Disease-associated antibody phenotypes and probabilistic seroprevalence estimates during the emergence of SARS-CoV-2

Xaquin Castro Dopico¹#, Leo Hanke¹*, Daniel J. Sheward¹*, Murray Christian¹, Sandra Muschiel², Nastasiya F. Grinberg³, Monika Adori¹, Laura Perez Vidakovich¹, Kim Chang II¹, Sharesta Khoenkhoin¹, Pradeepa Pushparaj¹, Ainhoa Moliner Morro¹, Marco Mandolesi¹, Mattias Forsell⁴, Jonathan Coquet¹, Martin Corcoran¹, Joanna Rorbach⁵,⁶, Soo Aleman⁷, Gordana Bogdanovic², Gerald McInerney¹, Tobias Allander¹,², Chris Wallace³,⁸, Ben Murrell¹*, Jan Albert¹,²*, Gunilla B. Karlsson Hedestam¹#

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

*Equal contribution
#Correspondence
Abstract

Serological studies are critical for understanding pathogen-specific immune responses and informing public health measures\(^1,2\). By developing highly sensitive and specific trimeric spike (S)-based antibody tests, we report IgM, IgG and IgA responses to SARS-CoV-2 in COVID-19 patients (\(n=105\)) representing different categories of disease severity. All patients surveyed were IgG positive against S. Elevated anti-SARS-CoV-2 antibody levels were associated with hospitalization, with IgA titers, increased circulating IL-6 and strong neutralizing responses indicative of intensive care status. Antibody-positive blood donors and pregnant women sampled during the pandemic in Stockholm, Sweden (weeks 14-25), displayed on average lower titers and weaker neutralizing responses compared to patients; however, inter-individual anti-viral IgG titers differed up to 1,000-fold. To provide more accurate estimates of seroprevalence, given the frequency of weak responders and the limitations associated with the dichotimization of a continuous variable\(^3,4\), we used a Bayesian approach to assign likelihood of past infection without setting an assay cut-off. Analysis of blood donors (\(n=1,000\)) and pregnant women (\(n=900\)) sampled weekly demonstrated SARS-CoV-2-specific IgG in 7.2% (95% Bayesian CI [5.1 – 9.5]) of individuals two months after the peak of spring 2020 COVID-19 deaths. Seroprevalence in these otherwise healthy cohorts increased steeply before beginning to level-off, following the same trajectory as the Stockholm region deaths over this time period.
Introduction

As SARS-CoV-2 only recently crossed the species barrier\textsuperscript{5}, populations globally were immunologically naïve to the virus. Characterizing the antibody response to nascent outbreaks is, therefore, central to optimizing approaches to tackle COVID-19 and future pandemics, furthering our basic understanding of human immunology\textsuperscript{6–8}.

To date, numerous SARS-CoV-2 studies have reported seroprevalence and disease-associated antibody phenotypes\textsuperscript{9–11}, isolated virus-neutralizing monoclonal antibodies\textsuperscript{12–15} and used convalescent individuals to define metrics for plasma therapy\textsuperscript{16}. However, consensus on several key issues remains outstanding. For instance, the majority of serology data are derived from commercial kits utilizing SARS-CoV-2 spike derivatives (e.g. RBD, S1 or S2 domains) or the nucleocapsid to detect pathogen-specific antibodies\textsuperscript{9,17,18}. Several of these assays suffer from epitope loss\textsuperscript{19}, increased cross-reactivity\textsuperscript{9,20} and lower sensitivity\textsuperscript{21}, highlighting the need for comparative studies to identify optimal assay formats for individual and population level analysis\textsuperscript{22–24}. Here, we developed a highly sensitive and specific ELISA assay based on native-like prefusion-stabilized spike trimers\textsuperscript{25}, as well as the receptor-binding domain (RBD), to accurately evaluate anti-viral antibody levels in COVID-19 patients and key community groups. To identify disease-associate antibody phenotypes, anti-viral IgM, IgG and IgA levels were analyzed alongside \textit{in vitro} virus neutralisation capacity and a descriptive set of clinical features, including intensive care status.

