b2 (12 BAU/mL per week), but had no effect on the half-life. Consistent with this, the 3-week dosing interval for BNT162b resulted in a lower peak level compared to an 8-week interval. Other studies also reported lower antibody levels at 2-3 weeks[165], 4 weeks[166], and 14-34 days[167] post-second BNT162b2 dose with 3-week vs extended dosing intervals. However, these studies only measured antibody levels at specific time points after the second vaccination, which may not be optimal for comparison, since I found that antibody levels were still increasing from 3-6 weeks post-second dose with the 3-week interval (**Figure *5.19***). I found a slightly longer half-life with 3-week vs. extended dosing intervals, which led to a similar duration of protection against new infections despite the lower peak levels.

Older individuals had lower post-second dose IgG peak levels following BNT162b2, but age-related differences were smaller for ChAdOx1. Females had a higher peak IgG level for both vaccines, consistent with widely reported enhanced immune responses in females[148,149,192–194]. Healthcare workers had higher IgG peak levels for both vaccines, potentially reflecting a "healthy worker" effect[195], ongoing occupational exposure, or undetected prior infection. I also found that those reporting non-white ethnicity had higher IgG peak levels. Non-white ethnicity has also been previously associated with higher antibody levels after natural infection[123,175,196]; these findings could be due to genetic or societal differences, or differential rates of undetected prior infection. Long-term health conditions were associated with lower peak levels and shorter half-lives for both vaccines. My results for BNT162b2 were consistent with a large-scale study from Israel, where older individuals, males, and those reporting health conditions had significantly lower IgG levels[197].

I estimated anti-spike IgG mean half-lives after second ChAdOx1 or BNT162b2 doses of 81 days and 52 days, respectively, in those without prior infection, and 94 days and 63 days in those vaccinated and previously infected. The half-life of BNT162b2 was consistent with estimates from previous studies at around 50 days[198–200]. Differences were also seen between the two vaccines, e.g. ChAdOx1 had a lower peak level but BNT162b2 had a shorter half-life, and factors generally had larger effects on BNT162b2 than ChAdOx1. Differences in vaccine response are expected given the differing design and mechanism of action of BNT162b2, an mRNA vaccine, and ChAdOx1, an adenovirus vector-based vaccine[201,202]. Alternative booster vaccines may reduce vaccine-specific differences[203].

After the first vaccination, the latent class analysis identified four distinct types of vaccine responses. Around ~6% of participants were low responders to both ChAdOX1 and BNT162b2 vaccines, with low responses independently associated with several long-term health conditions and more common in older participants, but comprising a similar percentage of those receiving the different vaccines. This low-responder group could be identified by a negative antibody result on assay from day 28 post-vaccine. Similar underlying latent classes were identified following single doses of the two vaccines, with different mean responses overall due to different percentages estimated to fall into the 'high' and 'medium' response classes for ChAdOx1 and BNT162b2. Further studies are also required to assess whether different degrees of response are associated with different rates of waning over time and protection against clinical outcomes. A study in 10,412 long-term care residents showed 65% and 68% protection against laboratory-confirmed SARS-CoV-2 infection 28-42 days after vaccination with the ChAdOx1 and BNT162b2 vaccines respectively, suggesting differences in antibody responses may have limited impact on outcomes, at least in the short term[204]. Similar short-term (6 week) protection against symptomatic infection, hospitalisation, and death with single doses of both vaccines was also seen in adults over 70 years in England[205]. However, non-response to a second ChAdOx1 or BNT162b2 dose was much smaller, <1%, suggesting that second doses can significantly boost an initial suboptimal response in most individuals.

Using data from all participants in the national survey, there was a 76% reduction in symptomatic SARS-CoV-2 infection following a first vaccine dose, rising to 95% after two doses, with no evidence of differences post-BNT162b2 and ChAdOx1 vaccinations[82]. A major outstanding question is the extent to which antibody titres (and other immunological measurements) are correlates of post-vaccination protection and its duration. Data from ChAdOx1 and mRNA-1273 vaccine trials showed higher levels of

binding and neutralizing antibodies at a limited number of timepoints were associated with a lower risk of infection and a higher vaccine efficacy[172,173]. Population studies also found an association between lower antibody level and higher risk of subsequent infection[206]. Correlating time-updated antibody measurements to protection from infection is important to inform the timing of boosters and other control measures. I found higher anti-spike IgG levels were associated with increased protection from infection, but that the level of protection associated with a given antibody level depended on the mechanism generating the antibodies, with natural infection resulting in lower measured antibody levels, but greater protection at a given antibody level, compared to vaccination. I did not see evidence of vaccine-specific differences, with the greater protection from BNT162b vs. ChAdOx1 explained by higher antibody levels, rather than increased protection at a given antibody level. Using 67% protection against infection as an example threshold, protection was relatively short-lived in those not previously infected and receiving ChAdOx1; mean levels fell below this threshold at 50-90 days post second dose, but protection was more prolonged following two doses of BNT162b, 160-230 days. In those without prior infection, 6 months after the second dose, only 10% of those with ChAdOx1 would maintain 67% protection, while 40-80% of those with BNT162b2 would still be above the threshold, indicating that a booster dose may be prioritised to those who had two doses of ChAdOx1, and would need to be individualised to optimise protection after two BNT162b2 doses. However, 9 months after the second dose of either vaccine, almost everyone without prior infection was below the threshold level associated with 67% protection. These results are consistent with several studies reporting waning vaccine effectiveness 3-6 months after the second BNT162b2 or ChAdOx1 vaccinations[207–211]. Estimated durations of protection were much longer following natural infection, ~1-2 years for those unvaccinated. The protection in those vaccinated with prior infection could

be even longer given they have a longer half-life; however, I was not able to quantify this due to limited data. The estimates of correlates of protection were based on Delta variant infections, and so account for the increased antibody levels needed for neutralisation of this compared with earlier variants[212,213]. However, further increases in antibody levels required for protection with other variants, such as Omicron, would substantially reduce the proportion of the population protected from infection. For example, one study found that neutralisation of Omicron was 8-fold lower than Delta following two BNT162b2 vaccinations[214]; if the relationship between antibody levels and levels of protection remained unchanged with Omicron, this would mean antibody levels after two vaccinations would not provide effective protection against Omicron infection **(Figure *5.41*)**. Nevertheless, protection against severe infection, including hospitalisation and death, is likely to last considerably longer and be potentially more robust[135]. Further long-term follow up data are essential for ongoing monitoring, as it is difficult to predict the emergence of and level of immune protection required for new variants.

Higher antibody levels post first SARS-CoV-2 vaccination in previously infected individuals have been reported[189,215–217]. Consistent with previous studies,[159,160,218] I also found that in previously infected participants, a single dose of ChAdOx1 or BNT162b2 vaccine led to high anti-spike IgG antibody positivity and quantitative levels. Where vaccine supplies are limited, this supports prioritising those without evidence of prior infection for vaccination, and in particular delaying or omitting second doses in those with robust serological evidence of prior infection in these scenarios. However, none of these previous studies estimated the trajectory of antibody response post first vaccination, overall or according to important characteristics such as age. I found that in those with prior infection, not only were antibody peak levels higher for all three vaccines compared to those without prior infection, by 200-400 BAU/mL, showing a substantial boosting

effect of prior infection, but the subsequent waning was also slower following BNT162b2 and mRNA-1273, supporting sustained protection from a single dose in previously infected individuals. The combination of prior infection with a single vaccination resulted in similar antibody levels regardless of vaccine type (**Figure *5.34***), despite single ChAdOx1 vaccination resulting in lower peak levels than BNT162b2 in those without prior infection. The higher peak levels from mRNA-1273 than BNT162b2 could be potentially explained by a higher spike protein delivery in mRNA-1273.

These results provide data that could be used to compare different approaches to global immunisation strategies, deploying limited resources in the most effective way to deliver maximum population immunity rapidly during a period when incidence remains high and vaccine access is not yet universal. Given the low percentage of vaccinated individuals and high SARS-CoV-2 seroprevalence in some settings[219–221], assuming widespread prior infection, a single vaccination could provide population-level protection as a short-term measure.

However, such a blanket approach would result in infection-naive individuals initially receiving one vaccination only, which provides suboptimal protection[170]. To reduce this risk, where prior infection is expected to be high, another option would be to stratify individual vaccination based on an affordable and rapid lateral flow immunoassay (LFIA) for antibody detection, and information on previous PCR positivity[222–226]. An extensive comparison of different fingerprick-based LFIA antibody tests reported high specificity (98.8–99.8%)[227], meaning that there would be few false positives in a high prevalence setting. Sensitivity was lower (69%-86%)[227], which would reduce efficiency, but not vaccine programme effectiveness. Although exact prices will vary by manufacturer and

may be negotiated, the cost can be as low as US$1-2 per test[228]. In comparison, a ChAdOx1 vaccine costs around $4 per dose excluding delivery and storage costs[229].

For an unvaccinated population of 10 million individuals willing to be vaccinated with a true seroprevalence of 50%, a LFIA with 99% specificity and 80% sensitivity costing $1.5 per test, and a cost of $5 per vaccination (including delivery costs and storage), an LFIA test performed at the first vaccination visit would correctly identify 4,050,000 individuals as having antibodies and eligible for a one-dose schedule – of which 50,000 would be incorrectly identified as having antibodies – the remaining 5,950,000 being invited for a second vaccination according to the agreed dosing interval. This would result in 4.05 million fewer vaccinations needed in the short-term to produce equivalent population immunity to a universal two-dose campaign. This approach delivers a cost saving of $5.25 million which could be re-invested in securing robust long-term vaccine access (for results with other settings regarding LFIA test sensitivity and specificity, population size, true antibody prevalence, costs per vaccination and LFIA, see https://herc.shinyapps.io/Serology_vaccine_prioritisation/). While this study provides some support for modelling studies that suggest that prioritisation based on seropositivity may substantially improve the efficiency of vaccine allocation where seroprevalence is high, basing vaccination policies on studies combining antibody trajectories with correlates of protection against infection and modelling studies alone is challenging.

By January 2022, data are still awaited to determine the impact of the Omicron variant, and the relationship between prior infection/vaccination and immunological protection from this and other new variants. Given this, all approaches to vaccine scheduling will need to remain under intense scrutiny, and the international focus must remain firmly on assuring equitable access to full vaccination in all population settings, which will also

mitigate the potential for further variants of concern to emerge. The antibody responses and correlates of protection against Omicron infections are investigated in the next Chapter.

Study strengths include the use of fully adjusted Bayesian (log)linear mixed interval-censored models, which account for antibody measurements above the limit of quantification (450 BAU/mL), to directly estimate changes in antibody levels after the first or second vaccination adjusting for multiple characteristics. This method is able to capture any waning above the threshold, exploiting information on the percentage of measurements at or above the threshold over time, making it possible to estimate a higher peak level than the upper limit and a more accurate half-life. I used simulation studies to demonstrate that the models were robust to the assay limit and could produce accurate estimates of antibody peak levels and half-lives (simulations including interval censoring at similar levels to my dataset showed good coverage and no/negligible bias, **Table *5.19***).

Study limitations include the fact that the study did not measure other immune responses, including T cell or innate immune responses which are also involved in protection against infection;[201] they may explain the greater protection afforded by natural infection than vaccination at the same antibody level. Anti-spike IgG antibody levels were only measured using a single assay; however, I calibrated antibody levels to WHO BAU/mL units for comparison with other studies. Anti-spike IgG antibody levels are only an indirect proxy for vaccine efficacy; however, assay results have been shown to correlate closely with neutralising activity and vaccine protection (correlates of protection), thus could be used to inform vaccine policies. Neutralising antibodies were not measured, but neutralisation titres were strongly correlated with anti-spike IgG.

Another limitation was I assumed that associations between protection and antibody levels were constant over time, to predict vaccine effectiveness from antibody levels beyond the study follow-up. Even though antibody levels decrease over time, other vaccine-induced immune mechanisms, including memory responses, may theoretically provide greater protection at lower antibody levels. There may also be differences in antibody levels and classes in peripheral blood compared to potential exposure sites such as the respiratory tract. Hence, the estimated duration of protection could potentially be underestimated if other immune responses confer longer protection. However, observations based on waning antibodies are consistent with the waning of vaccine effectiveness in this cohort[170] and other studies[207–211].

In those with a single vaccination and prior infection, older participants had slower antibody declines, which could be due to the study design, where only those who seroconverted were included, as only these participants could have their antibody declines analysed. Older participants who seroconverted may have had more sustained immune responses than younger participants overall. I did not have enough data to model antibody response in older participants who received mRNA-1273, as 98% of participants who received mRNA-1273 were <50 years. Most of the participants reported white ethnicity (92.6%), so wider generalizability to non-white ethnic groups is less well defined and the data were insufficient to model other ethnic groups separately. However, estimated durations of protection for non-white participants were broadly as long or longer than for white participants, with non-white ethnicity associated with higher peak levels after a single ChAdOX1 or BNT162b2 vaccination.

In summary, in this population-representative study of individuals in the UK, vaccination results in detectable SARS-CoV-2 anti-spike IgG in the majority of individuals after the

first vaccination. The second ChAdOx1 or BNT162b2 dose significantly boosts anti-spike IgG levels, and the dosing interval has a limited impact on antibody response. Peak levels were higher after BNT162b than ChAdOx1, but older individuals, males, and those with long-term health conditions have lower antibody levels with BNT162b2. Antibody levels remain above the threshold level associated with 67% protection for 2-3 months for ChAdOx1 and 5-8 months for BNT162b2 in those without prior infection, and for 1-2 years for those unvaccinated who seroconvert after natural infection. Those vaccinated with prior infection could be protected for >1 year. These results may inform vaccination strategies; a third boosting dose should be prioritised for ChAdOx1 recipients, groups with faster antibody declines, and more clinically vulnerable individuals.

Prior infection also significantly boosts antibody responses after a single ChAdOx1, BNT162b2, or mRNA-1273 vaccination, producing higher peak levels and/or longer half-lives, comparable to, or even better than, those obtained from two vaccinations in those without prior infection. Based on the positivity threshold and reported correlates of protection, those with prior infection could be protected from infection for >1 year after a single vaccination. While studies show that a third vaccination boosts antibody responses[230] and provides better protection against infection than two vaccinations[231], and two vaccinations plus prior infection provides better protection against re-infection than prior infection alone[170], a large part of the global population has not yet had the chance to get their first vaccination due to limited vaccine supplies. These results could stimulate further studies assessing the impact of a one-dose strategy for those with prior infection, together potentially providing an evidence base to optimise population immunity in the context of resource limitations, while international SARS-CoV-2 immunisation programmes are scaled up and secured.

Table 5.19 Simulation results to demonstrate the Bayesian interval-censored (log)linear mixed model is robust to the upper quantification limit.

Note: Three scenarios were modelled: negative correlation, no correlation, and positive correlation between intercept and slope. Within each scenario, two parameter sets were chosen to mimic the different proportion of censored measurements (the proportion of censored data was 11% for ChAdOx1, 19% for BNT162b2, and 38% for mRNA-1273). 1000 samples were used each model and 100 repetitions of each simulation were performed. Simulations results showed good coverage and no/negligible bias, indicating the model is robust to the upper limit of detection of the assay.

Scenario	Intercept	Slope	Corre-lation	Percentage censored (%)	Coverage in intercept	Bias in intercept (95%CI)	Coverage in slope	Bias in slope (95%CI)	Coverage in correlation	Bias in correlation (95%CI)
1	8.5	-0.01	0	13-18	94%	0 (-0.02, 0.02)	94%	0.0001 (-0.0004, 0.0006)	94%	0 (-0.06, 0.08)
2	9.0	-0.01	0	38-44	96%	0 (-0.03, 0.04)	94%	0 (-0.0008, 0.001)	93%	0 (-0.08, 0.1)
3	8.8	-0.02	-0.2	12-16	97%	0 (-0.02, 0.02)	96%	0 (-0.0005, 0.0005)	97%	0 (-0.05, 0.06)
4	9.0	-0.01	-0.5	33-43	95%	0 (-0.03, 0.04)	95%	-0.0001 (-0.0006, 0.001)	98%	0 (-0.06, 0.05)
5	8.8	-0.02	0.5	17-22	93%	0 (-0.02, 0.03)	95%	0 (-0.0004, 0.0005)	96%	-0.01 (-0.05, 0.04)
6	9.0	-0.01	0.2	39-45	94%	0 (-0.02, 0.03)	96%	-0.0001 (-0.0006, 0.0006)	95%	0.01 (-0.06, 0.05)

Chapter 6 Correlates of protection against SARS-CoV-2 Omicron variant and antibody responses after booster vaccination or breakthrough infection

This chapter is based on a previous publication: *Wei, J., Matthews, P. C., Stoesser, N. et al. Protection against SARS-CoV-2 Omicron BA.4/5 variant following booster vaccination or breakthrough infection in the UK. Nat Commun (2023).*

All data analysed and literature discussed in this chapter were up to September 2022, when the Omicron BA.4/5 variant was dominant in the UK.

6.1 Introduction

Multiple SARS-CoV-2 vaccines have been developed and have been highly effective at reducing infections[202] and associated hospitalisation and death[232–234]. However, waning of vaccine-induced immunity means optimal protection from vaccination may be relatively short-lived, with reduced effectiveness 3-6 months after the second vaccinations[174,235,236] leading to widespread use of booster vaccinations. Reductions in vaccine effectiveness with time have been exacerbated by changes in circulating variants, with lower levels of protection against Delta versus Alpha and further reductions against different Omicron variants[237]. However, alongside this, large numbers of "breakthrough" infections (i.e. natural infection in the context of previous vaccination) mean that increasing numbers have some existing immunity from earlier infections.

In those who have received a primary vaccination course (typically two doses), understanding the relative extent of protection against further infection from booster vaccination has implications for vaccine policy. One response to waning vaccine-induced immunity is repeated vaccination of entire populations. COVID-19 vaccination

programmes targeting entire (adult) populations were estimated to be cost-saving when first and second vaccinations were introduced[238], despite being financially and logistically resource intensive and often requiring diversion of other healthcare resources to deliver. However, with increasing proportions of the population having at least some level of immunity due to previous vaccinations and infections, combined with lower risks of severe outcomes from more recent SARS-CoV-2 variants[239,240], the potential reduced benefits and on-going high opportunity costs of vaccinating entire populations repeatedly should be carefully considered.

In contrast to vaccination, previous infection may offer longer-lasting protection[241]. Therefore, for low-risk populations, if the chance of harm from infection following initial vaccination is sufficiently small, frequently repeated vaccination paid from healthcare budgets may not be required and could potentially even generate harm considering the opportunity costs of not being able to spend this budget on other interventions that result in more gains in quality-adjusted life-years. For example, fourth (or fifth) vaccinations in the last quarter of 2022 were offered only for those aged 50y or older (fifth for 75y or older or clinically vulnerable) in the UK, meaning natural infection will become the main immunological boosting mechanism for younger adults and children. However, natural infection could also bring risks such as exposure of vulnerable populations, complications including long COVID even in low-risk populations, and economic consequences to society.

Whilst it is not yet possible to assess the impact of these fourth/fifth vaccinations, the substantial expansion of third/booster mRNA vaccinations from 16 September 2021 in the UK, in parallel with large numbers of breakthrough SARS-CoV-2 infections among those who had not yet received a third/booster vaccination, particularly with the emergence of

Omicron variants from mid-November 2021, provides an opportunity to compare their impact on antibody responses.

Therefore in this chapter, I aim to investigate the duration of anti-trimeric spike IgG antibody responses following "breakthrough" infection vs third vaccination in those who had previously received two vaccinations but without evidence of prior infection. Antibody levels are correlated with protection against infection in previous studies[135,242–244], however, there is little information about antibody correlates of protection against Omicron variants. I therefore also aim to determine antibody-based correlates of protection against Omicron infections from 17 May 2022 onwards (predominantly BA.4/BA.5 lineages, refer to as 'BA.4/5 epoch') and use these to estimate how long infection is likely to be prevented after a third/booster vaccination versus breakthrough infection.

6.2 Methods

Raw data were processed using Stata MP 17. All analyses were performed in R 3.6 using the following packages: tidyverse (version 1.3.0), mgcv (version 1.8-31), cenGAM (version 0.5.3), brms (version 2.14.0), splines (version 3.6.1), ggeffects (version 0.14.3), arsenal (version 3.4.0), cowplot (version 1.1.0), bayesplot (version 1.7.2).

Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and laboratory tests were already described in Chapter 3.

6.2.1 Correlates of protection analysis

For the analysis of correlates of protection for new Omicron variants, I used data from study assessments with antibody measured in the preceding 21-59 days from 17 May 2022 to 12 September 2022 inclusive due to the change of dilution on 29 April 2022 (therefore

only including antibody measurements with 1:1600 dilution). The outcome was therefore mainly Omicron BA.4/5 infections (see Results). I included participants aged 18y and over, due to different recommendations regarding booster and primary vaccinations in younger participants. Analyses were based on tests conducted at study assessments. The outcomes of the model were results of PCR tests of nose and throat swabs taken at each assessment conducted as part of the study independently of symptoms and other characteristics. To define prior infection, I grouped positive tests from the study, plus additional positive swab tests identified from the English national testing programme (national testing data were not available for Scotland, Wales, and Northern Ireland), and from self-reported positive swab tests in all participants, into episodes[245] because PCR-positive results might be observed at multiple assessments after infection, and included only the first new PCR-positive test from a study assessment in each infection episode as the outcome (all episodes considered as exposures, as were positive anti-spike IgG result (≥23 BAU/mL) any time before the first vaccination). Assessments occurring in the 21 days before and after each vaccination were excluded, as infection rates change in the run-up to vaccination, reflecting structural bias (whereby those known to be positive defer planned vaccinations)[82]. I further included a positive anti-spike IgG result (≥23 BAU/mL) any time before the first vaccination or a self-reported positive test as a prior infection.

I used separate logistic generalised additive models (GAMs) for three outcomes: any positive PCR study test; a positive PCR study test with a moderate to high viral load (Ct value <30); and a positive study PCR test with self-reported symptoms. As previously described[174], I considered the effect of the most recent antibody measurement obtained 21–59 days before the current assessment. I excluded more recent measurements to avoid changes in antibody levels arising from recent infection that might be detected only at the routine study assessment despite occurring before this. Furthermore, I excluded

assessments where antibody measurements were longer ago as these older values may not correlate well with the antibody levels shortly before the assessments at which the outcome is being assessed. The relationship between antibody levels and the outcome was modelled using thin-plate splines. Based on vaccination and infection history, I divided the population into vaccinated participants without evidence of previous infection (62,146 participants, 106,653 assessments) and vaccinated participants with evidence of previous infection (58,373 participants, 98,036 assessments). Vaccinated participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together, assuming that the effect of vaccination number was mediated through the resulting antibody levels. Unvaccinated participants were excluded due to insufficient data (1,151 participants, 1,840 assessments). Vaccinated participants with a previous infection were further split by dominant variant at the earliest positive test in each infection episode as: Pre-Alpha (up to 6 December 2020, 3,490 participants, 5,945 assessments), Alpha (07 December 2020-16 May 2021, 4,199 participants, 7,064 assessments), Delta (17 May 2021-12 December 2021, 9,225 participants, 15,745 assessments), Omicron BA.1 (13 December 2021-20 February 2022, 15,615 participants, 26,736 assessments), and Omicron BA.2 (21 February 2022-05 June 2022, 22,079 participants, 38,223 assessments) where dates were chosen as the first surveillance week (starting Monday) where >50% of positive tests matched the S-gene of the new variant (S-negative for Alpha, BA.1/4/5; S-positive for pre-Alpha, Delta and BA.2). Participants tested PCR-positive at 5,680 (5.3%) of assessments without known previous infection and at 215 (3.6%), 258 (3.7%), 465 (3.0%), 529 (2.0%), and 201 (0.5%) assessments after Pre-Alpha, Alpha, Delta, Omicron BA.1, or Omicron BA.2 infections respectively (**Table *6.1*).** I found that estimates from previous Pre-Alpha and Alpha infections, were similar, so I combined these assessments to increase power (**Table *6.1***,

Figure *6.1*). The final model therefore consisted of five groups: vaccinated participants without evidence of infection before the current assessment, vaccinated participants with a most recent Pre-Alpha or Alpha infection, vaccinated participants with a most recent Delta infection, vaccinated participants with a most recent Omicron BA.1 infection, and vaccinated participants with a most recent Omicron BA.2 infection. In the absence of an unvaccinated reference group, I used vaccinated participants without previous infection with a most recent anti-spike IgG measurement of 16 BAU/mL as the reference group, and estimated the relative protection compared with this reference group. 16 BAU/mL was chosen because it was the threshold previously identified to define non-responders to vaccination[174].

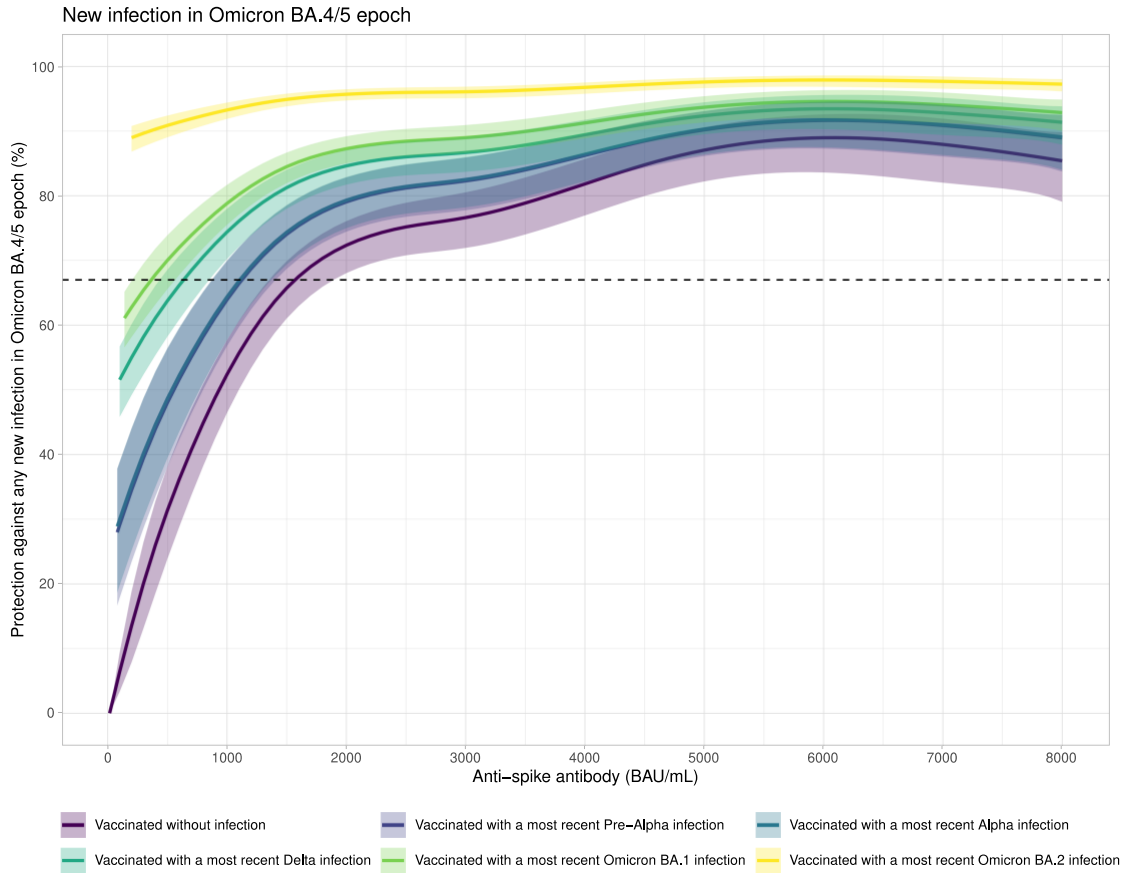


Figure 6.1 Association between anti-spike IgG levels and protection from new SARS-CoV-2 infection in the Omicron BA.4/5 epoch using the most recent antibody measurement obtained 21–59 days before the current assessment by infection variant.

