SARS-CoV-2 IgG seroprevalence surveys in blood donors before the vaccination campaign, France 2020-2021

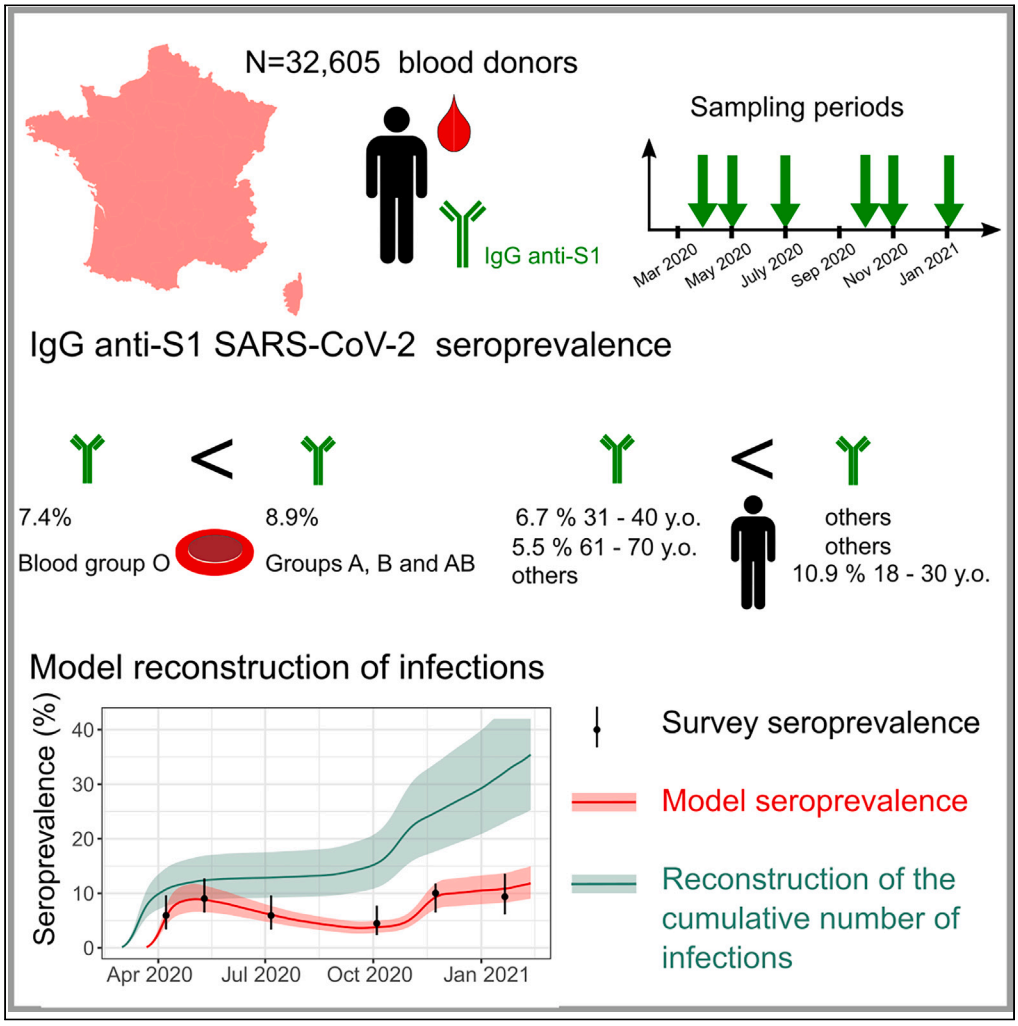
Pierre Gallian, Nathanaël Hozé, Nadège Brisbarre, Paola Mariela Saba Villarroel, Elif Nurtop, Christine Isnard, Boris Pastorino, Pascale Richard, Pascal Morel, Simon Cauchemez, et al.

iScience

CellPress
OPEN ACCESS

Article

SARS-CoV-2 IgG seroprevalence surveys in blood donors before the vaccination campaign, France 2020-2021



Pierre Gallian, Nathanaël Hozé, Nadège Brisbarre, ..., Pascal Morel, Simon Cauchemez, Xavier de Lamballerie

pierre.gallian@efs.sante.fr

Highlights

Immunity in France before the vaccination campaign was too low to achieve herd immunity

The infection/ hospitalization ratio increased with age

4 months after seroconversion, modeling estimated antibody decay to be ~50%

IgG seroprevalence was lower in individuals with O blood type than A, B, and AB

Gallian et al., iScience *26*, 106222
April 21, 2023
https://doi.org/10.1016/j.isci.2023.106222

iScience

CellPress
OPEN ACCESS

Article

SARS-CoV-2 IgG seroprevalence surveys in blood donors before the vaccination campaign, France 2020-2021

Pierre Gallian,[1,2,6,7,*] Nathanaël Hozé,[3,6] Nadège Brisbarre,[2,4] Paola Mariela Saba Villarroel,[2] Elif Nurtop,[2] Christine Isnard,[2,4] Boris Pastorino,[2] Pascale Richard,[1] Pascal Morel,[1,5] Simon Cauchemez,[3] and Xavier de Lamballerie[2]

SUMMARY

We conducted a cross-sectional study for SARS-CoV-2 anti-S1 IgG prevalence in French blood donors (n = 32605), from March-2020 to January-2021. A mathematical model combined seroprevalence with a daily number of hospital admissions to estimate the probability of hospitalization upon infection and determine the number of infections while correcting for antibody decay. There was an overall seroprevalence increase over the study period and we estimate that ~15% of the French population had been infected by SARS-CoV-2 by January-2021. The infection/hospitalization ratio increased with age, from 0.31% (18-30yo) to 4.5% (61-70yo). Half of the IgG-S1 positive individuals had no detectable antibodies 4 to 5 months after infection. The seroprevalence in group O donors (7.43%) was lower (p = 0.003) than in A, B, and AB donors (8.90%). We conclude, based on seroprevalence data and mathematical modeling, that a large proportion of the French population was unprotected against severe disease prior to the vaccination campaign.

INTRODUCTION

By mid-April 2022, the pandemic caused by SARS-CoV-2 has been responsible for more than 6.2 million deaths worldwide and more than 500 million people have been diagnosed with COVID-19. In France, since January 2020, over 27 million cases and 144,000 deaths have been recorded.[1] The pandemic continues to grow and hopes for curbing the dynamics of infections rely heavily on building up population immunity derived from natural infection or vaccines. Estimating the proportion of the population carrying antibodies is necessary to ascertain the degree of viral penetration, estimate the level of protection against infection and ascertain disease severity. Seroprevalence surveys in the blood donor population provide rapid, large-scale information on humoral immunity that can be used to map affected areas and monitor epidemic progression over time. Serological studies on SARS-CoV-2 are technically challenging but important given the proportion of asymptomatic forms.[2] Among patients, antibody levels appear to be higher and more persistent in those with a severe form of the infection compared to those with a moderate form. Reinfections are possible mainly in people with low levels of neutralizing antibodies.[3] Finally, anti-SARS-CoV-2 antibodies decrease or disappear over time and these kinetics differ according to the immunoglobulin and antigenic target sought.[4]

We report here a study of the seroprevalence of anti-SARS-CoV-2 IgG antibodies in the French blood donor population (called COVIDonneur) before the implementation of the general population vaccination campaign and before the circulation of antigenic variants. Using a mathematical model, we analyze these data jointly with age-stratified hospital admissions data to reconstruct epidemic dynamics in France.

RESULTS

SARS-CoV-2 anti-S1 IgG prevalence

A total of 32,605 donations (17,043 females and 15,562 males) from 31,922 donors (16,758 females (52.5%) and 15,164 males (47.5%)) were collected between the end of March 2020 and early February 2021 during six collection times (T1 –T6) over four survey's periods chosen according to the evolution of the COVID-19

[1]Établissement Français du Sang, La Plaine Saint Denis 93218, France

[2]Unité des Virus Émergents (UVE: Aix-Marseille University - IRD 190 - Inserm 1207), 13005 Marseille, France

[3]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université Paris Cité, UMR2000, CNRS, 75015 Paris, France

[4]Établissement Français du Sang Provence Alpes Côte d'Azur et Corse, 13005 Marseille France

[5]UMR RIGHT 1098, Inserm, Établissement Français du Sang, University of Franche-Comté, 25000 Besançon, France

[6]These authors contributed equally

[7]Lead contact

*Correspondence: pierre.gallian@efs.sante.fr

https://doi.org/10.1016/j.isci.2023.106222

epidemic in France (Table 1, Figure S1). SARS-CoV-2 anti-S1 IgG prevalence rates for each department according to the sampling period and evolution over time are shown in Figure 1. In January/February 2021 (survey time T6), no significant differences in rates of SARS-CoV-2 anti-S1 IgG according to gender could be found (Table 1 and Figure 2). Seroprevalence was highest ($p < 0.001$) in the 18-30 age group (10.90%; 95%CI 9.90-11.90) and lowest among the 31-40 (6.66%; 95%CI 5.60-7.72, $p = 0.004$) and 61-70 (5.46%; 95%CI 4.27-6.65, $p < 0.001$) age groups compared to other age groups (Figure 2). The time trends of the seroprevalence rate varied with age (Figure 3A). In blood donors over 40 years of age, seroprevalence rates ranged from 2.66% (95%CI 1.41-3.91) to 8.67% (95%CI 6.57-10.77) during the study. In the age group 31-40 years old, there was a decrease in seroprevalence over time and in the age group 18-30 years old seroprevalence increased from the summer of 2020.