A major consideration for antibody testing concerns setting the cut-off for positivity\textsuperscript{26–28}, which significantly affects seroprevalence estimates and individual clinical management. Currently employed approaches to the problem are severely limited by their high dependency on the representative nature of negative control values and the dichotomization of a continuous variable. Therefore, to obtain more accurate seroprevalence estimates in blood donors and pregnant women, we strictly controlled our assay with a large number of historical controls ($n=295$) and used our patient data to train probabilistic algorithms to handle ELISA data. We used our cut-off-independent approach to model population changes in seropositivity over time.

We sampled blood donors and pregnant women weekly throughout the outbreak. Blood donors serve as an important clinical resource, including for potential COVID-19 plasma therapies, while pregnant women require close clinical monitoring with respect to fetal-maternal health and are known to employ unique, yet poorly characterized immunological mechanisms that impact infectious pathology\textsuperscript{29–31}. As Sweden did not impose a strict lockdown in response to the pandemic and has reported relatively high per-capita morbidity and mortality (ECDC, EU), understanding
SARS-CoV2 seroprevalence in these cohorts helps plan clinical need and understand the development of immunity in the population.
Results

Anti-SARS-CoV-2 antibody phenotypes in COVID-19 patients, blood donors and pregnant women

To evaluate humoral immune responses to SARS-CoV-2, we developed ELISA protocols to profile IgM, IgG and IgA directed against a stabilized trimeric form of the spike (S) glycoprotein, its receptor binding domain (RBD) and the nucleocapsid protein. The two spike-based antigens were produced in mammalian cells in-house (Fig S1A) and S trimer conformation was confirmed by cryo-EM analysis.

The study as a whole, included sera from SARS-CoV-2 PCR-positive COVID-19 patients (n=105) representing different disease categories, current blood donors (n=1,000) and pregnant women (n=900) sampled during 2020, hereafter referred to as healthy donors, historical blood donors collected during spring 2019 (n=295) and individuals shown to be PCR+ for endemic coronaviruses (ECV+, n=20) (Table 1 and Materials and Methods). We used a subset of these samples as a validation set for assay optimization, comprising samples from COVID-19 patients (n=36) and two sets of negative controls, 2019 blood donors (n=75) and ECV+ individuals (n=20).

All COVID-19 patients in the validation set displayed IgG responses to the S antigen and almost all responded to RBD, albeit less strongly (Fig S1B and S2A). No reactivity was recorded for the negative control samples, except for two ECV+ individuals who displayed reproducible IgM reactivity to SARS-CoV-2 nucleocapsid and S, and two 2019 blood donors with weak anti-S IgM reactivity (Fig S1C and S2). Further investigation is required to establish the basis of these observations and whether cross-reactive memory or primordial lymphocyte lineages contribute to SARS-CoV-2 responses a priori.

Our assay revealed up to a 1000-fold difference in anti-viral antibody IgG titers between individuals when examining serially diluted sera (Fig S2). In patients, anti-viral IgG titers were comparable for S (EC50=3,064; 95% CI [1,197 - 3,626]) and nucleocapsid (EC50=2,945; 95% CI [543 - 3,936]) and lower for RBD [EC50=1,751; 95% CI 966 - 1,595]. A subset (approximately 10%) of patients did not have detectable IgG responses to the nucleocapsid (Fig S2), in agreement with previous studies.

We next screened all SARS-CoV-2 PCR+ COVID-19 patients in the study (n=105) and detected potent IgG responses against S in all individuals and against RBD in all but a 3% individuals (Fig 1A and S3A). When screening the healthy blood donor and pregnant cohorts, we observed a subset of individuals who were clearly seropositive against SARS-CoV-2, but the titers against S and RBD that were generally lower than those measured in COVID-19 patients (Fig 1A and S3A). However, IgG titers in this
group varied greatly, with some persons displaying levels similar to those measured in COVID-19 patients while others had titers similar to those observed at the higher end of the historical negative control range. This group of weak responders highlighted the challenge of using a pre-determined assay cut-off, a critical point that we address and present potential solutions for in Figure 4.

When examining IgM and IgA responses against S and RBD, we found them to be less potent and more variable between individuals than the IgG response (Fig 1B and S3B). Given their unique immunological roles and what is known about COVID-19 pathology, we sought to investigate whether isotype responses segregated with clinical features. Therefore, the COVID-19 patient cohort was classified according to clinical disease severity: Category 1 – non-hospitalized; Category 2 – hospitalized; Category 3 – intensive care (on mechanical ventilation) (Table 1).

