

Serological reconstruction of COVID-19 epidemics through analysis of antibody kinetics to SARS-CoV-2 proteins

Stéphane PELLEAU¹, Tom WOUDEBERG¹, Jason ROSADO^{1,2}, Françoise DONNADIEU¹, Laura GARCIA¹, Thomas OBADIA^{1,3}, Soazic GARDAIS¹, Yasmine ELGHARBAWY¹, Aurelie VELAY^{4,5}, Maria GONZALEZ⁶, Jacques Yves NIZOU⁷, Nizar KHELIL⁷, Konstantinos ZANNIS⁷, Charlotte COCKRAM⁸, Sarah Hélène MERKLING⁹, Annalisa MEOLA¹⁰, Solen KERNEIS^{11,12}, Benjamin TERRIER^{13,14}, Jerome de SEZE¹⁵, Delphine PLANAS¹⁶, Olivier SCHWARTZ¹⁶, François DEJARDIN¹⁷, Stéphane PETRES¹⁷, Cassandre von PLATEN¹⁸, Laurence AROWAS¹⁹, Louise PERRIN DE FACCI¹⁹, Darragh DUFFY²⁰, Clíona Ní CHEALLAIGH^{21,22}, Niall CONLON^{23,24}, Liam TOWNSEND^{21,22}, Heidi AUERSWALD²⁵, Marija BACKOVIC¹⁰, Bruno HOEN²⁶, Arnaud FONTANET^{27,28}, Ivo MUELLER^{1,29,30}, Samira FAFI-KREMER^{4,5}, Timothée BRUEL^{16,31}, Michael WHITE^{1*}

1. Malaria: Parasites and Hosts Unit, Department of Parasites and Insect Vectors, Institut Pasteur, Paris, France
2. Sorbonne Université, ED 393, F-75005 Paris, France
3. Hub de Bioinformatique et Biostatistique, Département Biologie Computationnelle, Institut Pasteur, Paris, France
4. CHU de Strasbourg, Laboratoire de Virologie, F-67091 Strasbourg, France
5. Université de Strasbourg, INSERM, IRM UMR_S 1109, Strasbourg, France
6. CHU de Strasbourg, Service de Pathologies Professionnelles, F-67091 Strasbourg, France
7. Institut Mutualiste Montsouris, Paris, France
8. Spatial Regulation of Genomes Unit, Department of Genomes and Genetics, Institut Pasteur, Paris, France
9. Insect-Virus Interactions Unit, Department of Virology and CNRS UMR 2000, Institut Pasteur, Paris, France
10. Structural Virology Unit, Department of Virology and CNRS UMR 3569, Institut Pasteur, Paris, France
11. Equipe Mobile d'Infectiologie, APHP Centre-Université de Paris, Paris, France
12. Epidemiology and Modelling of Bacterial Escape to Antimicrobials Unit, Department of Global Health, Institut Pasteur, Paris, France
13. Department of Internal Medicine, National Referral Center for Rare Systemic Autoimmune Diseases, Assistance Publique Hôpitaux de Paris-Centre (APHP-CUP), Université de Paris, Paris, France
14. PARCC, INSERM U970, Paris, France
15. Centre d'Investigation Clinique - INSERM CIC-1434, Strasbourg, France
16. Virus and Immunity Unit, Department of Virology, Institut Pasteur, Paris, France
17. Production and Purification of Recombinant Proteins Technological Platform, Center for Technological Resources and Research, Institut Pasteur, Paris, France
18. Center for Translational Science, Institut Pasteur, Paris, France
19. Investigation Clinique et Accès aux Ressources Biologiques (ICAReB), Center for Translational Research, Institut Pasteur, Paris, France.
20. Translational Immunology Lab, Institut Pasteur, Paris, France
21. Department of Infectious Diseases, St James's Hospital, Dublin, Ireland
22. Department of Clinical Medicine, School of Medicine, Trinity Translational Medicine Institute, Trinity College Dublin, Ireland
23. Department of Immunology, St James's Hospital, Dublin, Ireland
24. Department of Immunology, School of Medicine, Trinity College Dublin, Ireland
25. Virology Unit, Institut Pasteur du Cambodge, Institut Pasteur International Network, Phnom Penh, Cambodia
26. Direction de la Recherche Médicale, Institut Pasteur, Paris, France
27. Epidemiology of Emerging Diseases Unit, Department of Global Health, Institut Pasteur, Paris, France
28. Conservatoire National des Arts et Métiers, Paris, France.
29. Division of Population Health and Immunity, The Walter and Eliza Hall Institute, Melbourne, Australia
30. Department of Medical Biology, University of Melbourne, Melbourne, Australia
31. Vaccine Research Institute, Creteil, France

*Correspondence

Dr Michael White (michael.white@pasteur.fr)

Malaria: Parasites and Hosts Unit, Department of Parasites and Insect Vectors, Institut Pasteur, 25-28 Rue du Docteur Roux, Paris 75015, France

Abstract

Infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) induces a complex antibody response that varies by orders of magnitude between individuals and over time. Waning antibody levels lead to reduced sensitivity of serological diagnostic tests over time. This undermines the utility of serological surveillance as the SARS-CoV-2 pandemic progresses into its second year. Here we develop a multiplex serological test for measuring antibodies of three isotypes (IgG, IgM, IgA) to five SARS-CoV-2 antigens (Spike (S), receptor binding domain (RBD), Nucleocapsid (N), Spike subunit 2, Membrane-Envelope fusion) and the Spike proteins of four seasonal coronaviruses. We measure antibody responses in several cohorts of French and Irish hospitalized patients and healthcare workers followed for up to eleven months after symptom onset. The data are analysed with a mathematical model of antibody kinetics to quantify the duration of antibody responses accounting for inter-individual variation. One year after symptoms, we estimate that 36% (95% range: 11%, 94%) of anti-S IgG remains, 31% (9%, 89%) anti-RBD IgG remains, and 7% (1%, 31%) anti-N IgG remains. Antibodies of the IgM isotype waned more rapidly, with 9% (2%, 32%) anti-RBD IgM remaining after one year. Antibodies of the IgA isotype also waned rapidly, with 10% (3%, 38%) anti-RBD IgA remaining after one year. Quantitative measurements of antibody responses were used to train machine learning algorithms for classification of previous infection and estimation of time since infection. The resulting diagnostic test classified previous infections with 99% specificity and 98% (95% confidence interval: 94%, 99%) sensitivity, with no evidence for declining sensitivity over the time scale considered. The diagnostic test also provided accurate classification of time since infection into intervals of 0 – 3 months, 3 – 6 months, and 6 – 12 months. Finally, we present a computational method for serological reconstruction of past SARS-CoV-2 transmission using the data from this test when applied to samples from a single cross-sectional sero-prevalence survey.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causing coronavirus disease 2019 (COVID-19), has led to widespread morbidity and mortality since its emergence. The response to the SARS-CoV-2 pandemic is critically dependent on surveillance data, most notably numbers of COVID-19 associated hospital admissions and deaths recorded through health systems surveillance, as well as numbers of cases confirmed SARS-CoV-2 positive by PCR-based testing¹. Other tools are providing crucial complementary information, for example genomic surveillance has been key to tracking the emergence of novel SARS-CoV-2 variants². Serology, based on the detection of antibodies induced by previous infection with SARS-CoV-2, represents another category of surveillance information^{3,4}. Appropriately designed sero-prevalence studies can provide estimates of the proportion of a population who have been previously infected. Although no substitute for health systems surveillance, sero-prevalence studies have the advantage of accounting for asymptomatic cases, and symptomatic individuals who do not present to health systems. Sero-prevalence studies also provide information on the status of SARS-CoV-2 epidemics in situations where record keeping by health systems is not possible⁵.

Infection with SARS-CoV-2 induces diverse humoral and cellular immune responses⁶. Humoral immunity includes antibodies of several immunoglobulin isotypes targeting SARS-CoV-2 proteins, most notably Spike (S) and Nucleocapsid (N). The concentration of antibodies in blood varies substantially between individuals, and with time since infection^{6,7,8,9,10,11}. Studies of the duration of immunity to a range of coronaviruses demonstrated that antibodies remain detectable six years after infection, but continue to wane¹². Longitudinal follow-up of individuals infected with SARS-CoV-2 indicates a pattern of waning of antibody responses consistent with other coronaviruses^{13,14}. Within the first three months, antibody levels boost sharply and wane rapidly. Over a longer interval of eight months, antibody levels wane more slowly⁶. These observations can be explained by the bi-phasic nature of antibody kinetics¹⁵. In the first three months, antibodies are predominantly generated by short-lived plasma cells in secondary lymphoid organs. The long-term response is dominated by antibodies from long-lived plasma cells in the bone marrow. This pattern of bi-phasic waning is observed for infection- and vaccine-induced antibody responses to a wide range of pathogens^{16,17,18,19}.

Initial concerns that waning antibody responses would lead to reduced sensitivity of SARS-CoV-2 serological diagnostics over time have unfortunately been confirmed²⁰. Using self-administered lateral flow diagnostic tests, a sero-prevalence study of 365,000 adults after the first epidemic wave in the UK observed significantly declining sero-prevalence during July to September 2020⁴. Using a laboratory-based assay for measuring anti-N IgG antibodies, a sero-prevalence survey of blood donors from the Brazilian city of Manaus estimated that 26% of the population had previously been infected²¹. Application of a statistical adjustment to account for antibody waning led to an increased estimate of 76% previously infected. As the pandemic progresses, this problem of declining sensitivity of serological diagnostics is likely to get worse, potentially undermining the utility of sero-prevalence studies.

Serological diagnostics typically classify a sample as positive if a measured antibody level is greater than a defined cutoff. Analysis of quantitative rather than binary antibody levels provides additional information, for example antibody levels

are associated with time since infection, symptom severity, and sex²². However, large variation in antibody levels between individuals prevents this from having predictive value: detected antibodies could be from a recent infection, or due to immunological memory of an infection that cleared a year ago. This limitation has recently been overcome for a range of pathogens through the combination of multiplex assays and classification algorithms. Using machine learning algorithms to analyse quantitative antibody responses to multiple antigens, the time since previous infection can be estimated for *Plasmodium falciparum* malaria^{23,24}, *P. vivax* malaria²⁵, and cholera²⁶.

In this study, we use multiplex assays to measure antibodies to SARS-CoV-2 in health care workers and hospitalized patients followed for up to eleven months after infection, and apply mathematical models to characterize antibody kinetics in the first year following infection. Classification algorithms were developed that minimize the reduction in the sensitivity of serological tests over time, in addition to estimating time since previous infection from a single blood sample. Finally, we present a method for serological reconstruction of past SARS-CoV-2 transmission using samples from a single cross-sectional survey.

Methods

Samples

A panel of 407 negative control serum or plasma samples was assembled from pre-pandemic cohorts (before December, 2019) with ethical approval for broad antibody testing (Table 1). This included 258 serum samples from healthy adult French blood donors, 81 serum samples from Peruvian healthy adult donors, and 69 plasma samples from healthy adult donors from the Thai Red Cross. A panel of 407 positive control serum samples was assembled from individuals with recent SARS-CoV-2 infection. This included 72 samples from patients in Paris hospitals^{27,28}, 161 samples from Strasbourg healthcare workers, all confirmed positive by PCR-based testing²⁹. Also included were 174 samples from community members of Crépy-en-Valois, France, confirmed seropositive by flow cytometry based testing^{30,31}.

The duration of antibody responses following SARS-CoV-2 infection was studied in longitudinal cohorts of hospitalized patients and healthcare workers. 213 serum samples from 194 patients in Dublin hospitals were collected, with date post symptom onset extending to four months. 724 serum samples from 347 healthcare workers in Strasbourg hospitals were collected, with date post symptom onset extending to nine months³².

In April 2020, our team implemented a study of the sero-prevalence of SARS-CoV-2 in healthcare workers from Institut Mutualiste Montsouris, a hospital in Paris. Serum samples were collected from 769 healthcare workers, and tested with our multiplex assay (Appendix Figure 1). Healthcare workers who tested sero-positive in April 2020 were invited to present a second sample in January 2021. In total we obtained follow-up samples from 29 healthcare workers.