We assessed how anti-S1 IgG seroprevalence rates varied with the phenotypic expression of ABO blood group antigens in donations at the six surveillance times (Table 1). At T6, of all donations tested (n = 12,016), the prevalence of SARS-CoV-2 anti-S1 IgG in group O donors (7.43%; 95%CI 6.74-8.11) whose red blood cells do not carry A or B antigens was significantly lower ($p = 0.003$) than in groups A, B, and AB donors (8.90%; 95%CI 8.20-9.60). The seroprevalence in groups A and AB ("all A" antigen-carrying) donors was higher ($p = 0.011$) than in "non-A" donors (blood groups B and O): 8.93% (95%CI 8.15-9.71) versus 7.65% (95%CI 7.02-8.28). A similar difference ($p = 0.014$) was observed for "all B" groups: 9.81% (95%CI 8.33-11.29) versus 7.97% (95%CI 7.45-8.49) for "non-B." No significant difference was observed for the prevalence of anti-SARS-CoV-2 IgG according to Rhesus or Kell blood groups (Figure 2).

The proportion of neutralizing antibodies that tested positive (titer >10) in reactive samples (positive and indeterminate) increased during the studied period from 27.9% in March to April 2020 to 85.4% in early 2021 (Figure 3B).

Model of the dynamics of anti-SARS-CoV-2 IgG prevalence

We developed a mathematical model that combined the anti-SARS-CoV-2 IgG seroprevalence with the daily number of hospital admissions to estimate the probability of hospitalization upon infection and determine the number of infections while correcting for antibody decay. Using our mathematical model, we found that the proportion of population infected during the first wave (March-May 2020) was heterogeneous across France, consistent with previous studies, with seroprevalence ranging from 2.9% in Pays de la Loire to 15.8% in the Grand-Est region (Figure 4A).[5,6] The cumulative number of infected individuals remained stable during the summer and increased during the second wave that started in September 2020 (Figure 4A). In contrast, the seroprevalence was stable during the different sampling periods in most regions and even decreased in Grand-Est. In the model, we explained this difference between the number of infected and the seroprevalence by the decay of the antibody following seroconversion. The reconstructed seroprevalence data agreed well with the observations (Figure 4B), with 85% (145/170) of reconstructed values falling in the 95% confidence interval of the observations (in the surveys taken across age groups, sampling periods, and regions).

We estimated that the capacity of IgG anti-S1 assay to detect a past infection is 50% of its initial capacity after 3.4 months following seroconversion (95% CrI: 2.7-4.4 months) (Figure 5A). As a result, our model predicts that 90% (95% CrI: 84-95) of the individuals infected during the first wave of March-May 2020 were already seronegative by the time of the last sampling (January 2021). In an additional sensitivity analysis, we assumed that the decay rate could be different for those aged below and above 50 yo (Figure S2) and found no difference between the different age groups. This fast antibody decay leads to important differences between the proportion of infected individuals and the seroprevalence (Figure 5B), with an estimated proportion of infected individuals in the population being more than 2-fold higher than the observed seroprevalence (Figure 5B and Table S1). The infection hospitalization ratio increased with age, from 0.31% (95% CrI: 0.25-0.38) for the 18-30 age group, to 4.5% (95% CrI: 3.4-5.9) in the 61-70 age group (Figure 5C and Table S2). The patterns of infection-hospitalization ratio (IHR) were consistent with those derived from another seroprevalence survey (SAPRIS) that was conducted in three regions of metropolitan France in May 2020 (Figure 5C).[5,7,8]

In a sensitivity analysis, we applied our model to anti-nucleocapsid (anti-N) IgG seroprevalence data in three departments of metropolitan France at survey's times T3, T4, and T6. We found no major difference between the anti-S1 and anti-N IgG seroprevalence values (Figures S3A and S3B), both in the observations and the

Table 1. SARS-CoV-2 anti-S1 IgG prevalence by demographic characteristics for the 4 period of survey and detailed for the 6 survey's times

Survey 's period	Period 1 (Figure 1A)		Period 2 (Figure 1B)	Period 3 (Figure 1C)		Period 4 (Figure 1D)	Total
Survey 's time (T)	T1	T2	T3	T4	T5	T6	Total
Period of blood collection	2020, March 27 to April 4	2020, May 11-29	2020, July 1-23	2020, Sept 28 to Oct 29	2020, Nov 3-18	2021, Jan 14 to Feb 9	
Nb samples	2232	3241	2951	6009	6156	12016	32605
Nb of departments	5	8	7	12	12	27	30
Gender							
Male	1105; 85 (7.69%)	1510; 104 (6.89%)	1394; 72 (5.16%)	2773; 136 (4.90%)	2910; 212 (7.29%)	5870; 458 (7.80%)	15562; 1067 (6.86%)
Female	1127; 83 (7.36%)	1731; 130 (7.51%)	1557; 109 (7.00%)	3236; 185 (5.72%)	3246; 269 (8.29%)	6146; 528 (8.59%)	17043; 1304 (7.65%)
Age							
18–30	663; 33 (4.98%)	949; 55 (5.80%)	827; 41 (4.96%)	2049; 99 (4.83%)	1821; 197 (10.82%)	3762; 410 (10.90%)	10,071; 835 (8.29%)
31–40	436; 61 (13.99%)	581; 62 (10.67%)	535; 56 (10.47%)	1045; 85 (8.13%)	1308; 109 (8.33%)	2191; 146 (6.66%)	6096; 519 (8.51%)
41–50	527; 39 (7.40%)	692; 60 (8.67%)	610; 34 (5.57%)	1178; 77 (6.54%)	1305; 96 (7.36%)	2475; 181 (7.31%)	6787; 487 (7.18%)
51–60	408; 20 (4.90%)	618; 38 (6.15%)	593; 33 (5.56%)	1097; 43 (3.92%)	1095; 61 (5.57%)	2177; 172 (7.90%)	5988; 367 (6.13%)
61–70	198; 15 (7.58%)	401; 19 (4.74%)	386; 17 (4.40%)	640; 17 (2.66%)	627; 18 (2.87%)	1411; 77 (5.46%)	3663; 163 (4.45%)
ABO Blood type							
A	907; 80 (8.82%)	1371; 106 (7.73%)	1251; 77 (6.16%)	2308; 129 (5.59%)	2459; 198 (8.05%)	4810; 414 (8.61%)	13106; 1004 (7.66%)
B	205; 13 (6.34%)	298; 29 (9.73%)	291; 15 (5.15%)	625; 44 (7.04%)	624; 52 (8.33%)	1177; 103 (8.75%)	3220; 256 (7.95%)
O	1019; 66 (6.48%)	1499; 93 (6.20%)	1336; 85 (6.36%)	2885; 133 (4.61%)	2859; 213 (7.45%)	5656; 420 (7.43%)	15254; 1010 (6.62%)
AB	101; 9 (8.91%)	73; 6 (8.22%)	73; 4 (5.48%)	191; 15 (7.85%)	214; 18 (8.41%)	373; 49 (13.14%)	1025; 101 (9.85%)
Rhesus type							
Rh positive	1846; 139 (7.53%)	2657; 192 (7.23%)	2380; 155 (6.51%)	5043; 274 (5.43%)	5051; 395 (7.82%)	9851; 812 (8.24%)	26828; 1967 (7.33%)
Rh negative	386; 29 (7.51%)	584; 42 (7.19%)	571; 26 (4.55%)	966; 47 (4.87%)	1105; 86 (7.78%)	2165; 174 (8.04%)	5777; 404 (6.99%)
Kell type							
Kell positive	183; 18 (9.84%)	260; 17 (6.54%)	223; 8 (3.59%)	489; 27 (5.52%)	487; 28 (5.75%)	951; 64 (6.73%)	2593; 162 (6.25%)
Kell negative	2049; 150 (7.32%)	2981; 217 (7.28%)	2728; 173 (6.34%)	5520; 294 (5.33%)	5669; 453 (7.99%)	11065; 922 (8.33%)	30012; 2209 (7.36%)

[n; pos (%age)]. See also Figure S1.