Within-patient anti-S and anti-RBD responses were highly correlated for all three isotypes (Fig S3C). Furthermore, multivariate analyses showed that increased anti-viral IgM, IgG and IgA levels significantly correlated with disease severity (Fig 1C, S3D and S Table 1), in line with the lower titers observed in blood donors and pregnant women. This was most pronounced for pathogen-specific IgA, suggestive of advancing mucosal disease\(^3\). A more severe clinical picture was also strongly associated with elevated serum IL-6 (Fig 1C), a cytokine that feeds antibody production\(^3\). IL-6 is dysregulated in polygenic metabolic diseases\(^3\) and acute respiratory distress syndrome (ARDS)\(^4\), which are risk factors for COVID-19-associated mortality\(^4,4\). Notably, under univariate analysis, female patients showed lower anti-viral IgA levels than males in non-hospitalized and hospitalized groups, which was also true for anti-RBD IgA in multivariate analysis (Fig S4A)\(^4\). Overall, severity showed the most consistent relationship with any measure and was the primary driver of Ig levels.

To summarize patient antibody phenotypes, we generated a Spearman’s rank correlation matrix (Fig 2A), which highlights a negative association between patient date of birth and all dataset features, in-keeping with the worse prognosis for elderly COVID-19 patients, as well as a lack of association between serum IL-6 and anti-viral IgM levels, further supporting that levels of the cytokine and IgA mark a protracted, more severe clinical course. After accounting for the effect of age, sex and days since PCR+ test, anti-viral IgA titers were approximately 3-fold higher in intensive care vs. non-hospitalized patients, while IL-6 levels were ca. 10-fold increased (Fig 2B). Notably, we also detected low levels of anti-viral IgA when analyzing a small subset \((n=100)\) of healthy donors from this year (Fig S4B).

Critical future studies are required to establish the longevity of immunity to SARS-CoV-2, including the efficacy of cellular re-call responses, as detectable antibody titers
decline in the absence of antigen. Here, anti-viral IgG levels were maintained two months post-disease onset/positive PCR test, while IgM and IgA declined (Fig S4C), in agreement with their circulating t1/2 and reflecting their diagnostic utility. Patient serum was collected 6-61 days post-PCR (Fig S4D), with the median time from symptom onset to PCR test being 5 days. In longitudinal patient samples where we observed seroconversion, IgM, IgG and IgA peaked with similar kinetics when all three isotypes developed, although anti-viral IgA was not always generated in Category 1 and 2 patients (Fig 1D and S4E). We note that anti-viral IgM could be maintained two months post-PCR+/symptom onset (Fig S4E).

We next characterized virus neutralizing antibody responses in patients and healthy donors, a key parameter for understanding the potential for protective humoral immune responses and the selection of plasma therapy donors. Using an in vitro pseudotype virus neutralisation assay, we detected neutralizing antibodies in the serum of all COVID-19 patients, and in all except two healthy antibody-positive donors (from n=56). A range of neutralizing ID50 titers is apparent in patients' serum (geometric mean ID50=1,330; 95% CI [837 – 2,112], Fig 3A), with binding and neutralisation titers being highly correlated (Fig 2A). The strongest neutralizing responses were observed in samples from patients in intensive care (Category 3) (ID50=5,058; 95% CI [2,422 - 10,564]), in keeping with their elevated and presumably more developed antibody response (Fig 3B and 1C). Sera from healthy blood donors and pregnant women also displayed neutralizing responses, but similarly to the binding data, were generally less potent than that observed in COVID-19 patients (ID50=600; 95% CI [357 – 1,010] and ID50=350; 95% CI [228 - 538], respectively, Fig 3C). Across the two antigens and three isotypes, anti-RBD IgG was most strongly correlated with neutralization (Fig 3D).

Community seroprevalence estimations

As the Stockholm region is a busy metropolitan area and Sweden did not impose strict lockdown in response to the emergence of SARS-CoV-2, we sought to better understand the frequency and nature of anti-SARS-CoV-2 antibody responses in healthy blood donors and pregnant women during weeks 14-25 (March 30 to June 22 2020) (Fig 4A).

