Antibody responses to SARS-CoV-2 train machine learning to assign likelihood of past infection during virus emergence in Sweden

Xaquín Castro Dopico[1#], Leo Hanke[1°], Daniel J. Sheward[1°], Sandra Muschiol[2], Soo Aleman[3], Murray Christian[1], Nastasiya F. Grinberg[4], Monika Adori[1], Laura Perez Vidakovics[1], ChangIl Kim[1], Sharesta Khoenkhoen[1], Pradeepa Pushparaj[1], Ainhoa Moliner Morro[1], Marco Mandolesi[1], Marcus Ahl[3], Mattias Forsell[5], Jonathan Coquet[1], Martin Corcoran[1], Joanna Rorbach[6,7], Joakim Dillner[8], Gordana Bogdanovic[2], Gerald McInerney[1], Tobias Allander[1,2], Chris Wallace[4,9*], Ben Murrell[1*], Jan Albert[1,2*], Gunilla B. Karlsson Hedestam[1#]

Affiliations:
[1]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm 171 77, Sweden
[2]Department of Clinical Microbiology, Karolinska University Hospital, Stockholm 171 76, Sweden
[3] Department of Infectious Diseases, Karolinska Universitetssjukhuset, Huddinge 141 52, Sweden
[4]Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge, CB2 0AW
[5]Department of Clinical Microbiology, Umeå Universitet, Umeå 901 85, Sweden
[6]Department of Molecular Biochemistry & Biophysics, Karolinska Institutet, Stockholm 171 77, Sweden
[7]Max Planck Institute-Biology of Ageing, Karolinska Institutet Laboratory, Stockholm 171 77, Sweden
[8]Department of Laboratory Medicine, Division of Pathology, Karolinska Institutet, Huddinge 141 52, Sweden
[9]MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge CB2 0SR, United Kingdom

°*Equal contribution
#Correspondence

Abstract

Serology is critical for understanding pathogen-specific immune responses, but is fraught with difficulty, not least because the strength of antibody (Ab) response varies greatly between individuals and mild infections generally generate lower Ab titers[1–3]. We used robust IgM, IgG and IgA Ab tests to evaluate anti-SARS-CoV-2 responses in individuals PCR+ for virus RNA (*n*=105) representing different categories of disease severity, including mild cases. All PCR+ individuals in the study became IgG-positive against pre-fusion trimers of the virus spike (S) glycoprotein, but titers varied greatly. Elevated IgA, IL-6 and neutralizing responses were present in intensive care patients. Additionally, blood donors and pregnant women (*n*=2,900) sampled throughout the first wave of the pandemic in Stockholm, Sweden, further demonstrated that anti-S IgG titers differed several orders of magnitude between individuals, with an increase of low titer values present in the population at later time points[4,5]. To improve upon current methods to identify low titers and extend the utility of individual measures[6,7], we used our PCR+ individual data to train machine learning algorithms to assign likelihood of past infection. Using these tools that assigned probability to individual responses against S and the receptor binding domain (RBD), we report SARS-CoV-2-specific IgG in 13.7% of healthy donors five months after the peak of spring COVID-19 deaths, when mortality and ICU occupancy in the country due to the virus were at low levels. These data further our understanding of antibody responses to the virus and provide solutions to problems in serology data analysis.

Significance statement:

Antibody testing provides critical clinical and epidemiological information during an emerging disease pandemic. We developed robust SARS-CoV-2 IgM, IgG and IgA antibody tests and profiled COVID-19 patients and exposed individuals throughout the outbreak in Stockholm, Sweden, where full societal lockdown was not employed. As well as elucidating several disease immunophenotypes, our data highlight the challenge of identifying low IgG titer individuals, who comprise a significant proportion of the population following mild/asymptomatic infection, especially as antibody titers wane following peak responses. To provide a solution to this, we used SARS-CoV-2 PCR+ individual data to develop machine learning approaches that assigned *likelihood of past infection* to blood donors and pregnant women, improving the accuracy and utility of individual and population-level Ab measures.

Introduction

Characterization of the humoral response to nascent SARS-CoV-2 outbreaks is central to optimizing approaches to the pandemic and further our understanding of human immunology[8,9]. Here, we followed the first wave of the pandemic in Stockholm, Sweden, characterizing Ab responses in severely ill COVID-19 patients and exposed healthy individuals.

Despite the plethora of Ab testing and phenotyping for SARS-CoV-2[10–17], consensus on several key issues remains outstanding. For instance, the majority of data are derived from commercial, mass-produced kits utilizing spike derivatives (e.g. S1 or S2 domains) or the nucleocapsid to detect pathogen-specific antibodies[10,18,19]. Many of these assays suffer from epitope loss/modification[20], cross-reactivity[10,21] and suboptimal sensitivity[22–25]. Here, we developed highly sensitive and specific ELISA protocols based on in-house native-like pre-fusion-stabilized spike (S) trimers[26] and the smaller ACE-2 receptor-binding domain (RBD), and used them in tandem to evaluate anti-viral Ab responses. As S and the RBD are required for ACE-2-mediated cell entry, we also examined virus neutralisation capacity and isotype levels alongside a descriptive set of clinical features.

Aside from the target antigen, a major consideration for Ab testing concerns setting the assay cut-off for positivity. Commonly, 3 or 6 standard deviations (SD) from the mean of negative controls is used[27–29], which is highly dependent on a representative set of negative control sera - significantly affecting seroprevalence estimates and individual clinical management[30]. It is important to note that humoral responses within the population are not so much *positive* or *negative*, but rather represent a wide spectrum[2], highlighting the need for more quantitative tools to examine low titer individuals. To obtain accurate seroprevalence estimates in key community groups, we strictly controlled our assay with a large number of historical controls (*n*=595, repeatedly analyzed alongside test samples) and used tandem anti-S and RBD responses from SARS-CoV-2 PCR+ individuals to train machine learning (ML) algorithms on ELISA measurements.

Developing these approaches, that assigned likelihood of past infection to each data point (values and code openly available), we sampled 100 blood donors and 100 pregnant women per week throughout the outbreak in Stockholm. Both groups are a good proxy for population health. Furthermore, blood donors serve as an important clinical resource (including for COVID-19 plasma therapy), while pregnant women require close clinical monitoring with respect to fetal-maternal health and are known to employ poorly characterized immunological mechanisms impacting infectious pathology[31–33]. The study was terminated when new cases, mortality and ICU occupancy were at low levels. As Sweden did not impose a strict lockdown in response to the pandemic, these data provide a contrast to comparable settings where social lockdown was not imposed.

Results

Study samples are detailed in Table 1.

Antibody test development

We developed ELISA protocols to profile IgM, IgG and IgA specific for a stabilized spike (S) glycoprotein trimer[26], the RBD, and the internal viral nucleocapsid (N). The two S antigens were produced in mammalian cells (Fig. S1A) and the trimer conformation was confirmed in each batch by cryo-EM[34]. A representative subset of study samples (*n* = 230, Fig S1) was used for assay development (Fig. S1B). No IgG reactivity was recorded amongst the negative control samples during test development, although two individuals who were PCR-positive for endemic coronavirus-positive (ECV+) displayed reproducible IgM reactivity to SARS-CoV-2 N and S, and two 2019 blood donors had low anti-S IgM reactivity (Fig. S1C). Thus, further investigation is required to establish the contribution of cross-reactive memory and germline immunoglobulin alleles to SARS-CoV-2 responses[35]. We did not observe reproducible IgG reactivity to S or RBD across all 595 historical controls in the study.

Our assay revealed a greater than 10,000-fold difference in anti-viral IgG titers between Ab-positive individuals when examining serially diluted sera. In SARS-CoV-2 PCR+ individuals, anti-viral IgG titers were comparable for S (EC_{50}=3,064; 95% CI [1,197 - 3,626]) and N (EC_{50}=2,945; 95% CI [543 - 3,936]) and lower for RBD [EC_{50}=1,751; 95% CI 966 - 1,595]. A subset (*ca.* 10%) of the SARS CoV-2-confirmed individuals in test development did not have detectable IgG responses against N (Fig. S1D), as previously reported[10]. Therefore, we did not explore responses to N further.

Elevated anti-viral Ab titers are associated with increased disease severity

We next screened all SARS-CoV-2 PCR+ individuals collected for the study (*n*=105) and detected potent IgG responses against S in all participants, and against RBD in 97% of persons (Fig. 1A and S2A). In healthy blood donors and pregnant women, titers varied greatly but were generally lower (Fig. 1A and S2A). In PCR+ patients, IgM and IgA responses against S and RBD were generally weaker and more variable between individuals than the IgG response (Fig. 1B and S2B). Therefore, we sought to investigate whether isotype titers segregated with clinical features.

To achieve this, SARS-CoV-2 PCR+ individuals (*n*=105) were grouped with regards to their disease severity: Category 1 – *non-hospitalized (mild and asymptomatic infections)*; Category 2 – *hospitalized*; Category 3 – *intensive care* (on mechanical ventilation) (Table 1). In all PCR+ individuals, anti-S and anti-RBD responses were highly correlated (Fig. S2C) and multivariate analyses revealed that increased anti-viral IgM, IgG and IgA were positively correlated with disease severity (Supp. Table 1), in line with the lower titers observed in blood donors and pregnant women, who did not have signs or symptoms of COVID-19 when they were sampled.

Severe disease was most associated with virus-specific IgA, suggestive of mucosal disease[36], as well as elevated serum IL-6 (Fig. 1C and S2D), a cytokine that feeds Ab production[37–40]. IL-6 is dysregulated in several common non-communicable diseases[41–43] and during acute respiratory distress syndrome[44], risk factors for COVID-19-associated mortality[45,46]. Interestingly, we observed a lack of association between IL-6 and IgM levels, supporting that levels of the cytokine and IgA mark a protracted, severe clinical course of COVID-19. IgA anti-RBD responses were lower in non-hospitalized and hospitalized females as compared to males, trending similarly for S (Fig. S3A) and in line with females developing less severe disease[47].

In our study, PCR+ individual anti-viral IgG levels were maintained two months post-disease onset/positive PCR test, while IgM and IgA decreased, in agreement with their circulating $t_{1/2}$ and viral clearance (Fig. S3B). In longitudinal patient samples (sequential serum sampling of 10 PCR+ individuals in the study) where we observed seroconversion, IgM, IgG and IgA peaked with similar kinetics when all three isotypes developed, although IgA was not always generated in Category 1 and 2 individuals (Fig 1D). Overall, disease severity showed the most consistent relationship with any measure and was the primary predictor of Ab levels (Fig. S3C and D).