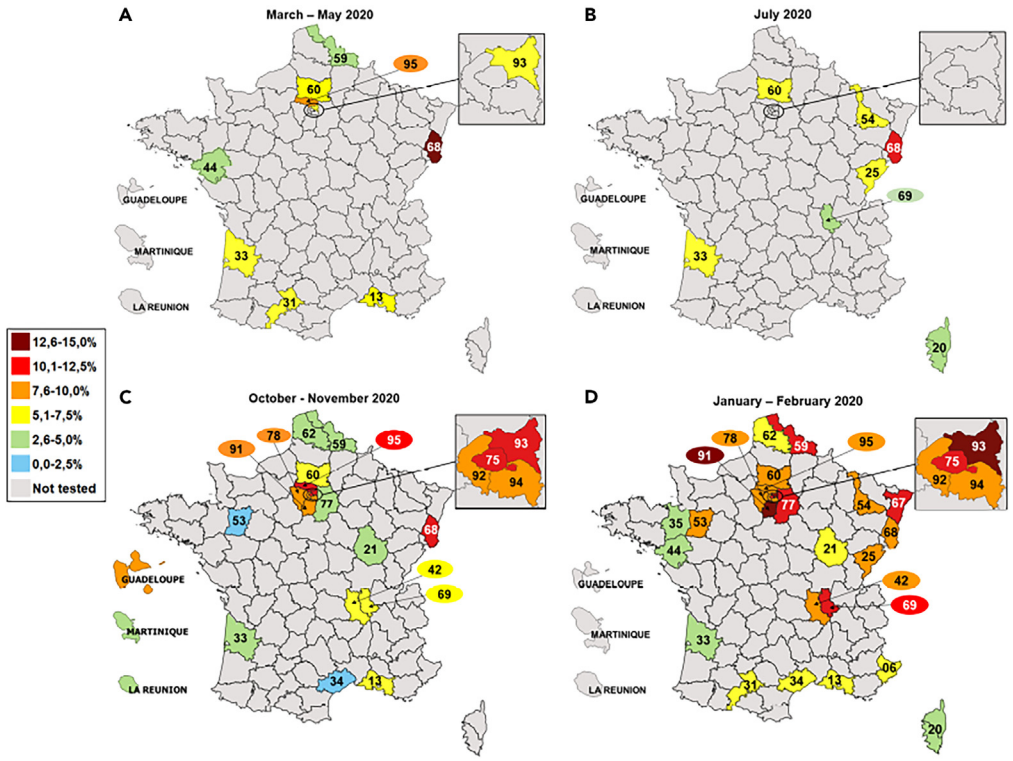


Figure 1. SARS-CoV-2 anti-S1 IgG seroprevalence in French departments tested at different times of the epidemics

(A) samples collected between March and May 2020; (B) samples collected in July 2020; (C) samples collected in October and November 2020 and (D) samples collected in January-February 2021. Numbers labeled on the maps referred to French departments: [06 = Alpes Maritimes], [13 = Bouches du Rhône], [20 = Corse], [21 = Côte d'Or], [25 = Doubs], [31 = Haute Garonne], [33 = Gironde], [34 = Hérault], [35 = Ille et Vilaine], [42 = Loire], [44 = Loire Atlantique], [53 = Mayenne], [54 = Meurthe et Moselle], [59 = Nord], [60 = Oise], [62 = Pas de Calais], [67 = Bas Rhin], [68 = Haut Rhin], [69 = Rhône], [75 = Paris], [77 = Seine et Marne], [78 = Yvelines], [91 = Essonne], [92 = Hauts de Seine], [93 = Seine Saint Denis], [94 = Val de Marne], [95 = Val d'Oise], [971 = Guadeloupe], [972 = Martinique] and [974 = La Réunion].Number of samples tested and the percentage of positive samples are detailed in Table S3.

model reconstructions when using Euroimmun assays and during the studied period. Moreover, we estimated that the decay of the sensitivity of the anti-S1 and anti-N IgG are not statistically different (Figure S3C), although decay is slightly faster for anti-N IgG assay: 3.9 months (95% CrI: 3.2-4.7) after seroconversion vs 4.7 months (95% CrI:3.8-5.8) for anti-S1 IgG assay (Figure S3C). More rapid waning of anti-N assays than anti-S assays has been reported in the literature.[9] Additionally, to test the robustness of the method we applied our mathematical model to anti-S1 IgG data from throughout France restricted to time periods T3, T4, and T6, and found similar estimates for the decay (half-life of 3.4 months (95% CrI: 3.2-3.6).

DISCUSSION

Our objective was to follow the anti-SARS-CoV-2 IgG prevalence rates at the onset of the outbreak in France and to estimate the seroprevalence before the roll out of vaccines. A total of six sero-epidemiological surveys were conducted from March 2020 to January 2021 using blood donor samples (n = 32,605) collected in several French departments.

We investigated seroprevalence at different time points and we deliberately included areas with high incidence. At the time of the first epidemic peak (March-April 2020), the seroprevalence rates observed in France among blood donors were as high as 10.6% (95%CI 8.30-12.90) (Table 2). Seroprevalence rates observed in blood donors were consistent with the number of cases observed in the general population, which was our original criterion for the selection of the departments to be investigated. These data were close to those observed in the Netherlands at the same period, which showed a strong geographical

Variables:	positive	negative	%age positive	Odds ratio [95% CI]
Female	528	5618	8.6%	1.11 [0.97, 1.27]
Male	458	5412	7.8%	0.90 [0.79, 1.03]
Age 18-30	410	3352	10.9%	1.63 [1.43, 1.86]
Age 31-40	146	2045	6.7%	0.76 [0.64, 0.92]
Age 41-50	181	2294	7.3%	0.86 [0.72, 1.01]
Age 51-60	172	2005	7.9%	0.95 [0.80, 1.13]
Age 61-70	77	1334	5.5%	0.62 [0.48, 0.78]
ABO blood group : A	414	4396	8.6%	1.09 [0.96, 1.25]
ABO blood group : B	103	1074	8.7%	1.08 [0.87, 1.34]
ABO blood group : AB	49	324	13.1%	1.73 [1.27, 2.35]
ABO blood group : O	420	5236	7.4%	0.82 [0.72, 0.94]
All A (A & AB)	463	4720	8.9%	1.18 [1.04, 1.35]
Non A (B & O)	523	6310	7.7%	0.84 [0.74, 0.96]
All B (B & AB)	152	1398	9,8%	1.26 [1.05, 1.51]
Non B (A & O)	834	9632	8.0%	0.80 [0.66, 0.96]
Blood group: O	420	5236	7.4%	0.82 [0.72, 0.94]
Non O (A, B & AB)	566	5794	8.9%	1.22 [1.07, 1.39]
Rhesus positive	812	9039	8.2%	1.03 [0.87, 1.22]
Rhesus negative	174	1991	8.0%	0.97 [0.82, 1.15]
Kell positive	64	887	6.7%	0.79 [0.61, 1.03]
Kell negative	922	10143	8.3%	1.26 [0.97, 1.64]

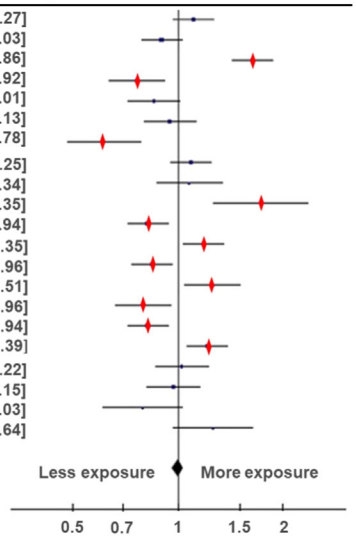


Figure 2. Demographics and risk factors associated with seropositivity

Forest plot of odds ratios and 95% confident interval reporting the association between positive IgG anti-SARS-CoV-2 and blood donation characteristics for the last survey (January/February 2021 - survey time T6). Odd ratio were estimated using the odd ratio calculator (MedCalc "https://www.medcalc.org/calc/odds_ratio.php). Significant differences ($p < 0.05$) are marked in red.

disparity with a maximum of 9.5% in the South of the country.[10] Such disparity has been reported in many blood donor populations regardless of the study period and the serological tests used and can be explained by the history of the introduction of the virus in the different territories studied, the chronology of implementation and the nature of the control measures decided by the health authorities.[10–15] In the early-2020 period, prevalence reported in Europe were well correlated with case data observed in the regions studied, both in areas of low (eg, Germany, South-East Italy, Denmark, Austria: 0.9%-2%) or high (eg, Lombardy in Italy: 19.7%) circulation.[11,13,15–17] In January 2021, at the start of the vaccination campaign, seroprevalence in metropolitan France ranged from 3.1% (95%CI 1.50-4.70) to 13.5% (95%CI 10.34-16.66) depending on the region (Table 2), comparable to that observed in other European blood donor populations: 5.4%-15%, again with regional disparities (Figure 1).[15,18,19]

Monitoring of serological results shows an increase in the proportion of neutralizing antibodies among IgG-positive individuals over time (Figure 3B). This may reflect a maturation of the neutralizing response over time and deserves further immunological characterization.