By surveying a large number of historical controls (n=295) during assay development, we identified a considerable number of samples with weak reactor-phenotypes, which must be taken into account when setting the assay cut-off. When using a small set of historical controls, such values may be missed, resulting in an incorrect cut-off that significantly increases the uncertainty of individual and population seroprevalence estimates. To illustrate this, we randomly sampled groups of 20 negative control samples (from 890 measures in 295 individuals) and calculated seroprevalence in
blood and pregnant women from weeks 17-19 based on a 6SD cut-off. Depending on
which 20 negative controls were sampled, the seroprevalence ranged from 5.7 to
8.7%, a 35% difference (Fig 4B). Weak responder status is likely influenced by many
factors, including genetic background, health status, total serum Ig and protein levels,
assay variability. Critically, test samples with low anti-viral titers may also fall into this
range, highlighting the need to better understand the assay boundary. Taking a one-
dimensional 6SD cut-off for anti-S IgG responses based on all 890 values from 295
negative donors, 7.7% of healthy donors tested positive two months after the peak of
deaths in the country (Fig 4C).

To exploit individual antibody titers and improve our estimates, we modelled the
probability that a sample is positive by training two parallel learners using our patient
and historical control anti-S and -RBD data, rather than setting a one-dimensional cut-
off. Our Bayesian approach for inferring seroprevalence without thresholds is
described in Christian et al\(^\text{47}\). Briefly, we exploit a class-balance corrected Bayesian
logistic regression to infer seroprevalence, using a modified Gaussian Process to
construct a prior over seroprevalence trajectories, sharing information between weeks.
Using this approach, which allowed us to model population changes over time, we
found the steep increase in positivity at the start of the pandemic to slow between
weeks 17-25, approximating 7.2% (95\% equal tailed Bayesian credible intervals [5.1
– 9.5]) at the last time point and suggesting that humoral immunity to the virus
develops slowly in these populations despite considerable virus spread in the
community.

To confirm our novel Bayesian approach, we compared it to various machine learning
algorithms, creating an ensemble SVM-LDA learner to maximize sensitivity, specificity
and consistency across different cross-validation strategies, where the proportion of
positives in the held-out test set was deliberately varied from the training set (see
Materials and Methods). We obtained convergent results between these tools and our
framework (Fig S5A). Strikingly, seroprevalence inferred using our Bayesian approach
exhibited the same trajectory as Stockholm County deaths when lagged by two and
half weeks (Fig S5B), allowing for the calculation of a seroprevalence case fatality rate
in appropriately powered cohorts.

Furthermore, such probabilistic approaches that annotate uncharacterized uncertainty
pave the way towards greater clinical utility for antibody measures. For example,
individuals 50% likely to be antibody-positive according to the ensemble learner (Fig
4E) can targeted for further investigation (e.g. re-testing). These tools have the
potential, therefore, to provide more qualitative individual antibody measurements.
Discussion

Serology remains the gold standard for estimating previous exposure to pathogens and benefits from a large historical literature. When robust, it indicates whether an individual has mounted an adaptive immune response against a specific agent and is a strong predictor of an anamnestic response and quicker recovery upon re-infection. Anti-viral antibody responses are central for protection against re-infection and for the protective efficacy of successful vaccines.

Although the concept of herd immunity is based upon the study of antibodies, worryingly, there is little-to-no standardization (or validation) of many of the available SARS-CoV-2 antibody tests. Globally, hospital staff and health authorities are struggling with test choices, negatively impacting individual outcomes and efforts to contain the virus. There is no single ideal test, although some are more accurate, affordable and scalable than others, and it seems reasonable that an international body undertake more basic research-private collaborations to advise on best practice, provide reagents, and further research in the area.

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 antibody responses in three cohorts of clinical interest. COVID-19 patients showed the highest anti-viral titers, developing augmented serum IgA and IL-6 with worsening disease, likely a consequence of age, sex and other risk factors, as well as prolonged immune stimulation, a higher infectious dose, and host genetics. Therefore, 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 therapeutic delivery.