We next characterized the virus neutralizing Ab response, a key parameter for understanding the potential for protective humoral immune responses and the selection of plasma therapy donors. Benefitting from a robust *in vitro* pseudotype virus neutralization assay[48], we measured serum inhibition of viral cell entry and detected neutralizing antibodies in the serum of all SARS-CoV-2 PCR+ individuals (*n*=48), and in all except two healthy Ab-positive donors screened (*n*=56). Neutralizing responses were not seen in samples before seroconversion (Fig. 1D) or negative controls. A large range of neutralizing ID_{50} titers was apparent, with binding and neutralizing Ab levels being highly correlated (Fig. S3D). The strongest neutralizing responses were observed in samples from patients on mechanical ventilation in intensive care (Category 3, g.mean ID_{50}=5,058; 95% CI [2,422 - 10,564]), in-keeping with their elevated Ab response (Fig 1E). Sera from healthy blood donors and pregnant women also displayed neutralizing responses, but consistent with the binding data were less potent than those observed in individuals with severe disease (ID_{50}=600; 95% CI [357 – 1,010] and ID_{50}=350; 95% CI [228 - 538], respectively, Fig. 1F). Across the two antigens and three isotypes, anti-RBD IgG levels were most strongly correlated with neutralization.

<u>Probabilistic seroprevalence estimates in blood donors and pregnant women</u>

As Stockholm is a busy urban area and Sweden did not impose strict lockdown in response to SARS-CoV-2 emergence, we sought to better understand the frequency and nature of anti-viral responses in healthy blood donors and pregnant women sampled throughout the first outbreak (March 30 - August 23rd 2020) (Fig. 2A). However, critical to accurate individual measures and seroprevalence estimates is the decision about whether a sample is defined as

positive or not. For example, current clinically approved tests use a ratio between a "representative" positive and negative serum calibrator to determine positivity, although we show here that these are highly variable.

To improve our understanding of the assay boundary, we repeatedly analyzed a large number of historical (SARS-CoV-2-negative) controls (blood donors from the spring of 2019, *n*=595) alongside test samples throughout the study. We considered the spread of negative values critical, since the use of a small and unrepresentative set of controls can lead to an incorrectly set threshold and errors in the seroprevalence measurement. This is illustrated by the random sub-sampling of non-overlapping groups of negative controls, resulting in a 40% difference in the seroprevalence estimate (Fig. S4A). Seroprevalence in the healthy cohorts according to conventional 3 and 6 SD cut-offs are shown in Fig. 2C.

The fact that many healthy donor test samples had optical densities between the 3 and 6 SD cut-offs for both or a single antigen (Fig. 2B and C), highlights the problem of assigning case to *low responder* values. Therefore, to exploit individual titers and improve our statistical estimates, we used the data from PCR+ individuals (Fig. 1) and our negative controls (Fig. S4B) to train machine learning (ML) algorithms to assign likelihood of past infection. A small cohort of seropositive individuals among Karolinska University Hospital staff (*n*=33) provided additional low titer training values four months post SARS-CoV-2+ PCR (Fig. S4C).

After comparing different methods for this purpose (Materials & Methods), we found that logistic regression (LOG) achieved the highest sensitivity, while linear discriminant analysis (LDA) showed the best specificity. LOG and LDA both model log odds of a sample being case as a linear equation with a resulting linear decision boundary, but differ in how the coefficients for the linear models are estimated from the data. When applied to the Ab response data, the output of LOG and LDA is the probability of each new sample being case. Therefore, we generated an equal-weighted ensemble learner (*ENS*) from the output of LOG and LDA that maximized sensitivity, specificity and consistency across different cross-validation strategies (Fig. 2D and S4D). While weekly rates varied (S Table 2), the *ENS* learner identified 13.7% seroprevalence in healthy blood donors and pregnant women at the last sampling week (Supp. Table 2). Importantly, *ENS* identified 155 (5.3%) blood donor and pregnant women measures to be associated with some degree of uncertainty, encouraging follow-up investigation in given cases (Fig. 2E and S4E).

Finally, to model population changes in seroprevalence over time, we developed and validated a cut-off-independent Bayesian ML framework able to share information between sampling weeks[49] (Fig. 2F and Materials & Methods). Using this model on the combined BD and PW data, we found an almost linear increase in seroprevalence since the start of the pandemic (Fig. 2E), consistent with continued virus spread in the Stockholm population during the study period. The results mirrored the results obtained using *ENS*, yielding a

seroprevalence of 13.2% (95% Bayesian CI [10.1-16.8]) at the end of the study period (Supp. Table 3). We propose that these tools and related approaches be used to facilitate future antibody measures and better characterize Ab test uncertainty at individual and population levels.

Discussion

Serology remains the *gold standard* for estimating previous exposure to pathogens and benefits from a large historical literature. Although the concept of herd immunity is based upon the study of antibodies, worryingly, there is no standardization for the many SARS-CoV-2 Ab tests currently available. Globally, hospital staff and health authorities are struggling with test choices, negatively impacting individual outcomes and efforts to contain the pandemic.

Benefitting from a robust antibody test developed alongside a diagnostic clinical laboratory responsible for monitoring sero-reactivity during the pandemic, we profiled SARS-CoV-2 Ab responses in three cohorts of clinical interest. COVID-19 patients receiving intensive care showed the highest anti-viral Ab titers, developing augmented serum IgA and IL-6 with worsening disease. Isotype-level measures may assist COVID-19 clinical management and determine, for example, whether all critically ill patients develop class-switched mucosal responses to SARS-CoV-2, potentially informing lung therapeutic delivery[50,51]. Our neutralization data showed that nearly all SARS-CoV-2 PCR+ individuals and healthy donors who seroconvert, develop neutralizing Ab capable of preventing S-mediated cell entry *in vitro*.

Outside of the severe disease setting, it is critical to accurately determine who and how many people have seroconverted. This is complicated by low titer values, which in some cases - and increasingly with time since exposure - overlap outlier values among negative control samples. Test samples with true low anti-viral titers will, therefore, fall into this range of *weak responders* as the B lymphocyte response contracts following viral clearance, highlighting the need to better understand the assay boundary in multiple dimensions. As future tests begin to survey individual Ab responses to a multitude of antigens in parallel, the ML approaches presented here will enable the identification of disease sub-types and facilitate longitudinal measures.

We applied these tools to blood donors and pregnant women, two good sentinels for population health, although they are not enriched for groups with high risk for SARS-CoV-2 infection, such as healthcare workers and public transportation employees, where seroprevalence may be higher. Blood donors are generally working age, active and mobile members of society with a good understanding of health, and pregnant women in Sweden will have been advised to take precautions against infectious diseases through their practitioners. Interestingly, in our study, both groups showed a similar seroprevalence during the time period analyzed. Tracking these cohorts over time, we modelled seroprevalence changes at the population level. We found the steep climb in Ab positivity at the start of the pandemic (as the virus emerged) to increase at a slower rate during subsequent weeks, reaching nearly 14% by five months from the peak of spring 2020 COVID-19 deaths in the country. These data indicate that serological herd immunity to the initial outbreak was not

achieved in these cohorts. We terminated the study in line with the decreasing caseload and number of fatalities in Sweden[52], despite on-going virus spread in the Stockholm population.

Given the uniqueness of the public health response to the pandemic in the country[53], these data may inform the management of this and future pandemics elsewhere. Our data also highlight high inter-individual variability in anti-viral Ab responses and offer solutions for how to handle this at individual and population levels.

Materials and methods

Human samples and ethical declaration

Samples from PCR+ individuals and admitted COVID-19 patients (*n*=105) were collected by the attending clinicians and processed through the Departments of Medicine and Clinical Microbiology at the Karolinska University Hospital. Samples were used in accordance with approval by the Swedish Ethical Review Authority (registration no. 2020-02811). All personal identifiers were pseudo-anonymized, and all clinical feature data were blinded to the researchers carrying out experiments until data generation was complete. PCR testing for SARS-CoV-2 RNA was by nasopharyngeal swab or upper respiratory tract sampling at Karolinska University Hospital. As viral RNA levels were determined using different qPCR platforms (with the same reported sensitivity and specificity) between participants, we did not analyze these alongside other features. PCR+ individuals (*n*=105) were questioned about the date of symptom onset at their initial consultation and followed-up for serology during their care, up to 2 months post-diagnosis. Serum from SARS-CoV-2 PCR+ individuals was collected 6-61 days post-test, with the median time from symptom onset to PCR being 5 days. In addition, longitudinal samples from 10 of these patients were collected to monitor seroconversion and isotype persistence.

Hospital workers at Karolinska University Hospital were invited to test for the presence of SARS-CoV-2 RNA in throat swabs in April 2020 and virus-specific IgG in serum in July 2020. We screened 33 PCR+ individuals to provide additional training data for ML approaches. All participants provided written informed consent. The study was approved by the National Ethical Review Agency of Sweden (2020-01620) and the work was performed accordingly.

Anonymized samples from blood donors (*n*=100/week) and pregnant women (*n*=100/week) were randomly selected from their respective pools by the department of Clinical Microbiology, Karolinska University Hospital. No metadata, such as age or sex information were available for these samples in this study. Pregnant women were sampled as part of routine for infectious diseases screening during the first trimester of pregnancy. Blood donors (*n*=595) collected through the same channels a year previously were randomly selected for use as negative controls. Serum samples from individuals testing PCR+ for endemic coronaviruses, 229E, HKU1, NL63, OC43 (*n*=20, ECV+) in the prior 2-6 months, were used as additional negative controls. The use of study samples was approved by the Swedish Ethical Review Authority (registration no. 2020-01807). Stockholm County death and Swedish mortality data was sourced from the ECDC and the Swedish Public Health Agency, respectively. Study samples are defined in Table 1.

Serum sample processing

Blood samples were collected by the attending clinical team and serum isolated by the department of Clinical Microbiology. Samples were anonymized, barcoded and stored at -

20°C until use. Serum samples were not heat-inactivated for ELISA protocols but were heat-inactivated at 56°C for 60 min for neutralization experiments.

SARS-CoV-2 antigen generation

The plasmid for expression of the SARS-CoV-2 prefusion-stabilized spike ectodomain with a C-terminal T4 fibritin trimerization motif was obtained from[26]. The plasmid was used to transiently transfect FreeStyle 293F cells using FreeStyle MAX reagent (Thermo Fisher Scientific). The ectodomain was purified from filtered supernatant on Streptactin XT resin (IBA Lifesciences), followed by size-exclusion chromatography on a Superdex 200 in 5 mM Tris pH 8, 200 mM NaCl.

The RBD domain (RVQ – QFG) of SARS-CoV-2 was cloned upstream of a Sortase A recognition site (LPETG) and a 6xHIS tag, and expressed in 293F cells as described above. RBD-HIS was purified from filtered supernatant on His-Pur Ni-NTA resin (Thermo Fisher Scientific), followed by size-exclusion chromatography on a Superdex 200. The nucleocapsid was purchased from Sino Biological.