Table 1: Panels of samples. Age and days post symptom onset are presented as median and ranges.

| cohort | status | participants | samples | PCR positive | age (years) | gender (% male) | days post symptoms |
|--|---|--------------|---------|--------------|-------------|-----------------|--------------------|
| Établissement Français du Sang 1 | pre-pandemic negative controls | 45 | 45 | | > 18 | | |
| Établissement Français du Sang 2 | pre-pandemic negative controls | 213 | 213 | | 42 (18,81) | 40% | |
| Thai Red Cross | pre-pandemic negative controls | 68 | 68 | | > 18 | | |
| Peruvian donors | pre-pandemic negative controls | 81 | 81 | | > 18 | | |
| Hôpital Bichat (Paris, France) | hospitalized patients | 2 | 8 | 2 | 31 (30, 32) | 100% | 14 (8, 24) |
| Hôpital Cochin (Paris, France) | hospitalized patients | 64 | 64 | 64 | 55 (25, 79) | 76% | 17 (10, 28) |
| Strasbourg hospitals 1 | infected healthcare workers | 161 | 161 | 161 | 32 (20, 62) | 31% | 24 (13, 39) |
| Crépy-en-Valois community | infected community members (flow cytometry positive) | 154 | 174 | 0 | 17 (15, 56) | 34% | |
| Dublin hospitals | hospitalized patients | 194 | 213 | 194 | 55 (21, 92) | 47% | 13 (1, 126) |
| Strasbourg hospitals 2 | follow-up of infected healthcare workers | 347 | 724 | 347 | 41 (21,74) | 23% | 132 (11, 284) |
| Institut Mutualiste Montsouris (Paris, France) | sero-prevalence survey in healthcare workers (unknown status) | 769 | 769 | 20 | 41 (18, 72) | 27% | |
| Institut Mutualiste Montsouris (Paris, France) | follow-up of sero-positive healthcare workers | 29 | 29 | 12 | 37 (24, 63) | 41% | 304 (285, 336) |

Serological assays

We set up a 9-plex bead-based assay allowing simultaneous detection of antibody responses to five SARS-CoV-2 antigens and four seasonal coronaviruses (Spike proteins of NL63, 229E, HKU1, OC43) in 1 μ L serum or plasma samples that were heat-inactivated at 56°C for 30 minutes. SARS-CoV-2 antigens were from Spike (whole trimeric Spike (S), its Receptor Binding Domain (RBD) and Spike subunit 2 (S2)), nucleocapsid protein (N) and a Membrane-Envelope fusion protein (ME). ME and S2 antigens were purchased from Native Antigen, Oxford, UK and all other antigens were produced as recombinant proteins at Institut Pasteur. The mass of proteins coupled on beads was optimized to generate a log-linear standard curve with a pool of 27 positive sera prepared from RT-qPCR-confirmed SARS-CoV-2 patients. We measured the levels of IgG, IgA and IgM of each sample in three separate assays. Briefly, serum was incubated with mixed antigen-coupled beads for 30 minutes at a 1/200 final dilution for IgG or 1/400 for IgA and IgM. Secondary antibodies conjugated to R-phycoerythrin (Jackson ImmunoResearch) were used at 1/120, 1/200 or 1/400 for detection of specific IgG, IgA and IgM respectively. All dilutions and cycles of washing steps were done in phosphate buffer saline supplemented with 1% bovine serum albumin and 0.05% (v/v) Tween-20. On each assay plate, two blanks (only beads, no serum) were included to control for background signal as well as a standard curve prepared from two-fold serial dilutions (1/50 to 1/102,400) of a pool of positive controls. Plates were read using a Luminex[®] MAGPIX[®] system and the median fluorescence intensity (MFI) was used for analysis. A 5-parameter logistic curve was used to convert MFI to relative antibody unit (RAU), relative to the standard curve performed on the same plate to account for inter-assay variations.

The data from our multiplex assay was compared against data from two different neutralization assays with live virus using a subset of serum samples. Firstly, we implemented an S-Fuse assay as described elsewhere³². Neutralization activity was measured as the reciprocal dilution required to obtain a 50% reduction in neutralization (IC50). Secondly, we implemented a foci reduction neutralization test (FRNT) based on the detection of neutralizing antibodies directed against SARS-CoV-2. This assay was performed under BSL-3 conditions as it facilitates infection of African green monkey kidney cells (VeroE6; ATCC CRL-1586) with live-virus of a Cambodian SARS-CoV-2 isolate (GISAID: EPI_ISL_411902). Infection is visualized 14-16h after inoculation by staining of infected cells with a SARS-CoV-2 specific antibody (# ABIN1030641), targeting the S2 subunit of the viral spike protein. Neutralizing antibody titers are expressed as the reciprocal serum/plasma dilution that induces 50% reduction of infection (FRNT50) and is calculated by log probit regression analysis.

Mathematical model of antibody kinetics

SARS-CoV-2 antibody kinetics are described using a previously published mathematical model of the immunological processes underlying the generation and waning of antibody responses following infection³³.

$$\begin{aligned}\frac{dP_s}{dt} &= \rho\beta - c_s P_s, \\ \frac{dP_l}{dt} &= (1 - \rho)\beta - c_l P_l, \\ \frac{dA}{dt} &= gP_s + gP_l - rA\end{aligned}$$

β denotes the boost in antibody secreting plasma cells. It is assumed that a proportion ρ of plasma cells are short-lived (P_s) waning at rate c_s , and a proportion $1 - \rho$ are long-lived (P_l) waning at rate c_l . Plasma cells generate antibodies (IgG, IgM or IgA) at rate g , and r is the rate of decay of antibody molecules.

Statistical inference was implemented within a mixed-effects framework allowing for characterisation of the kinetics within all individuals while also describing the population-level patterns (Statistical Appendix). On the population level, both the mean and variation in antibody kinetics are accounted for. Models were fitted in a Bayesian framework using Markov chain Monte Carlo methods with priors informed by estimates from long-term studies of antibody kinetics following infection with other coronaviruses²⁰.

Classification algorithms

Measurements of antibodies of three isotypes (IgG, IgM, IgA) to multiple SARS-CoV-2 antigens were used to create a training dataset. Samples where the time post symptom onset was ≤ 14 days or unknown were excluded. In total we had 407 samples from pre-pandemic negative controls and 1402 positive samples. In a first step, a Random Forests binary classification algorithm was developed. A threshold corresponding to 99% specificity was selected. In a second step, a Random Forests multiway classification algorithm was developed for categorizing samples into four classes: (i) negative; (ii) infected ≤ 3 months ago; (iii) infected 3 – 6 months ago; and (iv) infected 6 – 12 months ago. The algorithm was calibrated to have 99% specificity for correctly classifying negative samples. Positive samples were then classified according to the maximum number of votes. Uncertainty in classification performance was assessed via 1000-fold cross-validation with a training set comprising two thirds of the data and a disjoint testing set comprising a third of the data. Classification algorithms were implemented in R (version 3.4.3).

Statistical methods for serological surveillance

Imperfect diagnostic sensitivity causes a downwards bias in sero-prevalence estimates, whereas imperfect specificity causes an upwards bias in seroprevalence estimates. This bias can be corrected for using a Rogan-Gladen estimator, with incorporation of cross-validated uncertainty³⁴. Multiway classification algorithms provide estimates of the proportion of

a population infected during different time intervals. These estimates are subject to biases which can be adjusted for using a multivariate extension of the Rogan-Gladen estimator described in the Statistical Appendix.

Ethics

Serum samples were biobanked at the Clinical Investigation and Access to BioResources platform at Institut Pasteur (Paris, France). Samples were obtained from consenting individuals through the CORSER study (NCT04325646), directed by Institut Pasteur and approved by the Comité de Protection des Personnes Ile de France III, and the French COVID cohort (NCT04262921), sponsored by Inserm and approved by the Comité de Protection des Personnes Ile de France VI. Samples from French blood donors were approved for use by Etablissement Français du Sang (Lille, France) and approved through the CORSER study by the Comité de Protection des Personnes Ile de France VI. Sample collection in Hôpital Cochin was approved by the Research Ethics Commission of Necker-Cochin Hospital. Samples from healthcare workers in Strasbourg University Hospitals followed longitudinally were collected as part of an ongoing clinical trial (ClinicalTrials.gov Identifier: NCT04441684) which received ethical approval from the Comité de Protection des Personnes Ile de France III. Samples collected from patients in Dublin received ethical approval for study from the Tallaght University Hospital (TUH)/St James's Hospital (SJH) Joint Research Ethics Committee (reference REC 2020–03). Use of the Peruvian negative controls was approved by the Institutional Ethics Committee from the Universidad Peruana Cayetano Heredia (SIDISI 100873). The Human Research Ethics Committee at the Walter and Eliza Hall Institute of Medical Research and the Ethics Committee of the Faculty of Tropical Medicine, Mahidol University, Thailand, approved the use of the Thai negative control samples. Informed written consent was obtained from all participants or their next of kin in accordance with the Declaration of Helsinki.

Results

SARS-CoV-2 antibodies over time

Antibody responses of three isotypes (IgG, IgM, IgA) to nine coronavirus antigens were measured in 407 pre-pandemic serum samples, and 1402 serum samples from individuals with previous SARS-CoV-2 infection. 961/1402 of the positive samples were from individuals with SARS-CoV-2 infection confirmed by PCR-based testing with available data on time post symptoms. Figure 1 presents the IgG, IgM and IgA antibody responses to SARS-CoV-2 S, RBD, and N measured over time. Appendix Figures 2 and 3 presents the antibody responses to S2, ME, and the Spike proteins of the four human seasonal coronaviruses. Notably there is substantial inter-individual variation in antibody responses, with antibody levels varying by orders of magnitude between individuals. As a measure of functional immunity, we applied live virus neutralizing assays to a subset of samples. Substantial variation in neutralizing activity between individuals and over time (Appendix Figure 4).

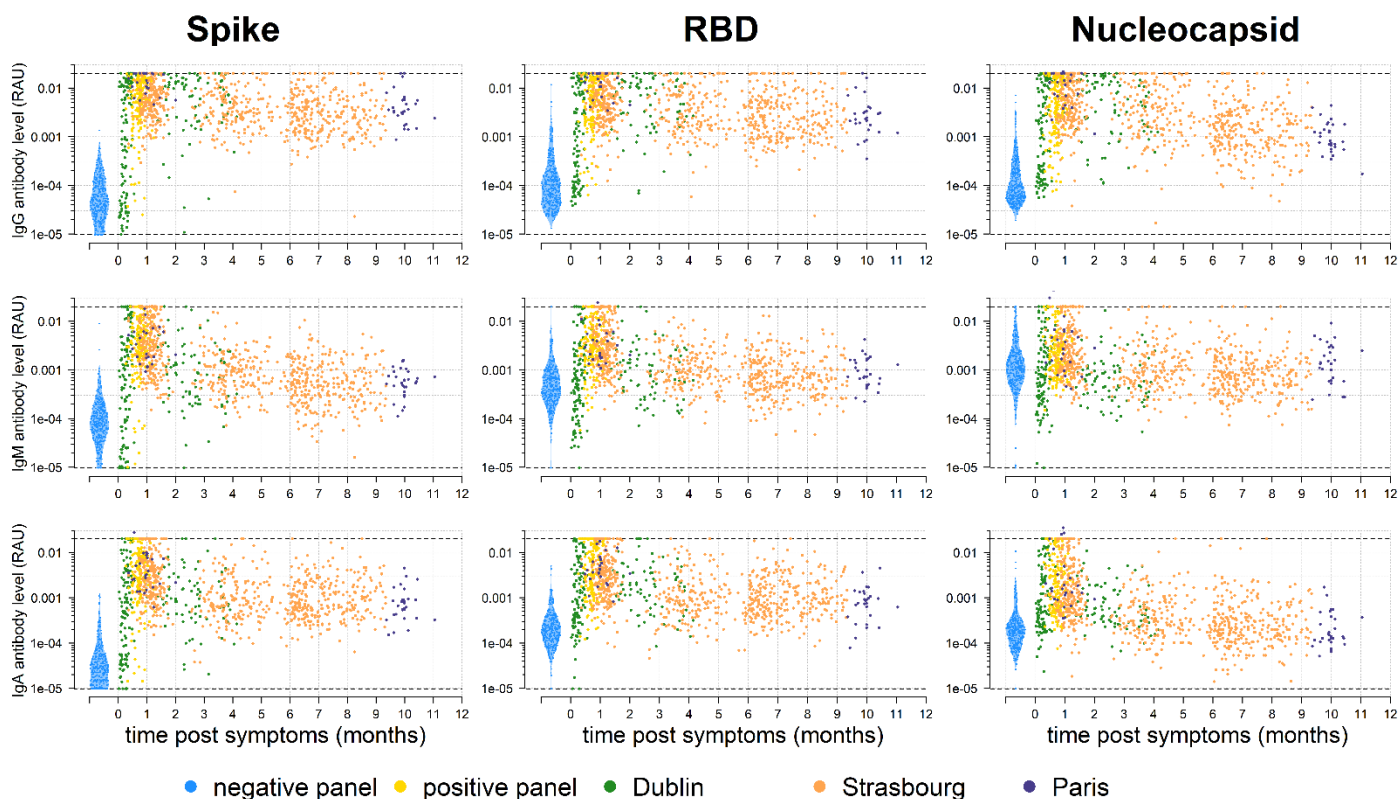


Figure 1: Antibody kinetics in the first year following infection with SARS-CoV-2. A bead-based multiplex Luminex assay used to measure antibodies of multiple isotypes (IgG, IgM, IgA) to multiple antigens in serum samples from individuals with PCR-positive SARS-CoV-2 infection and pre-pandemic negative controls.

Modelled SARS-CoV-2 antibody kinetics

Our data on measured SARS-CoV-2 antibody responses was supplemented with data from six other longitudinal studies of the SARS-CoV-2 antibody response, and one longitudinal study of the SARS-CoV-1 antibody response. Appendix Figure 5 provides a comparison of the measured antibody responses between the eight studies. The data have been rescaled so that the average antibody response for each study equals one at day 14 after symptom onset. In addition to the large inter-individual variation, there is notable variation in antibody levels between studies. For example, the large dynamic range in the study by Iyer *et al*⁸, reveals the boosting and subsequent decay of anti-S IgG antibodies in the first two months post infection, a phenomenon that is missed by assays without a comparably high upper limit of detection.

A mathematical model of antibody kinetics was simultaneously fitted to data from all eight studies. Appendix Figure 6 provides an overview of the fit of the model to the data. Figure 2a provides example fits to data from one individual from each of the eight studies. Figure 2b provides model predictions of the percentage antibody level remaining over the first two years post infection, where 100% is assumed to be the antibody response at day 14 following symptom onset. There are considerable differences in the pattern of waning between isotypes and between antigens. Table 2 summarises the duration of the antibody response. One year following symptom onset, we estimate that 36% (95% range: 11%, 94%) anti-S IgG remains, 31% (9%, 89%) anti-RBD IgG remains, and 7% (1%, 31%) anti-N IgG remains. The uncertainty represents the considerable degree of inter-individual variation in the duration of the antibody responses. Antibodies of the IgM isotype waned more rapidly, with 6% (0%, 27%) anti-S IgM remaining after one year, 9% (2%, 32%) anti-RBD IgM remaining after one year, and 15% (4%, 50%) anti-N IgM remaining after one year. Antibodies of the IgA isotype also waned rapidly, with 18% (4%, 67%) anti-S IgA remaining after one year, 10% (3%, 38%) anti-RBD IgA remaining after one year, and 3% (0%, 13%) anti-N IgA remaining after one year. We also observed comparable reductions in titres for viral neutralization over time (Appendix Figure 4), although the small number of samples prevented application of our model.