Note: The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. Participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together. Protection is defined as relative protection against a baseline protection afforded by 16 BAU/mL in those vaccinated without infection, which is the threshold for vaccine non-responders. Antibody measurements were plotted after the 1% percentile in each group (16, 80, 80, 100, 140, 200 BAU/mL, respectively). From the plot, estimates for Pre-Alpha and Alpha infection are similar (lines overlapped). Therefore, Pre-Alpha and Alpha infection are combined in the main model to increase power.

Table 6.1 Proportion of study assessments with positive PCR results in the correlates of protection analysis.

Note: Three outcomes are examined: all infection in the Omicron BA.4/5 epoch; infections with a moderate to high viral load (Ct<30) in the Omicron BA.4/5 epoch; infections with a self-reported symptom in the Omicron BA.4/5 epoch. In the main model, pre-Alpha and Alpha infections were combined because the estimates were similar.

		Test-negative assessments	Test-positive assessments	%	Combined test-negative assessments	Combined test-positive assessments	%
All infection	Vaccinated without infection	100,973	5,680	5.3			
	Vaccinated with a most recent Pre-Alpha infection	5,730	215	3.6	12,536	473	3.6
	Vaccinated with a most recent Alpha infection	6,806	258	3.7			
	Vaccinated with a most recent Delta infection	15,280	465	3.0			
	Vaccinated with a most recent Omicron BA.1 infection	26,207	529	2.0			
	Vaccinated with a most recent Omicron BA.2 infection	38,022	201	0.5			
Ct<30 infection	Vaccinated without infection	102,484	4,169	3.9			
	Vaccinated with a most recent Pre-Alpha infection	5,790	155	2.6	12,681	328	2.5
	Vaccinated with a most recent Alpha infection	6,891	173	2.4			
	Vaccinated with a most recent Delta infection	15,430	315	2.0			
	Vaccinated with a most recent Omicron BA.1 infection	26,381	355	1.3			
	Vaccinated with a most recent Omicron BA.2 infection	38,098	125	0.3			
Symptomatic infection	Vaccinated without infection	104,780	1,873	1.8			
	Vaccinated with a most recent Pre-Alpha infection	5,876	69	1.2	12,864	145	1.1
	Vaccinated with a most recent Alpha infection	6,988	76	1.1			
	Vaccinated with a most recent Delta infection	15,620	125	0.8			
	Vaccinated with a most recent Omicron BA.1 infection	26,604	132	0.5			
	Vaccinated with a most recent Omicron BA.2 infection	38,161	62	0.2			

I adjusted for the following confounders in all models: vaccination and infection history (as above), time from last vaccination or infection, age in years, geographic area (9 regions in England, or the devolved administrations Wales, Scotland or Northern Ireland), rural/urban classification of home address, sex, ethnicity (white versus non-white), household size, multi-generational household, deprivation, presence of long-term health conditions, working in a care home, having a patient-facing role in health or social care, direct or indirect contact with a hospital or care home, and smoking status. Calendar time and age were included using a tensor product spline[246], which was allowed to vary by region/country. Tensor product splines are used to model interactions of multiple covariates measured in different units. I also included tensor product spline interactions between antibody levels and age, and antibody levels and time since the last vaccination or infection, to examine the effects of age and time since the last event on the relationship between antibody and protection.

6.2.2 Antibody trajectory analysis

For the analysis of antibody trajectories, I included participants aged 18y and over who were eligible for third/booster vaccinations and had no evidence of previous infection. From 8th December 2020 to 12th September 2022, 259,561 participants received two vaccinations and had antibody measurements after the second vaccination. Of these, 32,810 participants were infected before the second vaccination and were excluded. 130,529 and 87,604 participants received two ChAdOx1 or BNT162b2 vaccinations for their primary course; 4,213 who received other vaccine types or mixed vaccine types were excluded. Among those who received two ChAdOx1 or BNT162b2 vaccinations, 194,679 received a third/booster vaccination (150,510 were boosted by BNT162b2, 43,257 were boosted by mRNA-1273). 912 received other vaccine types as a booster and were also excluded, and 25,544 had breakthrough infection post-second vaccination.

SARS-CoV-2 infection was defined as for the correlates analysis, namely as the earliest of a PCR-positive swab test in the study, a positive swab test in the English national testing programme (national testing data were not available for Scotland, Wales, and Northern Ireland), a self-reported positive swab test or a positive anti-spike IgG result (≥23 BAU/mL) any time before the first vaccination. Antibody measurements after a third/booster vaccination that happened after a post-second vaccination infection, and antibody measurements after infection that happened post-booster were excluded from the analyses. I limited the analyses to those whose primary vaccination course was homologous ChAdOx1 (with a dosing interval 6-13 weeks) or BNT162b2 (with a dosing interval 3-13 weeks). I excluded a small number of participants who were non-responders after the second vaccination, defined previously[174] as all antibody measurements being <16 BAU/mL after the second vaccination and having at least one antibody measurement 21 days after the second vaccination (N=416 excluded for ChAdOx1, N=174 excluded for BNT162b2). Age was truncated at 85y in all analyses to reduce the influence of outliers (1% of the population).

I used Bayesian linear mixed interval-censored models to estimate antibody levels and the effects of covariates on changes in antibody levels over time. I allowed the outcome to be interval censored in the model because anti-spike IgG measurements were censored at the upper limit of quantification. Models were built separately by primary vaccine course (ChAdOx1 or BNT162b2) and booster type (BNT162b2 or mRNA-1273) or infection. Analyses of primary ChAdOx1 vaccinations included 64,940, 21,960, and 10,830 participants boosted by BNT162b2, mRNA-1273 and infection, respectively. Analyses of primary BNT162b2 vaccinations included 44,197, 7,248, and 4,974 participants boosted by BNT162b2, mRNA-1273, and infection, respectively. Antibody trajectory consisted of three slopes: 1) the decline (waning) from 21 days post-second vaccination to the

third/booster vaccination or infection at t=0; 2) the increase from the third/booster vaccination or infection to the peak; 3) the decline (waning) after the peak post-booster or infection. Because there was variation in the time taken to reach peak antibody levels following different boosters/infection and in different study participants, the choice of the peak position could influence the estimate of the half-life post third/booster vaccination or infection, I separated the trajectories and built two models for each group: 1) a piecewise model: from 21 days post-second vaccination to 14 days post-third/booster vaccination or infection; 2) a decline model: from 42 days to 210 days post third/booster vaccination or infection to ensure that antibodies were waning in almost all participants from this timepoint onwards and thus better estimate the antibody decline (observed data shown in **Figure *6.2*, Figure *6.3***). I excluded measurement >210 days after the boosting event to ensure comparability of estimated declines post boosting event across groups; beyond 210 days there were insufficient numbers of measurements and/or measurements may have come from atypical participants who were boosted/infected much earlier (e.g. clinically vulnerable/at high risk of infection) that could potentially lead to bias.

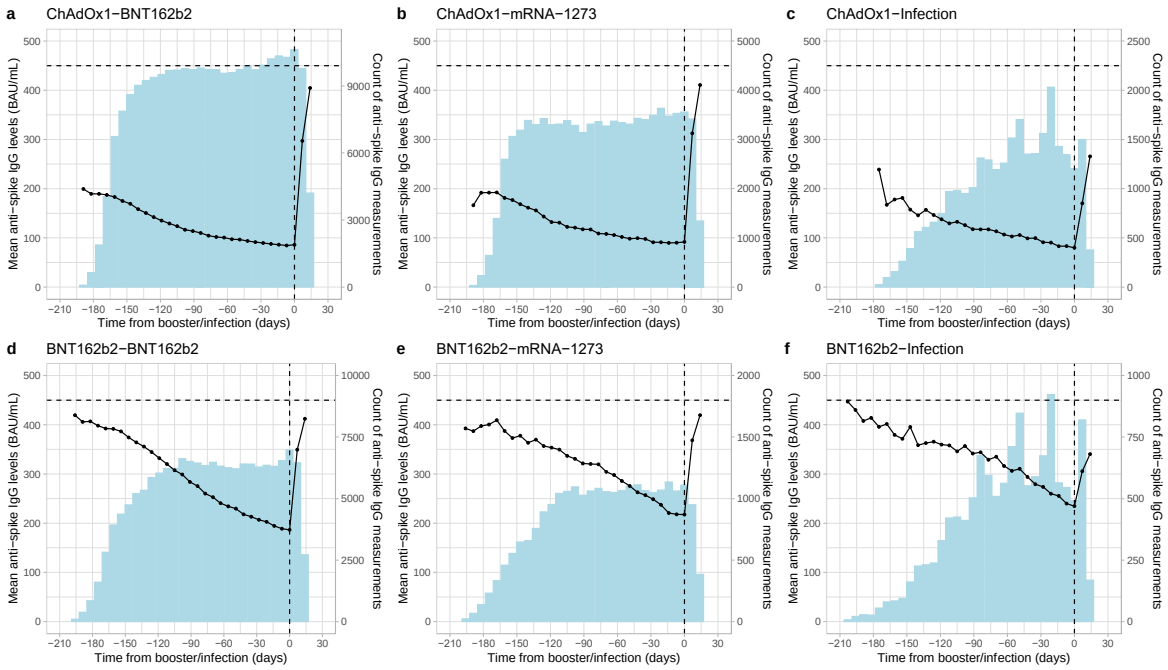


Figure 6.2 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day from 21 days after the second vaccination to 14 days after the booster/infection.

Note: Panels are separated by primary vaccination course (ChAdOx1, BNT162b2) and boosting event (BNT162b2, mRNA-1273, breakthrough infection). Time 0 indicates the date participant receiving a booster vaccination or the first date of a breakthrough infection. Horizontal dotted line indicates the upper limit of quantification in 1:50 dilution, which is 450 BAU/mL. Measurements above the upper limit of quantification of 450 BAU/mL are considered to be >450 BAU/mL in analyses.

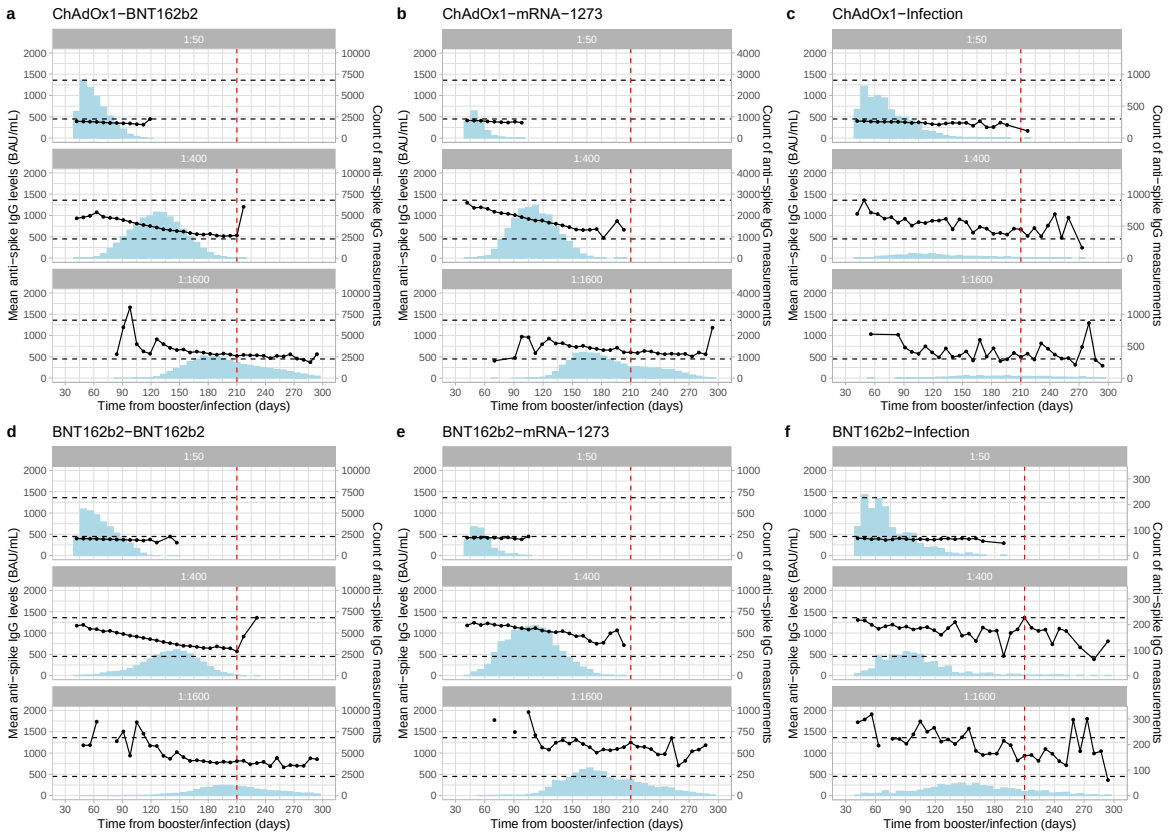


Figure 6.3 Mean anti-spike IgG levels (BAU/mL) and count of anti-spike IgG measurements by day from 42 days after the booster/infection for antibody decline modelling.

Note: Panels are separated by primary vaccination course (ChAdOx1, BNT162b2) and boosting event (BNT162b2, mRNA-1273, breakthrough infection). Measurements in different dilutions (1:50, 1:400, 1:1600) are plotted separately. Horizontal black dotted lines indicate the upper limits of quantification in 1:50 and 1:400 dilution, which are 450 and 1360 BAU/mL. Measurements above the upper limit of quantification of 450 BAU/mL in 1:50 dilution are considered to be >450 BAU/mL in analyses, and similarly for 1:400 and 1:1600 dilutions. I included data until 210 days after the booster/infection to ensure the same endpoint across groups (red dotted lines).

Population-level fixed effects, individual-level random effects for intercept and slopes, and correlation between random effects were included in all models. The outcome was modelled on the log2 scale and right-censored at 450, 1,360, and 8,000 BAU/mL for measurements obtained using a 1:50, 1:400, and 1:1600 dilution, respectively, reflecting interval censoring of IgG values at the upper limit of quantification (i.e. all measurements above the upper limit of quantification of 450 BAU/mL in 1:50 dilution were considered to be >450 BAU/mL in analyses, and similarly for 1:400 and 1:1600 dilutions).

To examine non-linearity in antibody declines after third/booster vaccination or infection, especially the assumption that the rate of antibody decline would flatten, I additionally fitted a model using three-knot splines for time (knots placed at 10th, 50th and 90th of included time points) and compared with the linear models for each group. For all six groups, the estimated trajectories were similar, so I retained the log-linear model for analysis (**Figure *6.4***).

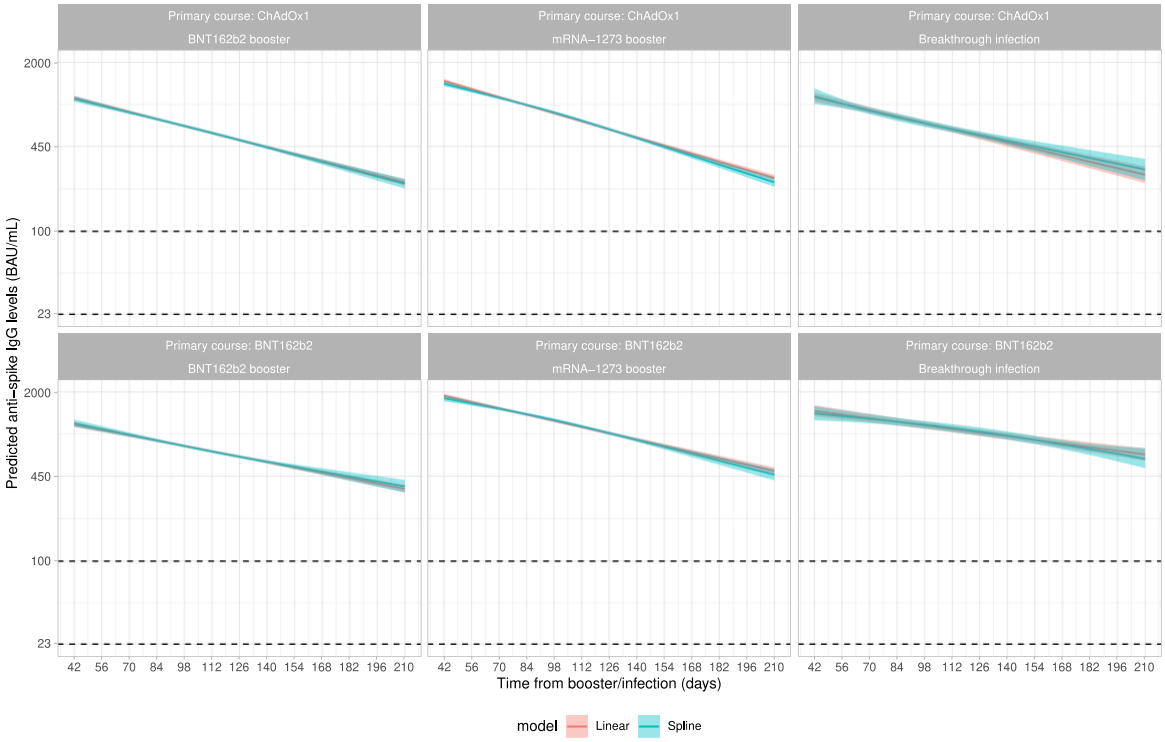


Figure 6.4 Comparison of linear exponential model with spline-based model in examining non-linearity of antibody decline.

Note: The estimated trajectory from the spline model (with 3 knots placed at 10th, 50th, and 90th of observed time points) is similar to the linear exponential model for all six groups, indicating that there was no evidence of antibody decline flattening over the period of observation (up to 210 days from booster/infection).

Multivariable models included the effect of age, time from the second vaccination to the third/booster vaccination or infection, sex, ethnicity (white vs non-white due to small numbers in the latter), reporting having a long-term health condition, and reporting working in healthcare, each on the intercept (main effect) and on the slope (interaction). I

used a 3-knot natural cubic spline for age (knots placed at 10th, 50th, and 90th percentile of unique integer ages) to allow non-linear effects with antibody rising and waning post third/booster vaccination or infection. I restricted the range of time from the second vaccination to the third/booster vaccination or infection between the 10th and 90th percentiles (**Table *6.2***). For infection models, I also examined the impact of infection type (Delta or Omicron BA.1) on antibody levels post-infection. For this analysis, infection type was defined primarily by sequencing data from CIS. If sequencing data was not available, Delta was defined as the first infection date in the most recent infection episode occurring from 10 April 2021 to 16 January 2022 and having at least one S-gene positive test (ORF1ab+N+S or ORF1ab+S or N+S) across the infection episode, Omicron BA.1 was defined as the first infection date in the most recent episode being after 29th November 2021 and not having any S-gene positive tests. If gene positivity was not available (primarily infections from the national testing programme), infections that happened between 14th June 2021 and 23rd November 2021 were considered as Delta and those that happened after 20th December 2021 were considered as Omicron BA.1. Infections with other variants or unknown variants were excluded from this model.

Table 6.2 Settings for each Bayesian linear mixed model.

Note: For each group, two separate models are fitted: 1) piecewise model on antibody decline after the second vaccination and subsequent increase after third/booster vaccination or infection; 2) antibody decline 42 days after the third/booster vaccination or infection. Only participants with a time between second vaccination to booster/infection from the 10th to 90th percentile were included in the models to avoid outlier influence. Age splines were included in the models to account for non-linearity in the association between age and antibody response.

	Median duration (days) between second vaccination and booster/infection (IQR)[range]	**Selected duration (days) between second vaccination and booster/infection (10^{th}+90^{th} percentile)**	**Median age (years) (IQR)[range] after using selected duration range**	**Age spline (years) (10^{th},50^{th},90^{th} percentile)**	**Number of participants (Total/Decline model)**
ChAdOx1-BNT162b2	189 (183-197) [40-439]	170-210	63 (53-70) [17-85]	45, 60, 75	64940/41165
ChAdOx1-mRNA-1273	187 (182-197) [72-437]	170-210	56 (49-62) [17-85]	45, 55, 65	21960/14748
ChAdOx1-infection	143 (98-176) [1-440]	60-200	50 (43-59) [17-85]	40, 50, 70	10830/4214
BNT162b2-BNT162b2	191 (184-202) [46-475]	150-220	66 (52-74) [16-85]	40, 65, 80	44197/24767
BNT162b2-mRNA-1273	180 (147-201) [67-473]	130-220	41 (34-60) [17-85]	30, 40, 70	7248/4408
BNT162b2-infection	140 (108-177) [1-483]	70-220	39 (32-55) [16-85]	30, 40, 70	4974/1857

For each Bayesian linear mixed model, weakly informative priors were used (**Table *6.3***). Markov chain Monte Carlo (MCMC) algorithms were used for posterior sampling. Six chains were run per model with 4,000 iterations and a warm-up period of 2,000 iterations to ensure convergence, which was confirmed visually and by ensuring the Gelman-Rubin statistic was <1.05 (**Table 6.4**). 95% credible intervals were calculated using the highest posterior density intervals.

Table 6.3 Priors used in the Bayesian linear mixed interval-censored models on estimating antibody decline 42 days after third/booster vaccination or infection.

Results remain similar when using weaker priors normal(0,2.5) for all coefficients.

Model term	Priors
Intercept	normal (10, 2)
Slope	normal (0, 0.5)
Coefficient for change in intercept (age)	normal (0, 1)
Coefficient for change in slope (age)	normal (0, 0.1)
Coefficient for change in intercept (duration)	normal (0, 1)
Coefficient for change in slope (duration)	normal (0, 0.1)
Coefficient for change in intercept (sex)	normal (0, 1)
Coefficient for change in slope (sex)	normal (0, 0.1)
Coefficient for change in intercept (ethnicity)	normal (0, 1)
Coefficient for change in slope (ethnicity)	normal (0, 0.1)
Coefficient for change in intercept (long-term health condition)	normal (0, 1)
Coefficient for change in slope (long-term health condition)	normal (0, 0.1)
Coefficient for change in intercept (healthcare worker)	normal (0, 1)
Coefficient for change in slope (healthcare worker)	normal (0, 0.1)
Random effect SD: intercept	normal (0, 1)
Random effect SD: slope	normal (0, 0.1)
Random effect intercept & slope covariance	lkj_corr_cholesky(1)

Table 6.4 Model coefficients and MCMC diagnostics for the Bayesian linear mixed models estimating antibody decline 42 days after the third/booster vaccination or infection.

	ChAdOx1 primary course				**BNT162b2 primary course**			
BNT162b2	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**
Population-level effects								
Intercept	10.1833	10.0771	10.2395	1.00	10.3093	10.2538	10.3647	1.01
T3	-0.0848	-0.0899	-0.0740	1.02	-0.0741	-0.0793	-0.0690	1.01
age1	-0.2053	-0.2760	-0.1368	1.00	-0.4625	-0.5886	-0.3377	1.00
age2	-0.4118	-0.4849	-0.3213	1.00	-0.2543	-0.3820	-0.1259	1.00
Male1	-0.3397	-0.3863	-0.2692	1.00	0.0048	-0.0551	0.0639	1.01
ethnicity1	-0.0139	-0.1193	0.0964	1.01	0.1252	-0.0204	0.2729	1.00
lthc1	-0.0454	-0.0963	0.0024	1.01	-0.0357	-0.0977	0.0272	1.00
hcw1	-0.0984	-0.3189	0.1162	1.00	-0.4434	-0.6552	-0.2210	1.03
dur30	0.2562	0.1458	0.3345	1.01	0.1283	0.0651	0.1908	1.00
T3:age1	0.0066	-0.0001	0.0131	1.01	0.0165	0.0054	0.0279	1.00
T3:age2	0.0269	0.0192	0.0330	1.00	0.0174	0.0069	0.0280	1.00
T3:Male1	0.0022	-0.0052	0.0065	1.00	-0.0071	-0.0124	-0.0019	1.01
T3:ethnicity1	0.0066	-0.0035	0.0171	1.01	0.0095	-0.0036	0.0228	1.00
T3:lthc1	-0.0025	-0.0067	0.0028	1.00	-0.0022	-0.0077	0.0033	1.00
T3:hcw1	0.0034	-0.0163	0.0216	1.01	0.0290	0.0112	0.0461	1.03
T3:dur30	-0.0085	-0.0155	0.0033	1.01	0.0030	-0.0027	0.0087	1.00

	ChAdOx1 primary course				BNT162b2 primary course			
Group-level effects								
sd(Intercept)	0.9699	0.0000	1.1859	1.01	1.0346	0.9935	1.0760	1.01
sd(T3)	0.0767	0.0679	0.1092	1.01	0.0782	0.0742	0.0823	1.02
cor(Intercept,T3)	-0.3021	-0.4485	-0.2519	1.02	-0.3572	-0.4079	-0.3039	1.01
sigma	0.5443	0.5039	0.7175	1.01	0.5205	0.5103	0.5307	1.01
mRNA-1273	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**
Population-level effects								
Intercept	10.5452	10.4897	10.6000	1.00	10.7881	10.6642	10.9161	1.00
T3	-0.1055	-0.1106	-0.1005	1.00	-0.0696	-0.0830	-0.0561	1.00
age1	0.0611	-0.0169	0.1383	1.00	-0.2368	-0.5039	0.0222	1.00
age2	-0.0279	-0.0898	0.0355	1.01	0.2146	0.0129	0.4160	1.00
Male1	-0.2279	-0.2794	-0.1739	1.00	0.2616	0.1636	0.3587	1.00
ethnicity1	-0.0165	-0.1629	0.1272	1.00	0.1780	-0.0658	0.4158	1.00
lthc1	0.0097	-0.0568	0.0758	1.00	0.0375	-0.0866	0.1630	1.00
hcw1	-0.1667	-0.4833	0.1602	1.00	0.3072	-0.2513	0.8681	1.00
dur30	0.0629	-0.0324	0.1584	1.00	0.0344	-0.0499	0.1184	1.00
T3:age1	-0.0088	-0.0160	-0.0015	1.00	-0.0077	-0.0356	0.0208	1.00
T3:age2	0.0017	-0.0040	0.0072	1.00	-0.0176	-0.0371	0.0021	1.00
T3:Male1	0.0114	0.0065	0.0162	1.00	-0.0043	-0.0143	0.0059	1.00
T3:ethnicity1	0.0020	-0.0113	0.0157	1.00	-0.0020	-0.0274	0.0235	1.00
T3:lthc1	-0.0107	-0.0169	-0.0046	1.00	-0.0155	-0.0280	-0.0033	1.00
T3:hcw1	0.0098	-0.0200	0.0392	1.00	-0.0051	-0.0583	0.0471	1.00
T3:dur30	0.0097	0.0013	0.0182	1.00	0.0088	-0.0001	0.0176	1.00
Group-level effects								
sd(Intercept)	1.1007	1.0681	1.1344	1.00	0.8530	0.7744	0.9298	1.00
sd(T3)	0.0650	0.0610	0.0689	1.02	0.0645	0.0545	0.0735	1.00
cor(Intercept,T3)	-0.3677	-0.4132	-0.3171	1.01	-0.2529	-0.3892	-0.0780	1.00
sigma	0.5614	0.5506	0.5722	1.01	0.6416	0.6158	0.6687	1.00
Infection	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**	**Estimate**	**l-95% CI**	**u-95% CI**	**Rhat**
Population-level effects								
Intercept	10.3100	10.1285	10.4874	1.00	10.8769	10.6669	11.0903	1.00
T3	-0.0816	-0.0938	-0.0693	1.00	-0.0536	-0.0681	-0.0389	1.00
age1	0.2739	0.0385	0.5067	1.00	-0.9581	-1.3349	-0.5763	1.00
age2	0.1077	-0.1161	0.3375	1.00	-0.6041	-0.9074	-0.3051	1.00
Male1	0.0986	-0.0522	0.2455	1.00	0.2569	0.0511	0.4635	1.00
ethnicity1	0.3474	-0.0027	0.6994	1.00	0.0710	-0.3371	0.4738	1.00
lthc1	0.0371	-0.1485	0.2202	1.00	-0.1520	-0.4081	0.1055	1.00
hcw1	-0.5064	-1.0172	0.0171	1.00	-0.3615	-0.7559	0.0423	1.00
dur30	0.2355	0.1720	0.2989	1.00	0.2124	0.1325	0.2922	1.00
T3:age1	0.0157	-0.0042	0.0357	1.00	0.0251	-0.0061	0.0565	1.00
T3:age2	0.0021	-0.0226	0.0267	1.00	0.0078	-0.0354	0.0505	1.00
T3:Male1	-0.0029	-0.0156	0.0095	1.00	-0.0111	-0.0256	0.0039	1.00
T3:ethnicity1	-0.0186	-0.0449	0.0073	1.00	0.0140	-0.0160	0.0426	1.00
T3:lthc1	-0.0040	-0.0201	0.0126	1.00	-0.0071	-0.0275	0.0131	1.00
T3:hcw1	0.0449	0.0045	0.0860	1.00	0.0443	0.0070	0.0811	1.00

	ChAdOx1 primary course				BNT162b2 primary course			
T3:dur30	-0.0068	-0.0115	-0.0021	1.00	-0.0055	-0.0112	0.0002	1.00
Group-level effects								
sd(Intercept)	1.6666	1.5770	1.7621	1.00	1.5171	1.4052	1.6374	1.00
sd(T3)	0.0555	0.0438	0.0668	1.01	0.0579	0.0463	0.0692	1.00
cor(Intercept,T3)	-0.8393	-0.9097	-0.7693	1.01	-0.7104	-0.7924	-0.6161	1.00
sigma	0.8476	0.8059	0.8911	1.01	0.7039	0.6596	0.7517	1.00

6.2.3 Estimation of protection

I combined the estimates of protection against infection by antibody level and of antibody declines to estimate the duration of protection against Omicron BA.4/5 infection. For those with a BNT162b2 or mRNA-1273 booster, I used estimates from the 'vaccinated participants without previous infection' group from the correlates model. For those with a breakthrough infection, since most earlier infections (>95%) were Delta and Omicron BA.1, I used estimates from the 'vaccinated participants with a most recent Delta infection" and "vaccinated participants with a most recent Omicron BA.1 infection' groups.