Anti-SARS-CoV-2 ELISA

96-well ELISA plates (Nunc MaxiSorp) were coated with SARS-CoV-2 S trimers, RBD or nucleocapsid (100 µl of 1 ng/µl) in PBS overnight at 4°C. Plates were washed six times with PBS-Tween-20 (0.05%) and blocked using PBS-5% no-fat milk. Human serum samples were thawed at room temperature, diluted (1:100 unless otherwise indicated), and incubated in blocking buffer for 1h (with vortexing) before plating. Serum samples were incubated overnight at 4°C before washing, as before. Secondary HRP-conjugated anti-human antibodies were diluted in blocking buffer and incubated with samples for 1 hour at room temperature. Plates were washed a final time before development with TMB Stabilized Chromogen (Invitrogen). The reaction was stopped using 1M sulphuric acid and optical density (OD) values were measured at 450 nm using an Asys Expert 96 ELISA reader (Biochrom Ltd.). Secondary antibodies (all from Southern Biotech) and dilutions used: goat anti-human IgG (2014-05) at 1:10,000; goat anti-human IgM (2020-05) at 1:1000; goat anti-human IgA (2050-05) at 1:6,000. All assays of the same antigen and isotype were developed for their fixed time and samples were randomized and run together on the same day when comparing binding between PCR+ individuals. Negative control samples were run alongside test samples in all assays and raw data were log transformed for statistical analyses.

In vitro virus neutralisation assay

Pseudotyped viruses were generated by the co-transfection of HEK293T cells with plasmids encoding the SARS-CoV-2 spike protein harboring an 18 amino acid truncation of the cytoplasmic tail[26]; a plasmid encoding firefly luciferase; a lentiviral packaging plasmid (Addgene 8455) using Lipofectamine 3000 (Invitrogen). Media was changed 12-16 hours post-transfection and pseudotyped viruses harvested at 48- and 72-hours, filtered through a 0.45

µm filter and stored at -80°C until use. Pseudotyped neutralisation assays were adapted from protocols validated to characterize the neutralization of HIV, but with the use of HEK293T-*ACE2* cells. Briefly, pseudotyped viruses sufficient to generate ~100,000 RLUs were incubated with serial dilutions of heat-inactivated serum for 60 min at 37°C. Approximately 15,000 HEK293T-ACE2 cells were then added to each well and the plates incubated at 37°C for 48 hours. Luminescence was measured using Bright-Glo (Promega) according to the manufacturer's instructions on a GM-2000 luminometer (Promega) with an integration time of 0.3s. The limit of detection was at a 1:45 serum dilution.

IL-6 cytometric bead array

Serum IL-6 levels were measured in a subset of PCR+ serum samples (*n*=64) using an enhanced sensitivity cytometric bead array against human IL-6 from BD Biosciences (Cat # 561512). Protocols were carried out according to the manufacturer's recommendations and data acquired using a BD Celesta flow cytometer.

Statistical analysis of SARS-CoV-2 PCR+ data

All univariate comparisons were performed using non-parametric analyses (Kruskal-Wallis, stratified Mann-Whitney, hypergeometric exact tests and Spearman rank correlation), as indicated, while multivariate comparisons were performed using linear regression of log transformed measures and Wald tests. For multivariate tests, all biochemical measures (IL-6, PSV ID50 neut., IgG, IgA, IgM) were log transformed to improve the symmetry of the distribution. As "days since first symptom" and "days since PCR+ test" are highly correlated, we cannot include both in any single analysis. Instead, we show results for one, then the other (Supp. Table 1).

Probabilistic seroprevalence estimations

Prior to analysis, each sample OD was standardized by dividing by the mean OD of "no sample controls" on that plate or other plates run on the same day. This resulted in more similar distributions for 2019 blood donor samples with 2020 blood donors and pregnant volunteers, as well as smaller coefficients of variation amongst PCR+ COVID patients for both SPIKE and RBD.

We employed two distinct probabilistic strategies for estimating seroprevalence without thresholds, each developed independently. Our machine learning approach consisted of evaluating different algorithms suited to ELISA data, which we compared through ten-fold cross validation (CV): logistic regression (LOG), linear discriminant analysis (LDA), and support vector machines (SVM) with a linear kernel. Logistic regression and linear discriminant analysis both model log odds of a sample being case as a linear equation with a resulting linear decision boundary. The difference between the two methods is in how the coefficients for the linear models are estimated from the data. When applied to new data, the output of logistic regression and LDA is the probability of each new sample being a case. Support vector

machines is an altogether different approach. We opted for a linear kernel, once again resulting in a linear boundary. SVM constructs a boundary that maximally separates the classes (i.e. the margin between the closest member of any class and the boundary is as wide as possible), hence points lying far away from their respective class boundaries do not play an important role in shaping it. SVM thus puts more weight on points closest to the class boundary, which in our case is far from being clear. Linear SVM has one tuning parameter *C*, a cost, with larger values resulting in narrower margins. We tuned *C* on a vector of values (0.001, 0.01, 0.5, 1, 2, 5, 10) via an internal 5-fold CV with 5 repeats (with the winning parameter used for the final model for the main CV iteration). We also note that the natural output of SVM are class labels rather than class probabilities, so the latter are obtained via the method of Platt[54].

We considered three strategies for cross-validation: i) random: individuals were sampled into folds at random, ii) stratified: individuals were sampled into folds at random, subject to ensuring the balance of cases:controls remained fixed and iii) unbalanced: individuals were sampled into folds such that each fold was deliberately skewed to under or over-represent cases compared to the total sample. We sought a method that worked equally well across all cross-validation schemes, as the true proportion of cases in the test data is unknown and so a good method should not be overly sensitive to the proportion of cases in the training data. We found most methods worked well and chose to create an ensemble (*ENS*) method combining the method with the highest sensitivity (LOG) with the highest specificity (LDA), defined as an unweighted average of the probabilities generated under both.

We trained the ensemble learner on all 733 training samples and predicted the probability of anti-SARS-CoV-2 antibodies in blood donors and pregnant volunteers sampled in 2020. The *ENS* learner had average sensitivity > 99.1% and average specificity >99.8%. We inferred the proportion of the sampled population with positive antibody status each week using multiple imputation. We repeatedly (1,000 times) imputed antibody status for each individual randomly according to the ensemble prediction, and then analyzed each of the 1,000 datasets in parallel, combining inference using Rubin's rules, derived for the Wilson binomial proportion confidence interval[55].

Our Bayesian approach is explained in detail in Christian *et al*[49]. Briefly, we used a logistic regression over anti-RBD and -S training data to model the relationship between the ELISA measurements and the probability that a sample is antibody-positive. We adjusted for the training data class proportions and used these adjusted probabilities to inform the seroprevalence estimates for each time point. Given that the population seroprevalence cannot increase dramatically from one week to the next, we constructed a prior over seroprevalence trajectories using a transformed Gaussian Process, and combined this with the individual class-balance adjusted infection probabilities for each donor to infer the posterior distribution over seroprevalence trajectories.

Data and code availability statement

Data generated as part of the study, along with custom code for statistical analyses, is openly available via our GitHub repositories: https://github.com/MurrellGroup/DiscriminativeSeroprevalence/ and https://github.com/chr1swallace/seroprevalence-paper.

Author contributions

GKH and XCD designed the study, analyzed the results and wrote the manuscript with input from co-authors. JA, TA, JD, SM, GB, MA and SA provided the study serum samples and clinical information. LH, LPV, AMM, DJS, KCI, BM and GM generated SARS-CoV-2 antigens and pseudotyped viruses. XCD and MF developed the ELISA protocols and XCD generated the data. DJS and BM developed and performed the neutralization assay. MCh and BM developed the Bayesian framework. CW and NFG assisted with patient data statistical analyses and executed machine learning approaches. MA, SK, PP, MM, JC, MCo and JR carried out wet lab experiments and assisted with data analysis.

Acknowledgments

We would like to thank the study participants and attending clinical teams. Secondly, we extend our thanks to Björn Reinius, Marc Panas, Julian Stark, Remy M. Muts and Darío Solis Sayago for their input and discussion. Funding for this work was provided by a Distinguished Professor grant from the Swedish Research Council (agreement 2017-00968) and NIH (agreement 400 SUM1A44462-02). CW and NFG are funded by the Wellcome Trust (WT107881) and MRC (MC_UP_1302/5).

Conflict of interest

The study authors declare no competing interests related to the work.

References

1. Altman, D., Douek, D. & Boyton, R. What policy makers need to know about COVID-19 protective immunity. *Lancet* (2020) doi:https://doi.org/10.1016/S0140-6736(20)30985-5.
2. Jonsson, S. *et al.* Identification of sequence variants influencing immunoglobulin levels. *Nat. Genet.* (2017) doi:10.1038/ng.3897.
3. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* (2020) doi:10.1038/s41591-020-0965-6.
4. 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* (2020) doi:10.1056/NEJMc2025179.

5. Choe, P. *et al.* Waning antibody responses in asymptomatic and symptomatic SARS-CoV-2 infection. *Emerg Infect Dis* **27**, DOI: 10.3201/eid2701.203515 (2020).
6. MacCallum, R. C., Zhang, S., Preacher, K. J. & Rucker, D. D. On the practice of dichotomization of quantitative variables. *Psychol. Methods* (2002) doi:10.1037//1082-989x.7.1.19.
7. Altman, D. G. & Royston, P. The cost of dichotomising continuous variables. *British Medical Journal* (2006) doi:10.1136/bmj.332.7549.1080.
8. Davis, M. M. A Prescription for Human Immunology. *Immunity* vol. 29 835–838 (2008).
9. Winter, A. K. & Hegde, S. T. The important role of serology for COVID-19 control. *The Lancet Infectious Diseases* (2020) doi:10.1016/S1473-3099(20)30322-4.
10. Long, Q. X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* (2020) doi:10.1038/s41591-020-0897-1.
11. Zhao, J. *et al.* Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clin. Infect. Dis.* (2020) doi:10.1093/cid/ciaa344.
12. Tan, W. *et al.* Viral Kinetics and Antibody Responses in Patients with COVID-19. *J Clin* **doi: 10.11**, (2020).
13. Robbiani, D. F. *et al.* Convergent Antibody Responses to SARS-CoV-2 Infection in Convalescent Individuals. *Nature* (2020) doi:10.1101/2020.05.13.092619.
14. Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science (80-.).* (2020) doi:10.1126/science.abc5902.
15. Ju, B. *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* (2020) doi:10.1038/s41586-020-2380-z.
16. Seydoux, E. *et al.* Analysis of a SARS-CoV-2-Infected Individual Reveals Development of Potent Neutralizing Antibodies with Limited Somatic Mutation. *Immunity* (2020) doi:10.1016/j.immuni.2020.06.001.
17. Duan, K. *et al.* Effectiveness of convalescent plasma therapy in severe COVID-19 patients. *Proc. Natl. Acad. Sci. U. S. A.* (2020) doi:10.1073/pnas.2004168117.
18. Jiang, S., Hillyer, C. & Du, L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends in Immunology* (2020) doi:10.1016/j.it.2020.03.007.
19. Van Elslande, J. *et al.* Diagnostic performance of seven rapid IgG/IgM antibody tests and the Euroimmun IgA/IgG ELISA in COVID-19 patients. *Clin. Microbiol. Infect.* (2020) doi:10.1016/j.cmi.2020.05.023.
20. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* (2020) doi:10.1016/j.cell.2020.02.058.
21. Cheng, M. *et al.* Cross-reactivity of antibody against SARS-coronavirus nucleocapsid protein with IL-11. *Biochem. Biophys. Res. Commun.* (2005) doi:10.1016/j.bbrc.2005.10.088.
22. Özçürümez, M. K. *et al.* SARS-CoV-2 antibody testing—questions to be asked. *J. Allergy Clin. Immunol.* (2020) doi:10.1016/j.jaci.2020.05.020.
23. GeurtsvanKessel, C. H. *et al.* Towards the next phase: evaluation of serological assays for diagnostics and exposure assessment. *Nat Comms* **doi: 10.10**, (2020).
24. Deeks, J. J. *et al.* Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst. Rev.* (2020) doi:10.1002/14651858.cd013652.
25. Traugott, M. *et al.* Performance of Severe Acute Respiratory Syndrome Coronavirus 2 Antibody Assays in Different Stages of Infection: Comparison of Commercial Enzyme-Linked Immunosorbent Assays and Rapid Tests. *J. Infect. Dis.* (2020)