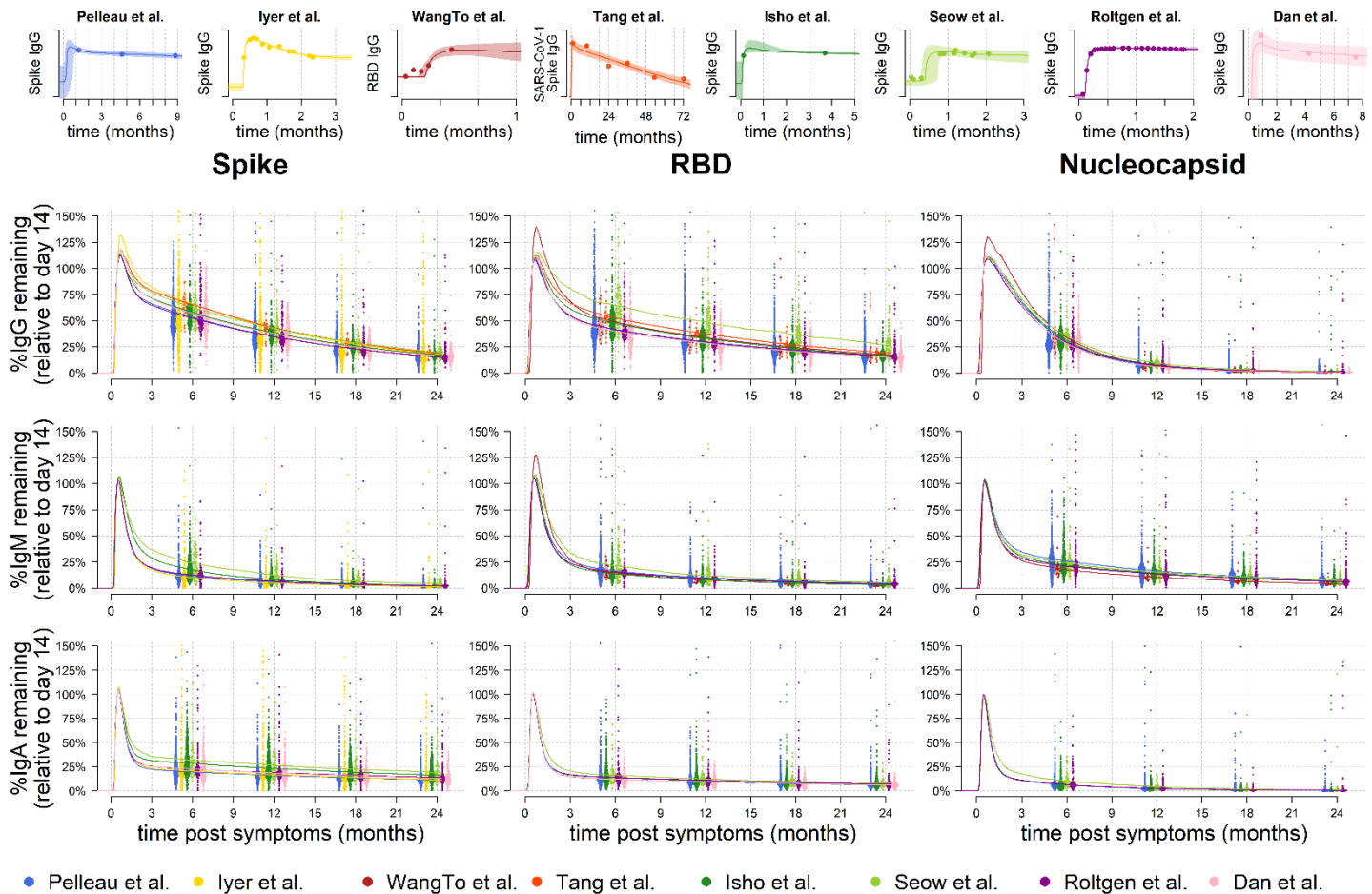


Figure 2: Modelled SARS-CoV-2 antibody kinetics. A mathematical model of SARS-CoV-2 antibody kinetics was simultaneously fitted to data from seven studies of SARS-CoV-2 and one study of SARS-CoV-1. **(a – top row)** Examples of the model fit to the data for one individual from each study. Data are represented as points, posterior median model prediction as lines, and 95% credible intervals as shaded areas. **(b – middle and bottom rows)** Model predicted duration of antibodies within the first 2 years following infection. Antibody levels are shown relative to the expected antibody level at day 14 following symptom onset. Each point represents the prediction from an individual at 6, 12, 18 and 24 months post symptoms. The median predictions for each of the eight studies are presented as lines.

Table 2: Estimated duration of antibody responses following SARS-CoV-2 infection. The percentage antibody level remaining over time is compared to the measured antibody level 14 days after symptom onset. Estimates are presented as the population median, with the 95% range due to inter-individual variation.

| | 6 months | 12 months | 24 months |
|------------------|-----------------|----------------|---------------|
| Spike IgG | 55% (16%, 100%) | 36% (11%, 94%) | 16% (5%, 55%) |
| RBD IgG | 43% (13%, 100%) | 31% (9%, 89%) | 16% (5%, 48%) |
| Nucleocapsid IgG | 30% (8%, 92%) | 7% (1%, 31%) | 0.8% (0%, 7%) |
| Spike IgM | 12% (1%, 52%) | 6% (0%, 27%) | 2% (0%, 9%) |
| RBD IgM | 16% (4%, 51%) | 9% (2%, 32%) | 4% (1%, 16%) |
| Nucleocapsid IgM | 23% (6%, 75%) | 15% (4%, 50%) | 7% (2%, 24%) |
| Spike IgA | 21% (4%, 82%) | 18% (4%, 67%) | 12% (3%, 47%) |
| RBD IgA | 12% (4%, 49%) | 10% (3%, 38%) | 6% (2%, 24%) |
| Nucleocapsid IgA | 6% (1%, 30%) | 3% (0%, 13%) | 0.6% (0%, 4%) |

Estimating time since previous SARS-CoV-2 infection

Using a dataset on measured IgG, IgM and IgA antibody levels from pre-pandemic negative controls and samples from individuals with SARS-CoV-2 infection confirmed by PCR-based testing and followed for up to eleven months after symptom onset, Random Forests binary classification algorithms were trained to identify individuals with previous SARS-CoV-2 infection (Appendix Figures 7 and 8). Next, a Random Forests multi-way classification was trained to simultaneously identify previous infection and estimate the time since infection (Figure 3). The diagnostic test identified samples from individuals with previous SARS-CoV-2 infection with 99% specificity and 98% (95% confidence interval: 94%, 99%) sensitivity (Figure 3a). The diagnostic test classified samples from individuals infected within the previous 3 months (Figure 3b). Notably, it was easier to distinguish recent infections (<3 months) from older infections (6 – 12 months), compared to infections that occurred 3 – 6 months ago. There was limited statistical signal to distinguish between infections that occurred 3 – 6 months ago, and older infections occurring more than 6 months ago (Figure 3c). A breakdown of classification performance by time since infection is provided in Figure 3D. The diagnostic test accurately classifies samples of all categories, with the exception of samples from individuals infected 3 – 6 months ago. Many of these samples were incorrectly classified in the neighbouring infection categories of 0 – 3 months or 6 – 12 months.

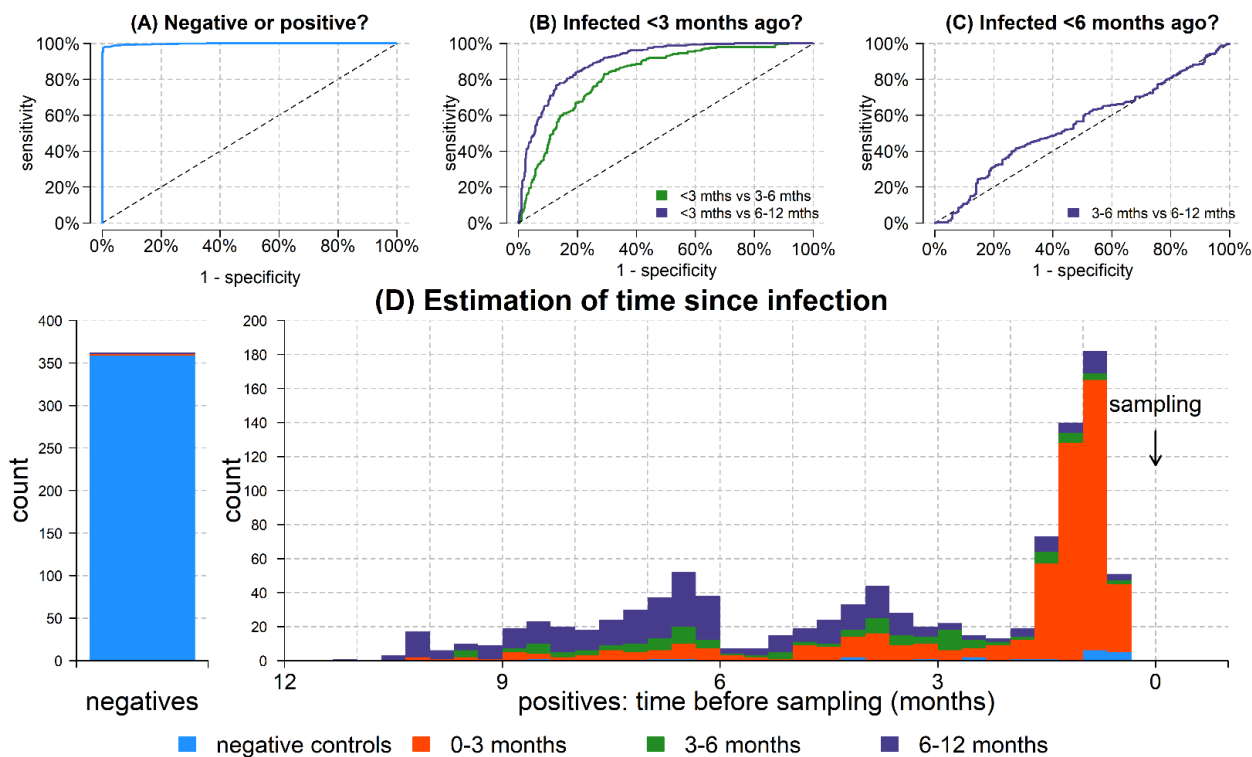


Figure 3: Classification of time since previous SARS-CoV-2 infection. A cross-validated multi-way classification algorithm was trained to estimate time since infection. **(A)** The algorithm can differentiate between positive and negative samples. **(B)** The algorithm can classify individuals infected within the previous 3 months. **(C)** There is limited diagnostic power to distinguish between infections that occurred 3-6 months ago versus 6-12 months ago. **(D)** Breakdown of classification performance according to time since previous infection. Colours represent model predicted classification. >99% of negative samples are correctly classified as negative (blue). For the positive samples, the distribution shows the time since previous infection. Samples with time since infection <3 months are mostly classified in the 0 – 3 month category (red). Samples with time since infection >6 months ago are mostly classified in the 6 – 12 month category (purple). There is a substantial degree of misclassification of samples with time since infection 3 – 6 months ago. This is due to the temporal imbalance in the training data.

Serological reconstruction of COVID-19 epidemics

When applied to samples from a cross-sectional sero-prevalence survey, a diagnostic test capable of estimating time since infection can provide a serological reconstruction of past SARS-CoV-2 epidemics. Figure 4 presents four simulated epidemics where 1,000 individuals were sampled, and 40% were previously infected. A range of scenarios were simulated, including a recent epidemic wave, an older epidemic wave, a double wave epidemic, and constant transmission. The algorithm was able to accurately reconstruct all epidemic profiles. For example, in the recent wave scenario 60% of the population were negative and the model estimated 59.8% (95% CI: 59.3%, 61.4%); 35.8% were infected 0 – 3 months ago and the model estimated 35.5% (95% CI: 25.7%, 40.3%); 4.2% were infected 3 – 6 months ago and the model estimated 4.2% (95% CI: 0.0%, 14.4%); and 0% were infected 6 – 12 months ago and the model estimated 0% (95% CI: 0.0%, 3.3%).

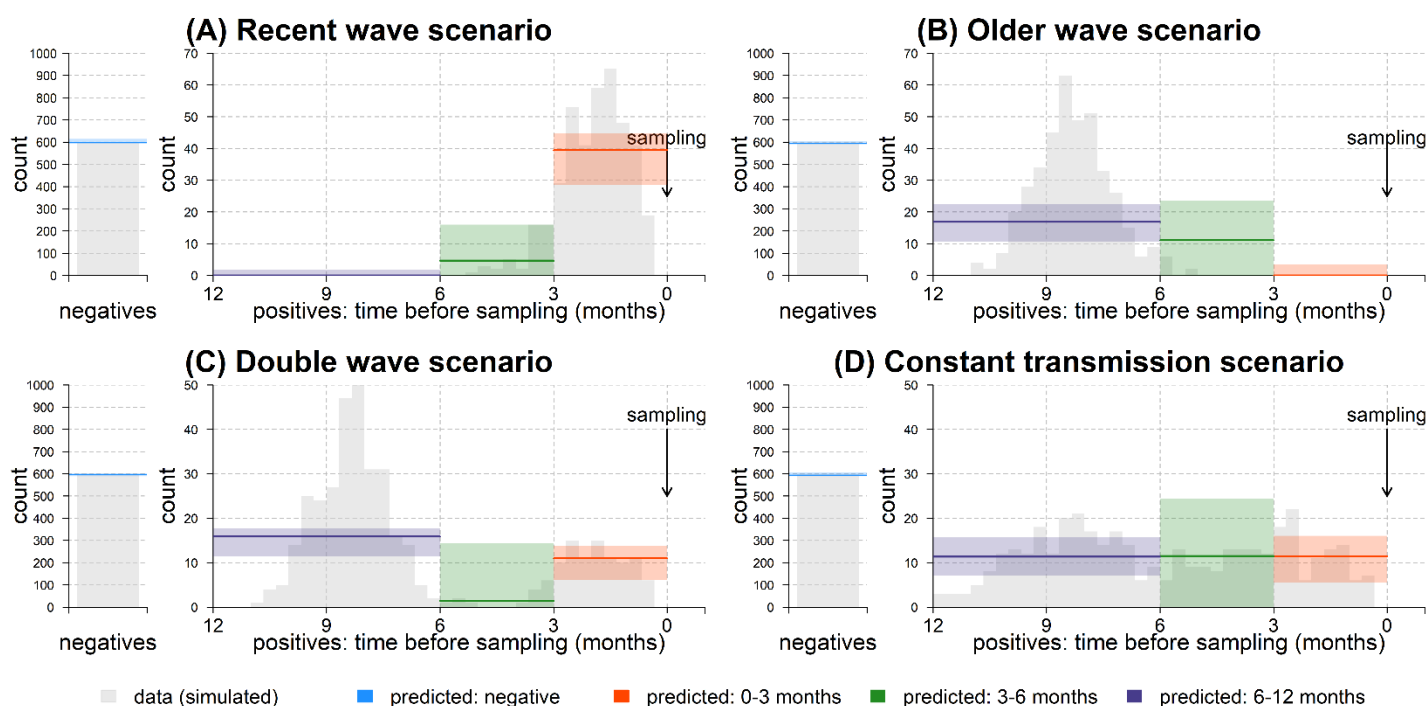


Figure 4: Serological reconstruction of previous COVID-19 epidemics. The grey histograms represent simulated data from a range of SARS-CoV-2 epidemic scenarios. Data were simulated from 1000 samples from a cross-sectional sero-prevalence study. It was assumed that 600 samples were negative, with 400 positive samples with varying time since infection. For previously infected individuals, the histograms represent the time since infection. Based on antibody levels sampled at a single time point, our computational method was used to provide a serological reconstruction of previous COVID-19 epidemics. The solid lines represent the model predictions for the categories of negative; infected 0 – 3 months ago; infected 3 – 6 months ago; and infected 6 – 12 months ago. Shaded regions represent 95% confidence intervals.