We also investigated the relationship between seroprevalence and blood groups. Previous studies have reported an association between ABO blood types and the risk of exposure to SARS-CoV-2 or the severity of COVID-19: individuals of group O appear to have a lower susceptibility to infection,[20–22] while those of group A (and in some studies those of groups B and/or AB) are more likely to be infected.[23] In a previous study, we observed a lower seroprevalence of neutralizing anti-SARS-CoV-2 antibodies in blood donors of group O compared to those of other groups (A, B, and AB).[24] These preliminary data are confirmed in the population of donors collected in 2021 (n = 12,016). The proportion of donors carrying anti-SARS-CoV-2 IgG from group O (7.43%; 95%CI 6.74-8.11) is lower ($p = 0.003$) than that of the other groups (8.90%; 95%CI 8.20-9.60) with an odds ratio of 0.82 [0.72-0.94] (Figure 2). Our results are in agreement with data reported in other blood donor surveys.[19,25,26] One previously proposed hypothesis to explain the protective effect observed in blood group O individuals is that the natural anti-A and anti-B antibodies present in group O individuals may limit the interaction between the viral spike protein and the ACE-2 receptor and thus play a role in susceptibility to infection.[27,28]

In addition, we investigated seroprevalence by age group. Among older donors, lower anti-SARS-CoV-2 IgG seroprevalence is frequently reported in the literature.[12,14,15,19,29] In our study, the age group 61-70

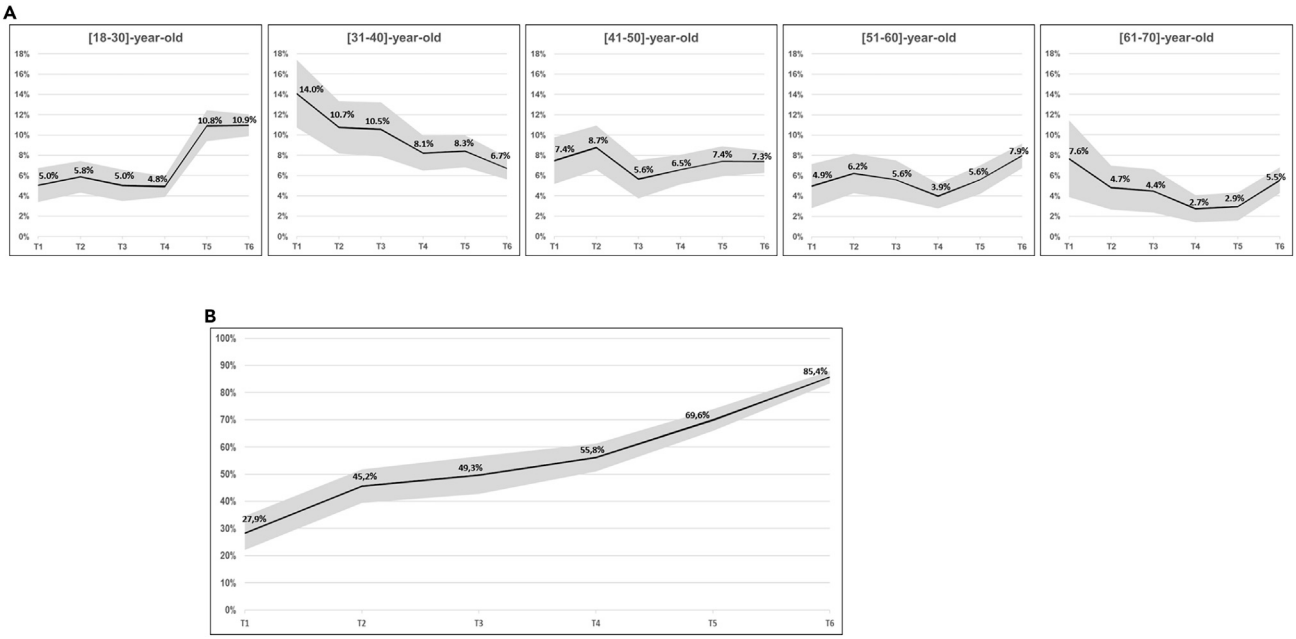


Figure 3. Follow-up of serological markers during the studied period

(A) Distribution of anti-S1 SARS-CoV-2 IgG seroprevalence in age groups during the survey. Data are represented as the percentage of prevalence and the envelopes represent the 95% confident interval. See also Table 1.

(B) Proportion of neutralizing antibodies (titers ≥ 20) in positive samples. Linear regression (dotted line) was calculated as y = 0,1048x + 0,1885 and coefficient of regression was R^2 = 0,9631. Data are represented as a percentage and the envelopes represent the 95% confident interval.

years (Figure 3A) had the lowest prevalence throughout the study, except during the first epidemic wave. In the first samples (March-April 2020), the presence of antibodies reflected early exposure to SARS-CoV-2, before the implementation of specific protective measures, including a strict lockdown between 17 March and 11 May 2020. Subsequently, lower prevalence rates among the elderly reflected adherence of this group to prevention measures implemented to protect populations most vulnerable to COVID-19. In those under 30 years old, we observed an increase in seroprevalence in the autumn of 2020. This reflects increased exposure during the summer and early autumn, which was probably linked to a relaxation of preventive measures during the summer and which occurred in a context of low detection of incident cases due to the frequency of a- or pauci-symptomatic forms in the youngest age group.

We extended our previous approach which used a mathematical model, to reconstruct the number of infections from hospital surveillance data and serological data.[5] This latter approach was based on a serological study from May 2020 where the waning of immunity was not observable. Here we accounted for the decay in the ability to detect IgG anti-S1 antibody by the serological assay used in our study. Sampling for the serological study performed at different time points—and during time period long enough—allowed us assessing the antibody decay. Our approach is similar to that of Chen et al.[30] who reconstructed the number of infections using serological surveys in blood donors across England and used mortality data to estimate the infection fatality rate. However, since mortality is low in the younger age groups we used hospitalization data. We assumed that the probability of hospitalization was stable during the period under review. Relaxing this assumption may lead to different parameter estimates.

Our analysis provided the proportion of infected adults by region and age group. Hospital admissions vary greatly between departments, partly because individuals may go to hospitals in the neighboring departments. In the hospitalization SI-VIC database, we were able to access the department where individuals were admitted to hospital but not their place of residence. By aggregating by region and not by department, we reduced this bias. We reported a very high rate of infections in the Grand-Est region, but the serological surveys were mainly performed in the Haut-Rhin department, which may not be representative of the virus circulation in the whole region. Otherwise, our data were consistent with the predictions made by SAPRIS, another large-scale serosurvey of the general population (Figure S4).[5] The lower proportion

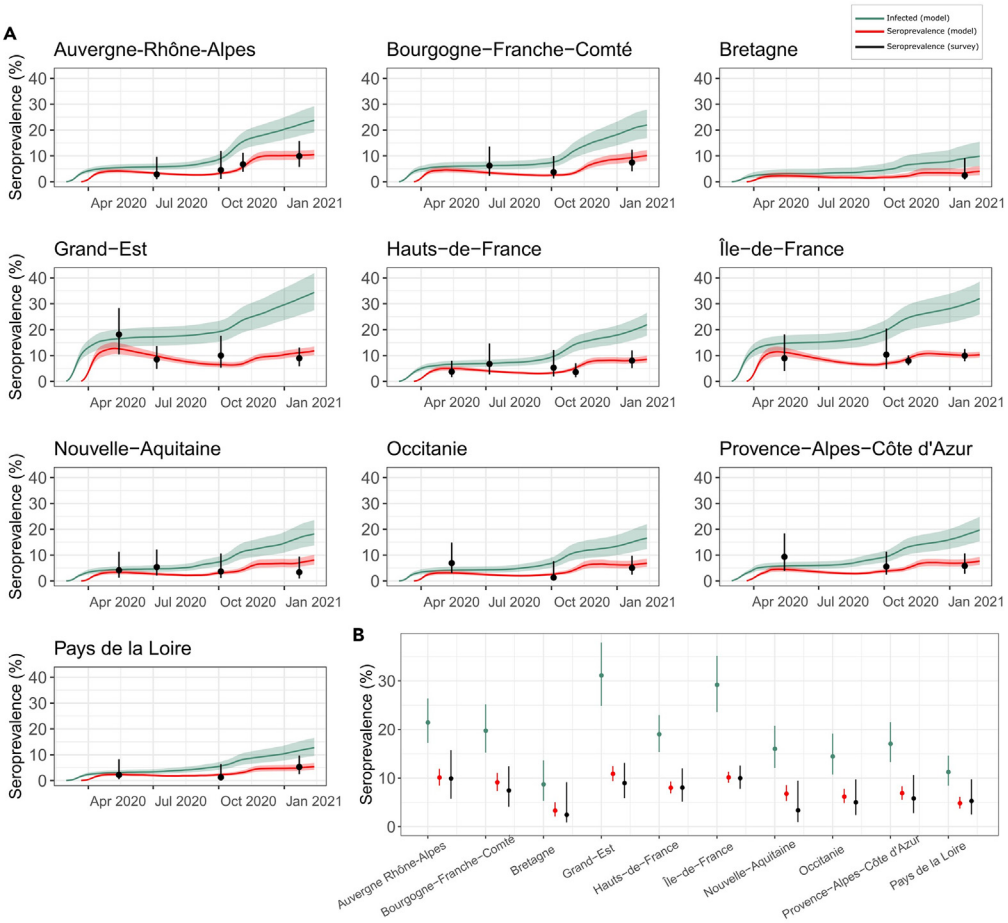


Figure 4. Model-based reconstruction of the cumulative proportion of adults infected

(A) Proportion of the population infected by SARS-CoV-2 (green) and reconstructed seroprevalence (red) in ten out of thirteen regions of metropolitan France. The solid lines and envelopes represent the mean posterior and 95% credible intervals, respectively. The dots and bars are the observed seroprevalence and 95% binomial confidence intervals, aggregated for the different age groups and departments within a region.