doi:10.1093/infdis/jiaa305.

26. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (80-.).* (2020) doi:10.1126/science.aax0902.
27. Saraswati, K., Phanichkrivalkosil, M., Day, N. P. J. & Blacksell, S. D. The validity of diagnostic cut-offs for commercial and in-house scrub typhus IgM and IgG ELISAs: A review of the evidence. *PLoS Negl. Trop. Dis.* (2019) doi:10.1371/journal.pntd.0007158.
28. Xu, H., Lohr, J. & Greiner, M. The selection of ELISA cut-off points for testing antibody to Newcastle disease by two-graph receiver operating characteristic (TG-ROC) analysis. *J. Immunol. Methods* (1997) doi:10.1016/S0022-1759(97)00128-2.
29. Greiner, M., Franke, C. R., Böhning, D. & Schlattmann, P. Construction of an intrinsic cut-off value for the sero-epidemiological study of Trypanosoma evansi infections in a canine population in Brazil: a new approach towards an unbiased estimation of prevalence. *Acta Trop.* (1994) doi:10.1016/0001-706X(94)90044-2.
30. Burgess, S., Ponsford, M. J. & Gill, D. Are we underestimating seroprevalence of SARS-CoV-2? *The BMJ* (2020) doi:10.1136/bmj.m3364.
31. Aghaeepour, N. *et al.* An immune clock of human pregnancy. *Sci. Immunol.* (2017) doi:10.1126/sciimmunol.aan2946.
32. Mor, G. & Cardenas, I. The Immune System in Pregnancy: A Unique Complexity. *American Journal of Reproductive Immunology* (2010) doi:10.1111/j.1600-0897.2010.00836.x.
33. Kourtis, A. P., Read, J. S. & Jamieson, D. J. Pregnancy and infection. *New England Journal of Medicine* (2014) doi:10.1056/NEJMra1213566.
34. Hanke, L. *et al.* An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *bioRxiv* (2020) doi:10.1101/2020.06.02.130161.
35. Herzenberg, L. A. & Herzenberg, L. A. Toward a layered immune system. *Cell* (1989) doi:10.1016/0092-8674(89)90748-4.
36. Cervia, C. *et al.* Systemic and mucosal antibody secretion specific to SARS-CoV-2 during mild versus severe COVID-19. *bioRxiv* (2020) doi:10.1101/2020.05.21.108308.
37. Eto, D. *et al.* IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce optimal follicular helper CD4 T cell (Tfh) differentiation. *PLoS One* **6**, e17739 (2011).
38. Dienz, O. *et al.* The induction of antibody production by IL-6 is indirectly mediated by IL-21 produced by CD4 + T cells. *J. Exp. Med.* (2009) doi:10.1084/jem.20081571.
39. Maeda, K., Mehta, H., Drevets, D. A. & Coggeshall, K. M. IL-6 increases B-cell IgG production in a feed-forward proinflammatory mechanism to skew hematopoiesis and elevate myeloid production. *Blood* (2010) doi:10.1182/blood-2009-07-230631.
40. Beagley, K. W. *et al.* Interleukins and IgA synthesis. Human and murine interleukin 6 induce high rate IgA secretion in IgA-committed B cells. *J. Exp. Med.* (1989) doi:10.1084/jem.169.6.2133.
41. Mauer, J. *et al.* Signaling by IL-6 promotes alternative activation of macrophages to limit endotoxemia and obesity-associated resistance to insulin. *Nat. Immunol.* (2014) doi:10.1038/ni.2865.
42. Kristiansen, O. P. & Mandrup-Poulsen, T. Interleukin-6 and diabetes: The good, the bad, or the indifferent? *Diabetes* (2005) doi:10.2337/diabetes.54.suppl_2.S114.
43. Ferreira, R. C. *et al.* Functional IL6R 358Ala Allele Impairs Classical IL-6 Receptor Signaling and Influences Risk of Diverse Inflammatory Diseases. *PLoS Genet.* **9**, (2013).

580 44. Park, W. Y. *et al.* Cytokine balance in the lungs of patients with acute respiratory
581 distress syndrome. *Am. J. Respir. Crit. Care Med.* (2001)
582 doi:10.1164/ajrccm.164.10.2104013.
583 45. Magro, G. SARS-CoV-2 and COVID-19: Is interleukin-6 (IL-6) the 'culprit lesion' of ARDS
584 onset? What is there besides Tocilizumab? SGP130Fc. *Cytokine: X* (2020)
585 doi:10.1016/j.cytox.2020.100029.
586 46. Chua, R. L. *et al.* COVID-19 severity correlates with airway epithelium–immune cell
587 interactions identified by single-cell analysis. *Nat. Biotechnol.* (2020)
588 doi:10.1038/s41587-020-0602-4.
589 47. *et al.* Viral epitope profiling of COVID-19 patients reveals cross-reactivity and
590 correlates of severity. *Science (80-.).* (2020) doi:10.1126/science.abd4250.
591 48. Bartosch, B. *et al.* In vitro assay for neutralizing antibody to hepatitis C virus: Evidence
592 for broadly conserved neutralization epitopes. *Proc. Natl. Acad. Sci. U. S. A.* (2003)
593 doi:10.1073/pnas.2335981100.
594 49. Christian, M. & Murrell, B. Discriminative Bayesian Serology: Counting Without
595 Cutoffs. *BioRxiv* (2020).
596 50. Larios Mora, A. *et al.* Delivery of ALX-0171 by inhalation greatly reduces respiratory
597 syncytial virus disease in newborn lambs. *MAbs* (2018)
598 doi:10.1080/19420862.2018.1470727.
599 51. Sheridan, C. Ablynx's nanobody fragments go places antibodies cannot. *Nat.*
600 *Biotechnol.* (2017) doi:10.1038/nbt1217-1115.
601 52. Intensivvårdsregistret, S. Antal som intensivvårdas med Covid-19 per dag. *Svenska*
602 *Intensivvårdsregistrets utdataportal*
603 https://portal.icuregswe.org/siri/report/corona.co (2020).
604 53. Paterlini, M. 'Closing borders is ridiculous': the epidemiologist behind Sweden's
605 controversial coronavirus strategy. *Nature* (2020) doi:10.1038/d41586-020-01098-x.
606 54. Platt, J. & others. Probabilistic outputs for support vector machines and comparisons
607 to regularized likelihood methods. *Adv. large margin Classif.* (1999).
608 55. Lott, A. & Reiter, J. P. Wilson Confidence Intervals for Binomial Proportions With
609 Multiple Imputation for Missing Data. *Am. Stat.* (2020)
610 doi:10.1080/00031305.2018.1473796.
611
612