Discussion

Infection with SARS-CoV-2 induces a complex, diverse immune response, that varies by orders of magnitude between individuals, and changes over time. This diversity is a challenge to immunologists and vaccinologists, but presents an opportunity to diagnostic developers armed with multiplex assays and machine learning algorithms. By quantifying the kinetics of different components of the humoral immune response, it is possible to provide classification of previous infection that minimizes the reduction of diagnostic sensitivity over time, and also allows estimation of the time since infection.

Based on data from a range of studies with up to eleven months follow up after symptom onset, we estimate that 31% (95% CrI: 9%, 89%) of anti-RBD IgG antibody levels remain one year after infection. Antibodies of other isotypes waned more rapidly, with 9% (95% CrI: 2%, 32%) of anti-RBD IgM antibody levels remaining after one year, and 10% (95% CrI: 3%, 38%) of anti-RBD IgA antibody levels remaining after one year. There was considerable variation in kinetic profiles between different SARS-CoV-2 antigens, with 7% (95% CrI: 1%, 31%) of anti-N IgG antibody levels remaining after one year. Although the determinants of the duration of antigen-specific antibody responses remain poorly understood²², the diversity in patterns of antibody kinetics can be quantified using epidemiological studies, yielding valuable information for serological diagnostics.

The majority of commercially available serological tests classify individuals as having previous SARS-CoV-2 infection if a measured antibody response is greater than a defined cutoff³⁵. Instead of reducing a complex antibody response to a binary data point, a more detailed serological signature based on quantitative measurements of multiple antibody responses provides two notable advantages²⁰. Firstly, the reduction in diagnostic sensitivity associated with waning antibodies is minimized – no reduction in diagnostic sensitivity over the eleven months of follow up was observed. Secondly, the time since previous infection can be estimated providing valuable additional epidemiological information. For the current assay, previous infections were categorised into intervals of 0 – 3 months, 3 – 6 months, and 6 – 12 months. More precise classification is possible in theory, but this must be balanced against the statistical signal. For example, there was little statistical signal to discriminate between infections that occurred 6 – 9 months ago, versus infections that occurred 9 – 12 months ago. It is anticipated that further improvements to the assay such as the incorporation of new antigens, more training samples with a range of time since infection, and better algorithms will lead to improvements in accuracy.

Existing sero-prevalence studies estimate the proportion of a population previously infected with SARS-CoV-2. The addition of a diagnostic tool capable of estimating time since infection allows for the serological reconstruction of past transmission trends. Thus, using samples from a sero-prevalence study collected at a single time point, we can discriminate between a scenario of constant SARS-CoV-2 transmission and a scenario where transmission occurs in distinct epidemic waves. This diagnostic technology has a range of possible applications. For countries that experienced double wave SARS-CoV-2 epidemics, it has been challenging to quantify the relative magnitude of these waves due to the time required to

scale up PCR-based diagnostic testing³⁶. Furthermore, there are many epidemiological settings where it is unknown if SARS-CoV-2 transmission was constant over time, or occurred as distinct epidemic waves⁵.

The widespread introduction of SARS-CoV-2 vaccines will lead to vaccine-induced immunity in many individuals. The majority of vaccines against SARS-CoV-2 target the Spike protein's RBD. As our diagnostic assay and algorithms include Spike and RBD antigens, serum samples from vaccinated individuals would likely be classified as positive. However, by measuring antibody responses to Nucleocapsid, Membrane, Envelope, and other viral proteins³⁷ it will be possible to modify the diagnostic test to provide three-way classification: negative; previously infected; or vaccinated. Identification of individuals who have been both infected and vaccinated will be challenging. Having classified an individual as previously infected or vaccinated, the diagnostic algorithm can then be used to estimate time since infection or time since vaccination. Such an approach could contribute to efforts to measure population-level immunity to SARS-CoV-2, whether induced by infection or vaccination.

There are several limitations to this study. Estimates of the duration of antibody responses are based on data from multiple studies, each using a unique immunoassay^{6,7,8,9,10,11,12}. Every immunoassay may differ in terms of background reactivity, cross-reactivity with other pathogens, protein formulation, dynamic range and reproducibility. We believe that the benefit of drawing on multiple data sources outweighs the benefit of having a smaller more homogeneous database, especially since the mathematical model of antibody kinetics is sufficiently flexible to incorporate data from multiple assays. Our selection of a mechanistic mathematical model of antibody kinetics is a potential limitation. The model is based on a mechanistic understanding of the immunological processes underlying the generation and persistence of antibodies, and imposes a flexible functional form on how antibody levels change over time. Although this approach has been validated in a range of applications^{16,17,18,19}, there will be instances where the model fails to capture the true pattern of antibody kinetics, for example in immune-deficient individuals. An advantage of a mechanistic model versus a non-parametric statistical model is the ability to make projections forward in time. We have provided predictions up to two years following infection, for example by estimating that 16% (5%, 48%) of anti-RBD IgG antibodies remain after two years. There is a risk to providing predictions beyond the timescale of the data, but these predictions can be easily falsified via continued longitudinal studies.

The diagnostic assay is not exempt from the challenges of antibody waning. Although no reductions in diagnostic sensitivity over time were observed, reduced sensitivity will likely be observed as we analyse additional samples over longer durations of follow up. There are substantial challenges in providing estimates of time since infection. Although the approach can reliably reconstruct the distribution of infection times across a population, there will be substantial uncertainty in estimates of the time since previous infection for individual samples.

Sero-prevalence studies are playing a critical role in monitoring the progress of the SARS-CoV-2 pandemic. In the early stages of the pandemic, immunoassays had the advantage of measuring high antibody levels in the initial months after infection. As the pandemic progresses, sero-prevalence will become more challenging to accurately measure due to

waning antibody responses and increased vaccine-induced immunity. Multiplex assays and algorithms accounting for how antibody levels change over time may be an important tool for ensuring the ongoing utility of sero-surveillance.

Declaration of interests

MTW, IM, JR, SPel, MB, and SPet are inventors on provisional patent PCT/US 63/057.471 on a serological antibody-based diagnostics of SARS-CoV-2 infection. TB and OS are coinventors on provisional patent PCT/US 63/020,063 entitled “S-Flow: a FACS-based assay for serological analysis of SARS-CoV-2 infection” submitted by Institut Pasteur. SK reports personal fees from Accelerate Diagnostics, Astellas, Merck Sharp & Dohme, Pfizer, and Menarini, and grants and personal fees from bioMérieux, outside of the submitted work.

Acknowledgements

The French COVID cohort is supported by the REACTing consortium and by the French Directorate General for Health. Jessica Vanhomwegen (Institut Pasteur) is thanked for provision of a MAGPIX machine. Jérôme Hadjadj and Laura Barnabei (Institut Imagine, Paris) are thanked for their work on the Hôpital Cochin study. Dionicia Gamboa (Cayetano Heredia University, Lima) is thanked for sharing negative control samples from Peru. Jetsumon Sattabongkot (Mahidol University, Bangkok) is thanked for sharing negative control samples from Thailand. Marie-Noelle Ungeheuer and Blanca Liliana Perlaza are thanked for processing samples at the Clinical Investigation and Access to BioResources platform in Institut Pasteur. We thank all patients and health-care workers who kindly agreed for samples to be used for medical research purposes.

Data availability

Data and code will be made available online following peer review of the article. In the meantime, requests for data or code should be made to the corresponding author.

Funding

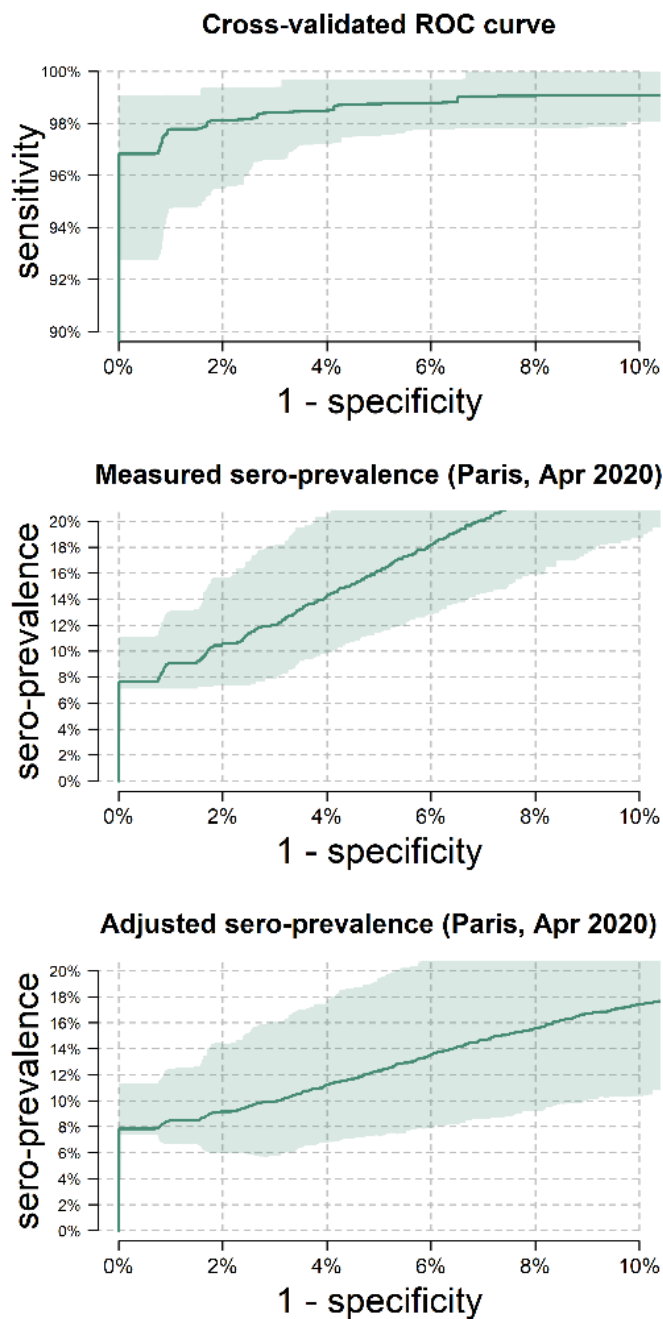
This work was supported by the European Research Council (MultiSeroSurv ERC Starting Grant 852373; MW), l’Agence Nationale de la Recherche and Fondation pour la Recherche Médicale (CorPopImm; MW), and the Institut Pasteur International Network (CoronaSeroSurv; MW). JR was supported by the Pasteur Paris University (PPU) International PhD Program. CC was supported by the European Research Council 771813. MB and AL were supported by the « URGENCE COVID-19 » fundraising campaign of Institut Pasteur. NC and CNC are part-funded by a Science Foundation Ireland (SFI) grant, Grant Code 20/SPP/3685. LT has been awarded the Irish Clinical Academic Training (ICAT) Programme, supported by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive, National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland (<https://icatprogramme.org/>). MB and AM were supported by Institut Pasteur TaskForce funding (TooLab project). SFK lab is funded by Strasbourg University Hospitals (SeroCoV-HUS; PRI 7782), the Agence Nationale de la Recherche (ANR-18-CE17-0028), Laboratoire d'Excellence TRANSPLANTEX (ANR-11-LABX-0070_TRANSPLANTEX), and Institut National de la Santé et de la Recherche Médicale (UMR_S 1109).