(B) Comparison of the proportion of the adult population infected on Jan, 15, 2021 (green), the corresponding model-based estimate of the seroprevalence (red), and the seroprevalence in the latest sampling January-February 2021 (black). For each administrative region (in italic), the following departments were included in the serostudy: *Auvergne-Rhône-Alpes*: Loire and Rhône ; *Bourgogne-Franche-Comté*: Côte d'Or and Doubs; *Bretagne*: Ille-et-Vilaine ; *Grand-Est*: Meurthe-et-Moselle, Bas-Rhin and Haut-Rhin ; *Hauts-de-France* : Nord, Oise and Pas de Calais; *Ile-de-France* : Paris, Seine-et-Marne, Yvelines, Essonne, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne, Val-d'Oise; *Nouvelle-Aquitaine*: Gironde; *Occitanie* : Haute-Garonne and Hérault, *Provence-Alpes-Côte d'Azur*: Bouches-du-Rhône and Alpes-Maritimes, *Pays de la Loire*: Loire-Atlantique and Mayenne. The dots and bars are the prevalence and 95% binomial confidence intervals. See also Figures S3 and S4.

of infected individuals in the SAPRIS survey can be explained by the fact that all adults were included, including those aged 70 and over who had a lower infection rate.

Overall, the use of this model has shown a high degree of consistency between the hospitalization data observed at a given time in the general population and the results of seroprevalence studies in blood donors conducted several weeks later. A similar finding was reported in a seroprevalence study of SARS-CoV-2 in adult donors, where the results were comparable to those obtained in household surveys targeting the general population.[31] Another advantage of the model is that it is possible to estimate anti- SARS-CoV-2 IgG seroprevalence rates from hospitalization data for territories where serological surveys have not been conducted. The use of this model, therefore, makes it possible to provide health authorities with rapid information on the proportion of the adult population that has been in contact with the virus.

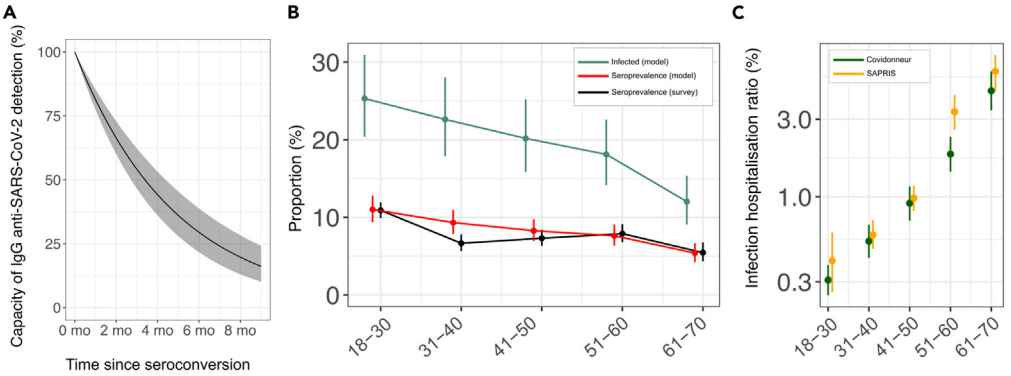


Figure 5. Estimates of the model parameters

(A) Capacity of the detection of the Euroimmun IgG anti-S1 assay as a function of time since seroconversion. In the model, we assume a 100% sensitivity at the time of seroconversion (21 days after infection). The solid line represents the mean and the envelope the 95% credible interval of the posterior distribution See also Figure S2.

(B) Comparison of the estimated cumulative exposure (green), the estimated seroprevalence accounting for antibody decay (red), and observed seroprevalence (black) for the different age groups at the final sampling time (January 2021). The dots and bars are the mean posterior and 95% credible interval for the model-estimated exposure and seroprevalence and the observed seroprevalence and 95% binomial confidence interval, respectively. See also Table S1.

(C) Model estimates of the infection-hospitalization ratio. The values of the IHR are represented on a log scale. The dots and bars represent the mean and the 95% credible interval of the posterior distribution (in green). We also represent the IHR obtained from a different serological survey (SAPRIS study, in orange). See also Table S2.

The impact of seroreversion or loss of IgG detection capacity is difficult to consider at any given time. Indeed, regions have been diversely affected by varying chronology of viral circulation and intensity over time. However, the model is able to mitigate these variations and provides an estimate of the decay time of IgG-S1 antibody detection by the Euroimmun test. The modeling data show that approximately 50% of individuals who are IgG-S1 positive have no detectable antibodies four months after infection (Figure 5A). In a study similar to ours conducted before the vaccination campaign in Italy (July 2020 to February 2021), the authors report that about 50% of anti-RBD IgG-positive individuals became negative after a median follow-up of 18 weeks (4.5 months).[19]

In conclusion, in this study, we showcased insights we could get from SARS-CoV-2 seroprevalence studies performed on blood donors in France. We observed differences in anti-SARS-CoV-2 IgG seroprevalence according to geographical distribution and over time. The seroprevalence in France before the start of the vaccination campaign was about 15%. The study showed that older people, who had the lowest seroprevalence rates, had complied with prevention measures. It identified an increase in exposure in the under-30s during the summer of 2020 that had gone unnoticed due to the high proportion of asymptomatic or paucisymptomatic infections in this population. The study of blood group distribution in anti-S1 IgG-positive individuals showed a slightly lower exposure to SARS-CoV-2 infection in group O donors compared to non-O donors. Over the course of the study, the proportion of neutralizing antibodies in IgG-positive samples increased, suggesting a maturation of the immune response. The use of a mathematical model showing a high degree of consistency in the different regions studied between general population hospitalization data and the results of seroprevalence studies in blood donors allows for rapid and extemporaneous quality estimates of the level of immunity acquired by the general population to be provided to health authorities.

Limitations of the study

Our study has several limitations: (i) the geographical areas investigated were chosen as representative of regions with different intensities of viral circulation and the study was not performed for the whole territory. Although the magnitude and dynamics of prevalence in each area are robust and of the same order of magnitude as data observed in other European countries, the study was not designed to provide an accurate overall seroprevalence. (ii) The blood donor population does not allow exploration of seroprevalence in the <18 and >70 age groups. (iii) Blood pre-donation selection measures exclude individuals with recent clinical signs suggestive of COVID-19 and those with persistent sequelae that may lead to an underestimation of rates. (iv) Individuals may be treated in hospitals in neighboring departments. In the SIVIC database,

Table 2. Weighted IgG anti-SARS-CoV-2 prevalence in studied areas

Departments: N° (name)	T1	T2	T3	T4	T5	T6
06 (Alpes Maritimes)						5.8%
13 (Bouches du Rhône)	5.6%	8.2%		5.7%		5.6%
21 (Cote d'Or)				4.1%		6.6%
25 (Doubs)			6.2%			8.2%
31 (Haute Garonne)		6.9%				6.9%
33 (Gironde)	5.9%	4.6%	5.0%	4.1%		3.3%
34 (Hérault)				1.4%		5.7%
35 (Ille et Vilaine)						3.6%
42 (Loire)					6.0%	8.8%
44 (Loire Atlantique)		2.4%				3.1%
53 (Mayenne)				1.4%		7.8%
54 (Meurthe et Moselle)			5.4%			8.4%
59 (Nord)		3.3%			3.4%	12.1%
60 (Oise)	7.8%	4.6%	7.3%	5.4%		8.0%
62 (Pas de Calais)					4.0%	6.4%
67 (Bas Rhin)						10.5%
68 (Haut Rhin)	10.6%	17.9%	11.3%	11.0%		9.3%
69 (Rhône)			3.6%	3.8%	7.8%	10.7%
75 (Paris)				10.8%	10.4%	11.6%
77 (Seine et Marne)					4.7%	12.1%
78 (Yvelines)					9.9%	9.7%
91 (Essonne)					8.4%	13.5%
92 (Hauts de Seine)					8.3%	7.6%
93 (Seine Saint-Denis)	6.4%				11.1%	12.7%
94 (Val de Marne)					9.1%	9.7%
95 (Val d'Oise)		10.0%			11.6%	8.3%

Weighted prevalence of SARS-CoV-2 anti-S1 IgG antibodies after correction for the number of residents and for the age distribution in the general population (26 departments in mainland France) using data from the Institut National de la Statistique et des Etudes Economiques (INSEE): https://www.insee.fr/fr/statistiques/fichier/1893198/estim-pop-dep-sexe-aq-1975-2022.xlsx

we have the department where individuals are admitted to the hospital but not their place of residence. In our study, the mathematical model allows us to determine the proportion of infected adults by region and age group. Aggregating by region and not by department reduced this bias.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Study design and participants
- METHOD DETAILS
  - Anti-SARS-CoV-2 IgG detection
  - Anti SARS-CoV-2 neutralizing antibodies titration
  - Blood grouping

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis
  - Hospitalization data
  - Reconstructing infection and seroprevalence
  - Deconvolution and infection curve reconstruction

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106222.