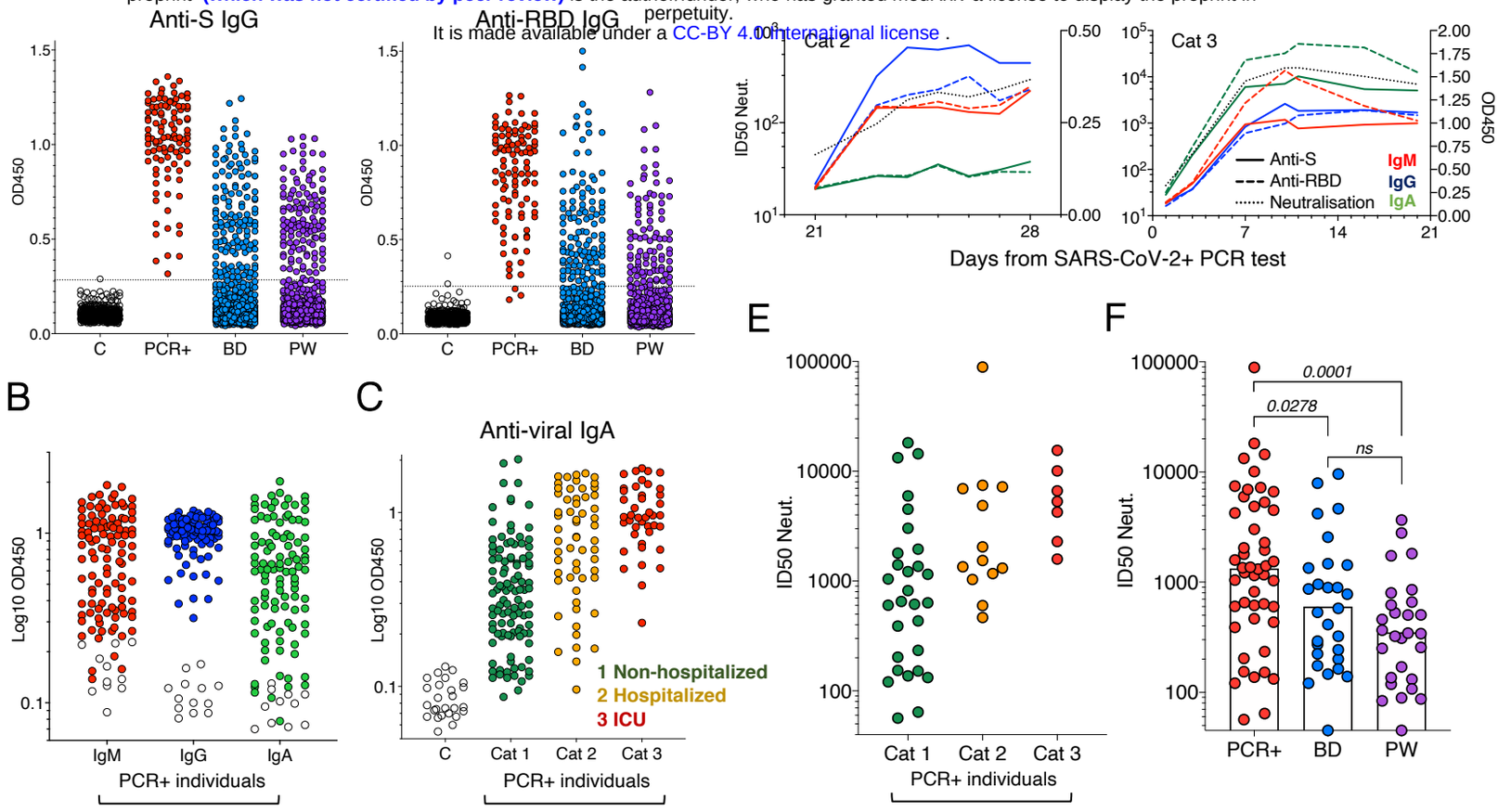

Figure 1: Anti-SARS-CoV-2 Ab phenotypes in COVID-19 patients, PCR+ individuals, blood donors and pregnant women (**A**) Raw optical density (OD) anti-S and -RBD IgG responses are shown in SARS-CoV-2 PCR+ individuals (*n*=105), blood donors (BD, *n*=1,500) and pregnant women (PW, *n*=1,400). Controls, C, represent *n*=595 blood donors from spring 2019. Conventional 6 SD cut-offs shown by dotted lines. (**B**) IgM, IgG and IgA responses against S and RBD in PCR+ individuals (*n*=105), with a limited number of controls for each assay represented by open circles. (**C**) Anti-viral Ab levels are associated with disease severity, most pronounced for IgA. COVID-19 patients in the ICU category were mechanically ventilated. Anti-S and RBD responses are graphed together. (**D**) Two discordant longitudinal profiles of seroconversion and neutralisation capacity are shown. (**E**) *In vitro* pseudotyped virus neutralization ID_{50} titers are associated with disease severity, with the highest titers observed in Cat 3 (ICU) patients. Forty-eight SARS-CoV-2 PCR+ individuals analyzed in duplicate. (**F**) Comparison of neutralization ID_{50} titers between PCR+ individuals (*n*=48), BD (*n*=28) and PW (*n*=28), all analyzed in duplicate. Bars represent the geometric mean and *P* values are from a two-tailed Mann-Whitney test.

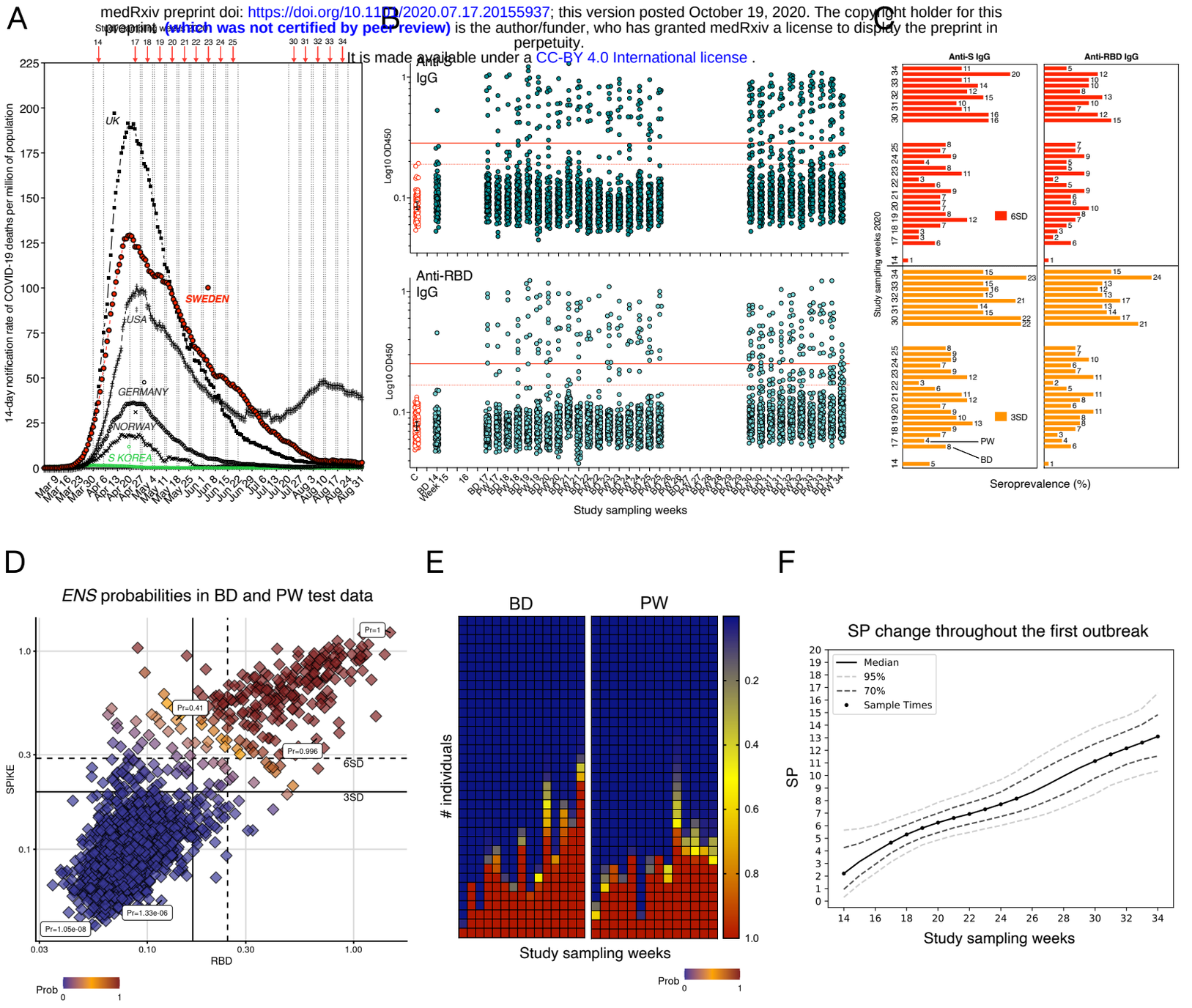

Figure 2: Probability-based seroprevalence estimates in Stockholm during the initial outbreak
(**A**) Study sample collection intervals are shown alongside the 14-day COVID-19 death rate per million inhabitants in Sweden and relevant countries for comparison. (**B**) Log-transformed un-normalized OD measurements from all BD and PW in the study. Conventional 3 (dotted red line) and 6 SD (solid red line) cut-offs are shown; calculated from *n*=595 historical controls; 100 random negative controls (C, with 95% CI of the median) are shown for each assay. (**C**) The percentage anti-S and -RBD IgG positive per sampling week in BD and PW show according to 3 or 6 SD cut-offs. (**D**) S and RBD responses from PCR+ individuals were used to train different machine learning algorithms to assign likelihood of past infection. We created an ensemble learner (*ENS*) from the output of logistic regression and linear discriminant analysis, providing a highly sensitive, specific and consistent multi-dimensional solution to the problem of weak reactors, and assigning each data point a probability of being positive. Conventional 3 and 6 SD cut-offs are shown for each antigen, with probabilities assigned to selected points. (**E**) Heatmap of assigned *ENS* probabilities for the top 35 BD and PW values per week, with each square representing an individual. (**F**) Seroprevalence (SP) estimates in Stockholm modelled over time in BD and PW using a cut-off-independent Bayesian framework.

Table 1 – Study samples

SARS-CoV-2 PCR+ individuals[§]	*n*=105
Females	*n*=44 (41.9%)
Males	*n*=61 (58.1%)
Median age (years)	53.0 (49-61)
Females	51.5 (48-56.2)
Males	55.0 (49-63)
Non-hospitalized	*n*=53
Females, males	28, 25
Hospitalized patients	*n*=31
Females, males	12, 17
Intensive care (ICU) patients	*n*=21
Females, males	3, 17
SARS-CoV2+ PCR	*n*=105
Sample collection dates	March-August 2020
SARS-CoV-2 PCR+ KI hospital staff	*n*=33
Sample collection dates	July 2020
Blood donors	*n*=1,500
Sample collection dates	Weeks 14-34 (March-August) 2020
Pregnant women	*n*=1,400
Sample collection dates	Weeks 17-34 (April-August) 2020
Historical blood donors	*n*=595
Sample collection dates	April-June 2019
ECV+ donors	*n*=20
Sample collection dates	July-December 2019

[§] *Under the care of Karolinska University Hospital*
No additional metadata available for any samples