To estimate the duration from third/booster vaccination or breakthrough infection to reaching the threshold level associated with 67% protection, I used Metropolis Hastings posterior sampling to generate 100 sets of model coefficients from the correlates of protection models and hence estimate 100 antibody levels associated with 67% protection for each age group. I then extracted 100 draws from the posterior predictions from the antibody trajectory models and hence estimate 100 antibody levels for each time point. I then combined both posterior draws, for each draw from the antibody model, I produced 100 timepoints where the antibody level reached the 67% protection, resulting in 10,000 predictions of time to reaching the 67% protection for each age group. Finally, I calculated the median and 95% credible interval from the 10,000 predictions.

To estimate the actual population-level protection by calendar time, I combined results from the correlates of protection models and antibody trajectory models to get the population-level estimates. I used Metropolis Hastings posterior sampling to generate 100 sets of model coefficients from the correlates of protection models and extracted 100 draws from the posterior predictions from the antibody trajectory models, resulting in 10,000 predictions of actual protection level for each participant at each antibody level. Median protection level and 95% credible intervals were calculated by age group and calendar time. Protection estimations were based on assumptions that participants did not have previous infection before the first vaccination, did not receive another vaccination and were not infected after their third/booster vaccination or breakthrough infection.

6.3 Results

6.3.1 Correlates of protection against Omicron BA.4/5 infection

I used data from 17 May 2022 (when lagged antibody measurements >1360 BAU/mL became available, see Methods, corresponding to the start of the BA.4/5 infection wave[247]) to 12 September 2022 to determine the relationship between anti-spike antibody levels and protection from infection while Omicron BA.4/5 variants were dominant in the UK. During this period, of 19,311 sequenced infections, 13,097 (67.8%) were BA.5/sub-lineages, 2,924 (15.1%) BA.4/sub-lineages, 3,268 (16.9%) BA.2/sub-lineages, and 22 (0.1%) other Omicron recombinants or Delta.

To determine correlates of protection against Omicron BA.4/5 infection and the effect of previous infection, the vaccinated population was divided into those without evidence of previous infection (62,146 participants, 106,653 assessments) and those with evidence of previous infection (58,373 participants, 98,036 assessments), defining previous infection by positive swabs in the study, or the national testing programme (England only) or self-

reported by participants. Unvaccinated participants were excluded due to insufficient data (1,151 participants, 1,840 assessments). Participants with one, two, three, and four vaccinations were combined, as most (83.9%) had received three vaccinations. Participants with previous infection were further divided into those with a most recent pre-Alpha/Alpha infection (combined as effects similar, **Figure *6.1***), a most recent Delta infection, a most recent Omicron BA.1 infection, or a most recent Omicron BA.2 infection (**Table *6.1***). Protection was defined relative to vaccinated participants without evidence of previous infection with antibody levels of 16 BAU/mL, the threshold previously identified for vaccine non-response[174]. Using logistic generalised additive models (GAMs) with new PCR swab test results from each study assessment as the outcome and the most recent antibody measurement obtained 21–59 days earlier, protection against new Omicron BA.4/5 infection increased with higher antibody levels in all groups; the increase was rapid for anti-spike IgG <2,000 BAU/mL and flattened after that (**Figure *6.5*a**). Higher antibody levels were needed to achieve the same level of protection in those without versus with previous infection. At the same antibody level, previous Delta/Omicron BA.1 infection afforded higher protection against new Omicron BA.4/5 infection than a previous Pre-Alpha/Alpha infection. For example, antibody levels associated with 67% protection against infection for vaccinated participants aged 60y without previous infection, with Pre-Alpha/Alpha, Delta, and Omicron BA.1 infection were 1560 (95% confidence interval [CI]1360-1900), 1120 (900-1340) BAU/mL, 640 (440-860), and 380 (200-580) BAU/mL, respectively. A previous Omicron BA.2 infection afforded the highest protection against Omicron BA.4/5 infection, with >80% of participants protected regardless of antibody levels (**Figure *6.5*b**). Protection against moderate to high viral load infections (cycle threshold (Ct) values <30) and symptomatic infection was similar (**Figure *6.5*b,c**). Results

remained similar restricting to those with only one prior infection (84,034 assessments, 90%).

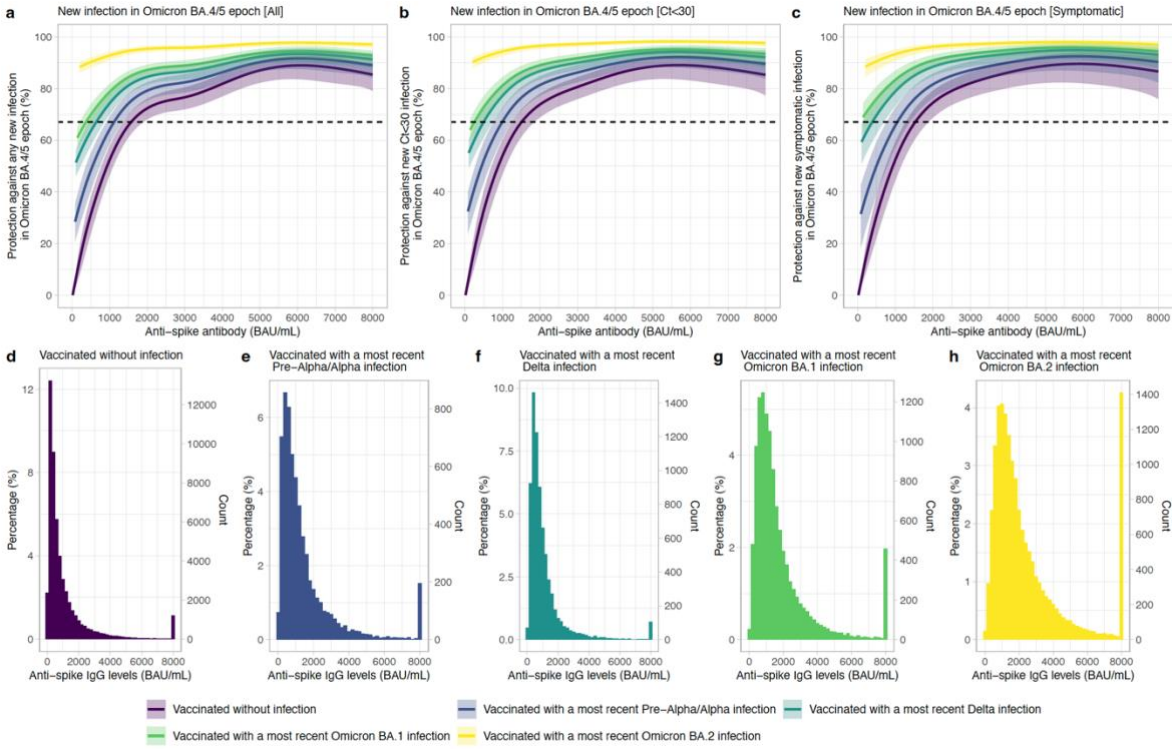


Figure 6.5 Association between anti-spike IgG levels and protection from new SARS-CoV-2 infection using the most recent antibody measurement obtained 21–59 days before the current assessment.

Note: a, Protection against any new infection in the Omicron BA.4/5 epoch. b, Protection against infection with a moderate to high viral load (Ct value < 30) in the Omicron BA.4/5 epoch. c, Protection against infection with self-reported symptoms in the Omicron BA.4/5 epoch. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. Five groups are investigated: vaccinated participants without evidence of prior infection, vaccinated participants with a most recent Pre-Alpha or Alpha infection, vaccinated participants with a most recent Delta infection, vaccinated participants with a most recent Omicron BA.1 infection, and vaccinated participants with a most recent Omicron BA.2 infection. Participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together. Protection is defined as relative protection against a baseline protection from 16 BAU/mL in those vaccinated without infection, which is the threshold for vaccine non-responders. In a-c, antibody measurements were plotted after the 1st percentile overall in each previous infection group (16, 80, 100, 140, 200 BAU/mL, respectively). Distribution and number of the most recent anti-spike IgG measurements for the four population groups are shown in d–h. Results remained similar restricting to those who had only one prior infection (84,034 assessments, 90%).

Older participants, especially those ≥70y, required higher antibody levels to reach the same level of protection. For example, in those without previous infection antibody levels

associated with 67% protection were 1200-1400 BAU/mL for those aged 20-50y but 2520 and 3380 BAU/mL at 75y and 80y. The differences were smaller between younger and older participants with a previous infection (**Figure *6.6***). For each variant, time from the last vaccination/infection had a limited impact on associations between antibody levels and protection, suggesting that time from last infection/vaccination was not independently associated with protection conditional on antibody levels (**Figure 6.7**). Results remained similar in a separate model examining time from last infection alone in prior infection groups (**Figure *6.8***).

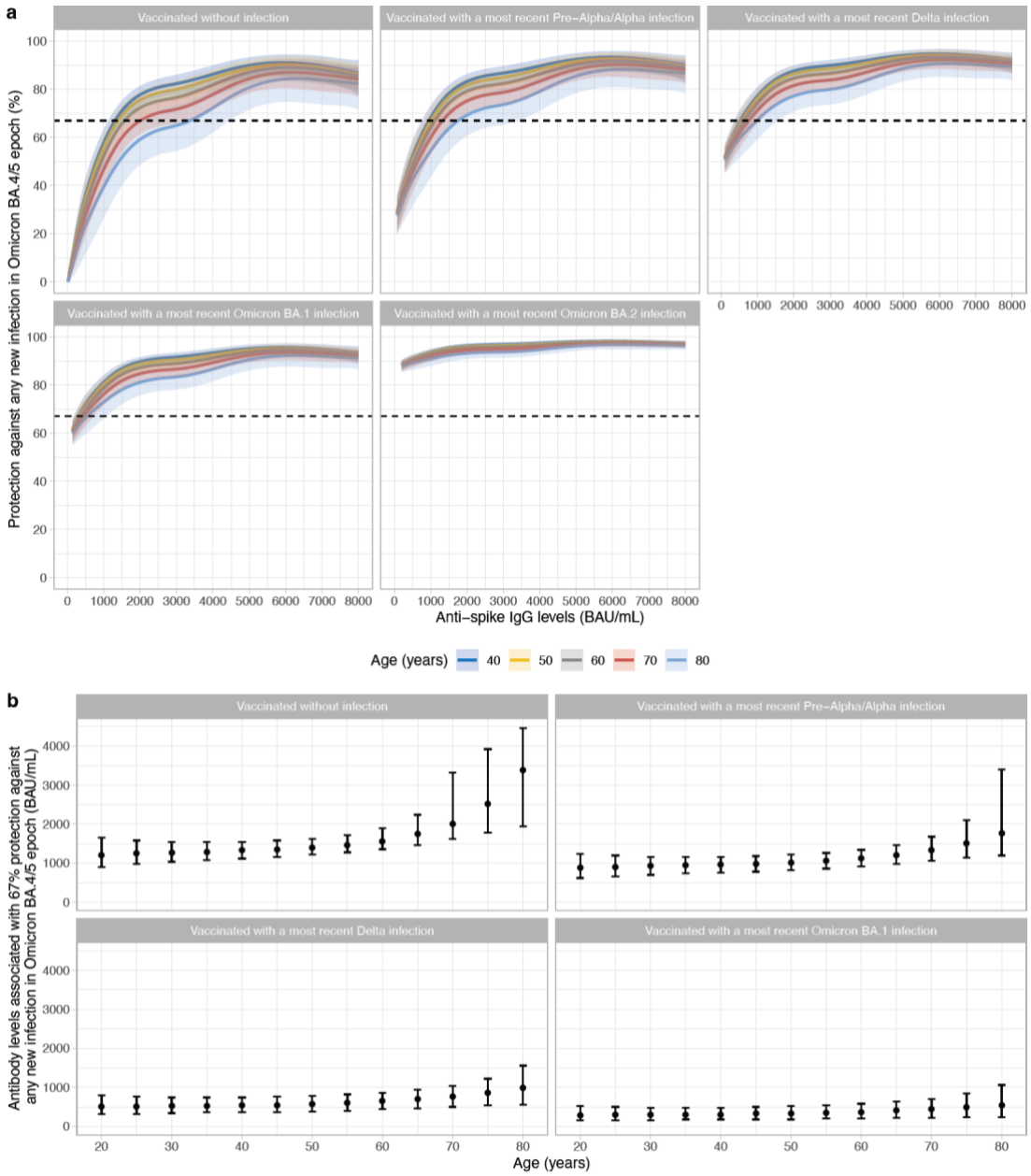


Figure 6.6 Effect of age on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels.

Note: a, Association between anti-spike IgG levels and protection from SARS-CoV-2 infection by age. Colour represents individual ages. Dotted line represents 67% protection against a baseline protection afforded by 16 BAU/mL in those vaccinated without infection, which is the threshold for vaccine non-responders. Five groups are investigated: vaccinated participants without evidence of prior infection, vaccinated participants with a most recent Pre-Alpha or Alpha infection, vaccinated participants with a most recent Delta infection, vaccinated participants with a most recent Omicron BA.1 infection, and vaccinated participants with a most recent Omicron BA.2 infection. Participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together. Antibody measurements were plotted after the 1st percentile overall in each previous infection group (16, 80, 100, 140, 200 BAU/mL, respectively). **b,** comparison of antibody levels associated with 67% protection against infection by age. The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction.

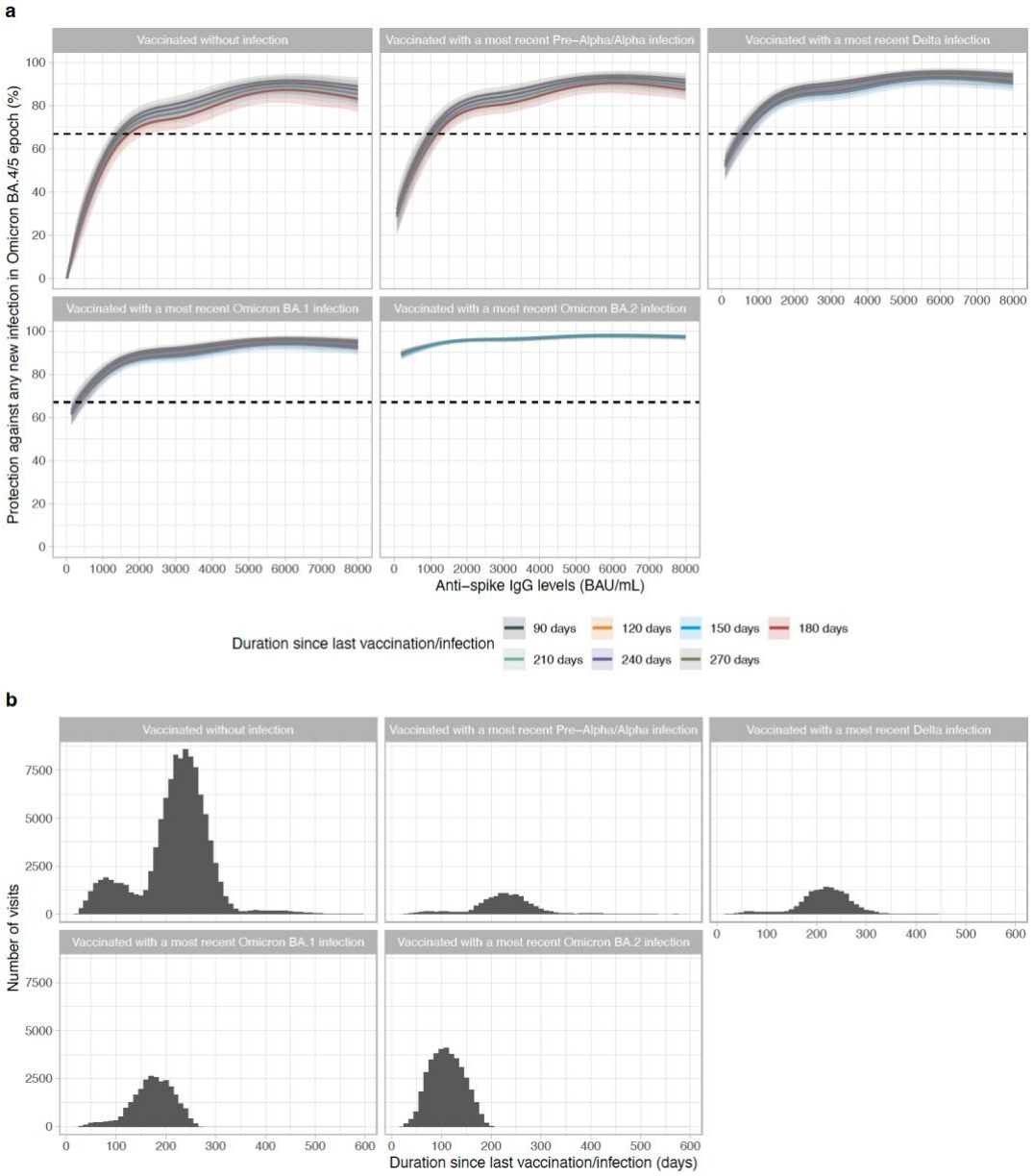


Figure 6.7 Effects from time since last vaccination or infection on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels.

Note: a, Association between anti-spike IgG levels and protection from SARS-CoV-2 infection by duration since last vaccination or infection to the current assessment. Colour represents duration since last vaccination or infection. Dotted line represents 67% protection against a baseline protection from 16 BAU/mL in those vaccinated without infection, which is the threshold for vaccine non-responders. Five groups are investigated: vaccinated participants without evidence of prior infection, vaccinated participants with a most recent Pre-Alpha or Alpha infection, vaccinated participants with a most recent Delta infection, vaccinated participants with a most recent Omicron BA.1 infection, and vaccinated participants with a most recent Omicron BA.2 infection. Participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together. Antibody measurements were plotted after the 1st percentile overall in each previous infection group (16, 80, 100, 140, 200 BAU/mL, respectively). The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. **b,** Distribution of duration from last vaccination or infection to the current assessment by group. The median duration was 230, 230, 220, 177, 110 days for the five groups.

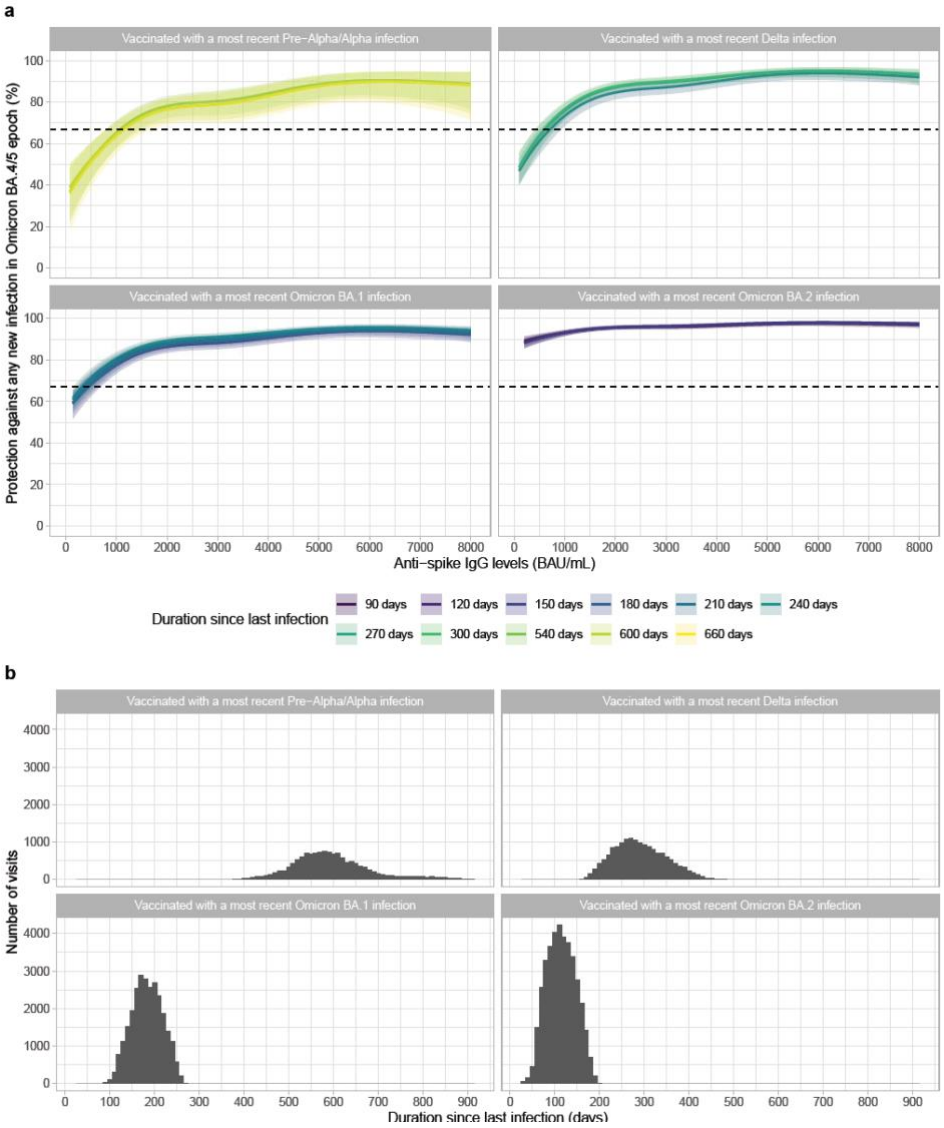


Figure 6.8 Effects from time since last infection on protection from new infection in the Omicron BA.4/5 epoch afforded by different antibody levels.

Note: a, Association between anti-spike IgG levels and protection from SARS-CoV-2 infection by duration since last infection to the current assessment, ignoring time from last vaccination in a separate model. Colour represents duration since last infection. Dotted line represents 67% protection against a baseline protection from 16 BAU/mL in those vaccinated without infection, which is the threshold for vaccine non-responders. Five groups are investigated: vaccinated participants without evidence of prior infection, vaccinated participants with a most recent Pre-Alpha or Alpha infection, vaccinated participants with a most recent Delta infection, vaccinated participants with a most recent Omicron BA.1 infection, and vaccinated participants with a most recent Omicron BA.2 infection. Only the four infected groups are shown here to show the effects of time since last infection. Participants with 1 (629 assessments, 0.3%), 2 (7,657 assessments, 3.7%), 3 (171,650 assessments, 83.9%), or 4 (24,753 assessments, 12.1%) vaccinations were grouped together. Antibody measurements were plotted after the 1st percentile overall in each previous infection group (80, 100, 140, 200 BAU/mL, respectively). The 95% CIs are calculated by prediction ± 1.96 × standard error of the prediction. b, Distribution of duration from last infection to the current assessment by group. The median duration was 585, 185, 183, 113 days for the four groups.

6.3.2 Antibody trajectories after third/booster vaccination and breakthrough infection

To estimate antibody trajectories, I used antibody measurements from 2 March 2021 to 12 September 2022. 154,149 participants aged ≥18y received two SARS-CoV-2 vaccinations with ChAdOx1 (with 6-13 week dosing intervals) or BNT162b2 (with 3-13 week dosing intervals) followed by BNT162b2 or mRNA-1273 third/booster vaccination or breakthrough infection (**Table *6.5***), and at least one antibody measurement after the second vaccination and no evidence of previous infection before the second vaccination.

I estimated rates of antibody decline from 21 days after the second vaccination and increases post-third/booster vaccination or infection using Bayesian piecewise linear interval-censored models[108,248]. Different models were fitted for each primary course (ChAdOx1 or BNT162b2) and boosting event (BNT162b2 or mRNA-1273 third/booster vaccination or infection). I adjusted for age, sex, long-term health conditions, ethnicity (white vs non-white), working in healthcare, and time from second vaccination to booster/infection. Overall, the median age at third/booster vaccination or infection was 60y (Interquartile range [IQR] 50-69), 84,080 (54.5%) participants were female, and 147,712 (95.8%) reported white ethnicity. 43,175 (28.0%) reported having a long-term health condition, and 4,198 (2.7%) were healthcare workers. The median time from second vaccination to third/booster vaccination was 6 months and to infection was 5 months. Characteristics of those included in an additional antibody decline model from 42 days after the third/booster vaccination or infection were similar to the overall population (**Table *6.6***).