ACKNOWLEDGMENTS

We are grateful to blood donors who participated in the study. We thank the regional French National Blood Service directors and their field teams for their highly contributive participation in blood collection: ETS Auvergne-Rhône-Alpes : Dr D. Legrand, Dr J. Courchelle, Dr V. Barlet; ETS Bourgogne-Franche-Comté : Dr C. Besiers; Dr C. Barisien; ETS Bretagne : Dr B. Danic, Dr A. Guillard; ETS Centre-Pays-de-la-Loire : Dr. F. Bigey, Dr C. Leclerc, Dr S. Le Cam; ETS Grand Est : Dr C. Gachet, Dr C. Claudel; ETS Guadeloupe-Guyane & ETS Martinique : Dr F. Maire, Dr M. Bordenet, Dr D. Arnaudova; ETS Hauts-de-France-Normandie : Dr R. Courbil, Dr N. Delemer, Dr C. Narboux; ETS Ile-de-France : Mr S. Noel, Dr A. Slimani; ETS Nouvelle-Aquitaine : Dr M. Jeanne, Dr F. Lassurguere; ETS Occitanie : Dr L. Bardiaux; Dr P. Lambert, Dr C. Maugard; ETS Provence-Alpes-Côte-d'Azur-Corse: Prof. J. Chiaroni, Dr. C. Lazaygues; ETS La-Réunion-Océan-Indien : Dr I. Delouane, Dr A. Le Tallec. We thank B. Bonneaudeau, A. Michaut, and K. Bornarel for their support and expert advice concerning ethical and juridic aspects.

This study was performed with the financial support of the Agence Nationale de la Recherche (ANR-20-COVI-0073-01, France), the Fondation pour la Recherche Médicale (Grant number: REA 202010012517, France) and the ANRS Maladie Infectieuses Émergentes (ANRS-MIE). SC acknowledges financial support from the Investissement d'Avenir program, the Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases program (grant ANR-10-LABX-62-IBEID); the INCEPTION project (grant PIA/ANR-16-CONV-0005); the European Union's Horizon 2020 research and innovation program (grants RECOVER 101003589 and VEO 874735); AXA; and Groupama.

AUTHOR CONTRIBUTIONS

X.d.L., P.M., and P.G. designed the study. N.B., C.I., P.M.S.V., E.N., and B.P. performed the sample acquisition, laboratory testing (Elisa and/or seroneutralization assays), and data analysis. P.G. and X.d.L. interpreted the virological findings. N.H. and S.C. made and interpreted the modeling studies. P.G. and N.H. drafted the article. P.M., P.R., and X.d.L. and S.C. reviewed the article. All co-authors contributed to the editing and final drafting of the article and figures. All authors approve the content and submission of the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 7, 2022
Revised: December 20, 2022
Accepted: February 11, 2023
Published: February 15, 2023

REFERENCES

1. (2022). COVID-19 dashboard by the center for systems science and engineering (CSSE) at Johns Hopkins university (JHU). https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.

2. Mizumoto, K., Kagaya, K., Zarebski, A., and Chowell, G. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship. Euro Surveill. *25*, 2000180. https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180.

3. Jeffery-Smith, A., Rowland, T.A.J., Patel, M., Whitaker, H., Iyanger, N., Williams, S.V., Giddings, R., Thompson, L., Zavala, M., Aiano, F., et al. (2021). Reinfection with new variants of SARS-CoV-2 after natural infection: a prospective observational cohort in 13 care homes in England. Lancet. Healthy Longev. *2*, e811–e819. https://doi.org/10.1016/S2666-7568(21)00253-1.

4. Pelleau, S., Woudenberg, T., Rosado, J., Donnadieu, F., Garcia, L., Obadia, T., Gardais, S., Elgharbawy, Y., Velay, A., Gonzalez, M., et al. (2021). Kinetics of the severe acute respiratory syndrome coronavirus 2 antibody response and serological estimation of time since infection. J. Infect. Dis. *224*, 1489–1499. https://doi.org/10.1093/infdis/jiab375.

5. Hozé, N., Paireau, J., Lapidus, N., Tran Kiem, C., Salje, H., Severi, G., Touvier, M., Zins, M.,

de Lamballerie, X., Lévy-Bruhl, D., et al. (2021). Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: a modelling study. Lancet Public Health *6*, e408–e415. https://doi.org/10.1016/S2468-2667(21)00064-5.

6. Warszawski, J., Bajos, N., Meyer, L., de Lamballerie, X., Seng, R., Beaumont, A.L., Slama, R., Hisbergues, M., Rahib, D., Lydié, N., et al. (2020). En mai 2020, 4,5 % de la population en France métropolitaine a développé des anticorps contre le SARS-CoV-2. Etudes et Résultats. Direction de la Recherche, des études, de l'évaluation et des statistiques *1167*, 1–6. https://drees.solidarites-sante.gouv.fr/sites/default/files/2020-10/er1167.pdf.

7. Carrat, F., de Lamballerie, X., Rahib, D., Blanché, H., Lapidus, N., Artaud, F., Kab, S., Renuy, A., Szabo de Edelenyi, F., Meyer, L., et al. (2021). Antibody status and cumulative incidence of SARS-CoV-2 infection among adults in three regions of France following the first lockdown and associated risk factors: a multicohort study. Int. J. Epidemiol. *50*, 1458–1472. https://doi.org/10.1093/ije/dyab110.

8. Lapidus, N., Paireau, J., Levy-Bruhl, D., de Lamballerie, X., Severi, G., Touvier, M., Zins, M., Cauchemez, S., and Carrat, F.; SAPRIS-SERO study group (2021). Do not neglect SARS-CoV-2 hospitalization and fatality risks in the middle-aged adult population. Infect. Dis. Now *51*, 380–382. https://doi.org/10.1016/j.idnow.2020.12.007.

9. Stone, M., Grebe, E., Sulaeman, H., Di Germanio, C., Dave, H., Kelly, K., Biggerstaff, B.J., Crews, B.O., Tran, N., Jerome, K.R., et al. (2022). Evaluation of commercially available high-throughput SARS-CoV-2 serologic assays for serosurveillance and related applications. Emerg. Infect. Dis. *28*, 672–683. https://doi.org/10.3201/eid2803.211885.

10. Slot, E., Hogema, B.M., Reusken, C.B.E.M., Reimerink, J.H., Molier, M., Karregat, J.H.M., IJlst, J., Novotný, V.M.J., van Lier, R.A.W., and Zaaijer, H.L. (2020). Low SARS-CoV-2 seroprevalence in blood donors in the early COVID-19 epidemic in The Netherlands. Nat. Commun. *11*, 5744. https://doi.org/10.1038/s41467-020-19481-7.

11. Fischer, B., Knabbe, C., and Vollmer, T. (2020). SARS-CoV-2 IgG seroprevalence in blood donors located in three different federal states, Germany, March to June 2020. Euro Surveill. *25*, 2001285. https://doi.org/10.2807/1560-7917.ES.2020.25.28.2001285.

12. Jones, J.M., Stone, M., Sulaeman, H., Fink, R.V., Dave, H., Levy, M.E., Di Germanio, C., Green, V., Notari, E., Saa, P., et al. (2021). Estimated US infection- and vaccine-induced SARS-CoV-2 seroprevalence based on blood donations. JAMA *326*, 1400–1409. https://doi.org/10.1001/jama.2021.15161.

13. Kaspersen, K.A., Hindhede, L., Boldsen, J.K., Mikkelsen, S., Vestergaard, L.S., Berthelsen, A.S.N., Moustsen-Helms, I.R., Holm, D.K., Nilsson, A.C., Sækmose, S.G., et al. (2022). Estimation of SARS-CoV-2 infection fatality rate by age and comorbidity status using antibody screening of blood donors during the COVID-19 epidemic in Denmark. J. Infect. Dis. *225*, 219–228. https://doi.org/10.1093/infdis/jiab566.

14. Stone, M., Di Germanio, C., Wright, D.J., Sulaeman, H., Dave, H., Fink, R.V., Notari, E.P., Green, V., Strauss, D., Kessler, D., et al. (2022). Use of US blood donors for national serosurveillance of severe acute respiratory syndrome coronavirus 2 antibodies: basis for an expanded national donor serosurveillance program. Clin. Infect. Dis. *74*, 871–881. https://doi.org/10.1093/cid/ciab537.

15. Weidner, L., Nunhofer, V., Jungbauer, C., Hoeggerl, A.D., Grüner, L., Grabmer, C., Zimmermann, G., Rohde, E., and Laner-Plamberger, S. (2021). Seroprevalence of anti-SARS-CoV-2 total antibody is higher in younger Austrian blood donors. Infection *49*, 1187–1194. https://doi.org/10.1007/s15010-021-01639-0.

16. Cassaniti, I., Percivalle, E., Sarasini, A., Cambiè, G., Vecchio Nepita, E., Maserati, R., Ferrari, A., Corcione, A., Di Martino, R., Bonetti, A., et al. (2021). Seroprevalence of SARS-CoV-2 in blood donors from the Lodi Red Zone and adjacent Lodi metropolitan and suburban area. Clin. Microbiol. Infect. *27*, 914.e1–914.e4. https://doi.org/10.1016/j.cmi.2021.01.030.

17. Di Stefano, M., Sarno, M., Faleo, G., Farhan Mohamed, A.M., Lipsi, M.R., De Nittis, R., Bruno, S.R., De Feo, L., Granato, T., Corso, G., et al. (2021). Low prevalence of antibodies to SARS-CoV-2 and undetectable viral load in seropositive blood donors from south-eastern Italy. Acta Haematol. *144*, 580–584. https://doi.org/10.1159/000515258.