References

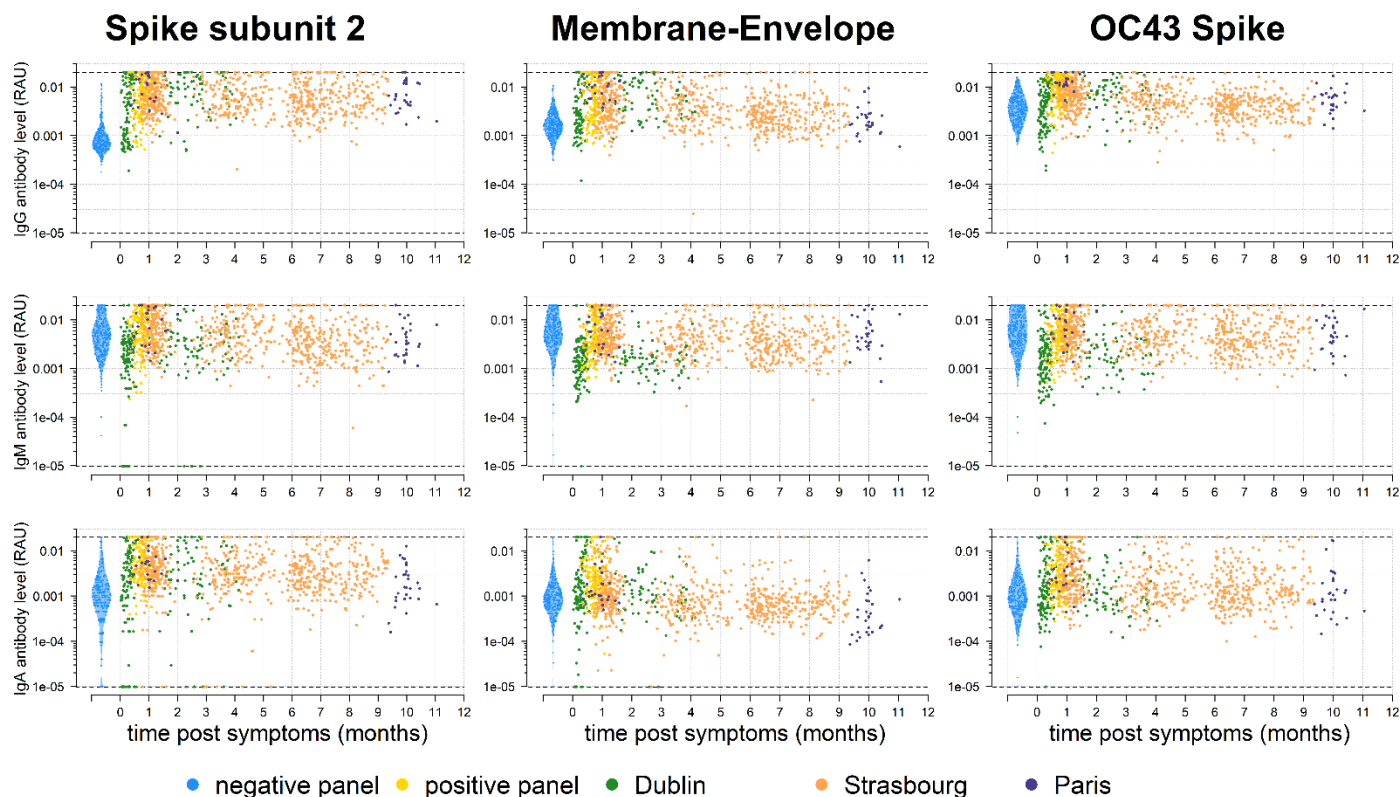
1. Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, *et al.* Estimating the burden of SARS-CoV-2 in France. *Science*. 2020; 369(6500): 208-211
2. Priesemann V, Balling R, Brinkmann MM, Ciesek S, Cypionka T, *et al.* An action plan for pan-European defence against new SARS-CoV-2 variants. *Lancet*. 2021; 397(10273): 469-470
3. Le Vu S, Jones G, Anna F, Rose T, Richard JP, *et al.* Prevalence of SARS-CoV-2 antibodies in France: results from nationwide serological surveillance. *medRxiv*. 2020; <https://doi.org/10.1101/2020.10.20.20213116>
4. Ward H, Cooke G, Atchison C, Whitaker M, Elliott, *et al.* Declining prevalence of antibody positivity to SARS-CoV-2: a community study of 365,000 adults. *medRxiv*. 2020; <https://doi.org/10.1101/2020.10.26.20219725>
5. Kagucia EW, Gitonga JN, Kalu C, Ochomo E, Ochieng B, *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies among truck drivers and assistants in Kenya. *medRxiv*. 2020; <https://doi.org/10.1101/2021.02.12.21251294>
6. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science*. 2021; 371(587)
7. To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, *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*. 2020; S1473-3099(20)30196-1.
8. Iyer AS, Jones FK, Nodoushania A, Kelly M, Beckera M, *et al.* Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Science Immunol*. 2020; 10.1126/sciimmunol.abe0367
9. Isho B, Abe KT, Zuo M, Jamal AJ, Rathod B, *et al.* Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Science Immunol*. 2020; 10.1126/sciimmunol.abe5511
10. Röltgen K, Powell AE, Wirz OF, Stevens BA, Hogan CA, *et al.* Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. *Science Immunol*. 2020; 10.1126/sciimmunol.abe0240
11. Seow J, Graham C, Merrick B, Acors S, Pickering S, *et al.* Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat Microbio*. 2020; 5: 1598-1607
12. Tang F, Quan Y, Xin ZT, Wrammert J, Ma MJ, Lv H, *et al.* Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J Immunol*. 2011; 186 (12) 7264-7268
13. Callow KA, Parry HF, Sergeant M, Tyrrell DA. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol Infect*. 1990; 105(2):435-46.
14. Choe PG, Perera RAPM, Park WB, Song KH, Bang JH, Kim ES, *et al.* MERS-CoV antibody responses 1 year after symptom onset, South Korea, 2015. *Emerg Infect Dis*. 2017; 23(7):1079-1084.
15. Slifka MK, Antia R, Whitmire JK, Ahmed R. Humoral immunity due to long-lived plasma cells. *Immunity*. 1998; 8: 363-72
16. Teunis PF, van Eijkeren JC, de Graaf WF, Marinović AB, Kretzschmar ME. Linking the seroresponse to infection to within-host heterogeneity in antibody production. *Epidemics* 2016; 16: 33-39.
17. Andraud M, Lejeune O, Musoro *et al.* Living on three time scales: the dynamics of plasma cell and antibody populations illustrated for hepatitis A virus. *PLoS Comput Biol* 2012; 8: e1002418.
18. White MT, Verity R, Griffin JT, *et al.* Immunogenicity of the RTS,S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial. *Lancet Infect Dis* 2015; 15: 1450-58.

19. White MT, Idoko O, Sow S, Diallo A, Kampmann B, *et al.* Antibody kinetics following vaccination with MenAfriVac: an analysis of serological data from randomised trials. *Lancet Inf Dis.* 2019; 19(3): 327-336
20. Rosado J, Pelleau S, Cockram C, Merklings SH, Demeret C, *et al.* Multiplex assays for the identification of serological signatures of SARS-CoV-2 infection: an antibody-based diagnostic and machine learning study. *Lancet Microbe.* 2020; 2(2): E60-E69
21. Buss LF, Prete Jr CA, Abraham CAA, Mendrone A, Salomon T, *et al.* Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science.* 2020; 371(6526): 288-292
22. Scepanovic P, Alanio C, Hammer C, Hodel F, Bergstedt J, *et al.* Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. *Genome Med.* 2018; 10(59)
23. Helb DA, Tetteh KK, Felgner PL, Skinner J, Hubbard A, Arinaitwe E, *et al.* Novel serologic biomarkers provide accurate estimates of recent *Plasmodium falciparum* exposure for individuals and communities. *Proc Natl Acad Sci USA.* 2015; 112(32):E4438-47
24. Yman V, Tuju J, White MT, *et al.* Distinct kinetics in antibody responses to 111 *Plasmodium falciparum* antigens identifies novel serological markers of recent malaria exposure. *medRxiv.* 2020; <https://doi.org/10.1101/2020.12.14.20242768>
25. Longley RJ, White MT, Takashima E, Brewster J, Morita M, Harbers M, *et al.* Development and validation of serological markers for detecting recent exposure to *Plasmodium vivax* infection. *Nature Med.* 2020; 26: 741-749
26. Azman AS, Lessler J, Luquero FJ, Bhuiyan TR, *et al.* Estimating cholera incidence with cross-sectional serology. *Sci Transl Med.* 2019; 11(480)
27. Lescure FX, Bouadma L, Nguyen D, Parisey M, Wicky PH, *et al.* Clinical and virological data of the first cases of COVID-19 in Europe: a case series. *Lancet Inf Dis.* 2020; 20(6): 697-706
28. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, *et al.* Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science.* 2020; 369(6504): 718-724
29. Fafi-Kremer S, Bruel T, Madec Y, Grant R, Tondeur L, *et al.* Serologic responses to SARS-CoV-2 infection among hospital staff with mild disease in eastern France. *EBioMed.* 2020; 59:102915
30. Grzelak L, Temmam S, Planchais C, Demeret C, Tondeur L, *et al.* A comparison of four serological assays for detecting anti-SARS-CoV-2 antibodies in human serum samples from different populations. *Sci Transl Med.* 2020; 12(559):eabc3103
31. Dufloo J, Grzelak L, Staropoli I, Madec Y, Tondeur L, *et al.* Asymptomatic and symptomatic SARS-CoV-2 infections elicit polyfunctional antibodies. *medRxiv.* 2020; <https://doi.org/10.1101/2020.11.12.20230508>
32. Planas D, Bruel T, Grzelak L, Guivel-Benhassine F, Staropoli I, *et al.* Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *medRxiv.* 2021; <https://doi.org/10.1101/2021.02.12.430472>
33. White MT, Griffin JT, Akpogheneta O, *et al.* Dynamics of the antibody response to *Plasmodium falciparum* infection in African children. *J Infect Dis* 2014; 210: 1115–22.
34. Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. *J Roy Stat Soc C.* 2020; 69: 1269-1283
35. National SARS-CoV-2 Serology Assay Evaluation Group. Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Inf Dis.* 2020; 20(12): 1390-1400
36. Pullano G, Di Domenico L, Sabbatini CE, Valdano E, Turbelin C, *et al.* Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature.* 2021; 590:134-139
37. Hachim A, Kavian N, Cohen CA, Chin AWH, Chu DKW, *et al.* ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nature Immunol.* 2020; 12: 1293-1301