Table 6.5 Characteristics of all participants included in the antibody trajectory analyses.

Note: Participants were divided into six groups based on the primary vaccine course (ChAdOx1 or BNT162b2) and the third/booster vaccination (BNT162b2 or mRNA-1273) or infection.

	ChAdOx1-BNT162b2 (N=64940)	ChAdOx1-mRNA-1273 (N=21960)	ChAdOx1-Infection (N=10830)	BNT162b2-BNT162b2 (N=44197)	BNT162b2-mRNA-1273 (N=7248)	BNT162b2-Infection (N=4974)	Total (N=154149)	p value
Months from second vaccination to booster/infection								
Median	6	6	5	6	6	5	6	
Q1, Q3	6, 7	6, 6	4, 6	6, 7	5, 7	4, 6	6, 7	
Age (years)								< 0.001
Median	63	56	50	66	41	39	60	
Q1, Q3	53, 71	49, 62	43, 59	52, 74	34, 60	32, 56	50, 69	
Sex								< 0.001
Female	34760 (53.5%)	11571 (52.7%)	5757 (53.2%)	25356 (57.4%)	3843 (53.0%)	2793 (56.2%)	84080 (54.5%)	
Male	30180 (46.5%)	10389 (47.3%)	5073 (46.8%)	18841 (42.6%)	3405 (47.0%)	2181 (43.8%)	70069 (45.5%)	
Ethnicity								< 0.001
Non-white	2363 (3.6%)	844 (3.8%)	504 (4.7%)	2043 (4.6%)	351 (4.8%)	332 (6.7%)	6437 (4.2%)	
White	62577 (96.4%)	21116 (96.2%)	10326 (95.3%)	42154 (95.4%)	6897 (95.2%)	4642 (93.3%)	147712 (95.8%)	
Report having a long-term health condition								
No	45834 (70.6%)	17428 (79.4%)	8684 (80.2%)	29429 (66.6%)	5635 (77.7%)	3964 (79.7%)	110974 (72.0%)	
Yes	19106 (29.4%)	4532 (20.6%)	2146 (19.8%)	14768 (33.4%)	1613 (22.3%)	1010 (20.3%)	43175 (28.0%)	
Healthcare worker								< 0.001
No	64184 (98.8%)	21817 (99.3%)	10701 (98.8%)	41365 (93.6%)	7181 (99.1%)	4703 (94.6%)	149951 (97.3%)	
Yes	756 (1.2%)	143 (0.7%)	129 (1.2%)	2832 (6.4%)	67 (0.9%)	271 (5.4%)	4198 (2.7%)	
Infection type							**Total (N=15804)**	
Delta			9696 (89.5%)			3194 (64.3%)	12890 (81.6%)	< 0.001
Omicron, BA.1			1119 (10.3%)			1596 (32.1%)	2715 (17.2%)	
Omicron, BA.2			13 (0.1%)			180 (3.6%)	193 (1.2%)	
Other			2 (0.0%)			0 (0.0%)	2 (0.0%)	
N-Miss			0			4	4	
Ct value								< 0.001
Median			20			21	20	

	ChAdOx1-BNT162b2 (N=64940)	ChAdOx1-mRNA-1273 (N=21960)	ChAdOx1-Infection (N=10830)	BNT162b2-BNT162b2 (N=44197)	BNT162b2-mRNA-1273 (N=7248)	BNT162b2-Infection (N=4974)	Total (N=154149)	p value
Q1, Q3			16, 26			17, 26	17, 26	
Symptom								< 0.001
No			2594 (24.1%)			1451 (29.4%)	4045 (25.8%)	
Yes, classic symptoms			4444 (41.3%)			1782 (36.1%)	6226 (39.6%)	
Yes, other symptoms			3735 (34.7%)			1701 (34.5%)	5436 (34.6%)	
N-Miss			57			40	97	

Table 6.6 Characteristics of participants included in the antibody decline analyses post third/booster vaccination or infection.

Note: Participants were divided into six groups based on the primary vaccine course (ChAdOx1 or BNT162b2) and the third/booster vaccination (BNT162b2 or mRNA-1273) or infection. Characteristics in this subgroup are similar to the overall population (**Table 6.5**).

	ChAdOx1-BNT162b2 (N=41152)	ChAdOx1-mRNA-1273 (N=14748)	ChAdOx1-Infection (N=4214)	BNT162b2-BNT162b2 (N=24749)	BNT162b2-mRNA-1273 (N=4403)	BNT162b2-Infection (N=1857)	Total (N=91123)	p value
Months from second vaccination to booster/infection								< 0.001
Median	6	6	4	6	6	4	6	
Q1, Q3	6, 7	6, 6	3, 5	6, 7	5, 7	3, 5	6, 7	
Age (years)								< 0.001
Median	63	57	50	66	46	39	61	
Q1, Q3	55, 71	51, 62	43, 59	55, 74	35, 63	33, 54	52, 70	
Sex								< 0.001
Female	22213 (54.0%)	7838 (53.1%)	2299 (54.6%)	14145 (57.2%)	2345 (53.3%)	1045 (56.3%)	49885 (54.7%)	
Male	18939 (46.0%)	6910 (46.9%)	1915 (45.4%)	10604 (42.8%)	2058 (46.7%)	812 (43.7%)	41238 (45.3%)	
Ethnicity								< 0.001
Non-white	1249 (3.0%)	507 (3.4%)	200 (4.7%)	958 (3.9%)	179 (4.1%)	121 (6.5%)	3214 (3.5%)	
White	39903 (97.0%)	14241 (96.6%)	4014 (95.3%)	23791 (96.1%)	4224 (95.9%)	1736 (93.5%)	87909 (96.5%)	
Report having a long-term health condition								< 0.001
No	29072 (70.6%)	11662 (79.1%)	3345 (79.4%)	16452 (66.5%)	3360 (76.3%)	1497 (80.6%)	65388 (71.8%)	
Yes	12080 (29.4%)	3086 (20.9%)	869 (20.6%)	8297 (33.5%)	1043 (23.7%)	360 (19.4%)	25735 (28.2%)	

	ChAdOx1-BNT162b2 (N=41152)	ChAdOx1-mRNA-1273 (N=14748)	ChAdOx1-Infection (N=4214)	BNT162b2-BNT162b2 (N=24749)	BNT162b2-mRNA-1273 (N=4403)	BNT162b2-Infection (N=1857)	Total (N=91123)	p value
Healthcare worker								< 0.001
No	40725 (99.0%)	14655 (99.4%)	4144 (98.3%)	23379 (94.5%)	4362 (99.1%)	1727 (93.0%)	88992 (97.7%)	
Yes	427 (1.0%)	93 (0.6%)	70 (1.7%)	1370 (5.5%)	41 (0.9%)	130 (7.0%)	2131 (2.3%)	
Infection type								< 0.001
Delta			3949 (93.7%)			1198 (64.5%)	5147 (84.8%)	
Omicron, BA.1			253 (6.0%)			552 (29.7%)	805 (13.3%)	
Omicron, BA.2			10 (0.2%)			106 (5.7%)	116 (1.9%)	
Other			2 (0.0%)			0 (0.0%)	2 (0.0%)	
N-Miss			0			1	1	
Ct value								< 0.001
Median			20			21	20	
Q1, Q3			16, 26			17, 26	17, 26	
Symptom								< 0.001
No			966 (23.0%)			519 (28.1%)	1485 (24.6%)	
Yes, classic symptoms			1889 (45.0%)			735 (39.8%)	2624 (43.4%)	
Yes, other symptoms			1341 (32.0%)			591 (32.0%)	1932 (32.0%)	
N-Miss			18			12	30	

As expected, given associations between antibody levels and infection risk[174], antibody levels after second vaccination were consistently lower in those who became infected than those who received a third/booster vaccination (before infection) (**Figure *6.9***; grey vs blue/red lines), as well as being lower in those receiving ChAdOx1 versus BNT162b2 primary courses (dashed vs solid lines). Mean antibody levels were significantly boosted regardless of primary course or type of boosting event (**Figure *6.9***), to levels higher than achieved after second vaccinations. The relatively lower antibody levels in those receiving a primary ChAdOx1 course were boosted more than those receiving a primary BNT162b2 course, such that post-booster both groups achieved similar antibody levels. mRNA-1273 boosters generated higher peak antibody levels than BNT162b2 boosters, but their subsequent waning was faster. Breakthrough infection boosted antibody levels to similar levels to BNT162b2 boosters, slightly lower than mRNA-1273 boosters. Estimated antibody declines after booster/infection were generally faster following primary ChAdOx1 courses, versus primary BNT162b2 courses. Following primary ChAdOx1 courses, estimated antibody declines were similar following infection or booster vaccinations, for example, the half-life was 96 (95% credible interval [CrI] 80-119) days after infection versus 78 (72-86) and 63 (61-66) days with BNT162b2 and mRNA-1273 boosters, respectively, for those aged 55y. Following primary BNT162b2 courses, the estimated declines after infection were slower than after boosters; for example, the half-life was 178 (117-382) versus 98 (91-106) and 93 (82-107) days with BNT162b2 and mRNA-1273 boosters, respectively, for those aged 55y, although the credible intervals were wider due to smaller sample sizes (**Figure *6.9*, Figure *6.10* a,b, Table *6.6***). 180 days after booster/infection, antibody levels were similar among participants with primary ChAdOx1 courses regardless of type of boosting event and were lower than participants with primary BNT162b2 courses boosted by mRNA-1273 or breakthrough infection (**Figure *6.11***).

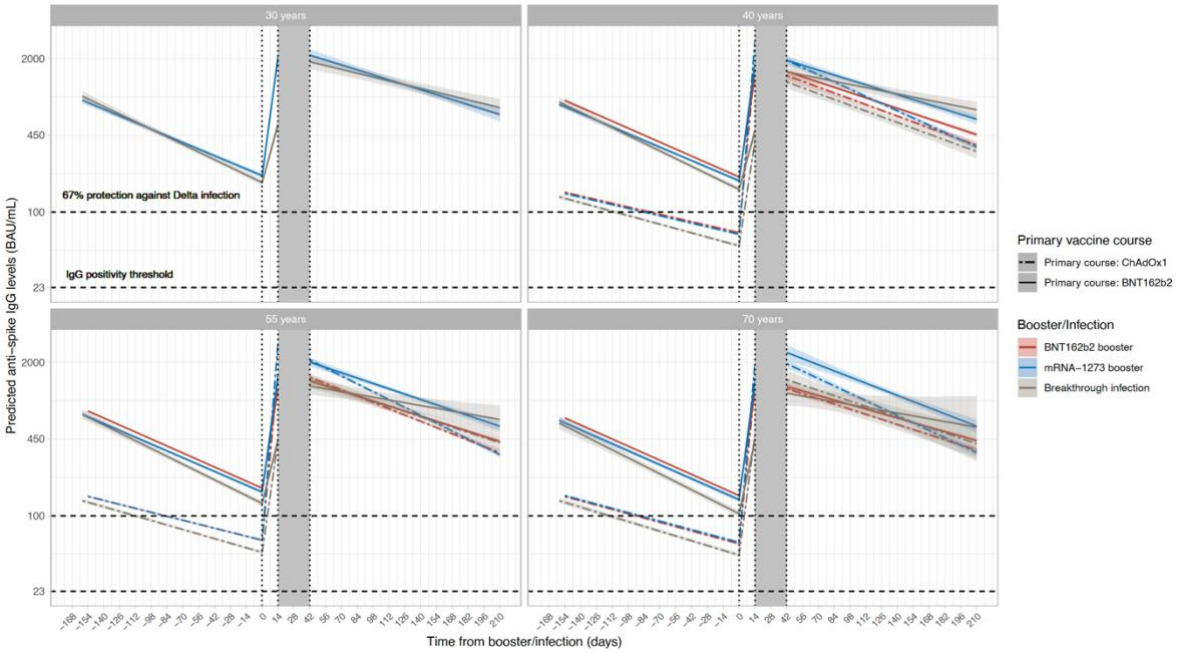


Figure 6.9 Posterior predicted mean trajectories (95%Crl) of anti-spike IgG levels from second vaccination through third/booster vaccination or infection using Bayesian linear mixed interval censored models.

Note: Time 0 indicates the earliest of the date each participant received their third/booster vaccination or their first breakthrough infection. For each group, two separate models were fitted: 1) piecewise model on antibody decline after the second vaccination and subsequent increase after third/booster vaccination or infection; 2) antibody decline 42 days after the third/booster vaccination or infection. Shaded area between 14- and 42-days post third/booster vaccination or infection represents different timepoints individuals reach peak antibody levels. Models are adjusted for age, sex, ethnicity, time from second vaccination to booster/infection, long-term health conditions, and healthcare role. Plotted at the reference categories (female, white ethnicity, 6 months between second vaccination and booster/infection, not reporting a long-term health condition, not working in healthcare). Line types indicate the primary vaccine course. Line colours indicate the booster type or infection. The ChAdOx1-BNT162b2 (red dotted line) and ChAdOx1-mRNA-1273 (blue dotted line) are overlapped before time 0. Plots are separated by age (30y only estimated for those who had BNT162b2 as primary and were boosted by mRNA-1273 or infection due to low numbers in other groups). Predicted values are plotted on a log scale. Black dashed lines indicate the correlate for 67% protection against Delta variant (100 BAU/mL) and the threshold of IgG positivity (23 BAU/mL).

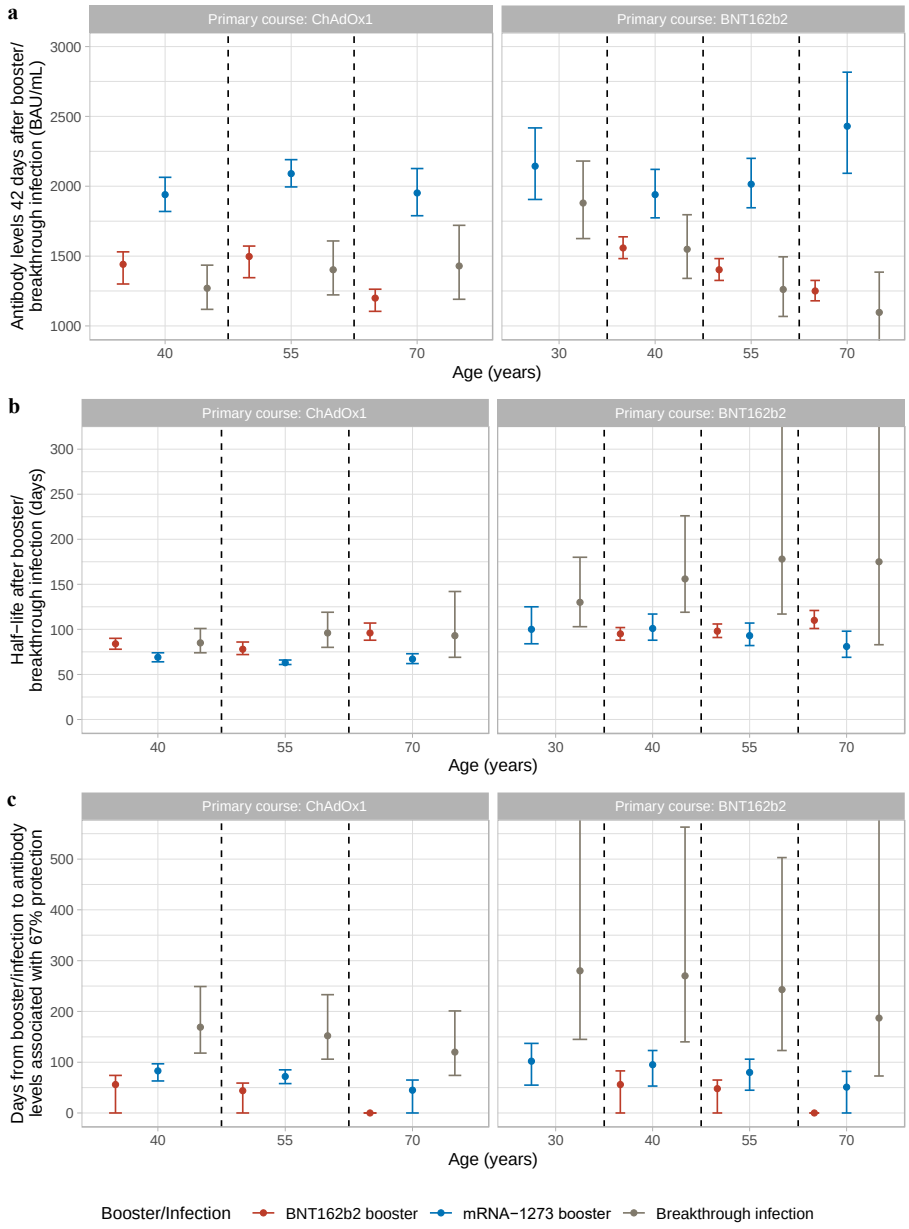


Figure 6.10 Comparisons of antibody levels 42 days post third/booster vaccination or infection, half-lives, and days from third/booster vaccination or infection to reaching antibody levels associated with 67% protection by primary vaccine course, third/booster vaccination or infection, and age.

Note: a, comparisons of antibody levels 42 days post third/booster vaccination or infection. b. comparisons of half-lives after third/booster vaccination or infection. **c,** comparisons of days from third/booster vaccination or infection to reaching antibody levels associated with 67% protection. Median values with 95% credible intervals are plotted. 95% credible interval in panel **c** are calculated from posterior simulations from the GAM model estimating correlates of protection and posterior predictions from the Bayesian linear mixed models estimating antibody levels. Predictions are on specific ages (30, 40, 55, 70 years). 30y is not plotted for ChAdOx1 primary course because majority of participants receiving ChAdOx1 primary course are >40y. Numbers are shown in **Table 6.6**. Plotted at the reference categories (female, white ethnicity, 6 months between second vaccination and booster/infection, not reporting a long-term health condition, not working in healthcare).

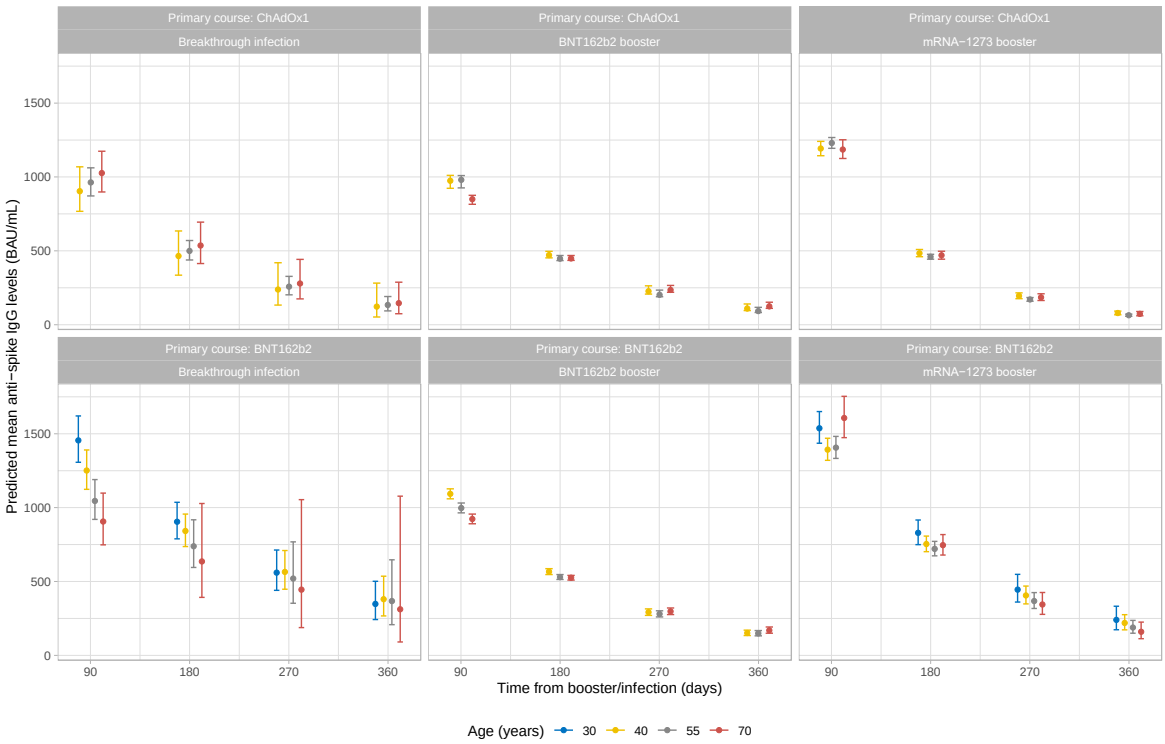


Figure 6.11 Predicted mean anti-spike IgG levels (95%Crl) at 90, 180, 270, and 360 days after the third/booster vaccination or breakthrough infection by ages.

Note: 30 years only estimated for those who had BNT162b2 as primary and were boosted by mRNA-1273 or infection due to low numbers in other groups.

Younger individuals generally generated higher antibody levels 42 days post-booster vaccination than older individuals, except those with primary BNT162b2 course and mRNA-1273 booster **(Figure *6.12*)**. Antibody levels were higher in individuals with longer times between second vaccinations and breakthrough infection or BNT162b2 booster vaccination, but not mRNA-1273 booster (**Figure *6.12***). There were relatively modest effects of other participant characteristics (**Figure *6.13***).

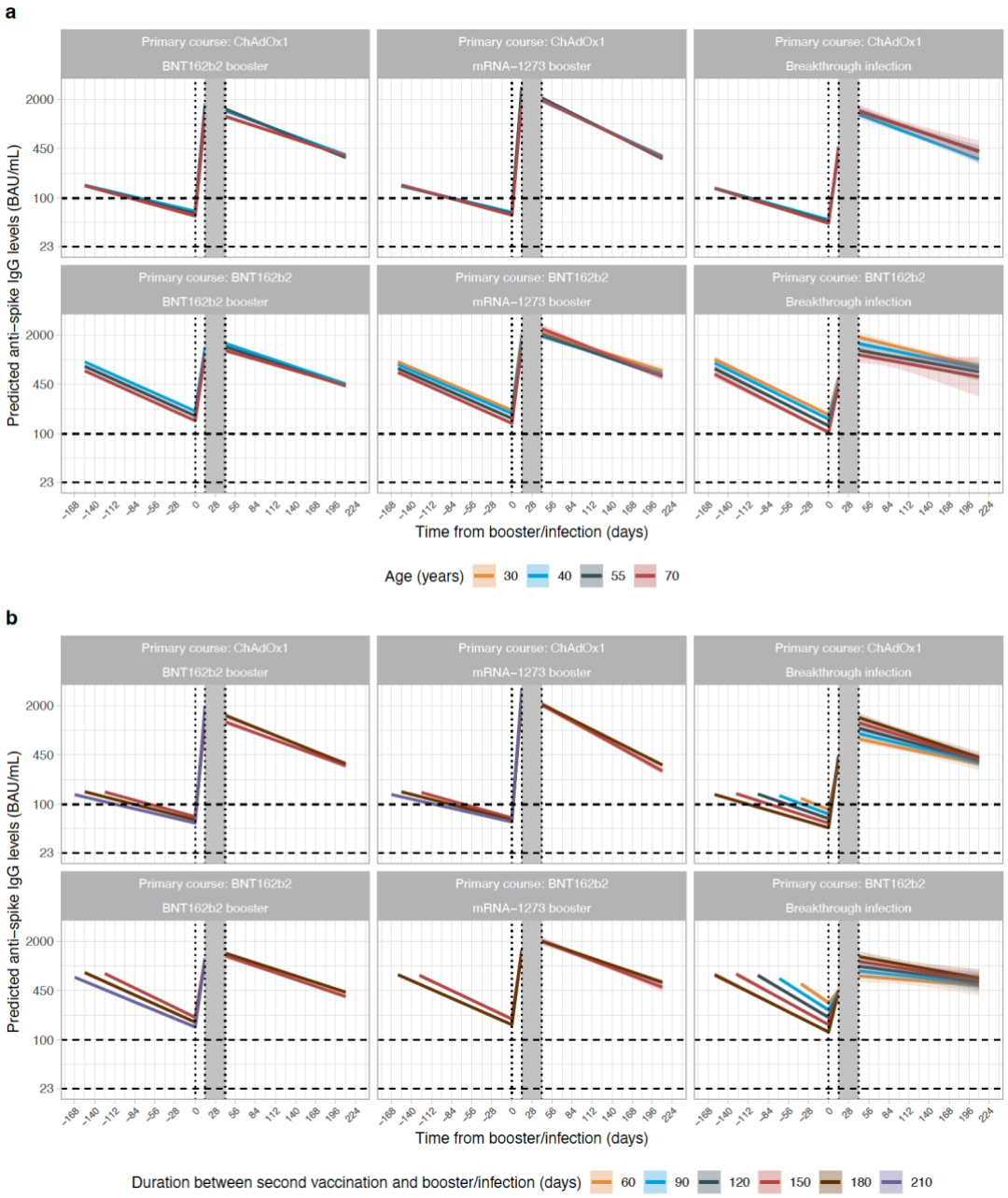


Figure 6.12 Posterior predicted mean trajectories (95%CrI) of anti-spike IgG levels from third/booster vaccination or infection.

Note: a, by age. b, by time from second vaccination to third/booster vaccination or infection. For each group, two separate models are fitted: 1) piecewise model on antibody decline after the second vaccination and subsequent increase after third/booster vaccination or infection; 2) antibody decline 42 days after the third/booster vaccination or infection. Shaded area between 14- and 42-days post third/booster vaccination or infection represents different timepoints individuals reach peak antibody levels. Models are adjusted for age, sex, ethnicity, time from second vaccination to booster/infection, long-term health conditions, and healthcare role. Plotted at the reference categories (female, white ethnicity, time from second vaccination to booster/infection 6 months, not reporting a long-term health condition, not working in healthcare). Plots are separated by primary vaccine courses and booster types or infection. Predicted values are plotted on a log scale. Black dashed lines indicate the correlate for 67% protection against Delta variant (100 BAU/mL) and the threshold of IgG positivity (23 BAU/mL).