Table 1

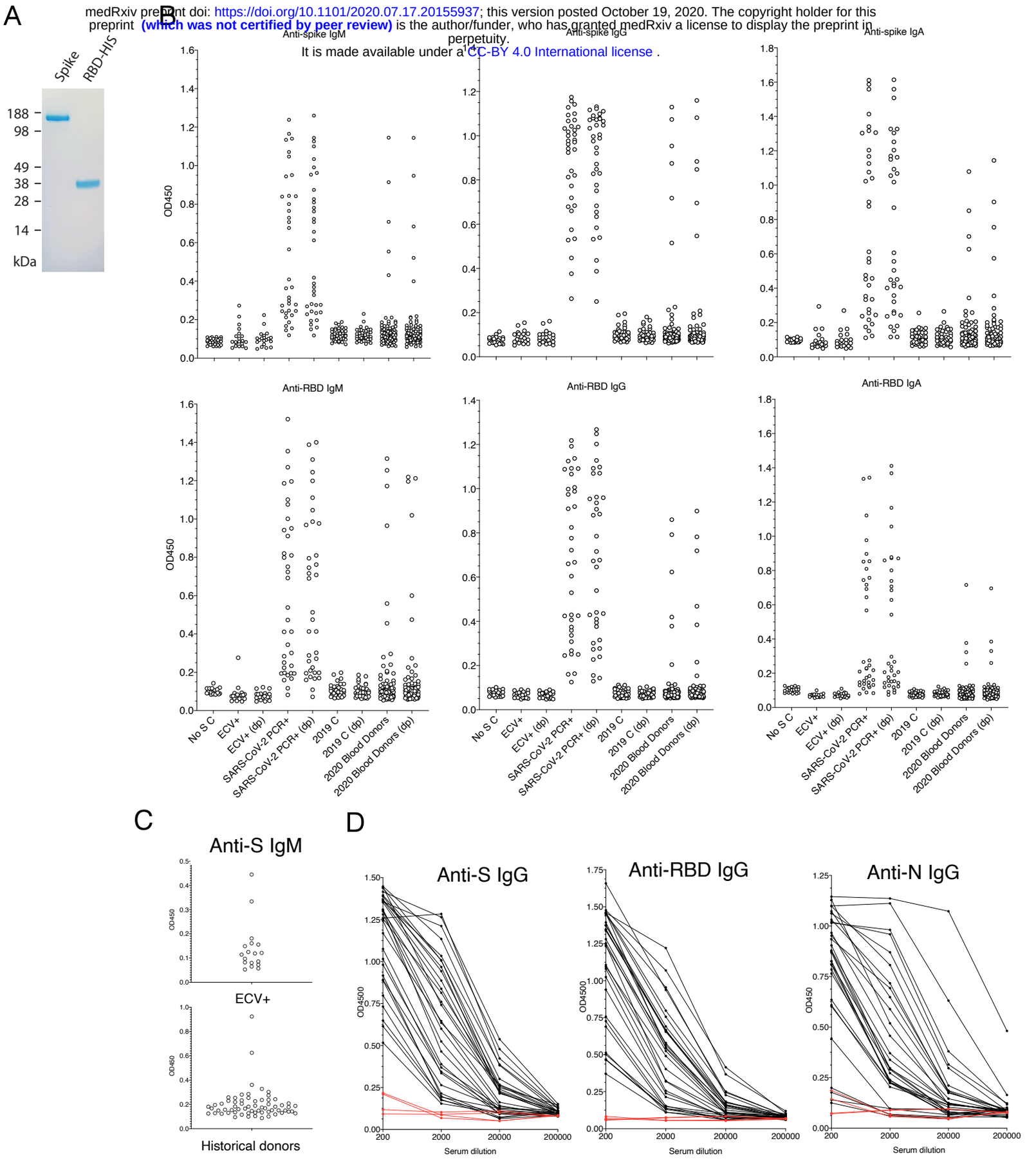

Figure S1: Anti-SARS-CoV-2 ELISA protocol development

(**A**) Trimeric S and RBD were expressed in 293F cells and purified as described. (**B**) A random subset of PCR+ individuals, negative controls and BD were used to validate the assays for the three isotypes, these were individuals with confirmed SARS CoV-2 infection (*n*=36, from a range of disease severity categories); blood donors from the spring of 2020 (*n*=100); two sets of negative controls, blood donors from the spring of 2019 (*n*=75) and individuals PCR+ for endemic coronaviruses (ECV+) (*n*=20). (**C**) Two ECV+ donors, K2 and K4, showed reproducible IgM binding to S. Testing of another subset of historical controls (*n*=75) for a similar observation, two additional individuals were found to show IgM binding to S. (**D**) Serial dilutions of PCR+ participant serum are shown (in a representative sample, *n*=40) of titrated individuals for anti-S and anti-RBD IgG.

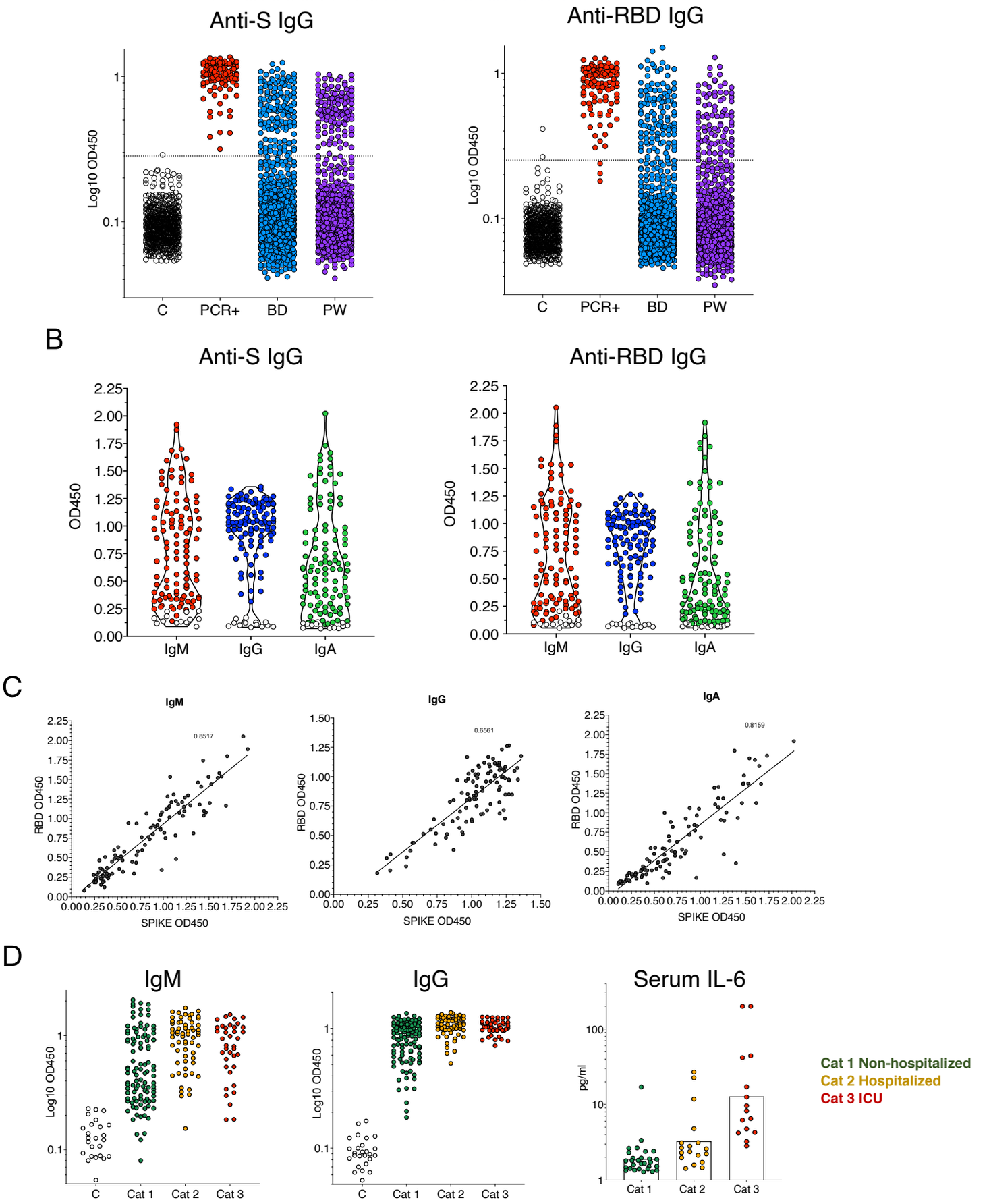

Figure S2: Antibody phenotypes in PCR+ individuals and healthy participants

(**A**) Log10 OD450 anti-S and -RBD IgG responses in SARS-CoV-2 PCR+ individuals and healthy donors, with a 6 SD cut-off shown calculated from all 595 negative control values. (**B**) Raw OD450 isotypic responses in PCR+ individuals. A limited number of negative controls are depicted by clear circles. (**C**) Anti-S vs -RBD responses in PCR+ individuals are highly correlated. (**D**) Elevated anti-viral Ab and serum IL-6 are associated with disease severity in PCR+ individuals, IgM and IgG.

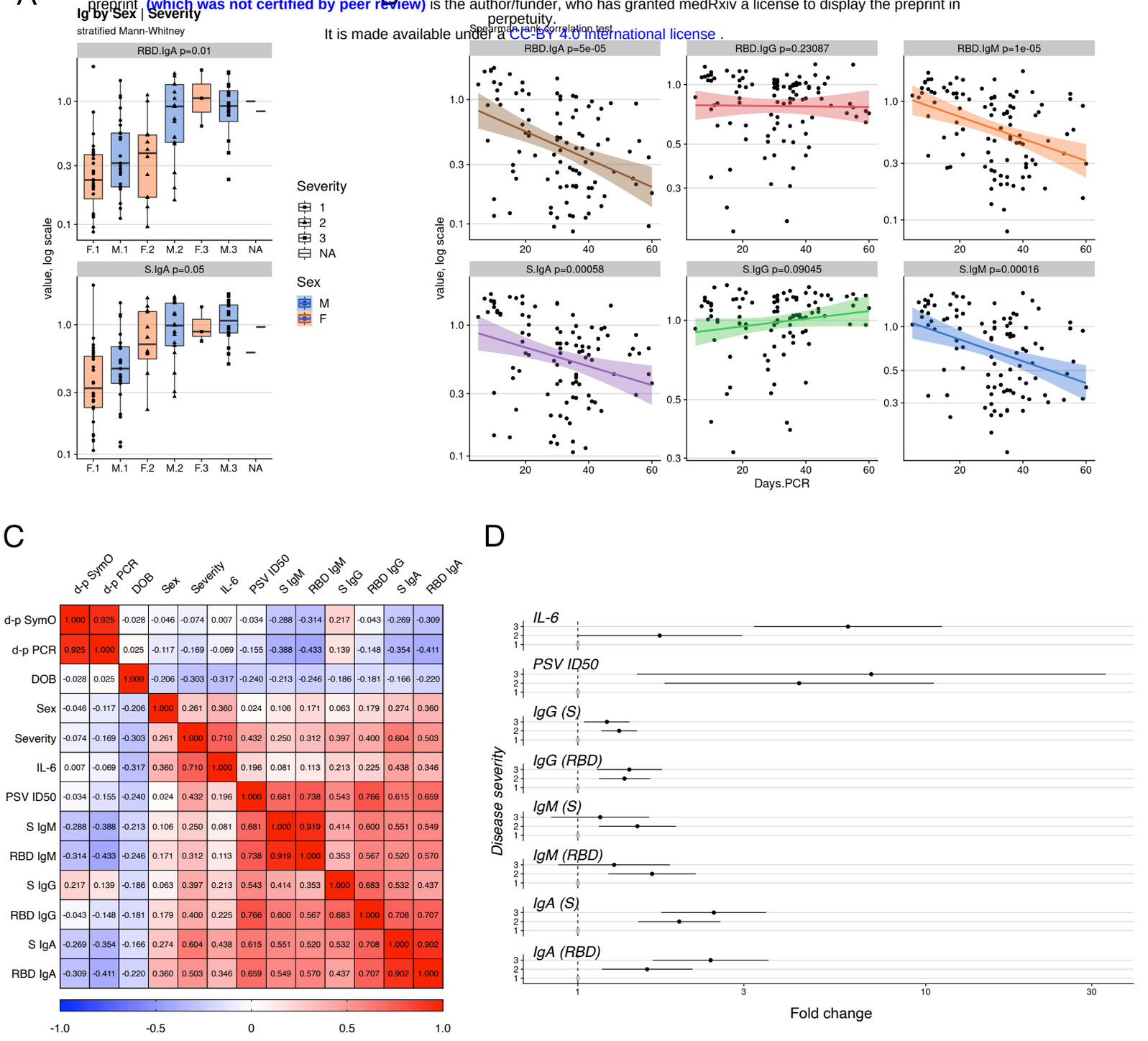

Figure S3: Antibody phenotypes in PCR+ individuals

(**A**) Antibody responses according to sex for anti-S and –RBD IgA. (**B**) IgM and IgA titers declined with time from first symptom/SARS-CoV-2+ PCR. IgG levels were maintained during this time. *P* values from a Spearman rank correlation test. (**C**) Spearman's rank correlation of PCR+ dataset features and antibody levels. DOB - *date of birth*; d-p SymO - *days post-symptom onset*; d-p PCR – *days post SARS-CoV-2+ PCR*; PSV ID50 – *neutralizing titer* (**D**) Adjusted fold-change for dataset features in PCR+ individuals compared to category 1. The effects of age (DOB), sex, days from PCR test were considered.

