International Journal of Epidemiology, 2021, 1–15
doi: 10.1093/ije/dyab110
Original Article

Original Article

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

Fabrice Carrat [1,2]* Xavier de Lamballerie,[3] Delphine Rahib,[4] Hélène Blanché,[5] Nathanael Lapidus,[1,2] Fanny Artaud,[6] Sofiane Kab,[7,8] Adeline Renuy,[7,8] Fabien Szabo de Edelenyi,[9] Laurence Meyer,[10,11] Nathalie Lydié,[4] Marie-Aline Charles [12] Pierre-Yves Ancel,[13,14] Florence Jusot,[15] Alexandra Rouquette,[10,11] Stéphane Priet,[3] Paola Mariela Saba Villarroel,[3] Toscane Fourié,[3] Clovis Lusivika-Nzinga,[1] Jérôme Nicol,[1] Stephane Legot,[7,8] Nathalie Druesne-Pecollo,[9] Younes Esseddik,[9] Cindy Lai,[16] Jean-Marie Gagliolo,[17] Jean-François Deleuze,[5] Nathalie Bajos,[18,†] Gianluca Severi,[6,19,†] Mathilde Touvier[9,†] and Marie Zins[7,8,†] for the SAPRIS and SAPRIS-SERO study groups

[1]Sorbonne Université, Inserm, Institut Pierre-Louis d'Epidémiologie et de Santé Publique, Paris, France, [2]Département de Santé Publique, APHP.Sorbonne Université, Paris, France, [3]Unité des Virus Émergents, UVE: Aix Marseille Univ, IRD 190, Inserm 1207, IHU Méditerranée Infection, Marseille, France, [4]Santé publique France, Saint-Maurice, France, [5]Fondation Jean Dausset-CEPH (Centre d'Etude du Polymorphisme Humain), CEPH-Biobank, Paris, France, [6]CESP UMR1018, Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, Villejuif, France, [7]Paris University, Paris, France, [8]Paris-Saclay University, Inserm UMS 11, Villejuif, France, [9]Sorbonne Paris Nord University, Inserm U1153, INRAE U1125, CNAM, Nutritional Epidemiology Research Team (EREN), Epidemiology and Statistics Research Center, University of Paris (CRESS), Bobigny, France, [10]Université Paris-Saclay, Inserm, CESP U1018, Le Kremlin Bicêtre, France, [11]Service de Santé Publique, APHP.Paris Saclay, Le Kremlin Bicêtre, France, [12]Ined, Inserm, EFS, UMS Elfe, Aubervilliers, France, [13]Obstetrical, Perinatal and Pediatric Epidemiology Research Team, Center for Epidemiology and Statistics Sorbonne Paris Cité, Inserm U1153, Paris Descartes University, Paris, France, [14]Clinical Research Unit, Center for Clinical Investigation P1419, Cochin Broca Hôtel-Dieu Hospital, Paris, France, [15]Université Paris-Dauphine, PSL-Research University, LEDa, Paris, France, [16]Institut de Santé Publique, Pôle Recherche Clinique, Institut National de la Santé et de la Recherche Médicale, Paris, France, [17]Inserm, Institut Thématique Santé Publique, Paris, France, [18]IRIS, Inserm/EHESS/CNRS, Aubervilliers, France[19]Department of Statistics, Computer Science and Applications, University of Florence, Florence, Italy

[†]Joint senior authors.

*Corresponding author. Institut Pierre-Louis d'Épidémiologie et de Santé Publique, 27 rue Chaligny, 75571 Paris Cedex 12, France. E-mail: fabrice.carrat@iplesp.upmc.fr

Accepted 4 May 2021

© The Author(s) 2021; all rights reserved. Published by Oxford University Press on behalf of the International Epidemiological Association

Abstract

Background: We aimed to estimate the seropositivity to anti-SARS-CoV-2 antibodies in May–June 2020 after the first lockdown period in adults living in three regions in France and to identify the associated risk factors.

Methods: Between 4 May 2020 and 23 June 2020, 16 000 participants in a survey on COVID-19 from an existing consortium of three general adult population cohorts living in the Ile-de-France (IDF) or Grand Est (GE) (two regions with high rate of COVID-19) or in the Nouvelle-Aquitaine (NA) (with a low rate) were randomly selected to take a dried-blood spot for anti-SARS-CoV-2 antibodies assessment with three different serological methods (ClinicalTrial Identifier #NCT04392388). The primary outcome was a positive anti-SARS-CoV-2 ELISA IgG result against the spike protein of the virus (ELISA-S). Estimates were adjusted using sampling weights and post-stratification methods. Multiple imputation was used to infer the cumulative incidence of SARS-CoV-2 infection with adjustments for imperfect tests accuracies.