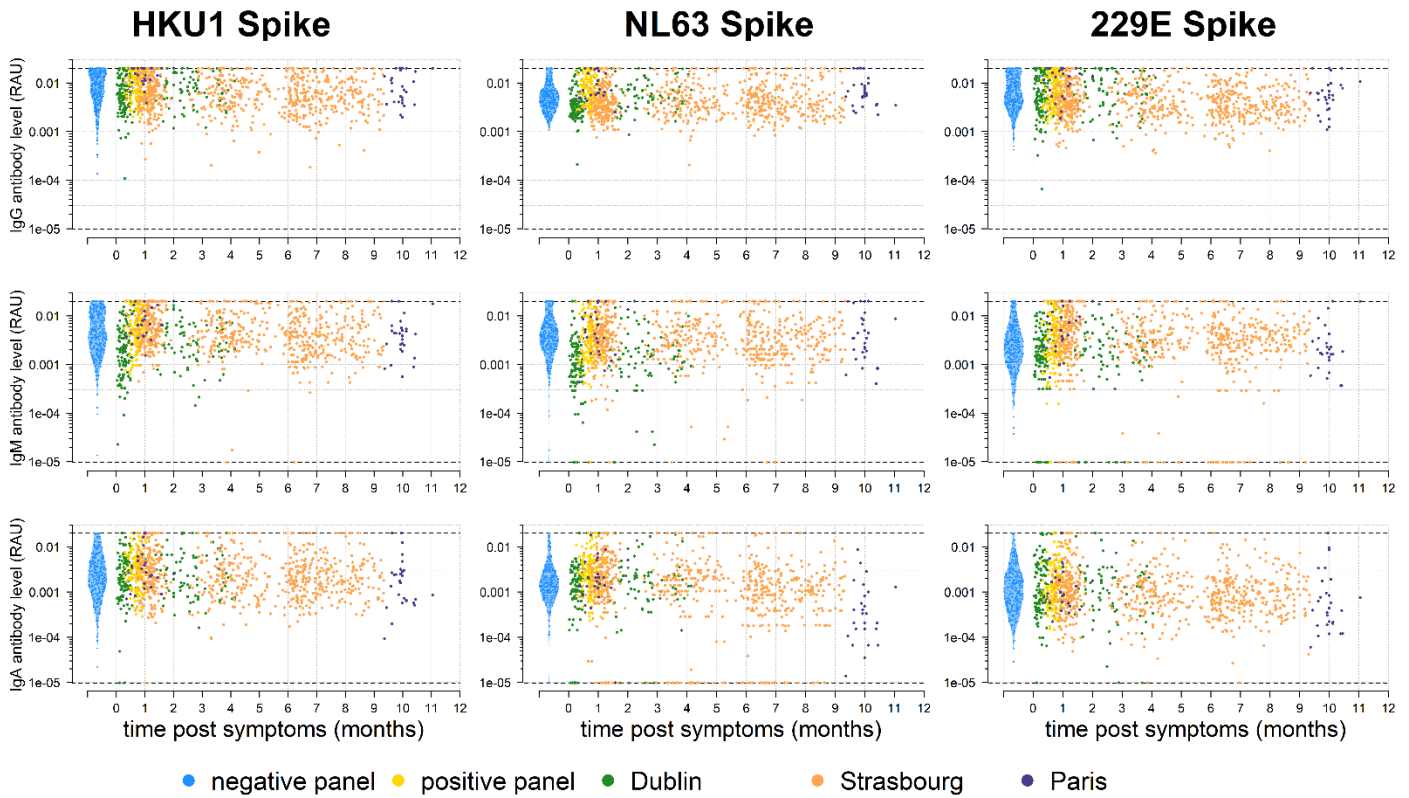
Appendix Figures



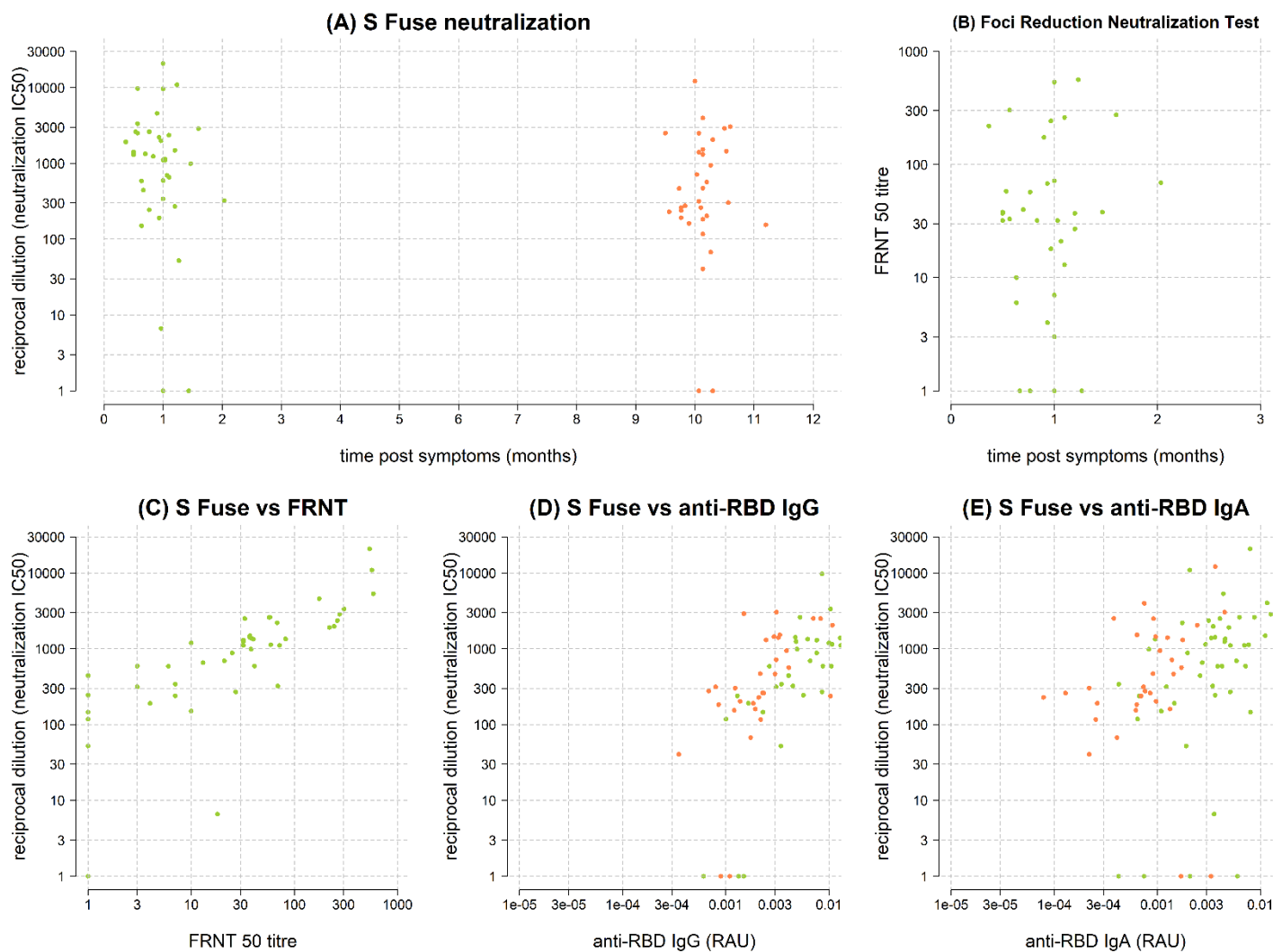
Appendix Figure 1: Application of multiplex assay to a sero-prevalence study in Institut Mutualiste Montsouris hospital in Paris. (A) Cross-validated receiver operating characteristic (ROC) curve showing the diagnostic performance of a multiplex assay for measuring IgG, IgM and IgA antibodies to SARS-CoV-2 antigens. This assay can achieve 98% sensitivity and 99% specificity. (B) Measured sero-prevalence in 769 healthcare workers across a range of diagnostic specificities. (C) Sero-prevalence adjusted with a Rogan-Gladen estimator. It can be shown that accuracy is maximized by selecting high specificity²⁰. At 99% specificity, this gives an estimated sero-prevalence of 8.7% (95 CI: 6.7%, 11.0%). In all panels, the green line represents the median prediction and the shaded regions represent the 95% confidence intervals.



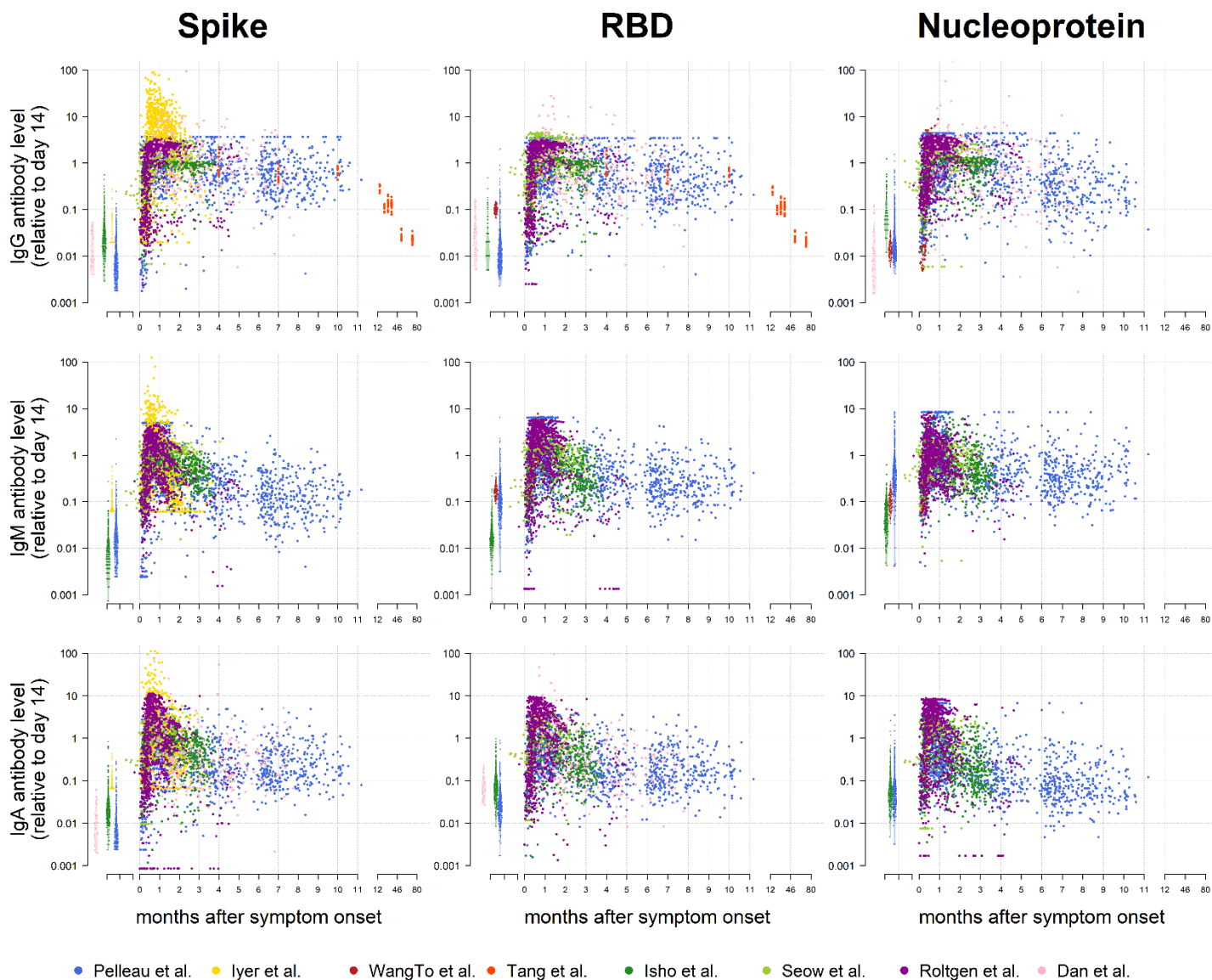
Appendix Figure 2: Antibody kinetics in the first year following infection with SARS-CoV-2. A bead-based multiplex Luminex assay was used to measure antibodies of multiple isotypes (IgG, IgM, IgA) to multiple antigens in serum samples from individuals with PCR-positive SARS-CoV-2 infection and pre-pandemic negative controls. Shown here are the antibody responses to the SARS-CoV-2 Spike subunit 2 (S2) and Membrane-Envelope fusion (ME) antigens, and the Spike protein of the OC43 human seasonal coronavirus.



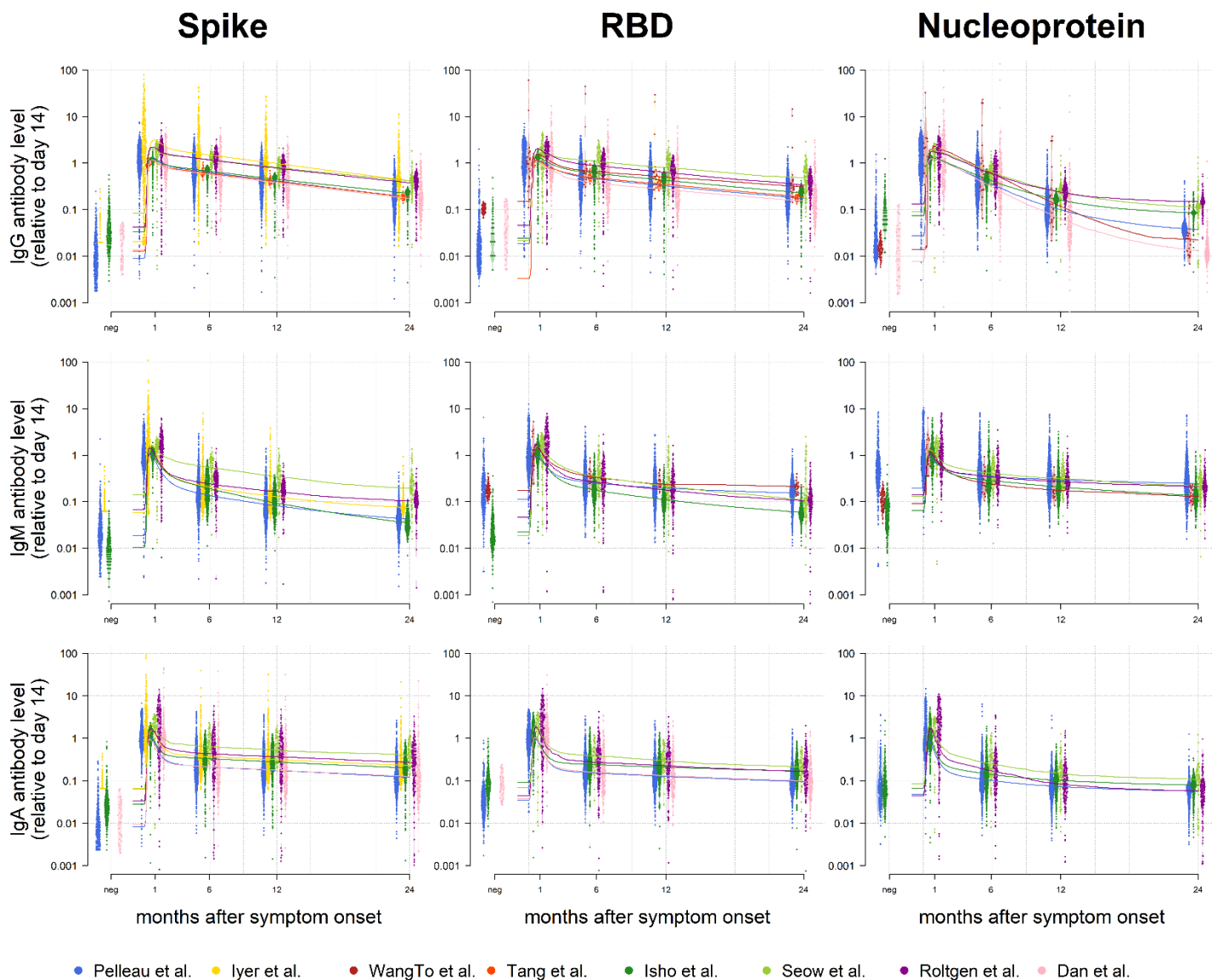
Appendix Figure 3: Antibody kinetics in the first year following infection with SARS-CoV-2. A bead-based multiplex Luminex assay was used to measure antibodies of multiple isotypes (IgG, IgM, IgA) to multiple antigens in serum samples from individuals with PCR-positive SARS-CoV-2 infection and pre-pandemic negative controls. Shown here are the antibody responses to the Spike proteins of the HKU1, NL63 and 229E human seasonal coronaviruses.