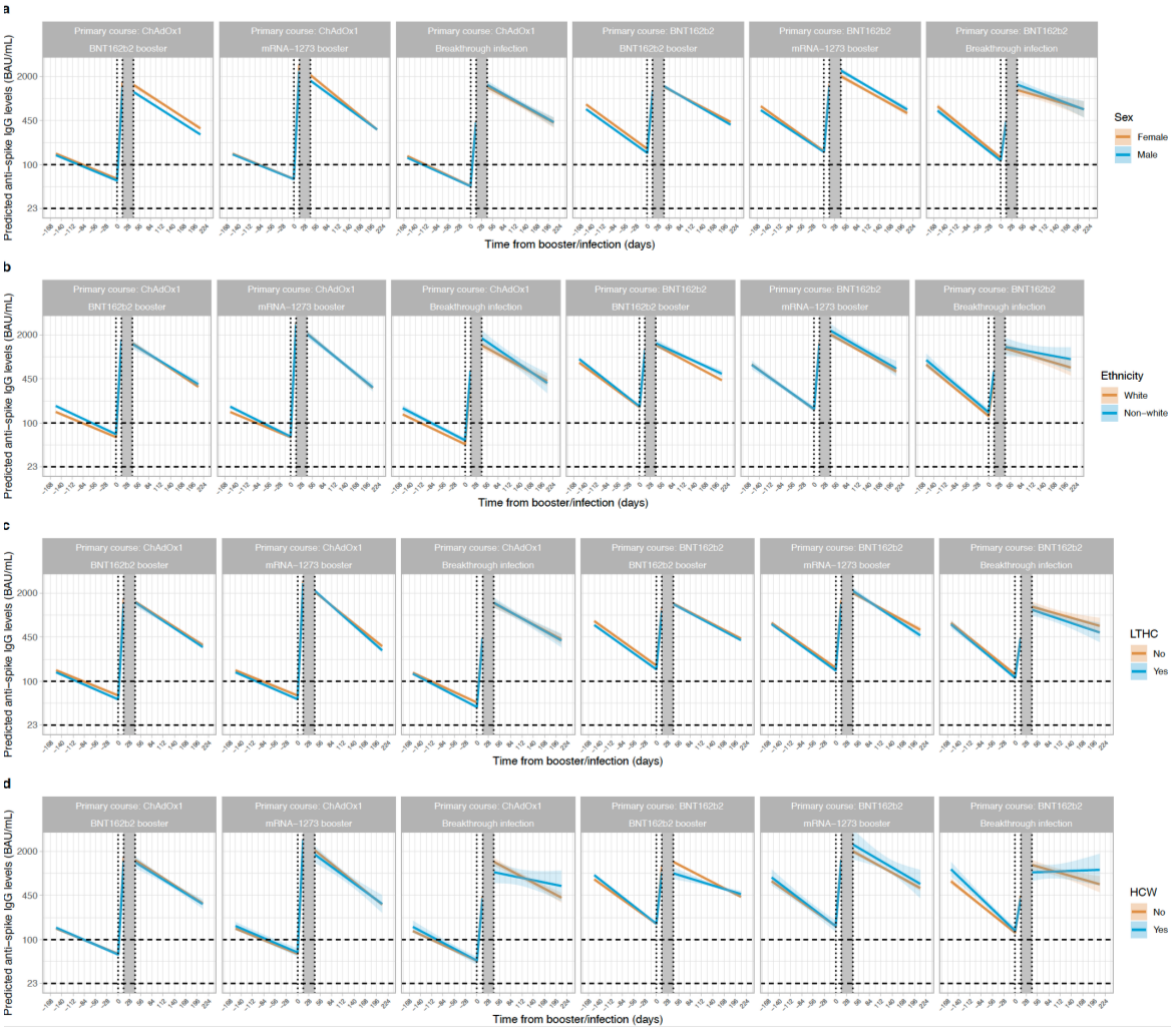


Figure 6.13 Posterior predicted mean trajectories (95%CrI) of anti-spike IgG levels from third/booster vaccination or infection by sex, ethnicity, long-term health condition, and healthcare role.

Note: Models are adjusted for age, sex, ethnicity, time from second vaccination to booster/infection, long-term health conditions, and healthcare role. Plotted at the reference categories (55y, 6 months from second vaccination to booster/infection, female, white ethnicity, not reporting long-term health condition, not working in healthcare). Plots are separated by primary vaccine course and booster types or infection. Predicted values are plotted on a log scale. Black dashed lines indicate the correlate for 67% protection against Delta variant (100 BAU/mL) and the threshold of IgG positivity (23 BAU/mL).

There was no evidence of differences in antibody peak levels or half-lives across breakthrough infections before a booster with different variants (Delta or Omicron BA.1, the dominant variants causing breakthrough infections in the study) (**Figure *6.14***).

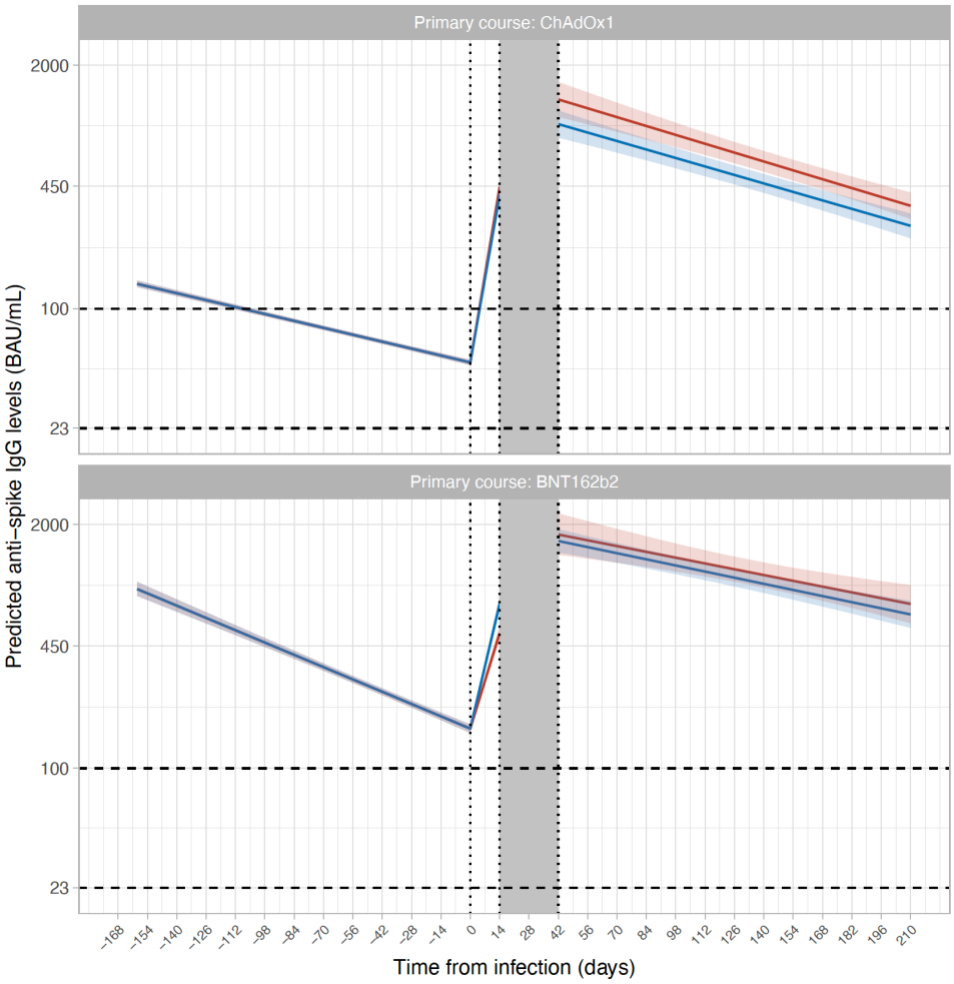


Figure 6.14 Effects of SARS-CoV-2 variant (Delta vs Omicron BA.1) on antibody responses after infection.

Note: Plotted by primary vaccine course (ChAdOX1 or BNT162b2). Plotted at the reference categories (40y, 6 months from second vaccination to booster/infection, female, white ethnicity, not reporting long-term health condition, not working in healthcare). Black dashed lines indicate the correlate for 67% protection against Delta variant (100 BAU/mL) and the threshold of IgG positivity (23 BAU/mL). Shaded area between 14- and 42-days post infection represents different timepoints individuals reach peak antibody levels. No differences on antibody levels and subsequent decline by infection variant.

6.3.3 Protection from the third/booster vaccination and breakthrough infection against new Omicron BA.4/5 infection

I combined the estimates of protection against infection by antibody level and of antibody declines to estimate the duration of protection against Omicron BA.4/5 infection, estimating the time from third/booster vaccination or infection to mean antibody levels reaching levels associated with 67% protection against infection relative to vaccinated

participants without evidence of previous infection with antibody levels 16 BAU/mL (**Figure *6.10*c**). BNT162b2 boosters did not provide this level of protection for participants aged >70y, while mRNA-1273 boosters provided <80 days of protection. For those aged 40y, antibody levels reached this threshold <80 days after BNT162b2 boosters, compared with 50-120 days for participants receiving mRNA-1273 boosters. For those aged 55y, BNT162b2 boosters provided <65 days of protection, while mRNA-1273 boosters provided 50-100 days of protection. For participants with breakthrough infection, antibody levels associated with 67% protection lasted for 120-170 days with ChAdOx1, and 180-280 days with BNT162b2, primary courses (**Figure *6.10*c**, **Table *6.7***). Variation was predominantly explained by the greater protection found at a given antibody level following infection versus boosters, with some contribution from slower antibody waning following primary BNT162b2 courses.

Table 6.7 Posterior predicted anti-spike IgG levels at 42 days from booster/infection (BAU/mL), half-lives after booster/infection (days), and time from booster/infection to reaching antibody levels associated with 67% protection.

Note: Results were separated by age (30, 40, 55, 70y), primary vaccine course (ChAdOx1 or BNT162b2), and booster/third vaccination or infection (BNT162b2, mRNA-1273, infection). Results were estimated at the reference category (female, white ethnicity, 6 months between second vaccination and booster/infection, not reporting a long-term health condition, not working in healthcare). Comparisons are also plotted in **Figure 6.10**. NE: Not estimable, due to the antibody levels not declining in the posterior median or upper credible interval. 95% credible intervals for days from booster/infection to reaching the 67% threshold level are calculated from posterior simulations from the GAM model estimating correlates of protection and posterior predictions from the Bayesian linear mixed models estimating antibody levels.

Age (years)	Primary vaccine	Booster/Infection	Median antibody levels at 42 days (95% Crl) (BAU/mL)	Median half-life (95%Crl) (days)	Median days from booster/infection to reaching 67% protection (95%Crl)
30	BNT162b2	mRNA-1273	2144 (1905-2418)	100 (84-125)	102 (55-137)
		Infection	1880 (1625-2180)	130 (103-180)	280 (145-611)
40	ChAdOX1	BNT162b2	1441 (1300-1530)	84 (80-99)	56 (0-74)
		mRNA-1273	1940 (1820-2064)	69 (64-74)	83 (63-97)
		Infection	1270 (1119-1435)	85 (74-101)	169 (118-249)
	BNT162b2	BNT162b2	1558 (1482-1638)	95 (88-102)	56 (0-83)
		mRNA-1273	1940 (1774-2121)	101 (88-117)	95 (53-123)
		Infection	1549 (1340-1796)	156 (119-226)	270 (140-563)
55	ChAdOX1	BNT162b2	1497 (1345-1572)	78 (75-90)	44 (0-59)
		mRNA-1273	2090 (1995-2190)	63 (61-66)	72 (58-85)
		Infection	1402 (1222-1608)	96 (80-119)	152 (106-233)
	BNT162b2	BNT162b2	1402 (1326-1482)	98 (91-106)	48 (0-65)
		mRNA-1273	2015 (1846-2200)	93 (82-107)	80 (45-106)
		Infection	1261 (1068-1495)	178 (117-382)	243 (123-503)
70	ChAdOX1	BNT162b2	1199 (1104-1263)	96 (91-110)	0 (0-0)
		mRNA-1273	1952 (1789-2127)	67 (62-73)	45 (0-65)
		Infection	1429 (1191-1720)	93 (69-142)	120 (74-201)
	BNT162b2	BNT162b2	1250 (1180-1326)	110 (101-121)	0 (0-0)
		mRNA-1273	2430 (2093-2817)	81 (69-98)	51 (0-82)
		Infection	1097 (865-1385)	175 (83-NE)	187 (73-NE)

I also estimated the proportion of participants with 67% protection 42, 90, 180, 270, and 360 days after third/booster vaccination or infection based on individual-level predictions, and assuming no further vaccination/infection (**Figure *6.15***). Following primary ChAdOx1 courses, 42 days after BNT162b2 and mRNA-1273 boosters, around 60% and 100% of those aged 40-55y, and around 20% and 80% of those aged 55-70y had antibody levels associated with ≥67% protection. However, no participant remained above this threshold level at 90 days. Following primary BNT162b2 courses, >80% of those aged <55y boosted with BNT162b2 and nearly everyone boosted with mRNA-1273 were above the threshold level 42 days later. At 90 days, almost every participant boosted with BNT162b2 fell below the level affording 67% protection, while >80% of those aged <55y and around 45% of those aged 55-70y boosted with mRNA-1273 remained above this threshold. No participant remained above the threshold at 180 days. For those with breakthrough infections, nearly all participants aged <70y were above the threshold 42 and 90 days after infection. Nearly all participants who received primary ChAdOX1 courses had antibody levels below the 67% protection threshold level by 270 days. However, for those who received primary BNT162b2 courses, the percentage decreased over time but remained high (>70%) for those <55y at 270 days. The decrease was greater for older participants and almost every participant >55y did not maintain ≥67% protection by 270 days, although credible intervals were wide due to small numbers.

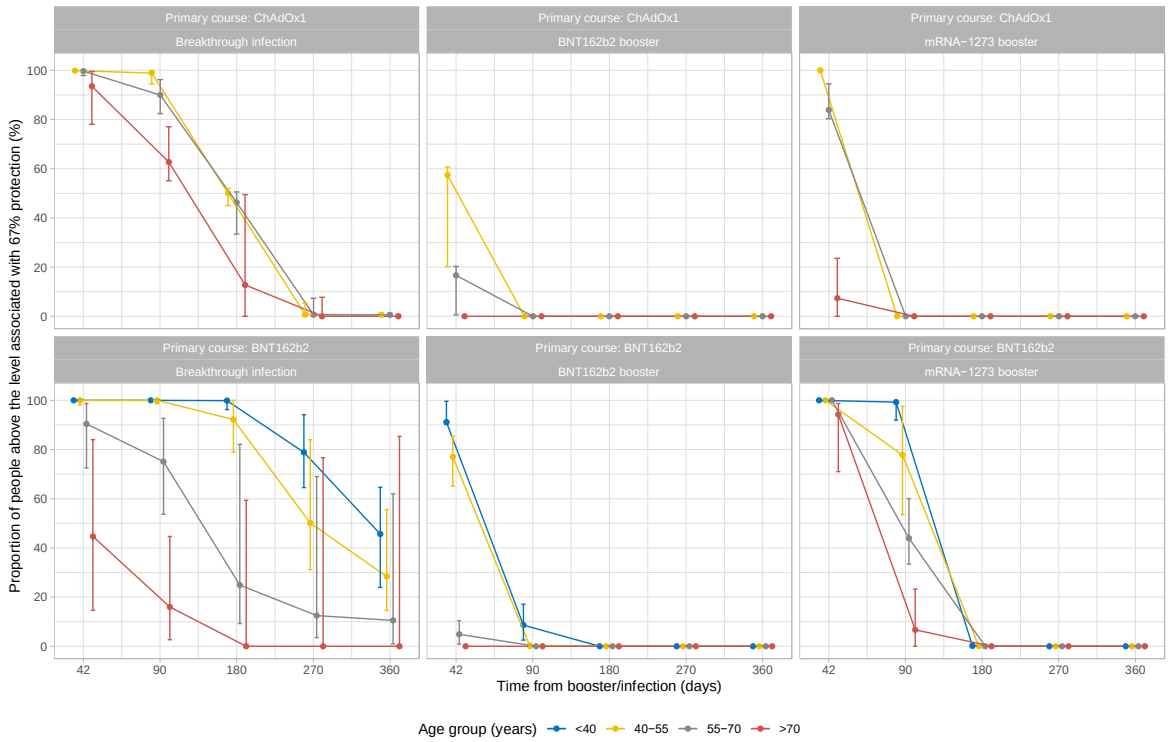


Figure 6.15 Proportion of participants above the anti-spike IgG antibody threshold level associated with 67% protection by time from third/booster vaccination or infection.

Note: Numbers of participants in each panel are [numbers in brackets represent <40, 40–55, 55-70 and >70 years]: ChAdOx1-Infection: n=4,214 [537, 2,167, 1,216, 294]; ChAdOx1-BNT162b2: n = 41,152 [1,295, 8,826, 19,232, 11,799]; ChAdOx1-mRNA-1273: n = 14,748 [738, 5,341, 7,859, 810]; BNT162b2-Infection: n = 1,857 [956, 438, 313, 150]; BNT162b2-BNT162b2: n = 24,749 [2,447, 3,613, 8,955, 9,734]; BNT162b2-mRNA-1273: n = 4,403 [1,791, 889, 1,409, 314]. '<40 year' group is not plotted for ChAdOx1 primary course because the vast majority of those receiving ChAdOx1 were 40 years of age or older. Median values with 95% credible intervals are plotted.

To estimate a lower bound for levels of protection in the UK population, had further vaccination campaigns not taken place and the virus not been circulating to further boost antibody levels via breakthrough infections, I calculated the median protection levels over calendar time to31 December 2022. In this scenario, the median level of protection in the UK would be 50-60% among those who had a breakthrough infection and 5-15% among those triple-vaccinated without infection, with younger individuals having higher levels of protection than older individuals (**Figure *6.16***).

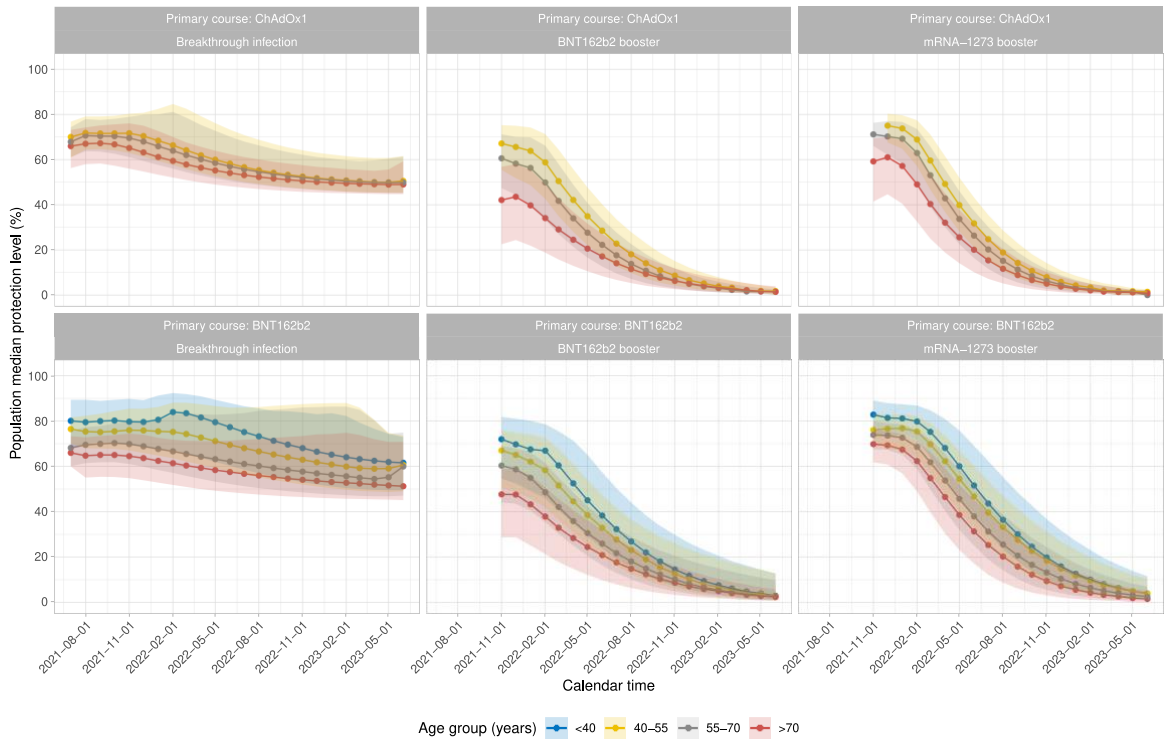


Figure 6.16 Median protection level by calendar time.

Note: Estimations were based on assumptions that participants did not receive another vaccination and were not infected after their third/booster vaccination or breakthrough infection. Numbers of participants in each panel are [numbers in brackets represent <40, 40–55, 55-70 and >70 years]: ChAdOx1-Infection: n=4,214 [537, 2,167, 1,216, 294]; ChAdOx1-BNT162b2: n = 41,152 [1,295, 8,826, 19,232, 11,799]; ChAdOx1-mRNA-1273: n = 14,748 [738, 5,341, 7,859, 810]; BNT162b2-Infection: n = 1,857 [956, 438, 313, 150]; BNT162b2-BNT162b2: n = 24,749 [2,447, 3,613, 8,955, 9,734]; BNT162b2-mRNA-1273: n = 4,403 [1,791, 889, 1,409, 314]. '<40 year' group is not plotted for ChAdOx1 primary course because the vast majority of those receiving ChAdOx1 were 40 years of age or older. 95% credible interval is calculated from posterior simulations from the GAM model estimating correlates of protection and posterior predictions from the Bayesian linear mixed models estimating antibody levels.

6.4 Discussion

In this study, I found that those with previous SARS-CoV-2 infections had a lower rate of observed new Omicron BA.4/BA.5 infections than those who were vaccinated but without evidence of any previous infection, despite both breakthrough infection and booster vaccination resulting in substantial increases in anti-spike IgG antibody levels, regardless of third/booster vaccine type or primary vaccine course. Breakthrough infections generated similar antibody levels to third/booster vaccinations, and subsequent declines in antibody

levels were similar or slightly slower than those after third/booster vaccinations. However, as antibody levels associated with the same level of protection against new Omicron BA.4/5 infections were lower in those with a previous infection than those without, the duration of protection after breakthrough infection was longer than after third/booster vaccination.

Studies of neutralising antibodies have shown significantly reduced protection against Omicron compared to wild type or preceding variants following two vaccinations[249–252]. I found that a third/booster vaccinations substantially increased anti-spike IgG levels and led to higher antibody levels post-booster than post-second vaccination, similar to previous studies[253–255]. This at least partially explains the effectiveness of third/booster vaccinations against Omicron infection compared to two vaccinations[234,256]. Although antibody levels post-second vaccination were much lower after ChAdOx1 than BNT162b2 primary courses, both groups were boosted to similar levels in my study, consistent with a previous VirusWatch study[257], i.e. mRNA boosters significantly increased the relatively lower antibody levels induced by an adenovirus-vectored primary vaccination course.

I found that mRNA-1273 boosters resulted in higher antibody levels 42 days after third/booster vaccinations than BNT162b2, consistent with widely reported higher antibody levels from mRNA-1273 vs. BNT162b2 after second[176,258,259] and third vaccinations[260,261]. This is likely explained by mRNA-1273 delivering 100 µg mRNA per dose (50 µg for booster doses), larger than BNT162b2 (30 µg mRNA). These results also potentially explain the higher vaccine effectiveness against new Omicron BA.1 infections reported 2-4 weeks after mRNA-1273 (70.1-73.9%) than BNT162b2 (62.4-67.2%)[237] boosters.

I found that breakthrough infections increased antibodies to similar levels to mRNA boosters, although the rate of increase was slower. Antibody declines after infection were similar to or slightly slower than after boosters, especially in those with primary BNT162b2 courses, suggesting more sustained immune responses post-infection than post-vaccination, consistent with my results in Chapter 4 and 5. Smaller studies have also found that neutralising activity was boosted by breakthrough infection after second vaccinations[258,262] similarly to after three vaccinations[263]. Taken together with the data on correlates of protection, my results indicate that breakthrough infection leads to longer-lasting immunity and thus offers more durable protection against future infections, both from the same and different variants. Sera from vaccinated individuals with breakthrough infections with pre-Omicron variants have been reported to cross-neutralise the Omicron variant, although less effectively than the Delta variant[264], and to a greater extent than sera from those without breakthrough infections[265–267].

Multiple studies have shown that antibody levels are a correlate of protection against infection, including several large studies or trials involving Alpha and Delta variant infections[174,243,244], and Omicron BA.1 infections in a small healthcare worker cohort[268]. To my knowledge, this study is the first to show higher anti-spike IgG levels are associated with increased protection against Omicron BA.4/5 infection. In Chapter 5 I estimated that the antibody level associated with 67% protection against new Delta infection was 100 BAU/mL in vaccinated individuals without prior infection[174]. Using the same 67% threshold, the antibody level required to provide the same level of protection against Omicron BA.4/5 infection was >1,000 BAU/mL in the same group, showing that much higher levels of antibodies against a wild-type trimeric spike antigen are needed to protect against new Omicron infection than new Delta infection, consistent with previous studies reporting lower neutralization of Omicron than Delta[269]. The level of protection associated

with a given antibody level was strongly affected by infection and vaccination history, with vaccination without prior infection resulting in the lowest protection at a given antibody level compared to those with both vaccination and prior/breakthrough infection. Among those with prior/breakthrough infection, protection was highest for those infected with a more recent variant. The strong association between time since last infection and variant of last infection make it difficult to determine the impact of both individually. However, given there was some variation within each variant and overlap between variants in time from previous infection, there was still no evidence of an effect of time since last infection on protection at a given antibody level having accounted for the most recent infecting variant, suggesting that protection is largely determined by the variant of the prior infection and the antibody levels achieved, rather than by time since last infection. Those with Omicron BA.2 breakthrough infections were estimated to have >80% protection against an Omicron BA.4/5 reinfection.

Using 67% protection against infection as a threshold, protection was short-lived following three vaccinations in individuals without previous infection. The population-average time from third/booster vaccination to the level associated with 67% protection was no more than 70 days for BNT162b2 boosters and 125 days for mRNA-1273 boosters. The estimated duration was longer following two vaccinations and breakthrough infection with Delta or Omicron BA.1, being 140-170 days and 180-315 days with primary ChAdOx1 and BNT162b2 courses respectively. This was partly because of slower antibody declines with primary BNT162b2 courses, but predominantly due to lower antibody levels associated with 67% protection after breakthrough infections. The proportion of individuals above the 67% threshold showed similar patterns. At 180 days, no participant with only a third/booster vaccination remained above the threshold, while all participants <55y with primary BNT162b2 courses and breakthrough infection maintained 67%

protection. This is consistent with previous studies in Qatar and Netherlands where higher protection was observed against new Omicron BA.1/BA.2 infections in people with both vaccination and previous infection than those who had only been vaccinated[270,271].

Based on this cohort, which is broadly representative of the UK population, assuming participants did not have further immune boosting events (vaccination or infection) after their third vaccination or breakthrough infection, the proportion in the UK who would remain protected at the 67% threshold would be <15% among those without a previous infection, and 50-60% among those with a previous infection by 31 December 2022. To increase population immunity, further booster vaccinations would be helpful for those who have not had a SARS-CoV-2 infection (25-35% of this cohort aged under 55y in September 2022, **Figure *6.17***), and those ≥55y with or without a previous infection. In the last quarter of 2022, a fourth vaccination is offered to those ≥50y in the UK (fifth for those ≥75y or clinically vulnerable). For younger individuals, relatively robust immunity is likely to have already been acquired from previous SARS-CoV-2 infections, and the clinical risks from new SARS-CoV-2 infections are much smaller[239,240,272].