Results: The analysis included 14 628 participants, 983 with a positive ELISA-S. The weighted estimates of seropositivity and cumulative incidence were 10.0% [95% confidence interval (CI): 9.1%, 10.9%] and 11.4% (95% CI: 10.1%, 12.8%) in IDF, 9.0% (95% CI: 7.7%, 10.2%) and 9.8% (95% CI: 8.1%, 11.8%) in GE and 3.1% (95% CI: 2.4%, 3.7%) and 2.9% (95% CI: 2.1%, 3.8%) in NA, respectively. Seropositivity was higher in younger participants [odds ratio (OR) = 1.84 (95% CI: 1.79, 6.09) in $<$40 vs 50–60 years old and OR = 0.56 (95% CI: 0.42, 0.74) in $\geq$70 vs 50–60 years old)] and when at least one child or adolescent lived in the same household [OR = 1.30 (95% CI: 1.11, 1.53)] and was lower in smokers compared with non-smokers [OR = 0.71 (95% CI: 0.57, 0.49)].

Conclusions: Seropositivity to anti-SARS-CoV-2 antibodies in the French adult population was $\leq$10% after the first wave. Modifiable and non-modifiable risk factors were identified.

Key words: SARS-CoV-2, COVID-19, general population, cohort, seroprevalence, risk factors

Key Messages

- The proportion of adults with a positive ELISA IgG test against the spike protein of SARS-CoV-2 in a general adult population remained low ($\leq$10%) in the French regions most affected by the first wave of SARS-CoV-2.
- Correcting for imperfect test accuracies, the cumulative incidence of SARS-CoV-2 infection after the first wave ranged from 3% to 11%.
- Seropositivity was higher in younger adults and in subjects living with children or adolescents in the household, whereas it was lower in active smokers.
- Fewer than 50% of the participants with a positive ELISA IgG against the spike protein were positive with the test detecting neutralizing antibodies or with the test detecting IgG against the nucleocapsid protein, and this result was strongly associated with the development of symptoms compatible with COVID-19.

Introduction

Serological surveys help to determine the extent of infection by a viral agent in a population and identify associated risk factors.[1] Since January 2020, serological surveys of SARS-CoV-2 have been performed in the general population on all continents. The median seroprevalence adjusted for imperfect test accuracy was estimated to be 3.2% (InterQuartile Range 1–6.4%) in 184 studies conducted in the general population.[2,3] Associations between a positive serological test (i.e. seropositivity) and demographic characteristics, such as a younger age, non-White ethnicity, occupational exposure, lower socio-economic status and higher population density, have been reported.[3–17] However, few studies have explored other determinants of

seropositivity related to the person's individual characteristics or behaviours.

In France, SARS-CoV-2-positive reverse transcription–polymerase chain reaction (RT–PCR) tests were first reported in imported cases in week 4 (24 January 2020), the generalized lockdown began in week 12 (17 March 2020) and emergency-room visits for COVID-19 peaked in week 13, decreasing thereafter. This led the French government to ease lockdown restrictions in week 20 (11 May 2020).

Our main goals were (i) to estimate the seropositivity to anti-SARS-CoV-2 antibodies in the French adult population at the end of first lockdown period in three regions and (ii) to identify the associated risk factors. The secondary objective was to estimate the cumulative incidence of SARS-CoV-2 infection with adjustments for imperfect test accuracies.

Participants and methods

Design

The present report combined data collected from questionnaires in the SAPRIS ('SAnté, Perception, pratiques, Relations et Inégalités Sociales en population générale pendant la crise COVID-19') survey in France, with serological results from the SAPRIS-SERO study.

The SAPRIS survey has been described elsewhere.[18] Briefly, the survey was created in March 2020 to evaluate the main epidemiological, social and behavioural challenges of the SARS-CoV2 epidemic in France in relation to social inequalities in health and healthcare. It is based on a consortium of prospective cohort studies including two child cohorts (not presented in this study) and three general-population-based adult cohorts (Supplementary Figure S1, available as Supplementary data at *IJE* online): (i) CONSTANCES, a 'general population' cohort including 204 973 adults aged 18–69 years at inclusion and randomly selected since 2012 to be a representative sample of the French adult population affiliated with French General Health Insurance Fund (the source population or ~85% of the total French population);[19] in CONSTANCES, 66 881 participants are followed using the internet, the rest through mailed questionnaires; (ii) E3N/E4N, a multigenerational adult cohort based on a community of families with 113 000 participants (including women recruited in 1990 and still actively followed up, their offspring and the fathers of these offspring) among whom 89 606 are followed using the internet;[20] (iii) NutriNet-Santé, a nutritional general-population-based internet cohort that started in 2009, with 170 000 included participants among whom 151 122 were still followed up in 2020.[21]