Our neutralisation data shows that patients and healthy donors who seroconvert develop antibody titers capable of preventing infection in vitro, demonstrating the beneficial role of immunoglobulins against SARS-CoV-2. Further research into the specific inflammatory molecules/pathways aggravating the disease course, however, would allow for more specific targeting and control of immunity, with IL-6 being a case-in-point (NCT04322773, NCT04359667), although the cytokine’s role in orchestrating adaptive responses may make its modulation a double-edged sword in some patients.

Outside of the severe disease setting, it is important to determine how many people have seroconverted to SARS-CoV-2. Blood donors and pregnant women are both good sentinels for population health, although they are not enriched for high-risk groups, such as public transportation employees, where rates of infection may be higher. Blood donors are generally working age, active and mobile members of society.
with a good understanding of health, and the majority of pregnant women in Sweden
would have been advised to take precautions against infectious diseases. However,
interestingly, both groups showed a similar seroprevalence.

By tracking these cohorts over time, we modelled seroprevalence changes at the
population level. We found a steep increase in antibody positivity at the start of the
pandemic, which increased, although at a slower rate, in blood donors and pregnant
women during subsequent weeks, in-line with a decreasing caseload and fatalities\textsuperscript{54}. Indeed, ICU occupancy and deaths are a better proxy for viral spread than PCR+
diagnoses, which are highly dependent on the number of tests carried out, and we
note the close approximation of our seroprevalence data to Stockholm County deaths
(Fig S5B), illustrating the power of mortality data to infer seroprevalence and vice
versa.

Together, our study defines key features of the humoral immune responses to
emerging beta-coronaviruses, delineating disease features, while our seroprevalence
data indicate that population immunity to SARS-CoV-2 develops slowly even in the
absence of lockdowns, suggesting that classical serological herd immunity following
natural infection will require a larger outbreak and higher clinical toll; notwithstanding
yet unknown contributions from T lymphocytes and other lineages. Given the
uniqueness of the approach in Sweden\textsuperscript{55}, these data may inform the management of
future pandemics. More widely, given the high inter-individual variability in antibody
responses to SARS-CoV-2, the genetic and environmental factors influencing
individual antibody responses will be important to elucidate.
Materials and methods

Human samples and ethical declaration

Samples from COVID-19 patients \((n=105)\) were collected through the Departments of Medicine and Clinical Microbiology at the Karolinska University Hospital and were handled and analyzed in accordance with approval by the Swedish Ethical Review Authority (registration no. 2020-02811). All personal identifiers were pseudo-anonymized, and clinical feature data were blinded to the researchers carrying out experiments until data generation was complete.

All patients in the study were confirmed PCR+ for SARS-CoV-2 by nasopharyngeal swab or upper respiratory tract sampling after being admitted to Karolinska University Hospital. As viral RNA CT values were determined using different qPCR platforms between patients, we did not analyze these alongside other available features. Patients 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. In addition, longitudinal samples from 10 of these patients were collected to monitor seroconversion and isotype persistence.

Anonymized samples from blood donors (100/week) and pregnant women (100/week) were randomly selected and obtained from the department of Clinical Microbiology, Karolinska University Hospital. No metadata, such as age or sex information was available for these samples. Pregnant women were sampled as part of routine for infectious diseases screening during the first trimester of pregnancy. Blood donors \((n=295)\) 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 \((n=20, \text{ ECV}+)\) in the prior 2-6 months, were used as additional negative controls. The use of these anonymized 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 Swedish Public Health Agency, respectively.

Study samples are defined in Table 1.

Serum sample processing

Blood samples were collected from patients by the attending clinical team and serum isolated by the department of Clinical Microbiology. Samples were 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\(^{25}\). 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-CoV2 ELISA

96-well ELISA plates (Nunc MaxiSorp) were coated with SARS-CoV-2 S, RBD or nucleocapsid (100 μl of 1 ng/μl) in PBS overnight at 4°C. Plates were washed six times with 300 ml PBS-Tween-20 (0.05%) and blocked using PBS-5% no-fat milk (Sigma). Human serum samples were thawed at room temperature, diluted (1:100 unless otherwise indicated), vortexed and incubated in blocking buffer for 1h 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 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 patients. All 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\(^{25}\); 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 RLU 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+ COVID-19 patient 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 patient 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 (Supplementary Table 1).