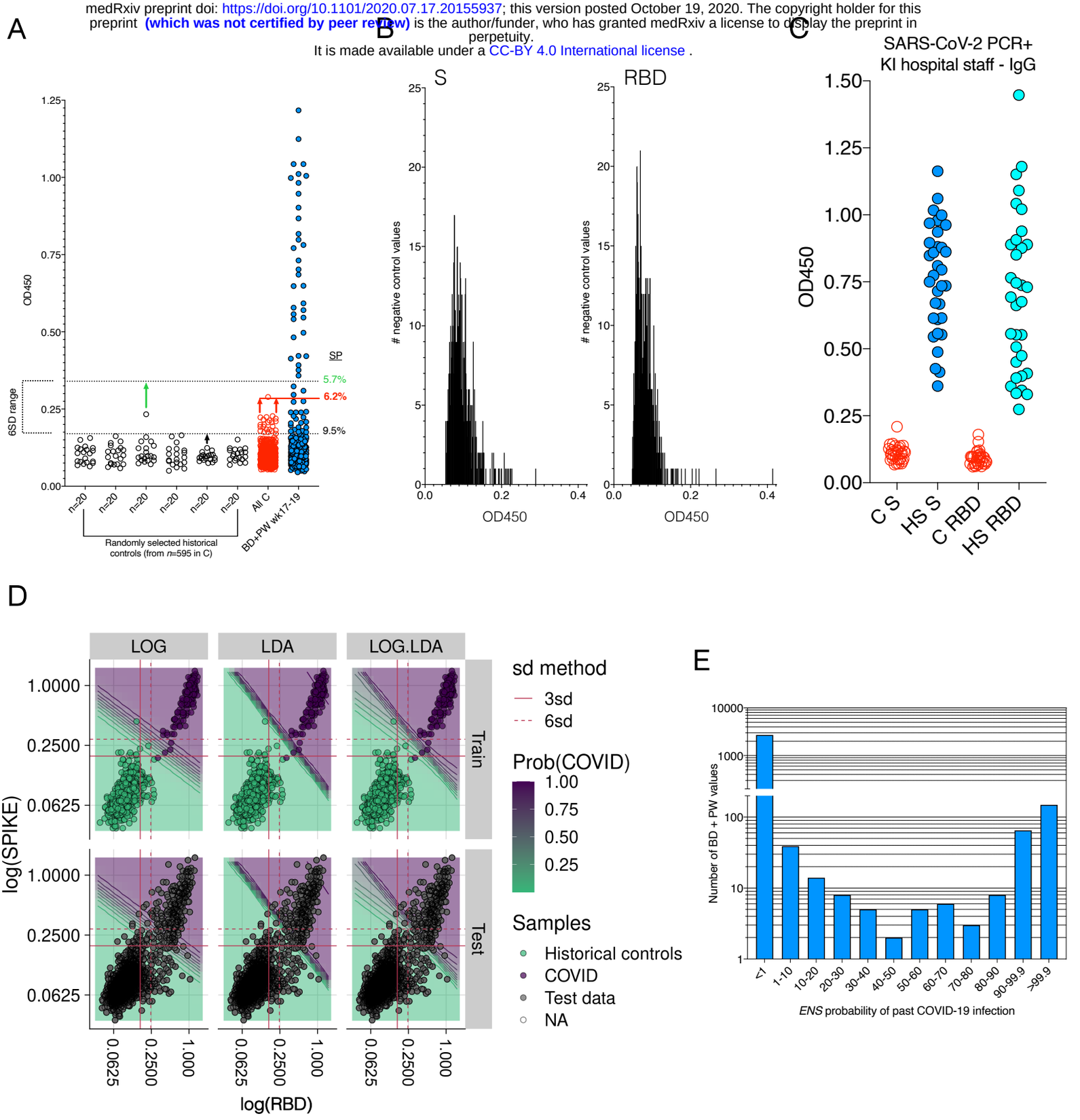

Figure S4: Considerations and methods for SP estimates in BD and PW

(**A**) Random sub-sampling of non-overlapping negative controls illustrates how the range of negative control values can influence the conventional test cut-off. Depending on the control values used to set the test threshold for positivity, seroprevalence (SP) estimates vary by 40%. 600 BD and PW values are used as an example. Anti-S IgG values are shown. (**B**) The distribution of negative control serum values for anti-S and –RBD IgG. (**C**) As ML methods improve with additional data, we analyzed a small cohort (*n*=33) of PCR+ Karolinska University hospital staff (HS) for S and RBD IgG responses. *n*=32 historical controls, C, were analyzed alongside. (**D**) Comparison of logistic regression, linear discriminant analysis and the *ENS* learner, showing the training data set and BD and PW test samples. *ENS* was trained using 595 negative control values and 138 PCR+ individuals. (**E**) *ENS* identified several study BD and PW to have uncertain measurements when S and RBD responses were considered, facilitating individual re-testing.

Supplementary Table 1 COVID-19 patient multivariate analysis

Days from first symptom

IL.6

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.0808524	0.7558497	1.4299833	0.1589412
Days.FFS	-0.0138028	0.0092132	-1.4981536	0.1403806
Severity2	0.6751029	0.2725780	2.4767330	0.0166860
Severity3	2.0818073	0.3139723	6.6305441	0.0000000
SexF	-0.0048214	0.2490798	-0.0193569	0.9846335
Years	0.0006194	0.0120141	0.0515519	0.9590912

ness with 2 outliers removed

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5938320	0.6085196	0.9758633	0.3340249
Days.FFS	-0.0048689	0.0073761	-0.6600921	0.5123499
Severity2	0.6314650	0.2134847	2.9578929	0.0047955
Severity3	1.5427576	0.2627737	5.8710497	0.0000004
SexF	-0.1881632	0.2020666	-0.9311939	0.3564134
Years	0.0047880	0.0095726	0.5001784	0.6192350

PSV.ID50

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.5012562	1.8338893	2.4544864	0.0193831
Days.FFS	-0.0150209	0.0222419	-0.6753408	0.5040267
Severity2	1.4888195	0.5062151	2.9410804	0.0058489
Severity3	1.4175205	1.0591033	1.3384157	0.1896430
SexF	-0.1325308	0.4944022	-0.2680627	0.7902707
Years	0.0484696	0.0266019	1.8220360	0.0772562

tness with 1 outlier removed

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.5186312	1.7467553	2.5868713	0.0142822

Days from SARS-CoV-2+ PCR test

IL.6

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.3337536	0.7900611	1.6881651	0.0972562
Days.PCR	-0.0179820	0.0091674	-1.9615105	0.0550817
Severity2	0.5419503	0.2784989	1.9459695	0.0569693
Severity3	1.7879936	0.3174311	5.6326988	0.0000007
SexF	-0.0418102	0.2496296	-0.1674892	0.8676226
Years	-0.0002995	0.0127464	-0.0234942	0.9813443

checking robustness with 2 outliers removed

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.7050756	0.6754753	1.0438215	0.3014906
Days.PCR	-0.0046541	0.0079825	-0.5830368	0.5624379
Severity2	0.5212971	0.2297047	2.2694227	0.0275018
Severity3	1.3678120	0.2739547	4.9928394	0.0000073
SexF	-0.1798193	0.2123891	-0.8466500	0.4011457
Years	0.0039051	0.0107259	0.3640784	0.7173046

PSV.ID50

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.2752610	1.6958032	3.1107742	0.0035845
Days.PCR	-0.0282459	0.0190912	-1.4795302	0.1474633
Severity2	1.4647193	0.4544978	3.2227205	0.0026504
Severity3	1.9419796	0.7912711	2.4542533	0.0189450
SexF	0.1871493	0.4244939	0.4408763	0.6618680
Years	0.0378735	0.0253337	1.4949856	0.1433991

checking robustness with 1 outlier removed

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.3454881	1.5958448	3.3496292	0.0019089

Days.FFS	-0.0063506	0.0215775	-0.2943131	0.7703621
Severity2	1.2195131	0.4986724	2.4455196	0.0199661
Severity3	1.4434429	1.0088452	1.4307873	0.1618932
SexF	-0.0004480	0.4750254	-0.0009431	0.9992532
Years	0.0403811	0.0256244	1.5758858	0.1245916

Days.PCR	-0.0197376	0.0183073	-1.0781255	0.2881490
Severity2	1.2106216	0.4404721	2.7484639	0.0092996
Severity3	1.9906995	0.7447805	2.6728673	0.0112300
SexF	0.3062324	0.4024577	0.7609059	0.4516688
Years	0.0297967	0.0240715	1.2378418	0.2237872

S.IgG

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4338183	0.1572226	-2.7592622	0.0070039
Days.FFS	0.0056162	0.0018643	3.0124437	0.0033555
Severity2	0.2636915	0.0582979	4.5231746	0.0000183
Severity3	0.1553421	0.0760363	2.0429981	0.0439450
SexF	-0.1061612	0.0537241	-1.9760431	0.0511786
Years	0.0025710	0.0026281	0.9782727	0.3305330

S.IgG

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.3943762	0.1612688	-2.4454589	0.0163490
Days.PCR	0.0055996	0.0019357	2.8928024	0.0047551
Severity2	0.2736918	0.0600135	4.5605006	0.0000155
Severity3	0.1924181	0.0764924	2.5155192	0.0136010
SexF	-0.0761660	0.0536223	-1.4204181	0.1588299
Years	0.0021211	0.0027262	0.7780124	0.4385365

RBD.IgG

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.5916830	0.2306770	-2.5649851	0.0119546
Days.FFS	0.0034375	0.0027353	1.2566928	0.2120811
Severity2	0.3166891	0.0855347	3.7024646	0.0003657
Severity3	0.3183027	0.1115605	2.8531842	0.0053580
SexF	-0.1630434	0.0788241	-2.0684471	0.0414354
Years	0.0021569	0.0038560	0.5593721	0.5772819