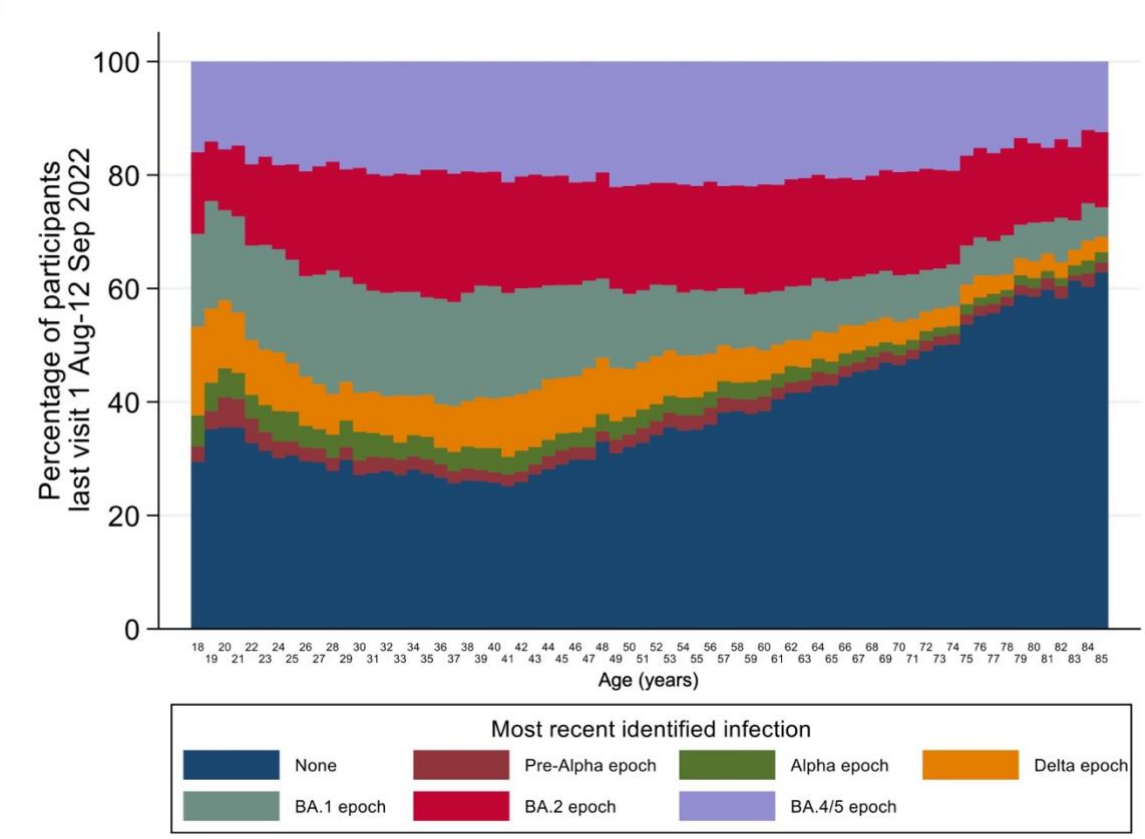


Figure 6.17 Percentages of participants' most recent previous infection by age using last assessment from 1 August 2022 to 12 September 2022.

Given these results, providing risks of hospitalisation/death and onward transmission to at-risk groups remain acceptably low, breakthrough infections may be an efficient mechanism to maintain immunity in healthy younger individuals without clinical vulnerability. However, there are still some risks associated with this approach, including ongoing circulation of variants that could put elderly and vulnerable populations at risk, and SARS-CoV-2 complications even in low-risk younger populations. Ongoing infection might also cause economic or societal consequences even with low morbidity and mortality. Nevertheless, continuing widespread use of booster vaccinations has substantial costs, both direct costs and opportunity costs from diversion of healthcare resources. Taken together with the lower effectiveness of current vaccines against Omicron infection than against earlier variants, continuing to vaccinate the whole population may have

limited benefits. For example, a cost-effectiveness analysis of COVID-19 vaccination in Kenya found vaccination of young adults may no longer be cost-effective[273]. New vaccines with higher effectiveness against Omicron variants or more sustained protection could be beneficial, but the current Omicron-specific vaccines did not elicit superior immune responses and only offered similar protection against infection to existing booster vaccines[274]. In many countries including the UK[275] and France[276], booster vaccinations are not being routinely offered to low-risk individuals. Therefore, for previously infected healthy young populations that have low risks of adverse consequences from infection, additional boosters may have limited benefits.

Study limitations include the fact that only anti-spike IgG were measured, and I assumed that associations between antibody levels and infection were constant over May-September 2022; other immune mechanisms may also provide protection against infections, including T cell and memory-based responses. I combined one, two, three, and four vaccinations together in the correlates of protection model, assuming the effect of number of vaccinations was mediated through the resulting antibody levels; power to detect heterogeneity by dose was very low as most participants had had 3 vaccinations before the study period. Residual confounding from other behavioural and epidemiological factors could exist when examining associations between previous infection and protection, e.g. those infected earlier might have a higher number of contacts thus have a higher risk of reinfection, and those with breakthrough infection might benefit from local herd immunity thus have a lower risk of reinfection. Neutralising antibody responses were not assayed in this study. Antibody levels were only measured in a single assay and models for antibody responses to booster vaccination/infection included measurements from three different dilutions to cover varying ranges of observed values over time. However, I used interval censored methods to account for different censoring thresholds, and models for correlates

of protection only used the highest dilution, in place when Omicron BA.4/5 infections dominated. Although the study design was to assess all participants every 28-42 days regardless of symptomatology, and most intervals between assessments were <45 days (**Figure *6.18***), a small number of infections could have been missed. However, to define previous infection, I used swab tests from the study, England's national testing programme, and self-reported tests to reduce misclassification. Some infections occurring between assessments may still have been missed, particularly among participants from Scotland, Wales, or Northern Ireland (14% of participants) where linked national testing data were not available, potentially leading to small overestimation of antibody levels and underestimation of protection in those with breakthrough infection. I could not model participants with a primary mRNA-1273 course due to insufficient data. I did not have data on hospitalisation or death, so I could not model the correlates of protection against more severe outcomes. Although not a limitation of my study per se, different levels of protection against infection at the same level of antibodies depending on prior infection status and the lineage of any prior infection, does potentially limit the practical application of antibodies as correlates of protection on an individual basis, particularly where the variant status of any previous infection is unknown.

In summary, both third/booster vaccination and infection post-second vaccination significantly increased anti-spike IgG levels, regardless of the primary vaccine course. Breakthrough infections had at least as strong boosting effects, and subsequent antibody declines were similar or slightly slower, than third/booster vaccinations. Based on correlates of protection against new Omicron BA.4/5 infections, protection was lower and shorter after third/booster vaccinations, but higher and longer after breakthrough infections, especially among younger individuals. These results could inform future vaccine strategies against current and potentially future Omicron variants. Providing risks

of hospitalisation/death, long-term complications, and onward transmission to at-risk groups remain acceptably low, breakthrough infections may offer good protection in healthy younger individuals without clinical vulnerability, but further research is needed to weigh the costs and benefits of booster vaccination considering my findings alongside the risk of hospitalisation, mortality, and long-term consequences of infection.

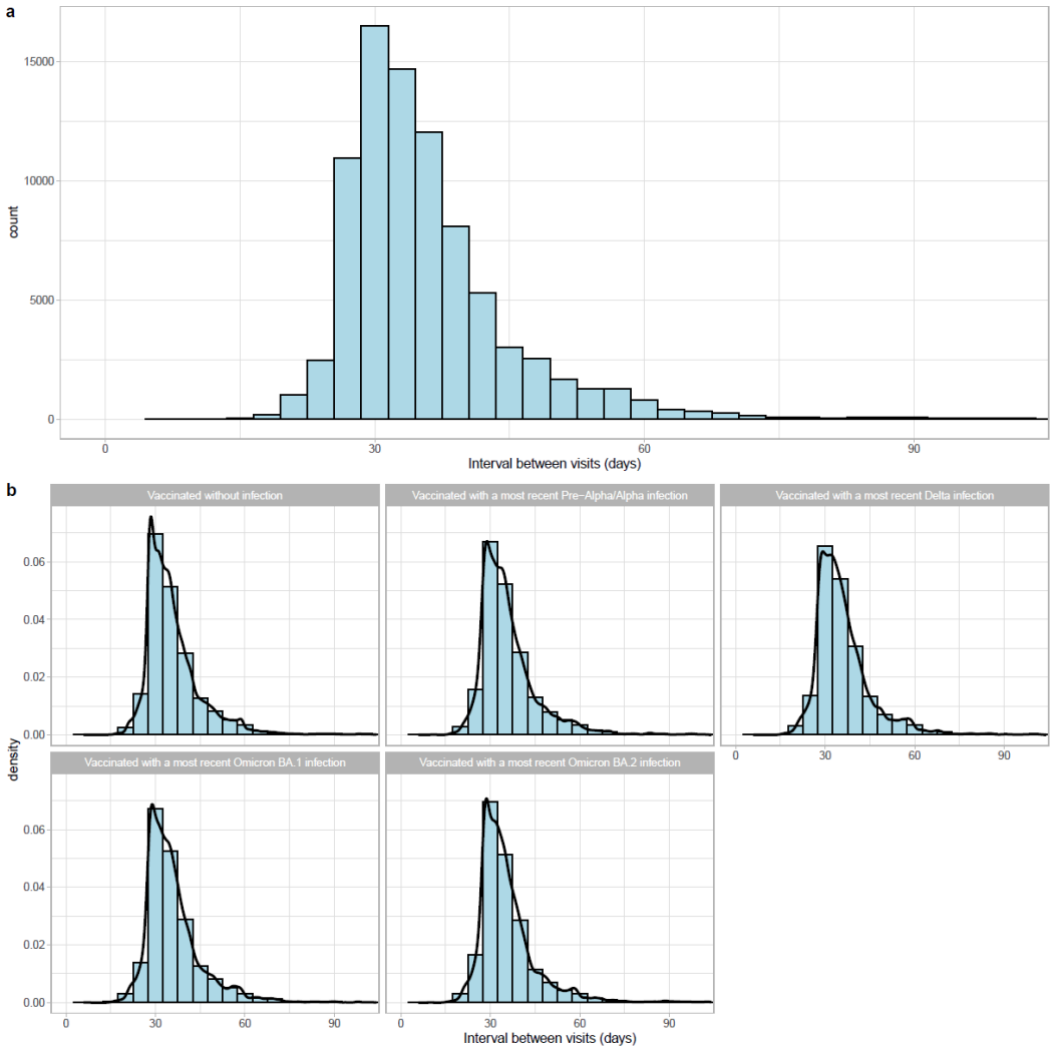


Figure 6.18 Distribution of the intervals between survey assessments included in the study.

Note: Assessments were planned on a 28–42-day cycle to achieve overall targets for swabs taken per month**.** Most intervals between assessments are <45 days, reducing the number of missed infections due to missed assessments (panel a). Additional information about positive tests taken outside the study was also used to classify previous infection status, see Methods. There was no evidence that this distribution varied by exposure groups (panel b).

Chapter 7 Conclusions and future work

Since the COVID-19 pandemic reached the UK in early 2020, more than 24 million confirmed cases of infection and more than 220,000 deaths have been reported up to July 2023[277]. In response to the pandemic, multiple vaccines have been developed, and much effort has been directed to understand both infection-induced and vaccine-induced immunity and protection against SARS-CoV-2 (re)infections.

The goal of this thesis is to examine the trajectories, levels, and durations of anti-spike IgG antibody responses after the SARS-CoV-2 natural infection and vaccination in the UK general population, examine the correlates of protection in different variant-dominant epochs, and predict the duration of protection following SARS-CoV-2 infection or vaccination against different variants.

Studies included in this thesis use data from the ONS COVID-19 Infection Survey (CIS). CIS is the largest regular survey globally of COVID-19 infections and antibodies. The overall purpose of this survey is to understand how many people of different ages across the UK have already had COVID-19, to help the government manage the pandemic and protect the NHS from being overwhelmed. It has been pivotal to the pandemic response since its initiation in April 2020. Until the survey paused in March 2023, it gathered and analysed more than 11 million swab tests and 3 million blood tests from hundreds of thousands of participants. This huge amount of serological and outcome data enabled me to carry out large-scale analyses on anti-spike IgG antibody responses. I choose several statistical approaches to serve this purpose. In Chapter 3, I introduced three models that I predominantly used for analysing antibody responses: generalised additive models (GAMs), Bayesian linear mixed models (BLMMs), and latent class mixed models (LCMMs).

GAMs allow a non-linear response to depend on the smooth functions of the predictor variables. In my analysis, I model the quantitative antibody responses over time, which show a non-linear relationship as they increase after the start of infection/vaccination, reach a peak level, and then gradually decline. GAMs can capture these non-linearities and are more flexible than other modelling techniques that require a parametric specification of the relationship between response and predictors.

LMMs are commonly used for modelling longitudinal data with repeated measurements. They can analyse unbalanced data sets with missing values at random and can model correlations between observations. LMMs are suitable for analysing antibody trajectories because antibody levels are measured in a longitudinal format with repeated measurements taken for each participant, but the number and intervals of measurements vary across different participants. Linear mixed models can include the population-level fixed effects for the whole participant population, as well as the individual-level random effects for each participant.

I used BLMMs instead of the frequentist LMMs because BLMMs have two major advantages over frequentist LMMs: 1) they can incorporate existing information by applying different prior distributions in estimating the model parameters, thus less rely on large sample sizes; 2) they can fit complex models with a large number of random variance components[248]. In my studies, the number of participants is large, while the number of antibody measurements per participant is relatively small, making it hard to fit frequentist LMMs. What is more, BLMM can incorporate interval censored methods. The antibody assay used in CIS is subject to the upper detection limit, for example, all measurements >450 BAU/mL are recorded as 450 BAU/mL in the 1:50 dilution. If I had treated all values >450 as 450 BAU/mL, antibody responses would have been

underestimated. BLMMs can add a variable to indicate whether the input value is censored or not, producing more accurate estimations of antibody levels.

LCMMs have been widely used in medical literature to uncover subpopulation structures for biomarker trajectories in longitudinal data[278–281]. Membership in latent classes is modelled with a multinomial logistic regression, and the biomarker trajectories within a latent class are described by a linear mixed model. In my studies, I hypothesize that there are distinct antibody trajectories within the population, as participants can respond differently to vaccines or infections depending on their characteristics. Therefore, LCMMs can be used to distinguish the latent subgroups of antibody responses. To my knowledge, LCMMs have not been used to model antibody levels in the literature, so using this method to model antibody trajectories after SARS-CoV-2 vaccination or infection is a novelty in my thesis.

By applying these methods to the longitudinally collected antibody data in CIS, I analysed infection-induced and vaccine-induced antibody responses in the UK general population. The main results and conclusions are summarised below.

First, there are distinct subgroups of antibody trajectories, especially a group of non-responders, after SARS-CoV-2 infection and vaccination. In Chapter 4, I analysed antibody responses after SARS-CoV-2 natural infection in the Pre-Alpha and Alpha epochs. Using an LCMM-based approach, I found that only 3 in 4 people in the UK responded to SARS-CoV-2 infection and generated a positive antibody response. The 24% non-responders could be partially attributed to the false-negative antibody results from a lower sensitivity of the antibody assay used[282], or false-positive PCR results[283]. More likely, this could represent a genuine non-response subpopulation, who were older with more long-term health conditions, fewer symptoms, and lower viral load, indicating an

impaired immunity or a milder infection. In Chapter 5, using the same approach I identified 5-6% of participants not responding to the first ChAdOx1 or BNT162b2 vaccination, who were again older and with more long-term health conditions. However, this suboptimal antibody response could be improved with a second vaccination, with only <1% of participants not generating a positive response after the second vaccination. This indicates that the second dose of the primary vaccination course is critical for people with impaired immunity to reliably seroconvert.

Second, vaccines are generally effective in generating antibody responses, with mRNA vaccines generating higher levels than adenovirus vaccines and natural infection, but antibodies after natural infection last longer. In Chapter 5, I examined the level and durability of antibody responses after the first and second SARS-CoV-2 vaccination combining results from three studies. Using GAM and BLMM approaches, I found that the first vaccination resulted in detectable antibody levels in most individuals, which were then significantly boosted by the second vaccination. BNT162b2 vaccines generated much higher peak levels than ChAdOx1 vaccines after the first and second vaccinations, although waned slightly faster. Potential mechanisms may involve the differences in the vaccine platform and the interaction with the immune system. Compared with what I found in Chapter 4 on natural infection, the peak levels after the second BNT162b2 or ChAdOx1 were both higher than that after the natural infection (959, 184 vs 113 BAU/mL), but the half-lives were shorter (51, 79 vs 184 days). These results indicate that the vaccine-induced antibody responses are stronger but short-lived, while infection-induced antibody responses are relatively lower but last longer.

The UK national booster campaign enabled me to further examine antibody responses after the booster vaccination vs. breakthrough infection post-second vaccination. In Chapter 6, I

found a substantial increase in antibody levels after booster vaccinations, regardless of booster vaccine type and primary vaccine course, indicating that the lower antibody response following a ChAdOx1 primary course can be significantly boosted by an mRNA booster. Breakthrough infection also significantly boosted antibody responses generated by a primary vaccination course, and the subsequent antibody declines were even slower than that after the boosters.

Third, multiple characteristics are associated with antibody responses, especially prior infection status. In Chapter 4, I found that females and people with non-white ethnicity had a stronger antibody response after infection, which was also found in Chapter 5 after the second vaccination. Older ages were found to be associated with a lower peak level after the second BNT162b2 vaccine, and the interval between the first and second vaccination was not associated with antibody responses post-second vaccination. The most significant effect came from prior infection status. Prior infection significantly boosted antibody responses after the first ChAdOx1, BNT162b2, or mRNA-1273 vaccine, producing higher peak levels and/or longer half-lives, comparable to or even better than two vaccinations without prior infection. After the second BNT162b2 or ChAdOX1 vaccination, individuals with prior infection again had much stronger and more sustained antibody responses than those without. Similar results were also found in other literature[284–286].

Fourth, antibody levels can predict the level of protection against future infection through 'correlates of protection'. To associate antibody responses with protection against infection, immune 'correlates of protection' need to be first determined. Correlates of protection are measurements of immune parameters that allow prediction of the degree of protection against infection. Immune correlates of protection can be used as surrogate endpoints for vaccine efficacy[287–289]. Antibody levels are commonly used because they are

easier to measure than cellular responses and are therefore more clinically useful. Usually, a protective threshold is developed to efficiently determine the proportion of the population protected. In Chapter 5, correlates of protection were determined for Delta infections, with the anti-spike IgG antibody threshold associated with 67% protection against infection to be 100 BAU/mL for both BNT162b2 and ChAdOx1 vaccines. Given antibody levels rose to higher levels following BNT162b2 than ChAdOx1, it explained the higher vaccine effectiveness after BNT162b vs. ChAdOx1. Although Chapter 4 found that natural infection generated much lower antibody levels than vaccination, infection was associated with the greatest protection at a given antibody level in the correlates of protection model.

However, the protection against variants that emerged later, such as the Omicron variants, can be much shorter due to the immune escape requiring higher antibody levels for the same level of protection[250]. Thus, the results and conclusions in Chapters 4 and 5 may not apply to the new variant and need to be updated. Therefore, in Chapter 6 I examined the correlates of protection using data from the Omicron BA.4/5 epoch and found that the antibody levels required to reach the same 67% protection were much higher than that in the Delta epoch among vaccinated individuals without prior infection, to be >1,000 BAU/mL. This proves the increased escape ability of the Omicron variant to vaccine-induced immunity and antibody neutralisation. It also shows that immune correlates of protection are variant-specific and need to be re-evaluated when new variants emerge[290]. Again, the protection was higher among those with both vaccination and infection given a certain antibody level, and the protection was higher with a more recent infection variant than an earlier variant.

Finally, protection from vaccines is relatively short-lived, while natural infection induces longer protection. Combining the antibody trajectory analyses with correlates of protection analyses in Chapters 4 and 5, I found that protection against infection was longer after the second BNT162b2 vaccination (5-8 months) than the second ChAdOx1 vaccination (2-3 months) and was the longest after natural infection (>1 year). Also, a single vaccination targeted to those previously infected may provide at least as good protection as to two-dose vaccination among those without prior infection. Similarly, in Chapter 6, I estimated that breakthrough infection provided longer-lasting protection against further Omicron infections than booster vaccinations after a primary vaccination course. This further indicates that a hybrid immunity, regardless of the order of infection or vaccination, could elicit stronger immune responses and provide better protection[291].

My results have generated several policy implications over the course of the pandemic. First, my finding that dosing interval has a limited impact on antibody levels supports the UK government's decision on extending the initial 3-4 week-doing schedule to 8-12 weeks to maximize initial vaccination coverage. Second, in terms of vaccine prioritization after a primary course, a third booster vaccination should be prioritized for ChAdOx1 recipients, more clinically vulnerable individuals, and individuals without prior infection. Third, in low- and middle-income countries where the infection prevalence is high while vaccine coverage is low, a single-dose strategy targeted to those previously infected could potentially be cost-effective in optimizing population immunity, although equitable access to full vaccination should remain as the international focus. Lastly, in terms of how to move forward in the Omicron epoch, for healthy younger individuals without clinical vulnerability, if they have had a breakthrough infection, further booster vaccinations may only have limited benefits given the opportunity costs associated with mass vaccination

programmes. This is also the current de facto policy in several countries such as the UK and France, with seasonal boosters only being given to the elderly and vulnerable groups.

The studies in my thesis are the largest studies in the UK to analyse antibody responses to SARS-CoV-2 vaccination and natural infection. These results have contributed to the growing understanding of humoral immune responses to the SARS-CoV-2 and provided evidence and implications to the UK's national pandemic planning and vaccination strategies. However, they are also subject to several limitations in general.

First, I only examined humoral anti-spike IgG antibody responses, but human immunity is more than one class of antibodies to a single antigen. T cell responses and innate immune responses also play an important role in protecting humans from infection, and while some variants like Omicron can escape humoral immunity induced by infection or vaccination because of multiple spike protein mutations, T cell responses are generally well-preserved, indicating that SARS-CoV-2 viral evasion from T cells may be limited[292,293]. Research on humoral antibody responses should be combined with research on cellular and innate immune responses to reveal the whole picture of immune protection against SARS-CoV-2 infection.

Although serum antibodies can help control viral spread within the body, other antibodies like mucosal antibodies, and particularly secretory IgA in the respiratory tract, can also help prevent the transmission of SARS-CoV-2 through the airway and may help prevent reinfection[294,295], but these types of antibodies were not analysed in my studies.

Second, the survey only collected data on SARS-CoV-2 infections and did not collect or link to the data on SARS-CoV-2 hospitalisation or death, which could inform the disease severity. I only estimated the association between antibody levels and protection against

infections, but the association between antibody levels and protection against hospitalisation or death was not reported.

Third, I did not examine the association between antibody levels and the long COVID. Longer-term consequences of COVID-19 affect a subset of the population and require further investigation.

Fourth, when estimating the correlates of protection, I assumed that the association between protection and antibody levels was constant over time, but antibodies only may not be perfectly predictive of protection, as persistent memory B-cells and T-cells may theoretically provide greater protection at lower antibody levels[296,297]. More studies will be required for understanding the mechanisms as well as the duration of memory immune responses and protective immunity to SARS-CoV-2.

Apart from these, these studies have some technical limitations. First, the antibody assay used in my studies had limited dynamic range, e.g. all measurements >450 BAU/mL were recorded as 450 BAU/mL in the 1:50 dilution. This could affect the model results and potentially underestimate the peak antibody levels. To address this, I used interval censored method in BLMMs to account for the censoring at the upper limit, which could produce more accurate results. Also, from the start of 2022 the assay dilution was changed to 1:400 and further to 1:1600 to allow for a wider dynamic range, so I could only use a subset of the data to model correlates of protection in the Omicron epoch. Second, because the study visits were performed on a monthly basis, the exact infection date could not be precisely identified. Participants could be infected between two consecutive visits but could only be detected on the date of the second visit. To make the infection date more accurate, I combined the data from the survey with the data from England's national testing programme (Test & Trace) and used the earlier recorded infection date, although

bias could still exist. Third, when defining prior infection status, I used data from the survey including PCR positive test results, positive antibody test results, and self-reported positive test results, but misclassification could exist, making it hard to perfectly identify those with prior infection, which may potentially lead to small overestimation of antibody levels and underestimation of protection in those with prior infection.

On 5 May 2023, WHO declared that SARS-CoV-2 is no longer a global health emergency[298]. However, while the world is ready to move forward, it is important to stress that the disease is not over, and new variants are still emerging, which require continuous efforts to protect the population especially the vulnerable groups. Therefore, in terms of future work, there are a few aspects for development.

First, studies in this thesis covered multiple SARS-CoV-2 infection waves in the UK, including Pre-Alpha, Alpha, Delta, Omicron BA.1, Omicron BA.2, and Omicron BA.4/5 variants. However, new Omicron subvariants are still emerging and circulating in the UK since October 2022, including BQ.1.1, which is a BA.5 sub-lineage, and XBB, which is a recombinant lineage derived from 2 previous Omicron sub-lineages[299]. Studies have shown that these subvariants had additional spike mutations causing impaired neutralization capabilities greater than earlier Omicron variants[300–303]. Given my results on higher antibody levels required for protection against Omicron than Delta, it is reasonable to speculate that these new variants may require even higher antibody levels to achieve protection. Therefore, it is important to continue monitoring SARS-CoV-2 variants, and future studies are needed to assess population immunity levels from both natural infection and vaccination against some of the newly emerged variants.

Second, since early 2022 new adapted bivalent vaccines have been developed. Pfizer-BioNTech and mRNA-1273 were adapted to encode both the ancestral wild-type spike

protein and BA.1 or BA.4/5 spike proteins, which were shown to have better neutralizing activity against the targeted Omicron variants than the original monovalent vaccines[304–307], and higher vaccine effectiveness against severe diseases[308]. However, some studies found the bivalent vaccine only had modest benefit[309] and did not generate robust neutralisation against the newly merged BQ.1.1 and XBB subvariants[301,303]. The bivalent vaccines have been approved in the UK and rolled out as part of the booster programme targeted at priority groups[310]. Further population-based studies are needed to examine the immunogenicity and effectiveness of those bivalent vaccines in the real world.

Third, this thesis only examined antibody responses based on population-level data, but studies associating the antibody responses to individual genetic variations may help better explain different immune responses and disease severity across populations. For example, a few human leukocyte antigen (HLA) alleles were found to be associated with antibody production after vaccination[311,312]. Future studies are needed to better understand the impact of human genetics on the immunogenicity of vaccines and disease outcomes, thus better informing vaccine strategies and development for existing and future variants.

Lastly, in terms of implications for future pandemics, I think CIS proves to be an important tool in assessing infection prevalence, monitoring new variants, and examining population immunity levels, which helps people understand COVID-19 and informs government decision-making throughout the pandemic. Therefore, carrying out a national survey like CIS could be beneficial to national pandemic responses if another pandemic happens in the future, although the cost of such surveys should also be taken into consideration, e.g. the cost of CIS is approximately £945 million[313]. What is more, pandemic research can be benefited from online data sharing and linkage. For example, my study did not examine correlates of protection against severe disease outcomes because mortality data was not

accessible. If mortality data could be shared, more rapid and comprehensive assessments of correlates of protection could be achieved.

In conclusion, this thesis has contributed to the understanding of the immunogenicity of SARS-CoV-2 natural infection and vaccination by analysing anti-spike IgG antibody responses in the UK general population. By examining the association between antibody responses and protection levels, I develop correlates of protection models in both Delta and Omicron periods, which are used to predict the protection level people acquired from infection or vaccination against future infections. These data have provided valuable evidence for the national and international decision-making on vaccination strategies throughout the pandemic, and have important implications to future pandemic response. Future studies should focus on continuous monitoring of the newly emerged and potentially future variants, understanding the underlying genetic factors associated with infection- or vaccine-induced immunity, and examining the immunogenicity and effectiveness of the new vaccines.

References

1. Coronavirus. https://www.who.int/health-topics/coronavirus#tab=tab_1.

2. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **382**, 727–733 (2020).

3. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **323**, 1239–1242 (2020).

4. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

5. Gorbalenya, A. E. *et al.* The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology 2020 5:4* **5**, 536–544 (2020).

6. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

7. Hu, B., Guo, H., Zhou, P. & Shi, Z. L. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology 2020 19:3* **19**, 141–154 (2020).

8. Wang, M. Y. *et al.* SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front Cell Infect Microbiol* **10**, 724 (2020).

9. Long-term effects of coronavirus (long COVID) - NHS. https://www.nhs.uk/conditions/coronavirus-covid-19/long-term-effects-of-coronavirus-long-covid/.

10. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. https://covid19.who.int/.

11. England Summary | Coronavirus (COVID-19) in the UK. https://coronavirus.data.gov.uk/.

12. Virology: Coronaviruses. *Nature* **220**, 650 (1968).

13. Chen, B. *et al.* Overview of lethal human coronaviruses. *Signal Transduction and Targeted Therapy 2020 5:1* **5**, 1–16 (2020).

14. Zhu, Z. *et al.* From SARS and MERS to COVID-19: A brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respir Res* **21**, 1–14 (2020).

15. Chan, J. F. W. *et al.* Middle East Respiratory syndrome coronavirus: Another zoonotic betacoronavirus causing SARS-like disease. *Clin Microbiol Rev* **28**, 465–522 (2015).

16. De Wit, E., Van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology 2016 14:8* **14**, 523–534 (2016).

17. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature 2020 579:7798* **579**, 265–269 (2020).

18. Jamison, D. A. *et al.* A comprehensive SARS-CoV-2 and COVID-19 review, Part 1: Intracellular overdrive for SARS-CoV-2 infection. *European Journal of Human Genetics 2022 30:8* **30**, 889–898 (2022).

19. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology 2020 19:3* **19**, 155–170 (2020).

20. Harrison, A. G., Lin, T. & Wang, P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol* **41**, 1100–1115 (2020).

21. SeyedAlinaghi, S. A. *et al.* Characterization of SARS-CoV-2 different variants and related morbidity and mortality: a systematic review. *Eur J Med Res* **26**, 1–20 (2021).

22. Variants: distribution of cases data, 20 May 2021 - GOV.UK. https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers/variants-distribution-of-cases-data#variant1.

23. Shahhosseini, N., Babuadze, G., Wong, G. & Kobinger, G. P. Mutation signatures and in silico docking of novel sars-cov-2 variants of concern. *Microorganisms* **9**, (2021).

24. Tracking SARS-CoV-2 variants. https://www.who.int/activities/tracking-SARS-CoV-2-variants.

25. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).

26. Chan, K. K., Tan, T. J. C., Narayanan, K. K. & Procko, E. An engineered decoy receptor for SARS-CoV-2 broadly binds protein S sequence variants. *Sci Adv* **7**, (2021).

27. Volz, E. *et al.* Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature 2021 593:7858* **593**, 266–269 (2021).

28. Meng, B. *et al.* Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep* **35**, (2021).

29. Davies, N. G. *et al.* Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, (2021).

30. Graham, M. S. *et al.* Changes in symptomatology, reinfection, and transmissibility associated with the SARS-CoV-2 variant B.1.1.7: an ecological study. *Lancet Public Health* **6**, e335 (2021).

31. COVID-19: The Second Wave May Not Be the Last – but Which One Will Be? – The Wire Science. https://science.thewire.in/health/covid-19-the-second-wave-may-not-be-the-last-but-which-one-will-be/.

32. Mishra, S. *et al.* Changing composition of SARS-CoV-2 lineages and rise of Delta variant in England. *EClinicalMedicine* **39**, 101064 (2021).

33. Dhawan, M. *et al.* Delta variant (B.1.617.2) of SARS-CoV-2: Mutations, impact, challenges and possible solutions. *Hum Vaccin Immunother* **18**, (2022).

34. Liu, C. *et al.* Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* **184**, 4220-4236.e13 (2021).

35. Mlcochova, P. *et al.* SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature 2021 599:7883* **599**, 114–119 (2021).

36. Wall, E. C. *et al.* Neutralising antibody activity against SARS-CoV-2 VOCs B.1.617.2 and B.1.351 by BNT162b2 vaccination. *The Lancet* **397**, 2331–2333 (2021).

37. Twohig, K. A. *et al.* Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis* **22**, 35 (2022).

38. Sheikh, A., McMenamin, J., Taylor, B. & Robertson, C. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *Lancet* **397**, 2461 (2021).

39. Fisman, D. N. & Tuite, A. R. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada. *CMAJ : Canadian Medical Association Journal* **193**, E1619 (2021).

40. Enhancing response to Omicron SARS-CoV-2 variant. https://www.who.int/publications/m/item/enhancing-readiness-for-Omicron-(B.1.1.529)-technical-brief-and-priority-actions-for-member-states.

41. Antonelli, M., Pujol, J. C., Spector, T. D., Ourselin, S. & Steves, C. J. Risk of long COVID associated with delta versus omicron variants of SARS-CoV-2. *Lancet* **399**, 2263 (2022).

42. Cele, S. *et al.* Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature 2021 602:7898* **602**, 654–656 (2021).

43. Iketani, S. *et al.* Antibody evasion properties of SARS-CoV-2 Omicron sublineages. *Nature 2022 604:7906* **604**, 553–556 (2022).

44. Yu, J. *et al.* Neutralization of the SARS-CoV-2 Omicron BA.1 and BA.2 Variants. *New England Journal of Medicine* **386**, 1579–1580 (2022).

45. Report 50 - Hospitalisation risk for Omicron cases in England | Faculty of Medicine | Imperial College London. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-50-severity-omicron/.

46. Shrestha, L. B., Foster, C., Rawlinson, W., Tedla, N. & Bull, R. A. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: Implications for immune escape and transmission. *Rev Med Virol* **32**, e2381 (2022).

47. Investigation of SARS-CoV-2 variants: technical briefings - GOV.UK. https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings.

48. Krammer, F. SARS-CoV-2 vaccines in development. *Nature 2020 586:7830* **586**, 516–527 (2020).

49. Mascellino, M. T., Di Timoteo, F., De Angelis, M. & Oliva, A. Overview of the Main Anti-SARS-CoV-2 Vaccines: Mechanism of Action, Efficacy and Safety. *Infect Drug Resist* **14**, 3459 (2021).

50. Li, Y. Der *et al.* Coronavirus vaccine development: from SARS and MERS to COVID-19. *Journal of Biomedical Science 2020 27:1* **27**, 1–23 (2020).

51. Watson, O. J. *et al.* Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect Dis* **22**, 1293–1302 (2022).

52. Barouch, D. H. Covid-19 Vaccines — Immunity, Variants, Boosters. *New England Journal of Medicine* **387**, 1011–1020 (2022).

53. Pardi, N., Hogan, M. J., Porter, F. W. & Weissman, D. mRNA vaccines — a new era in vaccinology. *Nature Reviews Drug Discovery 2018 17:4* **17**, 261–279 (2018).

54. Creech, C. B., Walker, S. C. & Samuels, R. J. SARS-CoV-2 Vaccines. *JAMA* **325**, 1318–1320 (2021).

55. Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **383**, 2603–2615 (2020).

56. Baden, L. R. *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine* **384**, 403–416 (2021).

57. 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. *The Lancet* **397**, 99–111 (2021).

58. Greinacher, A. *et al.* Thrombotic Thrombocytopenia after ChAdOx1 nCov-19 Vaccination. *New England Journal of Medicine* **384**, 2092–2101 (2021).

59. Schultz, N. H. *et al.* Thrombosis and Thrombocytopenia after ChAdOx1 nCoV-19 Vaccination. *New England Journal of Medicine* **384**, 2124–2130 (2021).

60. European Countries Suspend Use of AstraZeneca's COVID-19 Vaccine. https://pharmanewsintel.com/news/european-countries-suspend-use-of-astrazenecas-covid-19-vaccine.

61. Medicines and Healthcare products Regulatory Agency. Regulatory approval of Pfizer/BioNTech vaccine for COVID-19 - GOV.UK. https://www.gov.uk/government/publications/regulatory-approval-of-pfizer-biontech-vaccine-for-covid-19 (2020).

62. Medicines and Healthcare products Regulatory Agency. Oxford University/AstraZeneca COVID-19 vaccine approved - GOV.UK. https://www.gov.uk/government/news/oxford-universityastrazeneca-covid-19-vaccine-approved (2020).

63. Department of Health and Social Care. *UK COVID-19 vaccines delivery plan Contents*. https://www.gov.uk/government/publications/uk-covid-19-vaccines-delivery-plan (2021).

64. Optimising the COVID-19 vaccination programme for maximum short-term impact - GOV.UK. https://www.gov.uk/government/publications/prioritising-the-first-covid-19-vaccine-dose-jcvi-statement/optimising-the-covid-19-vaccination-programme-for-maximum-short-term-impact.

65. Vaccination roll out update - St Mary's Surgery. https://www.stmarysely.nhs.uk/2021/03/04/vaccination-roll-out-update/.

66. Young people aged 12 to 15 to be offered a COVID-19 vaccine - GOV.UK. https://www.gov.uk/government/news/young-people-aged-12-to-15-to-be-offered-a-covid-19-vaccine.

67. Most vulnerable to be offered COVID-19 booster vaccines from next week - GOV.UK. https://www.gov.uk/government/news/most-vulnerable-to-be-offered-covid-19-booster-vaccines-from-next-week.

68. JCVI advises a spring COVID-19 vaccine dose for the most vulnerable - GOV.UK. https://www.gov.uk/government/news/jcvi-advises-a-spring-covid-19-vaccine-dose-for-the-most-vulnerable.

69. JCVI statement on the COVID-19 booster vaccination programme for autumn 2022: update 3 September 2022 - GOV.UK. https://www.gov.uk/government/publications/covid-19-vaccines-for-autumn-2022-jcvi-advice-15-august-2022/jcvi-statement-on-the-covid-19-booster-vaccination-programme-for-autumn-2022-update-15-august-2022.

70. Vaccinations in the UK | Coronavirus in the UK. https://coronavirus.data.gov.uk/details/vaccinations (2021).

71. Eter, P., Elves, J. D. & Oitt, V. M. R. The Immune System. *https://doi.org/10.1056/NEJM200007063430107* **343**, 37–49 (2000).

72. Cox, R. J. & Brokstad, K. A. Not just antibodies: B cells and T cells mediate immunity to COVID-19. *Nature Reviews Immunology 2020 20:10* **20**, 581–582 (2020).

73. Qu, J. *et al.* Profile of Immunoglobulin G and IgM Antibodies Against Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Clin Infect Dis* **71**, 2255–2258 (2020).

74. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* **26**, 1200–1204 (2020).

75. Perreault, J. *et al.* Waning of SARS-CoV-2 RBD antibodies in longitudinal convalescent plasma samples within four months after symptom onset. *Blood* (2020) doi:10.1182/blood.2020008367.

76. Zhou, W. *et al.* The dynamic changes of serum IgM and IgG against SARS-CoV-2 in patients with COVID-19. *J Med Virol* jmv.26353 (2020) doi:10.1002/jmv.26353.

77. Lumley, S. F. *et al.* The Duration, Dynamics, and Determinants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Antibody Responses in Individual Healthcare Workers. *Clinical Infectious Diseases* **73**, e699–e709 (2021).

78. Ibarrondo, F. J. *et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *The New England journal of medicine* vol. 383 1085–1087 Preprint at https://doi.org/10.1056/NEJMc2025179 (2020).

79. COVID-19 Infection Survey participants thanked for 'huge contribution' to pandemic response - GOV.UK. https://www.gov.uk/government/news/covid-19-infection-survey-participants-thanked-for-huge-contribution-to-pandemic-response.

80. Coronavirus (COVID-19) Infection Survey: methods and further information - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/covid19infectionsurveypilotmethodsandfurtherinformation.

81. Pouwels, K. B. *et al.* Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. *Lancet Public Health* **6**, e30–e38 (2021).

82. Pritchard, E. *et al.* Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. *Nature Medicine 2021 27:8* **27**, 1370–1378 (2021).

83. Ainsworth, M. *et al.* Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis* **20**, 1390–1400 (2020).

84. Supasa, P. *et al.* Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell* **184**, 2201-2211.e7 (2021).

85. Dejnirattisai, W. *et al.* The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell* **184**, 2183-2200.e22 (2021).

86. Huo, J. *et al.* Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe* **28**, 445-454.e6 (2020).

87. Interim guidelines for detecting cases of reinfection by SARS-CoV-2 - PAHO/WHO | Pan American Health Organization. https://www.paho.org/en/documents/interim-guidelines-detecting-cases-reinfection-sars-cov-2.

88. Hastie, T. & Tibshirani, R. Generalized Additive Models. *https://doi.org/10.1214/ss/1177013604* **1**, 297–310 (1986).

89. Wood, S. N. Generalized additive models: An introduction with R, second edition. *Generalized Additive Models: An Introduction with R, Second Edition* 1–476 (2017) doi:10.1201/9781315370279/GENERALIZED-ADDITIVE-MODELS-SIMON-WOOD.

90. Pedersen, E. J., Miller, D. L., Simpson, G. L. & Ross, N. Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **2019**, (2019).

91. Wood, S. N. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. *J R Stat Soc Series B Stat Methodol* **73**, 3–36 (2011).

92. CRAN - Package mgcv. https://cran.r-project.org/web/packages/mgcv/index.html.

93. Wood, S. N. Thin plate regression splines. *J R Stat Soc Series B Stat Methodol* **65**, 95–114 (2003).

94. Wood, S. N. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. *Biometrics* **62**, 1025–1036 (2006).

95. Wood, S. N., Goude, Y. & Shaw, S. Generalized additive models for large data sets. *J R Stat Soc Ser C Appl Stat* **64**, 139–155 (2015).

96. Cnaan, A., Laird, N. M. & Slasor, P. TUTORIAL IN BIOSTATISTICS USING THE GENERAL LINEAR MIXED MODEL TO ANALYSE UNBALANCED REPEATED MEASURES AND LONGITUDINAL DATA. *Stat Med* **16**, 2349–2380 (1997).

97. Fox, J. Applied regression analysis & genealized linear models. *Sage* 5–816 (2016).

98. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* **67**, 1–48 (2015).

99. Proust-Lima, C., Philipps, V. & Liquet, B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. *J Stat Softw* **78**, (2015).

100. McCulloch, C. E., Lin, H., Slate, E. H. & Turnbull, B. W. Discovering subpopulation structure with latent class mixed models. *Stat Med* **21**, 417–429 (2002).

101. Lee, P. M. Bayesian statistics : an introduction. (2012).

102. Fornacon-Wood, I. *et al.* Understanding the Differences Between Bayesian and Frequentist Statistics. *Int J Radiat Oncol Biol Phys* **112**, 1076–1082 (2022).

103. van Zyl, C. J. J. Frequentist and Bayesian inference: A conceptual primer. *New Ideas Psychol* **51**, 44–49 (2018).

104. Sorensen, T., Hohenstein, S. & Vasishth, S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quant Method Psychol* **12**, 175–200 (2016).

105. Bürkner, P. C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* **80**, 1–28 (2017).

106. Lewandowski, D., Kurowicka, D. & Joe, H. Generating random correlation matrices based on vines and extended onion method. *J Multivar Anal* **100**, 1989–2001 (2009).

107. Hamra, G., MacLehose, R. & Richardson, D. Markov Chain Monte Carlo: an introduction for epidemiologists. *Int J Epidemiol* **42**, 627–634 (2013).

108. Burkner, P. Brms: An R package for Bayesian generalized linear mixed models using Stan. *J Stat Softw* (2013).

109. Roy, V. Convergence Diagnostics for Markov Chain Monte Carlo. *https://doi.org/10.1146/annurev-statistics-031219-041300* **7**, 387–412 (2020).

110. Overbaugh, J. Understanding protection from SARS-CoV-2 by studying reinfection. *Nat Med* **26**, 1680–1681 (2020).

111. Hall, V. J. *et al.* SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN). *The Lancet* **397**, 1459–1469 (2021).

112. Lumley, S. F. *et al.* Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. *New England Journal of Medicine* **384**, 533–540 (2021).

113. Bubar, K. M. *et al.* Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science (1979)* **371**, 916–921 (2021).

114. Suhandynata, R. T. *et al.* Longitudinal Monitoring of SARS-CoV-2 IgM and IgG Seropositivity to Detect COVID-19. *J Appl Lab Med* **5**, 908–920 (2020).

115. To, K. K. W. *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* **20**, 565–574 (2020).

116. Long, Q. X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* **26**, 845–848 (2020).

117. Crawford, K. H. D. *et al.* Dynamics of Neutralizing Antibody Titers in the Months After Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *J Infect Dis* **223**, 197–205 (2021).

118. Wang, K. *et al.* Longitudinal Dynamics of the Neutralizing Antibody Response to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection. *Clinical Infectious Diseases* **73**, e531–e539 (2021).

119. Oved, K. *et al.* Multi-center nationwide comparison of seven serology assays reveals a SARS-CoV-2 non-responding seronegative subpopulation. *EClinicalMedicine* **29–30**, (2020).

120. van Elslande, J. *et al.* Longitudinal follow-up of IgG anti-nucleocapsid antibodies in SARS-CoV-2 infected patients up to eight months after infection. *Journal of Clinical Virology* **136**, (2021).

121. Petersen, L. R. *et al.* Lack of Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in a Large Cohort of Previously Infected Persons. *Clinical Infectious Diseases* (2020) doi:10.1093/cid/ciaa1685.

122. Muecksch, F. *et al.* Longitudinal Serological Analysis and Neutralizing Antibody Levels in Coronavirus Disease 2019 Convalescent Patients. *J Infect Dis* **223**, 389–398 (2021).

123. Lumley, S. F. *et al.* The Duration, Dynamics, and Determinants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Antibody Responses in Individual Healthcare Workers. *Clinical Infectious Diseases* **73**, e699–e709 (2021).

124. Harvala, H. *et al.* Convalescent plasma therapy for the treatment of patients with COVID-19: Assessment of methods available for antibody detection and their correlation with neutralising antibody levels. *Transfusion Medicine* **31**, 167–175 (2021).

125. Ibarrondo, F. J. *et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *The New England journal of medicine* vol. 383 1085–1087 Preprint at https://doi.org/10.1056/NEJMc2025179 (2020).

126. Patel, M. M. *et al.* Change in Antibodies to SARS-CoV-2 over 60 Days among Health Care Personnel in Nashville, Tennessee. *JAMA - Journal of the American Medical Association* vol. 324 1781–1782 Preprint at https://doi.org/10.1001/jama.2020.18796 (2020).

127. Anna, F. *et al.* High seroprevalence but short-lived immune response to SARS-CoV-2 infection in Paris. *Eur J Immunol* **51**, 180–190 (2021).

128. Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e21 (2020).

129. Dan, J. M. *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science (1979)* **371**, (2021).

130. van Elslande, J., Gruwier, L., Godderis, L. & Vermeersch, P. Estimated half-life of SARS-CoV-2 anti-spike antibodies more than double the half-life of anti-nucleocapsid antibodies in healthcare workers. *Clinical Infectious Diseases* (2021) doi:10.1093/cid/ciab219.

131. Vanshylla, K. *et al.* Kinetics and correlates of the neutralizing antibody response to SARS-CoV-2 infection in humans. *Cell Host Microbe* **29**, 917-929.e4 (2021).

132. Pritchard, E. *et al.* Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. *Nature Medicine 2021 27:8* **27**, 1370–1378 (2021).

133. Proust-Lima, C., Séne, M., Taylor, J. M. G. & Jacqmin-Gadda, H. Joint latent class models for longitudinal and time-to-event data: A review. *Statistical Methods in Medical Research* vol. 23 74–90 Preprint at https://doi.org/10.1177/0962280212445839 (2014).

134. Antia, A. *et al.* Heterogeneity and longevity of antibody memory to viruses and vaccines. *PLoS Biol* **16**, (2018).

135. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat Med* 1–7 (2021) doi:10.1038/s41591-021-01377-8.

136. Baron, R. C. *et al.* Frequency of serological non-responders and false-negative RT-PCR results in SARS-CoV-2 testing: A population-based study. *Clin Chem Lab Med* **58**, 2131–2140 (2020).

137. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* **26**, 1200–1204 (2020).

138. Marklund, E. *et al.* Serum-IgG responses to SARS-CoV-2 after mild and severe COVID-19 infection and analysis of IgG non-responders. *PLoS One* **15**, (2020).

139. Rijkers, G. *et al.* Differences in antibody kinetics and functionality between severe and mild severe acute respiratory syndrome Coronavirus 2 infections. *Journal of Infectious Diseases* **222**, 1265–1269 (2020).

140. Pal, R. *et al.* Impaired anti-SARS-CoV-2 antibody response in non-severe COVID-19 patients with diabetes mellitus: A preliminary report. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews* **15**, 193–196 (2021).

141. Kennedy, N. A. *et al.* Anti-SARS-CoV-2 antibody responses are attenuated in patients with IBD treated with infliximab. *Gut* **70**, 865–875 (2021).

142. Chen, J. S. *et al.* Nonsteroidal Anti-inflammatory Drugs Dampen the Cytokine and Antibody Response to SARS-CoV-2 Infection. *J Virol* **95**, (2021).

143. Betrains, A. *et al.* Convalescent plasma treatment of persistent severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infection in patients with lymphoma with impaired humoral immunity and lack of neutralising antibodies. *British Journal of Haematology* vol. 192 1100–1105 Preprint at https://doi.org/10.1111/bjh.17266 (2021).

144. Yamamoto, S. *et al.* Antibody response to SARS-CoV-2 in people living with HIV. *Journal of Microbiology, Immunology and Infection* vol. 54 144–146 Preprint at https://doi.org/10.1016/j.jmii.2020.09.005 (2021).

145. Walker, A. S. *et al.* CT threshold values, a proxy for viral load in community sars-cov-2 cases, demonstrate wide variation across populations and over time. *Elife* **10**, (2021).

146. Wolf, A., Hulmes, J. & Hopkins, S. *Lateral flow device specificity in phase 4 (post marketing) surveillance*. https://www.ox.ac.uk/sites/files/oxford/media_wysiwyg/UK%20evaluation_PHE%20Porton%20Down%20%2 (2021).

147. Cohen, K. W. *et al.* Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep Med* **2**, 100354 (2021).

148. Takahashi, T. *et al.* Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature* **588**, 315–320 (2020).

149. Bunders, M. J. & Altfeld, M. Implications of Sex Differences in Immunity for SARS-CoV-2 Pathogenesis and Design of Therapeutic Interventions. *Immunity* vol. 53 487–495 Preprint at https://doi.org/10.1016/j.immuni.2020.08.003 (2020).

150. Grzelak, L. *et al.* Sex Differences in the Evolution of Neutralizing Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2. *J Infect Dis* **224**, 983–988 (2021).

151. Zeng, F. *et al.* A comparison study of SARS-CoV-2 IgG antibody between male and female COVID-19 patients: A possible reason underlying different outcome between sex. *J Med Virol* **92**, 2050–2054 (2020).

152. Shields, A. M. *et al.* Serological responses to SARS-CoV-2 following non-hospitalised infection: clinical and ethnodemographic features associated with the magnitude of the antibody response. *medRxiv* (2020) doi:10.1101/2020.11.12.20230763.

153. Garcia-Beltran, W. F. *et al.* COVID-19-neutralizing antibodies predict disease severity and survival. *Cell* **184**, 476-488.e11 (2021).

154. Pelleau, S. *et al.* Kinetics of the Severe Acute Respiratory Syndrome Coronavirus 2 Antibody Response and Serological Estimation of Time Since Infection. *J Infect Dis* **22**, 33 (2021).

155. Wheatley, A. K. *et al.* Evolution of immune responses to SARS-CoV-2 in mild-moderate COVID-19. *Nature Communications 2021 12:1* **12**, 1–11 (2021).

156. Chia, W. N. *et al.* Dynamics of SARS-CoV-2 neutralising antibody responses and duration of immunity: a longitudinal study. *Lancet Microbe* **2**, e240–e249 (2021).

157. Wajnberg, A. *et al.* Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science (1979)* **370**, 1227–1230 (2020).

158. Eyre, D. W. *et al.* Quantitative SARS-CoV-2 anti-spike responses to Pfizer-BioNTech and Oxford-AstraZeneca vaccines by previous infection status. *medRxiv* 2021.03.21.21254061 (2021) doi:10.1101/2021.03.21.21254061.

159. Ebinger, J. E. *et al.* Prior COVID-19 Infection and Antibody Response to Single Versus Double Dose mRNA SARS-CoV-2 Vaccination. *medRxiv* (2021) doi:10.1101/2021.02.23.21252230.

160. Ciccone, E. J. *et al.* SARS-CoV-2 seropositivity after infection and antibody response to mRNA-based vaccination. *medRxiv* 2021.02.09.21251319 (2021) doi:10.1101/2021.02.09.21251319.

161. Saadat, S. *et al.* Single Dose Vaccination in Healthcare Workers Previously Infected with SARS-CoV-2. *medRxiv* 2021.01.30.21250843 (2021) doi:10.1101/2021.01.30.21250843.

162. Subbarao, S. *et al.* Robust antibody responses in 70–80-year-olds 3 weeks after the first or second doses of Pfizer/BioNTech COVID-19 vaccine, United Kingdom, January to February 2021. *Eurosurveillance* **26**, 2100329 (2021).

163. Parry, H. M. *et al.* BNT162b2 Vaccination in People Over 80 Years of Age Induces Strong Humoral Immune Responses with Cross Neutralisation of P.1 Brazilian Variant. *SSRN Electronic Journal* (2021) doi:10.2139/ssrn.3816840.

164. Prioritising the first COVID-19 vaccine dose: JCVI statement - GOV.UK. https://www.gov.uk/government/publications/prioritising-the-first-covid-19-vaccine-dose-jcvi-statement.

165. Parry, H. *et al.* Extended interval BNT162b2 vaccination enhances peak antibody generation in older people. *medRxiv* 2021.05.15.21257017 (2021) doi:10.1101/2021.05.15.21257017.

166. Payne, R. P. *et al.* Immunogenicity of standard and extended dosing intervals of BNT162b2 mRNA vaccine. *Cell* **184**, 5699-5714.e11 (2021).

167. Amirthalingam, G. *et al.* Serological responses and vaccine effectiveness for extended COVID-19 vaccine schedules in England. *Nature Communications 2021 12:1* **12**, 1–9 (2021).

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

169. Bernal, J. L. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *New England Journal of Medicine* **385**, 585–594 (2021).

170. Pouwels, K. B. *et al.* Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nature Medicine 2021 27:12* **27**, 2127–2135 (2021).

171. Most vulnerable offered second dose of COVID-19 vaccine earlier to help protect against variants - GOV.UK. https://www.gov.uk/government/news/most-vulnerable-offered-second-dose-of-covid-19-vaccine-earlier-to-help-protect-against-variants (2021).

172. Feng, S. *et al.* Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nature Medicine 2021 27:11* **27**, 2032–2040 (2021).

173. Gilbert, P. B. *et al.* Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science (1979)* eab3435 (2021).

174. Wei, J. *et al.* Antibody responses and correlates of protection in the general population after two doses of the ChAdOx1 or BNT162b2 vaccines. *Nature Medicine 2022* 1–11 (2022) doi:10.1038/s41591-022-01721-6.

175. Wei, J. *et al.* Anti-spike antibody response to natural SARS-CoV-2 infection in the general population. *Nature Communications 2021 12:1* **12**, 1–12 (2021).