Probabilistic seroprevalence estimations

We employed two distinct probabilistic strategies for estimating seroprevalence without thresholds, each developed independently. Our Bayesian approach is explained in detail in Christian et al. Briefly, we use 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 adjust for the training data class proportions and use 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 construct a prior over seroprevalence trajectories using a transformed Gaussian Process, and combine this...
with the individual class-balance adjusted infection probabilities for each donor to infer
the posterior distribution over seroprevalence trajectories.

Our second approach was based on more traditional machine learning and statistical
approaches, which we compared through ten-fold cross validation (CV): logistic
regression, 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.

We considered three strategies for cross-validation:

- random: individuals were sampled into folds at random
- stratified: individuals were sampled into folds at random, subject to ensuring
  the balance of cases:controls remained fixed
- 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, although logistic regression was sensitive to
changes in the case proportion in the training data. We chose to create an ensemble
method combining that with the highest specificity (LDA) and the highest sensitivity
(SVM), defined as an unweighted average of the probabilities generated under SVM
and LDA. The ensemble learner had average sensitivity > 99.1% and average specificity 99.8%.

We then trained the ensemble learner on all 719 training samples and predicted the probability of anti-SARS-CoV-2 antibodies in blood donors and pregnant volunteers sampled in 2020. 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.

**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/](https://github.com/MurrellGroup/DiscriminativeSeroprevalence/) and [https://github.com/chriswallace/seroprevalence-paper](https://github.com/chriswallace/seroprevalence-paper).

**Author contributions**

GKH and XCD designed the study, analyzed the results and wrote the manuscript with input from co-authors. JA, TA, SM, GB 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 Dario 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 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. Mouquet, H. & Nussenzweig, M. C. Polyreactive antibodies in adaptive immune responses to viruses. *Cellular and Molecular Life Sciences* (2012) doi:10.1007/s00018-011-0872-6.

3. 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.

4. Altman, D. G. & Royston, P. The cost of dichotomising continuous variables. *British Medical Journal* (2006) doi:10.1136/bmj.332.7549.1080.

5. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2001017.

6. Davis, M. M. A Prescription for Human Immunology. *Immunity* vol. 29 835–838 (2008).

7. 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.

8. London, A. J. & Kimmelman, J. Against pandemic research exceptionalism. *Science* (2020) doi:10.1126/SCIENCE.ABC1731.

9. 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.

10. 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.

11. Tan, W. et al. Viral Kinetics and Antibody Responses in Patients with COVID-19. *J Clin* doi: 10.11, (2020).

12. 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.

13. 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.

14. Ju, B. et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* (2020) doi:10.1038/s41586-020-2380-z.

15. 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.

16. 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.

17. 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.

18. 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.*
19. 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.

20. 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.

21. 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.

22. GeurtsvanKessel, C. H. et al. Towards the next phase: evaluation of serological assays for diagnostics and exposure assessment. *Nat Comms* doi: 10.1038/s41467-020-17100-7.

23. 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.

24. 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.

25. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (80-. ).* (2020) doi:10.1126/science.1006461.

26. 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.

27. 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.

28. 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.

29. Aghaeepour, N. et al. An immune clock of human pregnancy. *Sci. Immunol.* (2017) doi:10.1126/sciimmunol.aan2946.

30. 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.

31. Kourtis, A. P., Read, J. S. & Jamieson, D. J. Pregnancy and infection. *New England Journal of Medicine* (2014) doi:10.1056/NEJMra1213566.

32. Hanke, L. et al. An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *bioRxiv* (2020) doi:10.1101/2020.06.02.130161.

33. 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.

34. 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).
35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. 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).

41. Park, W. Y. et al. Cytokine balance in the lungs of patients with acute respiratory distress syndrome. Am. J. Respir. Crit. Care Med. (2001) doi:10.1164/ajrccm.164.10.2104013.

42. Magro, G. SARS-CoV-2 and COVID-19: Is interleukin-6 (IL-6) the ‘culprit lesion’ of ARDS onset? What is there besides Tocilizumab? SGP130Fc. Cytokine: X (2020) doi:10.1016/j.cytox.2020.100029.

43. Chua, R. L. et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat. Biotechnol. (2020) doi:10.1038/s41587-020-0602-4.