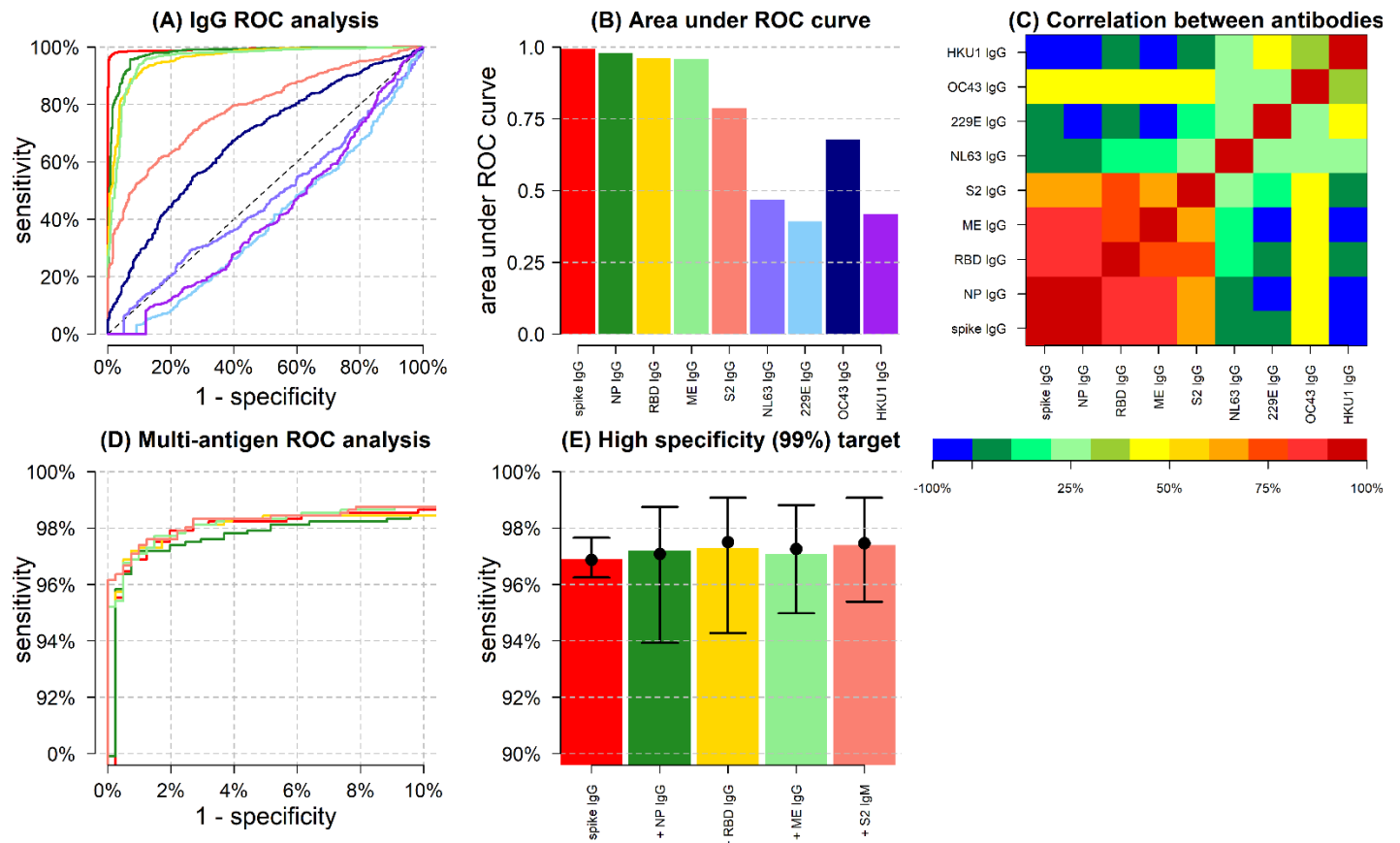
Appendix Figure 4: Association between measured antibody levels and viral neutralization. (A) Viral neutralization (denoted as reciprocal dilution for 50% neutralization) in serum samples from Institut Mutualiste Montsouris. **(B)** Foci reduction viral neutralization (denoted as reciprocal dilution for 50% neutralization) in serum samples from Institut Mutualiste Montsouris. **(C)** Comparison between two viral neutralization assays. **(D)** Comparison between viral neutralization and anti-RBD IgG. **(E)** Comparison between viral neutralization and anti-RBD IgA.



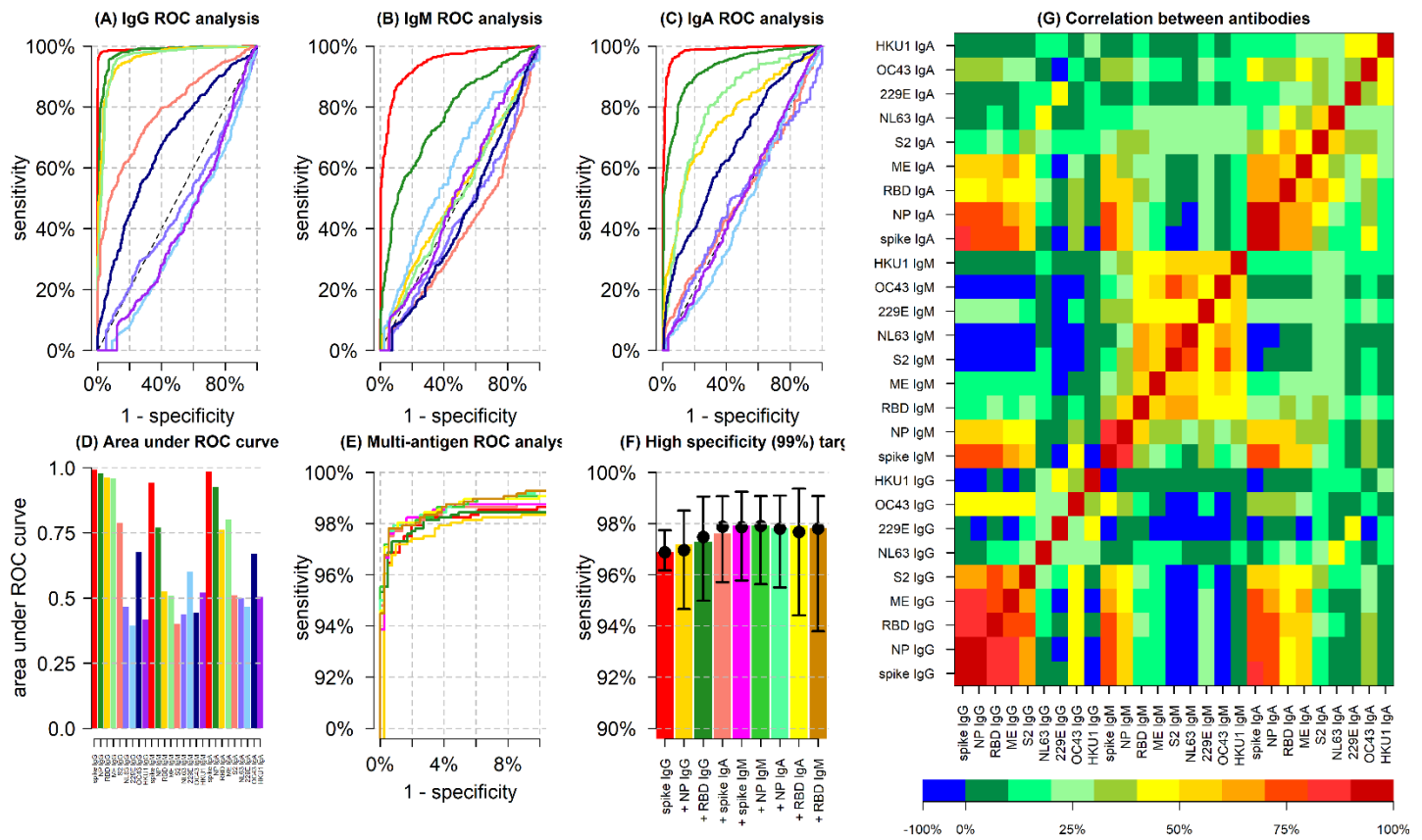
Appendix Figure 5: Data on SARS-Cov-2 antibody kinetics. Overview of data used for fitting mathematical model of antibody kinetics from seven studies with longitudinal follow-up of individuals after infection with SARS-CoV-2, and one study with longitudinal follow-up after infection with SARS-CoV-1. Note that the study by Tang *et al.* measured antibodies to SARS-CoV-1 over six years of follow-up, hence the cut in x-axis. There were considerable differences in the immunoassays used for each study, notably variation in antigens and different platforms. For each study, the data has been normalized so that at day 14 post symptom onset, the expected antibody level equals 1. Measured antibody levels from negative controls are shown in the left portion of the plot.



Appendix Figure 6: Modelled SARS-CoV-2 antibody kinetics. Overview of the fit of a mathematical model of antibody kinetics to data from seven studies with longitudinal follow-up of individuals after infection with SARS-CoV-2, and one study with longitudinal follow-up after infection with SARS-CoV-1. For each study, the data has been normalized so that at day 14 post symptom onset, the expected antibody level equals 1. Measured antibody levels from negative controls are shown in the left portion of the plot. Predicted antibody kinetics are shown for a period of 2 years. The solid line represents the median predicted antibody level for each study.



Appendix Figure 7: Overview of classification performance of assay and algorithm on IgG data. (A) Receiver Operating Characteristic (ROC) curves for IgG antibodies obtained by varying the cutoff for seropositivity. Colours correspond to antibodies against different antigens, as shown in panel B. **(B)** Area under the ROC curve (AUC) for individual IgG antibodies. **(C)** Spearman's correlation between measured antibody responses. **(D)** ROC curves for multiple biomarker classifiers generated using a random forests algorithm. Biomarkers are added sequentially with colours corresponding to panel E. **(E)** For a high specificity target (>99%), sensitivity increases with additional biomarkers, added sequentially. Sensitivity was estimated using a random forests classifier. Points and whiskers denote the median and 95% CIs from repeat cross-validation.



Appendix Figure 8: Overview of classification performance of assay and algorithm on IgG, IgM and IgA data. (A) Receiver Operating Characteristic (ROC) curves for IgG antibodies obtained by varying the cutoff for seropositivity. Colours correspond to antibodies against different antigens, as shown in panel D. **(B)** ROC curves for IgM antibodies obtained by varying the cutoff for seropositivity. **(C)** ROC curves for IgA antibodies obtained by varying the cutoff for seropositivity. **(D)** Area under the ROC curve (AUC) for individual IgG, IgM and IgA antibodies. **(E)** ROC curves for multiple biomarker classifiers generated using a random forests algorithm. Biomarkers are added sequentially with colours corresponding to panel F. **(F)** For a high specificity target (>99%), sensitivity increases with additional biomarkers, added sequentially. Sensitivity was estimated using a random forests classifier. Points and whiskers denote the median and 95% CIs from repeat cross-validation. **(G)** Spearman's correlation between measured antibody responses.

Statistical Appendix

Data on longitudinal antibody responses to SARS-CoV-2 proteins

Several teams have conducted studies of the duration of the antibody response against SARS-CoV-2 with varying durations of follow up. Our data analyses the duration of the SARS-CoV-2 antibody response over approximately eleven months. To better characterize the antibody kinetics over different time scales, we supplemented our database with data from six published studies on SARS-CoV-2 antibody kinetics. These data are summarized in Appendix Table 1. In addition to the studies on SARS-CoV-2 antibody kinetics, we also included data from one study of SARS-CoV-1 antibody kinetics in individuals with six years follow-up¹².

Appendix Table 1: Overview of studies incorporated in antibody kinetic model.

| study | assay | antigens | isotype | max follow up (days) | positive individuals | positive samples | negative samples |
|-------------------------------------|-------------------|--|---------------|----------------------|----------------------|------------------|------------------|
| Pelleau <i>et al.</i> (this study) | Luminex multiplex | Spike; RBD; S2; NP; ME; Spike (229E); Spike (NL63); Spike (OC43); Spike (HKU1) | IgG; IgM; IgA | | | | |
| Iyer <i>et al.</i> ⁸ | ELISA | Spike | IgG; IgM; IgA | 122 | 340 | 1088 | 1548 |
| Wang To <i>et al.</i> ⁷ | ELISA | RBD; NP | IgG; IgM | 28 | 22 | 104 | 93 |
| Tang <i>et al.</i> ¹² | ELISA | Spike (SARS-CoV-1) | IgG | 2160 | 23 | 147 | |
| Isho <i>et al.</i> ⁹ | ELISA | Spike; RBD; NP | IgG; IgM; IgA | 115 | 447 | 508 | 300 |
| Seow <i>et al.</i> ¹¹ | ELISA | Spike; RBD; NP | IgG; IgM; IgA | 94 | 96 | 302 | |
| Roltgen <i>et al.</i> ¹⁰ | ELISA | S1; RBD; NP | IgG; IgM; IgA | 150 | 165 | 841 | |
| Dan <i>et al.</i> ⁶ | ELISA | Spike; RBD; NP | IgG; IgA | 240 | 176 | 228 | 51 |

Mathematical model of antibody kinetics

SARS-CoV-2 antibody kinetics are described using a previously published mathematical model of the immunological processes underlying the generation and waning of antibody responses following infection or vaccination^{20,33}. For antibodies of three classes (IgG, IgM and IgA) targeting the studied antigens, the antibody kinetics are described by the following equations:

$$\begin{aligned}\frac{dP_s}{dt} &= \rho\beta - c_s P_s, \\ \frac{dP_l}{dt} &= (1 - \rho)\beta - c_l P_l, \\ \frac{dA}{dt} &= gP_s + gP_l - rA\end{aligned}$$

where β denotes the boost in antibody secreting plasma cells. It is assumed that a proportion ρ of plasma cells are short-lived (P_s) waning at rate c_s , and a proportion $1 - \rho$ are long-lived (P_l) waning at rate c_l . Plasma cells generate antibodies

(IgG, IgM or IgA) at rate g , and r is the rate of decay of antibody molecules. Assuming $P_s(0) = P_l(0) = 0$ and $A(0) = A_{bg}$, these equations can be solved analytically to give:

$$A(t) = A_{bg} + gB_0 \left(\frac{\rho(e^{-c_s(t-\delta)} - e^{-r(t-\delta)})}{(r - c_s)} + \frac{(1 - \rho)(e^{-c_l(t-\delta)} - e^{-r(t-\delta)})}{(r - c_l)} \right)$$

δ is the time after symptom onset when antibody levels start to increase. B_0 is the number of B cells, and g is the rate at which they secrete antibodies. As g and B_0 are not both identifiable without detailed and invasive experiments (e.g. bone marrow aspirates to measure antigen-specific plasma cells), we estimate $\beta = gB_0$. If r is the decay rate of antibody molecules, then we define $d_a = \frac{\log(2)}{r}$ to be the half-life of antibody molecules. Similarly, we define $d_s = \frac{\log(2)}{c_s}$ and $d_l = \frac{\log(2)}{c_l}$.

Note that this model of antibody kinetics follows the formulation outlined in White *et al.*³³ and not that outlined in Rosado *et al.*²⁰ which included an additional equation for modelling the proliferation of memory B cells. Following experiments on model identifiability with simulated data, it was found that there was limited ability to recover model parameters describing memory B cell proliferation.