18. Castro Dopico, X., Muschiol, S., Christian, M., Hanke, L., Sheward, D.J., Grinberg, N.F., Rorbach, J., Bogdanovic, G., Mcinerney, G.M., Allander, T., et al. (2021). Seropositivity in blood donors and pregnant women during the first year of SARS-CoV-2 transmission in Stockholm, Sweden. J. Intern. Med. *290*, 666–676. https://doi.org/10.1111/joim.13304.

19. Valenti, L., Pelusi, S., Cherubini, A., Bianco, C., Ronzoni, L., Uceda Renteria, S., Coluccio, E., Berzuini, A., Lombardi, A., Terranova, L., et al. (2021). Trends and risk factors of SARS-CoV-2 infection in asymptomatic blood donors. Transfusion *61*, 3381–3389. https://doi.org/10.1111/trf.16693.

20. Cai, H. (2020). Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir. Med. *8*, e20. https://doi.org/10.1016/S2213-2600(20)30117-X.

21. Cheng, Y., Cheng, G., Chui, C.H., Lau, F.Y., Chan, P.K.S., Ng, M.H.L., Sung, J.J.Y., and Wong, R.S.M. (2005). ABO blood group and susceptibility to severe acute respiratory syndrome. JAMA *293*, 1450–1451. https://doi.org/10.1001/jama.293.12.1450-c.

22. Dai, X. (2020). ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. Eur. J. Prev. Cardiol. *27*, 1436–1437. https://doi.org/10.1177/2047487320922370.

23. Pendu, J.L., Breiman, A., Rocher, J., Dion, M., and Ruvoën-Clouet, N. (2021). ABO blood types and COVID-19: spurious, anecdotal, or truly important relationships? A reasoned review of available data. Viruses *13*, 160. https://doi.org/10.3390/v13020160.

24. Gallian, P., Pastorino, B., Morel, P., Chiaroni, J., Ninove, L., and de Lamballerie, X. (2020). Lower prevalence of antibodies neutralizing SARS-CoV-2 in group O French blood donors. Antiviral Res. *181*, 104880. https://doi.org/10.1016/j.antiviral.2020.104880.

25. Butler, D., Coyne, D., Pomeroy, L., Williams, P., Holder, P., Carterson, A., Field, S., Waters, A., and O'Flaherty, N. (2022). Confirmed circulation of SARS-CoV-2 in Irish blood donors prior to first national notification of infection. J. Clin. Virol. *146*, 105045. https://doi.org/10.1016/j.jcv.2021.105045.

26. Vassallo, R.R., Dumont, L.J., Bravo, M.D., Hazegh, K., and Kamel, H. (2021). Progression and predictors of SARS-CoV-2 antibody seroreactivity in US blood donors. Transfus. Med. Rev. *35*, 8–15. https://doi.org/10.1016/j.tmrv.2021.07.003.

27. Deleers, M., Breiman, A., Daubie, V., Maggetto, C., Barreau, I., Besse, T., Clémenceau, B., Ruvoën-Clouet, N., Fils, J.F., Maillart, E., et al. (2021). Covid-19 and blood groups: ABO antibody levels may also matter. Int. J. Infect. Dis. *104*, 242–249. https://doi.org/10.1016/j.ijid.2020.12.025.

28. Zhao, J., Yang, Y., Huang, H., Li, D., Gu, D., Lu, X., Zhang, Z., Liu, L., Liu, T., Liu, Y., et al. (2021). Relationship between the ABO blood group and the coronavirus disease 2019 (COVID-19) susceptibility. Clin. Infect. Dis. *73*, 328–331. https://doi.org/10.1093/cid/ciaa1150.

29. Dodd, R.Y., Xu, M., and Stramer, S.L. (2020). Change in donor characteristics and antibodies to SARS-CoV-2 in donated blood in the US, June-August 2020. JAMA *324*, 1677–1679. https://doi.org/10.1001/jama.2020.18598.

30. Chen, S., Flegg, J.A., White, L.J., and Aguas, R. (2021). Levels of SARS-CoV-2 population exposure are considerably higher than suggested by seroprevalence surveys. PLoS Comput. Biol. *17*, e1009436. https://doi.org/10.1371/journal.pcbi.1009436.

31. Grant, R., Dub, T., Andrianou, X., Nohynek, H., Wilder-Smith, A., Pezzotti, P., and Fontanet, A. (2021). SARS-CoV-2 population-based seroprevalence studies in Europe: a scoping review. BMJ Open *11*, e045425. https://doi.org/10.1136/bmjopen-2020-045425.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Asymptomatic blood donors	France, this paper	Table 1
Critical commercial assays		
ELISA SARS-CoV-2 IgG assay	Euroimmun	Cat# EI 2606-9601 G
ELISA Anti-SARS-CoV-2 NCP (IgG)	Euroimmun	Cat# EI 2606-9601-2 G
Bacterial and virus strains		
D614G B.1 SARS CoV-2 European strain	courtesy of Pr C Drosten	BetaCoV/Germany/BavPat1/2020
Experimental models: Cell lines		
Vero E6/TMPRSS2	National Institute for Biological Standards and Control	Cat#100978
Chemicals, peptides, and recombinant proteins		
Dulbecco's minimal essential medium	Gibco	Cat#21090-022
Penicillin + streptomycin	Gibco	Cat#15070-063
Deposited data		
SI-VIC database	Agence du Numérique en Santé	https://www.data.gouv.fr/fr/datasets/r/08c18e08-6780-452d-9b8c-ae244ad529b3
Demographics by 5-year age groups in French administrative departments and regions	Institut national de la statistique et des études économiques	https://www.insee.fr/fr/statistiques/1893198
Software and algorithms		
Chi-Square Test Calculator	Social Science Statistics	https://www.socscistatistics.com/tests/chisquare/default2.aspx
Free statistical calculator, Odd ratio calculator	MedCalc	https://www.medcalc.org/calc/odds_ratio.php
IBM-SPSS Statistics	v 24.0.0.0 software	N/A
Rstan (Markov Chain Monte Carlo)	R package version 2.19.3	http://mc-stan.org [2020]
Custom made R scripts	R version 4.1.2	https://doi.org/10.5281/zenodo.7544548

RESOURCE AVAILABILITY

Lead contact

Further information and requests for ressources should be directed to and will be fulfilled by the lead contact, Pierre Gallian pierre.gallian@efs.sante.fr.

Materials availability

This study did not generate new unique materials.

Data and code availability

The paper used aggregated data and individual participant data will not be made available. Data are provided in a format that maintains anonymity of the participants. For each individual, we provide the result of the anti-S1 and anti-N IgG assay when available, the age group (aggregated in groups 18–30, 31–40, 41–50, 51–60, 61–70), department, survey period and median date of the survey. Code and data are available on https://doi.org/10.5281/zenodo.7544548.

Any additional information required to reanalyze the date reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study design and participants

This is a prospective repeated cross-sectional study for SARS-CoV-2 anti-S1 IgG prevalence in French blood donor population. Sampling was performed at six different times (T1 = end March–early April, T2 = May, T3 = July, T4 = October, T5 = November 2020 and T6 = mid-January–early February 2021) during the epidemic period prior to the implementation of the vaccination campaign and the circulation of antigenic variants (Figure S1).

Voluntary unpaid blood donors (18–70 years old) were recruited according to French regulatory requirements. The pre-donation medical questionnaire included the reporting of influenza-like clinical symptoms. During the COVID-19 epidemic, donors reporting a history of clinical and/or biological diagnosis or suspicion of SARS-CoV-2 infection were deferred for 28 days after recovery. Only data authorized to be recorded in the national blood bank software and pertinent with the aim of the study were analyzed: sex, age, date of donation and department of residence.

Our study was carried out on donors living in mainland France and in three overseas administrative areas: Guadeloupe and Martinique in the Americas and La Réunion in South-West Indian Ocean. The French territory is divided into administrative areas called 'departments', corresponding to NUTS3 level (96 on the mainland and five overseas). Geographical distribution of anti-SARS-CoV-2 seropositivity was assessed at the departmental level. When reconstructing the dynamics of infection with a model, we aggregated the serological data at the regional level (NUTS1 level).

The evolution of the epidemic was systematically monitored in three departments, two of which were selected because significant viral circulation was observed at the start of the epidemic (Oise (60) in Hauts-de-France and Haut Rhin (68) in eastern France and one with low viral circulation (Gironde (33)) in South West France. In these three departments, samples were tested for presence of both anti-S1 and anti-N antibodies at survey's times, T3 (n=1298), T4 (n=1338) and T6 (n=1329). In order to allow for a follow-up, during the last epidemiological survey (T6: January/February 2021), most departments previously studied in mainland France (n=24) were tested.

Donors recruited in the study were assigned to one of five age groups (18–30, 31–40, 41–50, 51–60, and 61–70). The information made available for analysis included date of donation, sex, age, and results for ABO, Rhesus, and Kell blood phenotyping.

Effective, sex and age repartition of blood donors for each studied period are summarized in Table 1.