176. Montoya, J. G. *et al.* Differences in IgG Antibody Responses following BNT162b2 and mRNA-1273 SARS-CoV-2 Vaccines. *Microbiol Spectr* **9**, (2021).

177. *SARS-CoV-2 variants of concern and variants under investigation in England*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1025827/Technical_Briefing_25.pdf (2021).

178. Noori, M. *et al.* Potency of BNT162b2 and mRNA-1273 vaccine-induced neutralizing antibodies against severe acute respiratory syndrome-CoV-2 variants of concern: A systematic review of in vitro studies. *Rev Med Virol* e2277 (2021) doi:10.1002/RMV.2277.

179. Ward, H. *et al.* REACT-2 Round 5: increasing prevalence of SARS-CoV-2 antibodies demonstrate impact of the second wave and of vaccine roll-out in England. *medRxiv* 2021.02.26.21252512 (2021) doi:10.1101/2021.02.26.21252512.

180. Abu Jabal, K. *et al.* Impact of age, ethnicity, sex and prior infection status on immunogenicity following a single dose of the BNT162b2 mRNA COVID-19 vaccine: real-world evidence from healthcare workers, Israel, December 2020 to January 2021. *Eurosurveillance* **26**, (2021).

181. Prendecki, M. *et al.* Effect of previous SARS-CoV-2 infection on humoral and T-cell responses to single-dose BNT162b2 vaccine. *Lancet* **0**, (2021).

182. Chang, W. H. A review of vaccine effects on women in light of the COVID-19 pandemic. *Taiwanese Journal of Obstetrics and Gynecology* vol. 59 812–820 Preprint at https://doi.org/10.1016/j.tjog.2020.09.006 (2020).

183. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nature Reviews Immunology* vol. 16 626–638 Preprint at https://doi.org/10.1038/nri.2016.90 (2016).

184. Zhong, D. *et al.* Durability of Antibody Levels After Vaccination With mRNA SARS-CoV-2 Vaccine in Individuals With or Without Prior Infection. *JAMA* **326**, 2524–2526 (2021).

185. Urbanowicz, R. A. *et al.* Two doses of the SARS-CoV-2 BNT162b2 vaccine enhance antibody responses to variants in individuals with prior SARS-CoV-2 infection. *Sci Transl Med* **13**, 847 (2021).

186. Buonfrate, D. *et al.* Antibody response induced by the BNT162b2 mRNA COVID-19 vaccine in a cohort of health-care workers, with or without prior SARS-CoV-2 infection: a prospective study. *Clin Microbiol Infect* **27**, 1845–1850 (2021).

187. Fraley, E. *et al.* The Impact of Prior Infection and Age on Antibody Persistence After Severe Acute Respiratory Syndrome Coronavirus 2 Messenger RNA Vaccine. *Clinical Infectious Diseases* (2021) doi:10.1093/CID/CIAB850.

188. Vicenti, I. *et al.* The second dose of the BNT162b2 mRNA vaccine does not boost SARS-CoV-2 neutralizing antibody response in previously infected subjects. *Infection* **1**, 1 (2021).

189. Krammer, F. *et al.* Antibody Responses in Seropositive Persons after a Single Dose of SARS-CoV-2 mRNA Vaccine. *New England Journal of Medicine* **384**, 1372–1374 (2021).

190. Ebinger, J. E. *et al.* Antibody responses to the BNT162b2 mRNA vaccine in individuals previously infected with SARS-CoV-2. *Nature Medicine 2021 27:6* **27**, 981–984 (2021).

191. Claro, F., Silva, D., Rodriguez, M., Rangel, H. R. & de Waard, J. H. Immunoglobulin G antibody response to the Sputnik V vaccine: previous SARS-CoV-2 seropositive individuals may need just one vaccine dose. *International Journal of Infectious Diseases* **111**, 261–266 (2021).

192. Terpos, E. *et al.* Age-dependent and gender-dependent antibody responses against SARS-CoV-2 in health workers and octogenarians after vaccination with the BNT162b2 mRNA vaccine. *Am J Hematol* **96**, E257–E259 (2021).

193. Amodio, E. *et al.* Antibodies Responses to SARS-CoV-2 in a Large Cohort of Vaccinated Subjects and Seropositive Patients. *Vaccines 2021, Vol. 9, Page 714* **9**, 714 (2021).

194. Ward, H. *et al.* Vaccine uptake and SARS-CoV-2 antibody prevalence among 207,337 adults during May 2021 in England: REACT-2 study. *medRxiv* 2021.07.14.21260497 (2021) doi:10.1101/2021.07.14.21260497.

195. Li, C.-Y. & Sung, E.-C. *A review of the healthy worker effect in occupational epidemiology*. *Occup. Med* vol. 49 https://academic.oup.com/occmed/article/49/4/225/1387118 (1999).

196. Shields, A. M. *et al.* Serological responses to SARS-CoV-2 following non-hospitalised infection: clinical and ethnodemographic features associated with the magnitude of the antibody response. *BMJ Open Respir Res* **8**, e000872 (2021).

197. Levin, E. G. *et al.* Waning Immune Humoral Response to BNT162b2 Covid-19 Vaccine over 6 Months. *New England Journal of Medicine* **385**, e84 (2021).

198. Favresse, J. *et al.* Antibody titres decline 3-month post-vaccination with BNT162b2. *Emerg Microbes Infect* **10**, 1495–1498 (2021).

199. Van Praet, J. T. *et al.* Dynamics of the Cellular and Humoral Immune Response After BNT162b2 Messenger Ribonucleic Acid Coronavirus Disease 2019 (COVID-19) Vaccination in COVID-19-Naive Nursing Home Residents. *J Infect Dis* **224**, 1690–1693 (2021).

200. Maeda, K. *et al.* Correlates of neutralizing/SARS-CoV-2-S1-binding antibody response with adverse effects and immune kinetics in BNT162b2-vaccinated individuals. *Scientific Reports 2021 11:1* **11**, 1–10 (2021).

201. Sadarangani, M., Marchant, A. & Kollmann, T. R. Immunological mechanisms of vaccine-induced protection against COVID-19 in humans. *Nature Reviews Immunology 2021 21:8* **21**, 475–484 (2021).

202. Tregoning, J. S., Flight, K. E., Higham, S. L., Wang, Z. & Pierce, B. F. Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nature Reviews Immunology 2021 21:10* **21**, 626–636 (2021).

203. Liu, X. *et al.* Safety and immunogenicity of heterologous versus homologous prime-boost schedules with an adenoviral vectored and mRNA COVID-19 vaccine (Com-COV): a single-blind, randomised, non-inferiority trial. *The Lancet* **398**, 856–869 (2021).

204. Shrotri, M. *et al.* Vaccine effectiveness of the first dose of ChAdOx1 nCoV-19 and BNT162b2 against SARS-CoV-2 infection in residents of Long-Term Care Facilities (VIVALDI study). *medRxiv* 2021.03.26.21254391 (2021) doi:10.1101/2021.03.26.21254391.

205. Bernal, J. L. *et al.* Early effectiveness of COVID-19 vaccination with BNT162b2 mRNA vaccine and ChAdOx1 adenovirus vector vaccine on symptomatic disease, hospitalisations and mortality in older adults in England. *medRxiv* 2021.03.01.21252652 (2021) doi:10.1101/2021.03.01.21252652.

206. Cheetham, N. J. *et al.* Antibody levels following vaccination against SARSCoV-2: associations with post-vaccination infection and risk factors in two UK longitudinal studies. *Elife* **12**, (2023).

207. Tartof, S. Y. *et al.* Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. *The Lancet* **398**, 1407–1416 (2021).

208. Goldberg, Y. *et al.* Waning Immunity after the BNT162b2 Vaccine in Israel. *New England Journal of Medicine* **385**, e85 (2021).

209. Eyre, D. W. *et al.* Effect of Covid-19 Vaccination on Transmission of Alpha and Delta Variants. *New England Journal of Medicine* (2022) doi:10.1056/NEJMOA2116597.

210. Bergwerk, M. *et al.* Covid-19 Breakthrough Infections in Vaccinated Health Care Workers. *New England Journal of Medicine* **385**, 1474–1484 (2021).

211. *SARS-CoV-2 variants of concern and variants under investigation in England*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1044481/Technical-Briefing-31-Dec-2021-Omicron_severity_update.pdf (2021).

212. Jalkanen, P. *et al.* COVID-19 mRNA vaccine induced antibody responses against three SARS-CoV-2 variants. *Nature Communications 2021 12:1* **12**, 1–11 (2021).

213. Planas, D. *et al.* Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature 2021 596:7871* **596**, 276–280 (2021).

214. Nemet, I. *et al.* Third BNT162b2 Vaccination Neutralization of SARS-CoV-2 Omicron Infection. *N Engl J Med* (2021) doi:10.1056/NEJMC2119358/SUPPL_FILE/NEJMC2119358_DISCLOSURES.PDF.

215. Manisty, C. *et al.* Antibody response to first BNT162b2 dose in previously SARS-CoV-2-infected individuals. *The Lancet* **397**, 1057–1058 (2021).

216. Saadat, S. *et al.* Binding and Neutralization Antibody Titers After a Single Vaccine Dose in Health Care Workers Previously Infected With SARS-CoV-2. *JAMA* **325**, 1467–1469 (2021).

217. Bradley, T. *et al.* Antibody Responses after a Single Dose of SARS-CoV-2 mRNA Vaccine. *New England Journal of Medicine* **384**, 1959–1961 (2021).

218. Krammer, F. *et al.* Antibody Responses in Seropositive Persons after a Single Dose of SARS-CoV-2 mRNA Vaccine. *N Engl J Med* (2021) doi:10.1056/NEJMc2101667.

219. Fryatt, A. *et al.* Community SARS-CoV-2 seroprevalence before and after the second wave of SARS-CoV-2 infection in Harare, Zimbabwe. *EClinicalMedicine* **41**, 101172 (2021).

220. Barrie, M. B. *et al.* SARS-CoV-2 antibody prevalence in Sierra Leone, March 2021: a cross-sectional, nationally representative, age-stratified serosurvey. *BMJ Glob Health* **6**, e007271 (2021).

221. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. https://covid19.who.int/.

222. Focosi, D., Franchini, M. & Maggi, F. Modified Hemagglutination Tests for COVID-19 Serology in Resource-Poor Settings: Ready for Prime-Time? *Vaccines 2022, Vol. 10, Page 406* **10**, 406 (2022).

223. Kruse, R. L. *et al.* A Hemagglutination-Based Semiquantitative Test for Point-of-Care Determination of SARS-CoV-2 Antibody Levels. *J Clin Microbiol* **59**, e0118621 (2021).

224. Kruse, R. L. *et al.* A rapid, point-of-care red blood cell agglutination assay detecting antibodies against SARS-CoV-2. *Biochem Biophys Res Commun* **553**, 165–171 (2021).

225. Redecke, V. *et al.* A rapid and affordable point of care test for antibodies against SARS-CoV-2 based on hemagglutination and artificial intelligence interpretation. *Scientific Reports 2021 11:1* **11**, 1–14 (2021).

226. Townsend, A. *et al.* A haemagglutination test for rapid detection of antibodies to SARS-CoV-2. *Nature Communications 2021 12:1* **12**, 1–12 (2021).

227. Moshe, M. *et al.* SARS-CoV-2 lateral flow assays for possible use in national covid-19 seroprevalence surveys (React 2): diagnostic accuracy study. *BMJ* **372**, (2021).

228. The £1.25 COVID-19 antibody test | Deloitte UK. https://www2.deloitte.com/uk/en/pages/annual-report-2020/stories/the-1-25-covid-19-antibody-test.html.

229. Poorer countries hit with higher price tag for Oxford/AstraZeneca vaccine – POLITICO. https://www.politico.eu/article/astrazeneca-vaccine-cost-higher-in-poorer-countries-coronavirus/ (2021).

230. Munro, A. P. S. *et al.* Safety and immunogenicity of seven COVID-19 vaccines as a third dose (booster) following two doses of ChAdOx1 nCov-19 or BNT162b2 in the UK (COV-BOOST): a blinded, multicentre, randomised, controlled, phase 2 trial. *The Lancet* **398**, 2258–2276 (2021).

231. Patalon, T. *et al.* Odds of Testing Positive for SARS-CoV-2 Following Receipt of 3 vs 2 Doses of the BNT162b2 mRNA Vaccine. *JAMA Intern Med* (2021) doi:10.1001/JAMAINTERNMED.2021.7382.

232. Cerqueira-Silva, T. *et al.* Effectiveness of CoronaVac, ChAdOx1 nCoV-19, BNT162b2, and Ad26.COV2.S among individuals with previous SARS-CoV-2 infection in Brazil: a test-negative, case-control study. *Lancet Infect Dis* (2022) doi:10.1016/S1473-3099(22)00140-2.

233. Bernal, J. L. *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**, (2021).

234. Tseng, H. F. *et al.* Effectiveness of mRNA-1273 against SARS-CoV-2 Omicron and Delta variants. *Nature Medicine 2022* 1–1 (2022) doi:10.1038/s41591-022-01753-y.

235. Goldberg, Y. *et al.* Waning Immunity after the BNT162b2 Vaccine in Israel. *New England Journal of Medicine* **385**, e85 (2021).

236. Tartof, S. Y. *et al.* Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. *The Lancet* **398**, 1407–1416 (2021).

237. Andrews, N. *et al.* Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *N Engl J Med* (2022) doi:10.1056/NEJMOA2119451.

238. Sandmann, F. G. *et al.* The potential health and economic value of SARS-CoV-2 vaccination alongside physical distancing in the UK: a transmission model-based

future scenario analysis and economic evaluation. *Lancet Infect Dis* **21**, 962–974 (2021).

239. Ward, I. L. *et al.* Risk of covid-19 related deaths for SARS-CoV-2 omicron (B.1.1.529) compared with delta (B.1.617.2): retrospective cohort study. *BMJ* **378**, (2022).

240. Nyberg, T. *et al.* Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *The Lancet* **399**, 1303–1312 (2022).

241. Pilz, S., Theiler-Schwetz, V., Trummer, C., Krause, R. & Ioannidis, J. P. A. SARS-CoV-2 reinfections: Overview of efficacy and duration of natural and hybrid immunity. *Environ Res* **209**, 112911 (2022).

242. Wei, J. *et al.* Antibody responses and correlates of protection in the general population after two doses of the ChAdOx1 or BNT162b2 vaccines. *Nature Medicine 2022 28:5* **28**, 1072–1082 (2022).

243. Feng, S. *et al.* Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nature Medicine 2021* 1–9 (2021) doi:10.1038/s41591-021-01540-1.

244. Gilbert, P. B. *et al.* Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science (1979)* **375**, 43–50 (2022).

245. Coronavirus (COVID-19) Infection Survey, characteristics of people testing positive for COVID-19, UK - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveycharacteristicsofpeopletestingpositiveforcovid19uk/30march2022.

246. Wood, S. N. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics* **62**, 1025–1036 (2006).

247. Coronavirus (COVID-19) Infection Survey, UK - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/16march2023.

248. Sorensen, T., Hohenstein, S. & Vasishth, S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quant Method Psychol* **12**, 175–200 (2016).

249. Schubert, M. *et al.* Human serum from SARS-CoV-2-vaccinated and COVID-19 patients shows reduced binding to the RBD of SARS-CoV-2 Omicron variant. *BMC Med* **20**, 102 (2022).

250. Dejnirattisai, W. *et al.* SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. *Cell* **185**, 467-484.e15 (2022).

251. Edara, V. V. *et al.* mRNA-1273 and BNT162b2 mRNA vaccines have reduced neutralizing activity against the SARS-CoV-2 omicron variant. *Cell Rep Med* **3**, 100529 (2022).

252. Carreño, J. M. *et al.* Activity of convalescent and vaccine serum against SARS-CoV-2 Omicron. *Nature* **602**, (2022).

253. Garcia-Beltran, W. F. *et al.* mRNA-based COVID-19 vaccine boosters induce neutralizing immunity against SARS-CoV-2 Omicron variant. *Cell* **185**, 457-466.e4 (2022).

254. Pajon, R. *et al.* SARS-CoV-2 Omicron Variant Neutralization after mRNA-1273 Booster Vaccination. *New England Journal of Medicine* (2022) doi:10.1056/NEJMC2119912/SUPPL_FILE/NEJMC2119912_DISCLOSURES.PDF.

255. Cheng, S. M. S. *et al.* Neutralizing antibodies against the SARS-CoV-2 Omicron variant BA.1 following homologous and heterologous CoronaVac or BNT162b2 vaccination. *Nature Medicine 2022 28:3* **28**, 486–489 (2022).

256. Lauring, A. S. *et al.* Clinical Severity and mRNA Vaccine Effectiveness for Omicron, Delta, and Alpha SARS-CoV-2 Variants in the United States: A Prospective Observational Study. *medRxiv* (2022) doi:10.1101/2022.02.06.22270558.

257. Yavlinsky, A. *et al.* Anti-spike antibody trajectories in individuals previously immunised with BNT162b2 or ChAdOx1 following a BNT162b2 booster dose. *medRxiv* 2022.02.07.22270451 (2022) doi:10.1101/2022.02.07.22270451.

258. Evans, J. P. *et al.* Neutralizing antibody responses elicited by SARS-CoV-2 mRNA vaccination wane over time and are boosted by breakthrough infection. *Sci Transl Med* (2022) doi:10.1126/SCITRANSLMED.ABN8057.

259. Markewitz, R. *et al.* Kinetics of the Antibody Response to Boostering With Three Different Vaccines Against SARS-CoV-2. *Front Immunol* **13**, 26 (2022).

260. Sablerolles, R. S. G. *et al.* Immunogenicity and Reactogenicity of Vaccine Boosters after Ad26.COV2.S Priming. *N Engl J Med* (2022) doi:10.1056/NEJMOA2116747.

261. Atmar, R. L. *et al.* Homologous and Heterologous Covid-19 Booster Vaccinations. *N Engl J Med* (2022) doi:10.1056/NEJMOA2116414.

262. Collier, A. Y. *et al.* Characterization of immune responses in fully vaccinated individuals following breakthrough infection with the SARS-CoV-2 delta variant. *Sci Transl Med* (2022) doi:10.1126/SCITRANSLMED.ABN6150.

263. Tober-Lau, P. *et al.* Cross-Variant Neutralizing Serum Activity after SARS-CoV-2 Breakthrough Infections. *Emerg Infect Dis* **28**, (2022).

264. Suntronwong, N. *et al.* COVID-19 Breakthrough Infection after Inactivated Vaccine Induced Robust Antibody Responses and Cross-Neutralization of SARS-CoV-2

Variants, but Less Immunity against Omicron. *Vaccines 2022, Vol. 10, Page 391* **10**, 391 (2022).

265. Bates, T. A. *et al.* Antibody Response and Variant Cross-Neutralization After SARS-CoV-2 Breakthrough Infection. *JAMA* **327**, 179–181 (2022).

266. Miyamoto, S. *et al.* Vaccination-infection interval determines cross-neutralization potency to SARS-CoV-2 Omicron after breakthrough infection by other variants. *Med* **3**, 249-261.e4 (2022).

267. Kitchin, D. *et al.* Ad26.COV2.S breakthrough infections induce high titers of neutralizing antibodies against Omicron and other SARS-CoV-2 variants of concern. *Cell Rep Med* **3**, 100535 (2022).

268. Dimeglio, C. *et al.* Antibody Titers and Protection against Omicron (BA.1 and BA.2) SARS-CoV-2 Infection. *Vaccines 2022, Vol. 10, Page 1548* **10**, 1548 (2022).

269. Muik, A. *et al.* Neutralization of SARS-CoV-2 Omicron by BNT162b2 mRNA vaccine-elicited human sera. *Science (1979)* **375**, 678–680 (2022).

270. Altarawneh, H. N. *et al.* Effects of Previous Infection and Vaccination on Symptomatic Omicron Infections. *N Engl J Med* **387**, 21–34 (2022).

271. Andeweg, S. P. *et al.* Protection of COVID-19 vaccination and previous infection against Omicron BA.1, BA.2 and Delta SARS-CoV-2 infections. *Nature Communications 2022 13:1* **13**, 1–9 (2022).

272. Thompson, E. J. *et al.* Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. *Nature Communications 2022 13:1* **13**, 1–11 (2022).

273. Orangi, S. *et al.* Epidemiological impact and cost-effectiveness analysis of COVID-19 vaccination in Kenya. *BMJ Glob Health* **7**, 9430 (2022).

274. Offit, P. A. Bivalent Covid-19 Vaccines — A Cautionary Tale. *https://doi.org/10.1056/NEJMp2215780* (2023) doi:10.1056/NEJMP2215780.

275. Covid booster: Who can get another jab this spring? - BBC News. https://www.bbc.co.uk/news/health-55045639.

276. Autumn Covid booster recommended only for ‘at risk’ groups in France. https://www.connexionfrance.com/article/French-news/Covid-19/Autumn-Covid-booster-recommended-only-for-at-risk-groups-in-France.

277. The United Kingdom: WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. https://covid19.who.int/region/euro/country/gb.

278. Deneer, R. *et al.* Detecting patients with PMI post-CABG based on cardiac troponin-T profiles: A latent class mixed modeling approach. *Clin Chim Acta* **504**, 23–29 (2020).

279. McCulloch, C. E., Lin, H., Slate, E. H. & Turnbull, B. W. Discovering subpopulation structure with latent class mixed models. *Stat Med* **21**, 417–429 (2002).

280. Li, C. *et al.* Trajectories of Perioperative Serum Tumor Markers and Colorectal Cancer Outcomes: A Retrospective, Multicenter Longitudinal Cohort Study. *EBioMedicine* **74**, (2021).

281. Brombin, C. *et al.* Modelling the impact of protein-kinase R allelic variant on HIV biomarkers trajectories by means of latent class mixed models. *Scientific Reports 2022 12:1* **12**, 1–9 (2022).

282. Fox, T. *et al.* Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst Rev* **11**, (2022).

283. Dinnes, J. *et al.* Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* **3**, (2021).

284. LeMaster, C., Geanes, E. S., Fraley, E. R., Selvarangan, R. & Bradley, T. Vaccination After SARS-CoV-2 Infection Increased Antibody Avidity Against the Omicron Variant Compared to Vaccination Alone. *J Infect Dis* **226**, 1712–1716 (2022).

285. Pérez-Alós, L. *et al.* Modeling of waning immunity after SARS-CoV-2 vaccination and influencing factors. *Nature Communications 2022 13:1* **13**, 1–11 (2022).

286. Gobbi, F. *et al.* Antibody Response to the BNT162b2 mRNA COVID-19 Vaccine in Subjects with Prior SARS-CoV-2 Infection. *Viruses* **13**, (2021).

287. Krammer, F. A correlate of protection for SARS-CoV-2 vaccines is urgently needed. *Nature Medicine 2021 27:7* **27**, 1147–1148 (2021).

288. And, C. & Immunology, V. Correlates of Protection Induced by Vaccination. *Clinical and Vaccine Immunology* **17**, 1055–1065 (2010).

289. Perry, J. *et al.* Does a humoral correlate of protection exist for SARS-CoV-2? A systematic review. *PLoS One* **17**, e0266852 (2022).

290. Nilles, E. J. *et al.* Tracking immune correlates of protection for emerging SARS-CoV-2 variants. *Lancet Infect Dis* **23**, 153–154 (2023).

291. Crotty, S. Hybrid immunity. *Science (1979)* **372**, 1392–1393 (2021).

292. Geers, D. *et al.* SARS-CoV-2 variants of concern partially escape humoral but not T-cell responses in COVID-19 convalescent donors and vaccinees. *Sci Immunol* **6**, (2021).

293. Keeton, R. *et al.* T cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature* **603**, 488–492 (2022).

294. Röltgen, K. & Boyd, S. D. Antibody and B cell responses to SARS-CoV-2 infection and vaccination. *Cell Host Microbe* **29**, 1063–1075 (2021).

295. Russell, M. W., Moldoveanu, Z., Ogra, P. L. & Mestecky, J. Mucosal Immunity in COVID-19: A Neglected but Critical Aspect of SARS-CoV-2 Infection. *Front Immunol* **11**, 3221 (2020).

296. Cox, R. J. & Brokstad, K. A. Not just antibodies: B cells and T cells mediate immunity to COVID-19. *Nature Reviews Immunology 2020 20:10* **20**, 581–582 (2020).

297. Brasu, N. *et al.* Memory CD8+ T cell diversity and B cell responses correlate with protection against SARS-CoV-2 following mRNA vaccination. *Nature Immunology 2022 23:10* **23**, 1445–1456 (2022).

298. Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic.

299. COVID-19 variants identified in the UK – latest updates - GOV.UK. https://www.gov.uk/government/news/covid-19-variants-identified-in-the-uk-latest-updates.

300. Arora, P. *et al.* Neutralisation sensitivity of the SARS-CoV-2 XBB.1 lineage. *Lancet Infect Dis* **23**, 147–148 (2023).

301. Wang, Q. *et al.* Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* **186**, 279-286.e8 (2023).

302. Imai, M. *et al.* Efficacy of Antiviral Agents against Omicron Subvariants BQ.1.1 and XBB. *New England Journal of Medicine* **388**, 89–91 (2023).

303. Kurhade, C. *et al.* Low neutralization of SARS-CoV-2 Omicron BA.2.75.2, BQ.1.1 and XBB.1 by parental mRNA vaccine or a BA.5 bivalent booster. *Nat Med* **29**, (2023).

304. Winokur, P. *et al.* Bivalent Omicron BA.1–Adapted BNT162b2 Booster in Adults Older than 55 Years. *New England Journal of Medicine* **388**, 214–227 (2023).

305. Zou, J. *et al.* Neutralization of BA.4–BA.5, BA.4.6, BA.2.75.2, BQ.1.1, and XBB.1 with Bivalent Vaccine. *New England Journal of Medicine* (2023) doi:10.1056/NEJMC2214916/SUPPL_FILE/NEJMC2214916_DISCLOSURES.PDF.

306. Scheaffer, S. M. *et al.* Bivalent SARS-CoV-2 mRNA vaccines increase breadth of neutralization and protect against the BA.5 Omicron variant in mice. *Nat Med* **29**, 247–257 (2023).

307. Chalkias, S. *et al.* A Bivalent Omicron-Containing Booster Vaccine against Covid-19. *N Engl J Med* **387**, 1279–1291 (2022).

308. Lin, D.-Y. *et al.* Effectiveness of Bivalent Boosters against Severe Omicron Infection. *New England Journal of Medicine* (2023)

doi:10.1056/NEJMC2215471/SUPPL_FILE/NEJMC2215471_DISCLOSURES.PDF.

309. Collier, A. Y. *et al.* Immunogenicity of BA.5 Bivalent mRNA Vaccine Boosters. *New England Journal of Medicine* (2023) doi:10.1056/NEJMC2213948/SUPPL_FILE/NEJMC2213948_DISCLOSURES.PDF.

310. Most vulnerable to be offered spring COVID-19 booster - GOV.UK. https://www.gov.uk/government/news/most-vulnerable-to-be-offered-spring-covid-19-booster.

311. Mentzer, A. J. *et al.* Human leukocyte antigen alleles associate with COVID-19 vaccine immunogenicity and risk of breakthrough infection. *Nature Medicine 2023 29:1* **29**, 147–157 (2022).

312. Crocchiolo, R. *et al.* Polymorphism of the HLA system and weak antibody response to BNT162b2 mRNA vaccine. *HLA* **99**, 183–191 (2022).

313. COVID-19 Infection Survey cost - Office for National Statistics. https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/covid19infectionsurveycost.