44. Jin, J. M. et al. Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. Front. Public Heal. (2020) doi:10.3389/fpubh.2020.00152.

45. Takaki, A. et al. Cellular immune responses persist and humoral responses decrease two decades after recovery from a single-source outbreak of hepatitis C. Nat. Med. (2000) doi:10.1038/75063.

46. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. J. Allergy Clin. Immunol. (2010) doi:10.1016/j.jaci.2009.09.046.

47. Christian, M. & Murrell, B. Discriminative Bayesian Serology: Counting Without Cutoffs. BioRxiv (2020).

48. Behring & Kitasato. Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Thieren. Dtsch. Medizinische Wochenschrift (1890) doi:10.1055/s-0029-1207589.

49. Plotkin, S. A. Vaccines: Correlates of Vaccine-Induced Immunity. Clin. Infect. Dis. (2008) doi:10.1086/589862.

50. Larios Mora, A. et al. Delivery of ALX-0171 by inhalation greatly reduces respiratory syncytial virus disease in newborn lambs. MAbs (2018) doi:10.1080/19420862.2018.1470727.

51. Sheridan, C. Ablynx’s nanobody fragments go places antibodies cannot. Nat. Biotechnol. (2017) doi:10.1038/nbt1217-1115.
52. Hurst, S. M. et al. IL-6 and its soluble receptor orchestrate a temporal switch in the pattern of leukocyte recruitment seen during acute inflammation. *Immunity* **14**, 705–714 (2001).

53. Pedersen, B. K. IL-6 signalling in exercise and disease. in *Biochemical Society Transactions* (2007). doi:10.1042/BST0351295.

54. Intensivvårdsregistret, S. Antal som intensivvårdas med Covid-19 per dag. *Svenska Intensivvårdsregistrets utdataportal* https://portal.icuregswe.org/siri/report/corona.co (2020).

55. Paterlini, M. ‘Closing borders is ridiculous’: the epidemiologist behind Sweden’s controversial coronavirus strategy. *Nature* (2020) doi:10.1038/d41586-020-01098-x.

56. Platt, J. & others. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv. large margin Classif.* (1999).

57. Lott, A. & Reiter, J. P. Wilson Confidence Intervals for Binomial Proportions With Multiple Imputation for Missing Data. *Am. Stat.* (2020) doi:10.1080/00031305.2018.1473796.
Figure 1: Patient, blood donor and pregnant women anti-SARS-CoV-2 antibody phenotypes

(A) Log transformed anti-S and -RBD IgG responses are shown in COVID-19 patients (n=105), blood donors (BD, n=1,000) and pregnant women (PW, n=900). Controls (C) represent n=295 individuals analyzed repeatedly (to n=890). (B) IgM, IgG and IgA responses against S and RBD in COVID-19 patients (n=105), with a limited number of controls for each assay represented by open circles. (C) COVID-19 patient (n=105) anti-viral antibody levels and serum IL-6 levels according to clinical disease severity. S and RBD responses are graphed together, alongside a limited number of controls. (D) Three longitudinal patient profiles of seroconversion and neutralisation capacity are shown in hospitalized individuals.
Figure 2: Antibody-clinical feature associations in COVID-19 patients

(A) Spearman’s rank correlation of patient 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 (B) Adjusted fold change for dataset features in COVID-19 patients compared to category 1 individuals. The effects of age (DOB), sex, days from PCR test were considered.
Figure 3: In vitro virus neutralizing responses
(A) Pseudotyped virus neutralization by serial dilutions of serum from COVID-19 patients (n=48, red), blood donors (blue) and pregnant women (purple). (B) Neutralization ID₅₀ titers according to disease severity category. (C) Comparison of neutralization ID₅₀ titers between patients, blood donors and pregnant women. (D) Correlation between anti-S and -RBD IgM, IgG and IgA antibodies and neutralisation titers.
Figure 4: Probability-based seroprevalence in Stockholm during the height of the pandemic