Methodology for statistical inference

The model was fitted to longitudinal antibody level measurements from all participants. Mixed effects methods were used to capture the natural variation in antibody kinetics between individual participants, whilst estimating the average value and variance of the immune parameters across the entire population of individuals. The models were fitted in a Bayesian framework using Markov Chain Monte Carlo (MCMC) methods. Mixed effects methods allow individual-level parameters to be estimated for each participant separately, with these individual-level (or mixed effects) parameters being drawn from global distributions.

We consider a scenario where we have data from M studies, each of which has data from N_m individuals. For example, for each participant n in study m the half-life of the short-lived ASCs may be estimated as $d_s^{m,n}$ (an individual-level parameter). The $N = \sum_{m=1}^M N_m$ estimates of the local parameters $d_s^{m,n}$ will be drawn from a probability distribution. A log-Normal distribution is suitable as it has positive support on $[0, \infty)$. Thus we have $\log(d_s^{m,n}) \sim N(\mu_s, \sigma_s^2)$. The mean d_s and the variance Σ_s^2 of the estimates of $d_s^{m,n}$ are given by $d_s = e^{\mu_s + \frac{\sigma_s^2}{2}}$ and $\Sigma_s^2 = (e^{\sigma_s^2} - 1)e^{2\mu_s + \sigma_s^2}$.

Model likelihood

For individual n from study m we have data on observed antibody levels $A^{m,n} = \{a_1, \dots, a_j\}$ at times $T^{m,n} = \{t_1, \dots, t_j\}$. We denote $D^{m,n} = (A^{m,n}, T^{m,n})$ to be the vector of data for individual n from study m . For individual n from study m , the

parameters $\theta^{m,n} = (A_{bg}^{m,n}, \beta^{m,n}, \delta^{m,n}, d_s^{m,n}, d_l^{m,n}, d_a^{m,n}, \rho^{m,n})$ are estimated. The model predicted antibody levels will be $\{A(t_1), A(t_2), \dots, A(t_j)\}$. We assume log-Normally distributed measurement error such that the difference between $\log(a_j)$ and $\log(A(t_j))$ is Normally distributed with variance $\sigma_{obs,m}^2$. The parameters for observational variance are specific for each study due to differences in the assays utilised. For model predicted antibody levels $A(t_j)$ the data likelihood for individual n from study m is given by

$$L_{mod}^{m,n}(\theta^{m,n} | D^{m,n}) = \prod_{j \in J} \frac{e^{-\frac{(\log(a_j) - \log(A(t_j)))^2}{2\sigma_{obs,m}^2}}}{a_j \sigma_{obs,m} \sqrt{2\pi}}$$

Mixed effects likelihood

As described above, for each individual there are seven parameters to be estimated. These individual-level parameters are drawn from population-level distributions. Some parameters are assumed not to depend on the assay in each study (the antibody kinetic parameters: $\delta^{m,n}, d_b^{m,n}, d_s^{m,n}, d_l^{m,n}, d_a^{m,n}, \rho^{m,n}$). Each of these parameters are assumed to be drawn from the same distribution across all studies. Some parameters are assay dependent, notably $A_{bg}^{m,n}, \beta^{m,n}$. These parameters are drawn from distributions specific for each study. The mixed effects likelihood can be written as follows:

$$L_{mix}^{m,n}(\theta^{m,n} | D^{m,n}) = \left(\frac{e^{-\frac{(\log(A_{bg}^{m,n}) - \mu_{A,m})^2}{2\sigma_{A,m}^2}}}{\sqrt{2\pi} A_{bg}^{m,n} \sigma_{A,m}} \right) \left(\frac{e^{-\frac{(\log(\beta^{m,n}) - \mu_{\beta,m})^2}{2\sigma_{\beta,m}^2}}}{\sqrt{2\pi} \beta^{m,n} \sigma_{\beta,m}} \right) \left(\frac{e^{-\frac{(\log(\delta^{m,n}) - \mu_{\delta})^2}{2\sigma_{\delta}^2}}}{\sqrt{2\pi} \delta^{m,n} \sigma_{\delta}} \right) \\ \left(\frac{e^{-\frac{(\log(d_s^{m,n}) - \mu_s)^2}{2\sigma_s^2}}}{\sqrt{2\pi} d_s^{m,n} \sigma_s} \right) \left(\frac{e^{-\frac{(\log(d_l^{m,n}) - \mu_l)^2}{2\sigma_l^2}}}{\sqrt{2\pi} d_l^{m,n} \sigma_l} \right) \left(\frac{e^{-\frac{(\log(d_a^{m,n}) - \mu_a)^2}{2\sigma_a^2}}}{\sqrt{2\pi} d_a^{m,n} \sigma_a} \right) \left(\frac{e^{-\frac{(\log(\frac{\rho^{m,n}}{1-\rho^{m,n}}) - \mu_{\rho})^2}{2\sigma_{\rho}^2}}}{\sqrt{2\pi} \rho^{m,n} (1 - \rho^{m,n}) \sigma_{\rho}} \right) L_{mod}^{m,n}(\theta^{m,n} | D^{m,n})$$

As the proportion of the ASCs that are long-lived must be bounded by 0 and 1, the individual-level parameters $\rho^{m,n}$ are assumed to be drawn from logit-Normal distributions.

Total model likelihood

Denote $D = \{D^1, \dots, D^N\}$ to be the vector of data for all N participants from all studies. We denote $\theta^{m,n} = (\mu_{A,m}, \sigma_{A,m}, \mu_{\beta,m}, \sigma_{\beta,m}, \mu_{\delta}, \sigma_{\delta}, \mu_s, \sigma_s, \mu_l, \sigma_l, \mu_a, \sigma_a, \mu_{\rho}, \sigma_{\rho}, \theta^1, \dots, \theta^N)$ to be the combined vector of population-level parameters and individual-level parameters to be estimated. The total likelihood is obtained by multiplying the likelihood for each participant

$$L_{total}(\theta|D) = \prod_{n \in N} L_{mix}^{m,n}(\theta^n|D^n)$$

Markov Chain Monte Carlo parameter update

The model was fitted to the data using Markov Chain Monte Carlo (MCMC) methods. A three stage parameter update regimen was utilised with a Metropolis-within-Gibbs sampler with sequential updating of individual-level parameters, population-level parameters, and observational variance parameters. A' indicates an attempted update.

1. Individual-level parameter Metropolis-Hastings update.

- For each participant n in study m :
- Update local parameters: $\theta^{m,n'} = (A_{bg}^{m,n'}, \beta^{m,n'}, \delta^{m,n'}, d_s^{m,n'}, d_l^{m,n'}, d_a^{m,n'}, \rho^{m,n'})$
 - Calculate updated mixed effects likelihood $L_{mix}^{m,n}(\theta^{m,n'}|D^{m,n})$
 - Accept the parameter update with probability $\min\left(1, \frac{L_{mix}^{m,n}(\theta^{m,n'}|D^{m,n})}{L_{mix}^{m,n}(\theta^{m,n}|D^{m,n})}\right)$

2. Population-level parameter Gibbs update.

- For each of the $i \in \{A_{bg}^m, \beta^m\}$ we obtain new estimates of the population level parameters μ_i' and $\tau_i' = \frac{1}{\sigma_i'}$ as follows:

$$\mu_i' \sim N\left(\frac{\tau_{i,0}\mu_{i,0} + \tau_i \sum_{N_m} \log(y_i^n)}{\tau_{i,0} + \tau_i N_m}, \frac{1}{\sqrt{\tau_{i,0} + \tau_i N_m}}\right)$$

$$\tau_i' \sim \Gamma\left(\frac{N_m}{2} + k, \left(\frac{1}{\theta} + \frac{1}{2} \sum_{N_m} (\log(y_i^n) - \mu_i')^2\right)^{-1}\right)$$

where $\mu_{i,0}$ and $\tau_{i,0}$ parameterise the Normal prior distribution on the mean, and k and θ parameterise the Gamma prior distribution on the precision (inverse standard deviation). y_i^n is the individual-level parameter i in individual n .

- For each of the $i \in \{\delta, s, l, a\}$ we obtain new estimates of the population level parameters μ_i' and $\tau_i' = \frac{1}{\sigma_i'}$ as follows:

$$\mu_i' \sim N\left(\frac{\tau_{i,0}\mu_{i,0} + \tau_i \sum_m \sum_n \log(y_i^n)}{\tau_{i,0} + \tau_i N}, \frac{1}{\sqrt{\tau_{i,0} + \tau_i N}}\right)$$

$$\tau_i' \sim \Gamma\left(\frac{N}{2} + k, \left(\frac{1}{\theta} + \frac{1}{2} \sum_m \sum_n (\log(y_i^n) - \mu_i')^2\right)^{-1}\right)$$

where $\mu_{i,0}$ and $\tau_{i,0}$ parameterise the Normal prior distribution on the mean, and k and θ parameterise the Gamma prior distribution on the precision (inverse standard deviation). y_i^n is the individual-level parameter i in individual n .

- For ρ we obtain new estimates of the population level parameters μ_{ρ}' and $\tau_{\rho}' = \frac{1}{\sigma_{\rho}'}$ as follows:

$$\mu_{\rho}' \sim N\left(\frac{\tau_{\rho,0}\mu_{\rho,0} + \tau_{\rho} \sum_m^M \sum_n^{N_m} \text{logit}(\rho^n)}{\tau_{\rho,0} + \tau_{\rho}N}, \frac{1}{\sqrt{\tau_{\rho,0} + \tau_{\rho}N}}\right)$$

$$\tau_{\rho}' \sim \Gamma\left(\frac{N}{2} + k, \left(\frac{1}{\theta} + \frac{1}{2} \sum_m^M \sum_n^{N_m} (\text{logit}(\rho^n) - \mu_{\rho}')^2\right)^{-1}\right)$$

where $\mu_{\rho,0}$ and $\tau_{\rho,0}$ parameterise the Normal prior distribution on the mean, and k and θ parameterise the Gamma prior distribution on the precision (inverse standard deviation).

3. Observational variance parameter Metropolis-Hastings Update.

- For each study m , update the observational variance parameter $\sigma_{obs,m}'$
- Calculate updated total likelihood $L_{total}(\theta'|D)$ and the updated prior probability density $P(\theta')$
- Accept the parameter update with probability $\min\left(1, \frac{L_{total}(\theta'|D)P(\theta')}{L_{total}(\theta|D)P(\theta)}\right)$

The MCMC algorithm was implemented in C++ compiled in Microsoft Visual Studio. The covariance of the multivariate-Normal proposal distributions for Metropolis-Hastings updates were adaptively tuned using the estimated posterior distributions during a burn-in phase of 1 million MCMC iterations. The magnitude of the proposed step size was calibrated using a Robbins-Munro algorithm to achieve an acceptance rate of approximately 23%. The total number of MCMC iterations was 10,000,000. The effective number of iterations was calculated using the effectiveSize routine in the R library coda and the effective size was checked to be $> 1,000$ for all parameters. The statistical inference procedure was repeated twice to allow for assessment of chain convergence using Gelman-Rubin convergence diagnostics R_c . For all population-level chains, we ensure $R_c < 1.05$. For all individual-level chains, we ensure $R_c < 1.1$.

Serological surveillance algorithms

A ROC curve obtained from a training data set consisting of positive and negative samples is described by a sequence of estimated sensitivities and specificities $\{E(se_i, sp_i)\}$, where E denotes an estimator. N -fold cross-validation generates samples of sensitivity $\{se_{i1}, \dots, se_{iN}\}$ for each $E(sp_i)$ and samples of specificity $\{sp_{i1}, \dots, sp_{iN}\}$ for each $E(se_i)$. Following a previously outlined approach³⁴, for each pair i of sensitivity and specificity, we obtain N estimates of the measured seroprevalence M_{in} in a scenario with true seroprevalence T as follows:

$$M_{in} = Tse_{in} + (1 - T)(1 - sp_{in})$$

Following the Rogan-Gladen estimator approach, this equation can be rearranged to give an adjusted estimate of true seroprevalence:

$$E(T_{in}) = \frac{M_{in} + E(sp_i) - 1}{E(se_i) + E(sp_i) - 1}$$

with $E(T_{in}) = 0$ if $M_{in} < 1 - E(se_i)$. Both M_i and $E(T_i)$ are summarized as medians with 95% ranges.

Define α^K as a K -way classification algorithm that can classify a sample into one of K categories. Let C_0 be the category for a sample from an individual not previously infected with SARS-CoV-2. Let C_k be the category for individuals previously infected with SARS-CoV-2 in one of $K - 1$ time intervals defined by $(0, t_1, \dots, t_{K-1})$. α^K can be any of a range of algorithms from Random Forests, Support Vector Machines, or ordinal logistic regression. α^K can be trained on a data set, with classification performance assessed using N -fold cross-validation. For a serum sample, let P_k denote the classification score (analogous to a probability) that this sample belongs to category k .

Given a vector of classification scores P_k , we can assign a serum sample to a category. A frequently used approach is to select k according to the maximum of P_k . In order to retain better control over diagnostic specificity, we select an alternative two step approach:

- if $P_0 \geq \gamma_{99}$ then the sample is classified as negative
- if $P_0 < \gamma_{99}$ then the sample is classified into category k according to the maximum of P_1, \dots, P_{K-1}

Here γ_{99} is selected based on the training data to ensure a target specificity of 99%. The classification performance of this approach can be assessed using a confusion matrix C_{ij} defined by the proportion of samples of category i classified as category j . For the example with $K = 4$ explored in more detail here, we define three intervals: 0 – 3 months; 3 – 6 months; 6 – 12 months.

Following the approach of the Rogan-Gladen estimator, define \mathbf{M} to be the vector of measured prevalences in each of the K categories. An estimator of the true sero-prevalence is provided by $E(\mathbf{T}) = \mathbf{C}^{-1} * \mathbf{M}$ where \mathbf{C}^{-1} is the inverse of the confusion matrix. Depending on the values of the confusion matrix, this approach may have $T_k < 0$. In this instance we assign $T_k = 0$ and transfer the predictions proportionally to the other categories.