RBD.IgG

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4864143	0.2323367	-2.0935746	0.0390193
Days.PCR	0.0026259	0.0027887	0.9416332	0.3488201
Severity2	0.3090881	0.0864603	3.5749131	0.0005579
Severity3	0.3412088	0.1102011	3.0962391	0.0025904
SexF	-0.1373444	0.0772525	-1.7778632	0.0786949
Years	0.0009867	0.0039277	0.2512189	0.8021991

S.IgM

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.8300628	0.3562796	-2.3298074	0.0220285
Days.FFS	-0.0117670	0.0042247	-2.7852943	0.0065067
Severity2	0.4216830	0.1321079	3.1919588	0.0019404
Severity3	0.1256475	0.1723047	0.7292168	0.4677418
SexF	-0.1603918	0.1217434	-1.3174576	0.1909928
Years	0.0144308	0.0059556	2.4230696	0.0173698

S.IgM

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.7023442	0.3495024	-2.0095550	0.0473764
Days.PCR	-0.0140527	0.0041950	-3.3498589	0.0011693
Severity2	0.3945068	0.1300616	3.0332313	0.0031359
Severity3	0.1480218	0.1657746	0.8929099	0.3742101
SexF	-0.0547174	0.1162104	-0.4708481	0.6388518
Years	0.0113959	0.0059083	1.9287769	0.0568080

RBD.IgM

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.0387929	0.4130592	-2.5148767	0.0136629
Days.FFS	-0.0138623	0.0048980	-2.8302083	0.0057240
Severity2	0.5057259	0.1531617	3.3019087	0.0013732
Severity3	0.1944795	0.1997646	0.9735434	0.3328641
SexF	-0.2399048	0.1411454	-1.6996999	0.0926033
Years	0.0171226	0.0069047	2.4798392	0.0149844

RBD.IgM

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.8655000	0.3973594	-2.178129	0.0319255
Days.PCR	-0.0172033	0.0047694	-3.606981	0.0005008
Severity2	0.4913299	0.1478708	3.322698	0.0012758
Severity3	0.2406897	0.1884739	1.277045	0.2047656
SexF	-0.1394645	0.1321230	-1.055566	0.2939009
Years	0.0133309	0.0067174	1.984546	0.0501421

S.IgA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2343625	0.3626807	-0.6461952	0.5197793
Days.FFS	-0.0097902	0.0043006	-2.2764744	0.0251627
Severity2	0.7529383	0.1344814	5.5988283	0.0000002
Severity3	1.0042292	0.1754004	5.7253524	0.0000001
SexF	-0.2661040	0.1239307	-2.1472001	0.0344358
Years	-0.0057710	0.0060626	-0.9519015	0.3436685

S.IgA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.3012762	0.3728826	-0.8079654	0.4211724
Days.PCR	-0.0098957	0.0044757	-2.2110117	0.0294865
Severity2	0.6714710	0.1387621	4.8390072	0.0000052
Severity3	0.9000684	0.1768642	5.0890358	0.0000019
SexF	-0.2758111	0.1239844	-2.2245634	0.0285297
Years	-0.0041292	0.0063036	-0.6550579	0.5140475

RBD.IgA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2710503	0.4094717	-0.6619513	0.5096753
Days.FFS	-0.0180531	0.0048554	-3.7181129	0.0003466
Severity2	0.5614106	0.1518315	3.6975911	0.0003718
Severity3	0.9648132	0.1980296	4.8720661	0.0000046
SexF	-0.4683261	0.1399195	-3.3471098	0.0011886
Years	-0.0023814	0.0068447	-0.3479188	0.7287046

RBD.IgA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2267285	0.4122576	-0.5499679	0.5836598
Days.PCR	-0.0196414	0.0049483	-3.9693438	0.0001418
Severity2	0.4589973	0.1534149	2.9918687	0.0035499
Severity3	0.8788819	0.1955404	4.4946297	0.0000200
SexF	-0.4433399	0.1370767	-3.2342474	0.0016890
Years	-0.0029664	0.0069692	-0.4256455	0.6713496

Supplementary Table 2 *ENS* learner estimates

Type	Week	N	Estimate	lower.ci	upper.ci
Blood donors	Wk14	100	0.0105199999999998	0.00186267413688987	0.0571124452130694
Blood donors	Wk17	100	0.0599300000000007	0.0277067084436435	0.124818422697639
Blood donors	Wk18	100	0.0302500000000004	0.0103369986968445	0.0852198537817097
Blood donors	Wk19	100	0.0819899999999993	0.0420758061055538	0.153692847546347
Blood donors	Wk20	100	0.0888300000000009	0.046518592804669	0.163045245940735
Blood donors	Wk21	100	0.0606200000000007	0.0279295559709151	0.126590435953148
Blood donors	Wk22	100	0.0518599999999995	0.0223967600811795	0.115503146139443
Blood donors	Wk23	100	0.103139999999999	0.0568755500979782	0.179860591957947
Blood donors	Wk24	100	0.0491799999999993	0.0208070659558649	0.111824427597925
Blood donors	Wk25	100	0.0645900000000004	0.0299472449446712	0.133780545972474
Blood donors	Wk30	100	0.14169	0.0843021051364242	0.22839998832092
Blood donors	Wk31	100	0.088070000000001	0.0460906749148218	0.161798798516266
Blood donors	Wk32	100	0.13474	0.07991254449231	0.218260850763384
Blood donors	Wk33	100	0.122090000000001	0.0711556354845674	0.201571777003771
Blood donors	Wk34	100	0.181310000000001	0.116915378458736	0.270315226797433
Pregnant volunteers	Wk17	100	0.02785	0.00854309329510716	0.0869623409679927
Pregnant volunteers	Wk18	100	0.0593100000000001	0.0263543448650153	0.128056041293278
Pregnant volunteers	Wk19	100	0.0779199999999988	0.0392832595316614	0.148676738526427
Pregnant volunteers	Wk20	100	0.0613500000000007	0.028401055595685	0.12750757406586
Pregnant volunteers	Wk21	100	0.090440000000002	0.0482644667169825	0.163152425697498
Pregnant volunteers	Wk22	100	0.0205599999999997	0.00565556676979313	0.0719026340509295
Pregnant volunteers	Wk23	100	0.0725799999999993	0.0355650368403294	0.142429663079253
Pregnant volunteers	Wk24	100	0.0902100000000021	0.0481585199364663	0.162703789795822
Pregnant volunteers	Wk25	100	0.0705999999999996	0.0338472442954616	0.141419016311214
Pregnant volunteers	Wk30	100	0.13676	0.0802505417653763	0.22339634766703
Pregnant volunteers	Wk31	100	0.10084	0.0549469530519067	0.17785049838699
Pregnant volunteers	Wk32	100	0.1025	0.055496584850798	0.181656768852696
Pregnant volunteers	Wk33	100	0.0988799999999996	0.0537530085744113	0.174890088243812
Pregnant volunteers	Wk34	100	0.0921200000000004	0.0482559484435156	0.168784935290148

Combined	Wk14	100	0.0105199999999998	0.00186267413688987	0.0571124452130694
Combined	Wk17	200	0.0438900000000003	0.0227358905689	0.0830537591523687
Combined	Wk18	200	0.0447800000000002	0.0231970950185039	0.084702634433189
Combined	Wk19	200	0.0799549999999993	0.0494561100547252	0.126754068380355
Combined	Wk20	200	0.0750900000000004	0.0456186256545027	0.121183433120669
Combined	Wk21	200	0.075530000000001	0.0462139889421297	0.121081899484158
Combined	Wk22	200	0.0362099999999997	0.01767721923418	0.0727338989258544
Combined	Wk23	200	0.0878600000000001	0.0555074115333125	0.136347182613079
Combined	Wk24	200	0.0696949999999998	0.0417896481831635	0.11401710633583
Combined	Wk25	200	0.067595	0.0396097669899812	0.113025523482377
Combined	Wk30	200	0.139225	0.0959384551539447	0.197769606826778
Combined	Wk31	200	0.0944549999999999	0.0605819855281929	0.144356990235433
Combined	Wk32	200	0.11862	0.07965512726089	0.173061192324389
Combined	Wk33	200	0.110484999999999	0.0737258058315149	0.162359527188935
Combined	Wk34	200	0.136715	0.0949178112379659	0.192992997104021

Supplementary Table 3

Bayesian framework esimates

Weeks	low_95 CI	low_70 CI	median	high_70 CI	high_95 CI
14	0.003601532073810209	0.010015687500615324	0.02436939594002145	0.043456945152414495	0.05616568875771716
15	0.013309857927117485	0.020660967591371423	0.032365176467805026	0.04686039254667125	0.057377820643647284
16	0.02043808572781734	0.029802245381973035	0.040635466265145775	0.05175183854151783	0.060108358798232966
17	0.028769918494488296	0.03854036071416008	0.0474550179123166	0.056517830868095725	0.06401279669503077
18	0.037835374202155374	0.04509305165735621	0.052891120720793466	0.061614679337302775	0.06926247627139652
19	0.045462175043307455	0.050609840723507125	0.05808259302264415	0.0667054486902604	0.07438983764874765
20	0.04989302616185007	0.054706576021972896	0.06245970619115278	0.07129624957262747	0.07830821520536643
21	0.0521467423382186	0.05791442812480507	0.06609758175045291	0.07540895724676397	0.08350881740670542
22	0.055047827319195435	0.06082934394956948	0.06969771188215743	0.07893154646022037	0.08722346458242437
23	0.05711768032568647	0.0637732596810297	0.07335796512117096	0.08312494379587673	0.09144734645121028
24	0.06014606597769269	0.06715508580434398	0.07787976883612677	0.08778705973980627	0.09682553258559098
25	0.06299793120194423	0.07069250124677944	0.0826946388601972	0.09317717560756934	0.10369128433164672
26	0.06602989547112409	0.07508000767946306	0.08853855012377168	0.09994937727503311	0.11182458833601834
27	0.06985908821576527	0.08017185227745025	0.0943598228711556	0.10855805993002561	0.11926306095722854
28	0.07476922509955214	0.08558359989814068	0.10022639992693635	0.11541477895810452	0.1271732638528928
29	0.08082697654221173	0.09275025575500671	0.10674528554019073	0.1217751009883986	0.1344831422782898
30	0.08646878995507364	0.09813325123006669	0.11267222418259133	0.12725420939287801	0.13926953632738423
31	0.09151610720807343	0.10385128682754727	0.11794062384310447	0.13236742750330485	0.14463687104722203
32	0.09644471072863017	0.10818117279519804	0.12286691608479247	0.1371665138984982	0.1507955221582018
33	0.09949317715060758	0.11140308251196819	0.12768748654593065	0.14278530800957942	0.15741777628373163
34	0.10113899378615351	0.11440405570419197	0.1318447325503226	0.14907986575631751	0.16784211995757845

S