Only volunteer blood donors were included. They were informed that samples might be used for epidemiological studies and more specifically for the current study. The study was approved by the scientific direction of the EFS and an ethic committee (CPP (Comité de Protection des Personnes) Ouest 6-CPP 1286/MS1, N°IDRCB: 2020-A01200-39). All data used for epidemiological studies were de-identified.

METHOD DETAILS

Anti-SARS-CoV-2 IgG detection

ELISA detection of SARS-CoV-2 specific IgG was performed using the ELISA SARS-CoV-2 IgG assay (Euroimmun, Lubeck, Germany) targeting the S1 sub-unit of the spike protein that include the RBD (receptor binding domain) region and the ELISA Anti-SARS-CoV-2 NCP (IgG) assay (Euroimmun, Lubeck, Germany) targeting the nucleocapsid antigen. Serology tests were performed according to manufacturer's instruction. For each sample, we calculated the ratio sample OD/cut-off OD. A result was considered positive if the sample ratio was $\geq$ 1.1.

Anti SARS-CoV-2 neutralizing antibodies titration

All reactive samples with ELISA positive or indeterminate results were tested using a virus neutralization test (VNT) . This VNT consisted of (i) bringing into contact a serum sample dilution with a fixed amount of a given variant of SARS-CoV-2, (ii) inoculating this mixture onto a cell culture (for 5 days at 37°C under 5% CO2) and (iii) reading cytophatic effect (CPE) at day 5. Dilutions of samples were performed in Dulbecco's minimal essential medium with 100 I.U./mL penicillin and 100 μg/mL streptomycin. Dilutions tested

were 20, 40, 80 and 160. The contact step, during which specific antibodies are fixed on infectious virus particles and can subsequently inhibit virus growth, was performed under a 220 µL volume for one hour at 37°C using a D614G B.1 European strain (BetaCoV/Germany/BavPat1/2020, courtesy of Pr C Drosten). It corresponded to 0.5 TCID50 per µL of final serum dilution. 100µL of this mixture were inoculated onto Vero E6 TMPRSS2+ cells in a 96-well microplate format. The test was automated in a NSB3 laboratory for all dilution and dispensing steps and for CPE reading (the presence or absence of a cytopathic effect was evaluated by an artificial intelligence program after training on manually assessed image banks). The test was used in a duplicate format. Neutralisation titers >10 were considered positive.

Blood grouping

ABO/Rh/Kell phenotypes were determined as a part of the routine blood bank procedures, not as a research procedure.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Association of the presence of anti- SARS-CoV-2 IgG antibodies with biological factors (blood groups) and demographic data (sex, age, place of residence) was analyzed using the Chi-square test ("Chi-Square Test Calculator" https://www.socscistatistics.com/tests/chisquare/default2.aspx) and the odd ratio calculator (MedCalc "https://www.medcalc.org/calc/odds_ratio.php). Results were considered statistically significant when p-value was lower than 0.05. Univariate and multivariate analysis were performed using the IBM-SPSS Statistics v 24.0.0.0 software (Chicago, IL, USA).

Hospitalization data

The number of hospitalized patients was obtained from the SI-VIC database (système d'information pour le suivi des victimes - patient monitoring information system), used to monitor hospitalizations in the event of exceptional health situations. This database is maintained by the French ANS (Agence du Numérique en Santé) and provides real time information on COVID-19 patients hospitalized in public and private French hospitals. Data, including age, hospitalization date, outcome and region, are sent daily to Santé publique France, the French national public health agency. All COVID-19 cases registered are either biologically confirmed or present with a computed tomographic image highly suggestive of SARS-CoV-2 infection.

Reconstructing infection and seroprevalence

We designed a statistical framework that combines the observed seroprevalence at different time points with a model of antibody decay to retrieve the total number of infections. Using the number of hospital admissions, we reconstructed the dynamics of the infection using a deconvolution approach, assuming an infection-to-hospitalization delay of 11 days. Since only a fraction of the infections result in hospitalization, we introduced a probability of hospitalization to obtain an estimate of the number of infections. The vast majority of infected individuals (with severe as well as asymptomatic forms of the disease) have developed IgG approximately 21 days post infection. We used the cumulative number of infections divided by the total adult population within a region as a measure of the proportion infected. A model of exponential decay of the sensitivity of the assay was then applied to explain the observed seroprevalence at the different time points.

More specifically, we denote x_i and n_i the number of seropositive and number of samples within group i with shared characteristics (at a given date, age class, location). The infection-to-hospitalization delay distribution is modelled with a gamma distribution (mean = 11 and sd = 3.3). The number of hospitalizations on day t is given by the convolution of the delay with incident infections

$$H(t) = p_H \sum_{j=0}^{Tmax} I(t-j)D(j),$$

where p_H is the hospitalization probability, $D(j)$ is the probability that infection-to-hospitalization delay is j days, and $Tmax$ is 400 days. This relation is inverted to reconstruct the daily number of infections $I(t)$ using the deconvolution procedure that we describe below.

Since antibody levels decrease over time, we modelled the expected seroprevalence as the cumulative number of infections modulated by the time since seroconversion (conservatively estimated at 21 days

post-infection). We assumed an exponential decay of the sensitivity of the serological assay. The seroprevalence on day t is given by

$$s(t) = \frac{1}{N}\sum_{j=0}^{Tmax} I(t - 21 - j)e^{-\delta j},$$

where N is the population size. The likelihood function is obtained over all locations, age groups, and sampling periods:

$$L = \prod_i f(x_i|\, s_i.n_i\,),$$

where s_i is the seroprevalence for group i on the day corresponding to the median date of the blood collection for the given sampling period (see Table 1) and $f(.|\lambda)$ is a Poisson distribution of mean λ. In addition, we assumed regional differences in the probability of hospitalization.

A Bayesian approach was chosen to estimate the model parameters. In a Bayesian inference framework, parameters of the model are treated as random variables and the uncertainty in the parameter estimation is treated explicitly. Data and prior knowledge about the range of possible parameter values are fed into the inference framework to provide a range of credible possible values. Bayes' theorem expresses that the conditional probability of a parameter value given observed data (the posterior distribution) is proportional to the probability of this data given the parameters (the likelihood) and to prior belief on the values of the parameters (the prior).

Markov Chain Monte Carlo (MCMC) methods are employed to numerically approximate the posterior distribution, by sampling from the posterior distribution and exploring the parameter space. We used such approach to estimate jointly the probability of hospitalization upon infection p_H for age groups 18–30, 31–40, 41–50, 51–60, 61–70 as well as the dynamics of the sensitivity of the antibody assay expressed by the parameter δ. The MCMC algorithm was the no-U-turn sampler, the adaptive variant of the Hamiltonian Monte Carlo algorithm, and was implemented in Rstan (Stan Development Team. RStan: the R interface to Stan. R package version 2.19.3. http://mc-stan.org [2020]).

We run four chains of 10,000 iterations each and remove 50% for burn-in. Flat priors were chosen for all parameters. We used 2.5% and 97.5% percentiles from the resulting posterior distributions for 95% credible intervals (CrI) for the parameters.

Due to a number of uncertainties, we applied the model to data from all sampling periods except for the first (T1) that happened during the first wave in early April 2020. Assuming a fixed 21-day delay between infection and seroconversion, and an 11-day delay between infection and hospitalizations, this sampling date corresponds to early hospitalizations during the first epidemic wave. At this time, dating and reporting of hospital admissions was not yet stabilized.

Deconvolution and infection curve reconstruction

We used a deconvolution approach to recover the unobserved curve of the daily number of infections from joint analysis of the daily number of hospitalizations and the infection-to-hospitalization delay distribution.

Denote $(\lambda_1, ..., \lambda_N)$ and $(H_1, ..., H_N)$ the number of infections and hospitalizations, respectively, on days $1,...,N$. Denote $(d_1, ..., d_n)$ the infection-to-hospitalisation delay distribution, with d_j the probability that an infected individual is admitted to hospital j days after infection, and $\sum_{j=1}^{n} d_i = 1$. We set $n = 40$ and $d_i = 0$ if $i \leq 0$ or $i > n$. Denote p the probability that an infected individual is hospitalized.

The expected number of hospitalizations on day i is given by the convolution

$$H_i = p\sum_{j<i} d_{i-j}\lambda_j.$$

The deconvolution procedure reconstructs the daily number of infections iteratively by using an expectation maximization algorithm. The algorithm starts from an initial guess $(\lambda_1{}^0, ..., \lambda_N{}^0)$ and the incidence vector is updated with the formula

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{q_j} \sum_{i>j} \frac{d_{i-j} H_i}{H_i^n},$$

where $H_i^n = \sum_{j<i} d_{i-j} \lambda_j^n$ is the expected number of hospitalisations that occur on day i, based on the vector of incidence at the nth iteration and $q_j = \sum_{-j+1 \le i \le N-j} d_i$ is a normalization factor that represents the probability that an hospitalisation resulting from an infection on day j will be observed during the interval $1...N$. The iteration is stopped when the normalized χ^2 statistic, given by $\chi^2 = \frac{1}{N} \sum_i \frac{(H_i^n - H_i)^2}{H_i^n}$, is below 1 for the first time. The initial vector of incidence $(\lambda_1^0, ..., \lambda_N^0)$ is set to be the unshifted hospitalization curve.