(A) Study sample collection intervals are shown alongside the rolling weekly average COVID-19 deaths per million inhabitants in relevant countries for comparison. (B) By randomly selecting six groups of 20 negative controls from a pool of 890 values (from n=295 individuals), the range of the 6SD cut-off is shown to significantly affect positivity estimates in healthy donors (HD) from weeks 17-19 2020. (C) Using a conventional 6SD cut-off, we found 6.2, 6.5 and 7.7% of healthy donors to be antibody-positive in weeks 17-19, 20-22 and 23-25, respectively. (D) Seroprevalence estimates in healthy donor cohorts through time using a cut-off-independent Bayesian approach. (E) An ensemble SVM-LDA learner trained on our patient data illustrates the problem of weak responders and provides a multi-dimensional solution by providing individual datapoints with a likelihood of positivity, a feature with potential for clinical use.
| Table 1 – Study patient, blood donor and pregnant women characteristics |
|---------------------------------------------------------------|
| **COVID-19 patients (n=)**                                    | 105 |
| Females                                                      | 44 (41.9%) |
| Males                                                        | 61 (58.1%) |
| Median age (years)*                                          | 53.0 (49-61) |
| Females                                                      | 51.5 (48-56.2) |
| Males                                                        | 55.0 (49-63) |
| Non-hospitalized                                             | 53 (50.5%) |
| Females, males                                               | 28, 25 |
| Hospitalized                                                 | 31 (29.5%) |
| Females, males                                               | 12, 17 |
| Intensive care (ICU)                                         | 21 (20%) |
| Females, males                                               | 3, 17 |
| Disease category unknown                                     | 2 (1.9%) |
| Females, males                                               | 1, 1 |
| SARS-CoV2+ PCR                                               | 105 |
| Sample collection dates                                      | February-June 2020 |
| **Blood donors (n=)**                                        | 1,000 |
| Median age (years)                                           | N/A |
| Sex                                                          | N/A |
| SARS-CoV2+ PCR                                               | N/A |
| Sample collection dates                                      | Weeks 14-25 2020 |
| **Pregnant women (n=)**                                      | 900 |
| Age (years)                                                  | N/A |
| SARS-CoV2+ PCR                                               | N/A |
| Sample collection dates                                      | Weeks 17-25 2020 |
| **Historical blood donors (n=)**                             | 295 |
| Median age (years)                                           | N/A |
| Sex                                                          | N/A |
| SARS-CoV2+ PCR                                               | No |
| Sample collection dates                                      | Spring 2019 |
* LQR-UQR given in brackets
**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 subset of patients and negative controls was used to validate the assays for the three isotypes. (C) Two ECV+ donors, K2 and K4, showed reproducible IgM binding to S, please refer to their increased responses to S and nucleocapsid in S2. Testing another subset of historical controls (n=75) for a similar observation, two additional individuals were found to show IgM binding to S.
**Figure S2: Patient binding titrations**

A subset (n=40) PCR+ COVID-19 patients were titrated in our assay. Shown are these responses against the three antigens and isotypes in the study. Four ECV+ (red lines) were in a similar dilution series alongside.
Figure S3: Healthy donor and patient antibody phenotypes

(A) Raw OD450 anti-S and -RBD IgG responses are shown for patients and healthy donors. (B) Raw OD450 isotypic responses in patients against S and RBD. A limited number of negative controls are depicted by clear circles. (C) Anti-S vs -RBD responses in COVID-19 patients are highly correlated. (D) Log transformed patient anti-S and -RBD levels according to disease severity by isotype. P values from a Cuzick test for trend.
Figure S4: Patient antibody phenotypes
(A) Antibody responses according to sex are shown (left panel) and severity (right panel). *P* values from a Kruskal Wallis or stratified Mann-Whitney, as shown. (B) Lower levels of IgA were observed in a subset (*n*=100) of blood donors and pregnant women. Patient (P) and healthy donor (HD) responses to S and RBD are shown. (C) IgM and IgA titers declined with time from first symptom/SARS-CoV-2+ PCR and admittance to care. IgG levels were maintained. *P* values from a Spearman rank correlation test. (D) Patient sampling according to date since PCR+ test shown for anti-viral IgM responses. (E) Longitudinal profiles of category 1 and 2 patients after seroconversion.
Figure S5: Comparison of cut-off-independent machine learning methods

(A) Overlay and agreement of SVM-LDA ensemble learner predictions with our Bayesian framework. (B) Seroprevalence changes over time follow a very similar trajectory to Stockholm County deaths when lagged by 2.5 weeks.