All participants from the original cohorts with regular access to electronic (internet) questionnaires and who were still being actively followed up on 1 April 2020 ($n = 279\,478$) were invited to participate in the SAPRIS survey. There was no other restriction on eligibility criteria in the survey. Two self-administered questionnaires covering the lockdown and the post-lockdown periods were sent as of 1 April 2020 and returned before 27 May 2020. Variables collected in the first questionnaire included socio-demographics; household size and composition; history of COVID-19 diagnosis; cough or fever from the beginning of the year; a detailed description of the participant's symptoms in the previous 2 weeks; co-morbidities, healthcare use and treatment; employment; daily life; childcare; alcohol, tobacco and cannabis use; social and sexual life; preventive measures; risk perception and beliefs. The second questionnaire updated all these variables (except sexual life), collected the history of SARS-CoV-2 RT–PCR testing and added the skin lesions to the list of symptoms to report in the previous 2 weeks. Participant representatives were involved in testing the readability and understanding of the questionnaires. A total of 120 500 out of 279 478 (43%) adults who were invited to participate in the survey completed the first questionnaire, 114 595 (39%) the second questionnaire and 102 001 (37%) completed both.

The SAPRIS-SERO study (#NCT04392388) included participants enrolled in the SAPRIS survey to assess their serological status (Supplementary Figure S1, available as Supplementary data at *IJE* online). We randomly selected 16 000 of the participants from the SAPRIS survey for this study who were residents from one of the three French administrative regions: Ile-de-France (IDF) or Grand Est (GE), i.e. the two regions with the highest reported cumulated rates of hospitalization for or deaths from COVID-19 at the end of the lockdown period (>250/100 000), or Nouvelle-Aquitaine (NA), a region with a low reported rate (<50/100 000).[22] Self-sampling dried-blood spot kits were mailed to each participant including material (a card, lancets, pad), detailed printed instructions on how to perform the test and a self-addressed stamped padded envelope to be returned with the card to the centralized biobank (CEPH-Biobank, Paris, France). Kits were received, then blood spots were visually assessed, registered, punched and stored in 2D FluidX 96-Format 0.5-mL tubes (Brooks) in –30°C freezers. Tubes were sent to the virology laboratory (Unité des virus émergents, Marseille, France) for serological analysis. Eluates were processed using a commercial ELISA test (Euroimmun®, Lübeck, Germany) to detect anti-SARS-CoV-2 antibodies (IgG) directed against the S1 domain of the spike protein of the virus (ELISA-S). All samples with an ELISA-S test optical density ratio of ≥0.7 were also tested with an ELISA test to detect

IgG antibodies against the SARS-CoV-2 nucleocapsid protein (ELISA-NP) (Euroimmun®, Lübeck, Germany) and with an in-house micro-neutralization assay to detect neutralizing anti-SARS-CoV-2 antibodies (SN), as described elsewhere.[23] The agreement of the results of ELISA SARS-CoV-2 assays performed using dried-blood spot samples compared with conventional serum was excellent: with serum as the gold standard, the dried-blood spot showed 98.1–100% sensitivity and 99.3–100% specificity.[24,25] Dried-blood spot vs serum was also validated internally before the beginning of the study.

Ethical approval and written or electronic informed consent were obtained from each participant before enrolment in the original cohort. The SAPRIS survey was approved by the Inserm ethics committee (approval #20–672 dated 30 March 2020). The SAPRIS-SERO study was approved by the Sud-Mediterranée III ethics committee (approval #20.04.22.74247) and electronic informed consent was obtained from all participants for dried-blood spot testing.

Outcomes

The main outcome was seropositivity to the ELISA-S test. In accordance with the manufacturer's instructions, a test was considered to be ELISA-positive with an optical density ratio of ≥1.1, ELISA-indeterminate between 0.8 and 1.1, and ELISA-negative at <0.8. The sensitivity and specificity of the ELISA-S test at the 1.1 threshold (considering indeterminate results as negative) in outpatients and plasma donors were reported to be 87% and 97.5%, repectively.[26] The secondary outcomes were a positive ELISA-NP (using the same thresholds) and positive SN defined as a titre of ≥40.

We used a multiple imputation (MI) method to infer the probability of infection among participants and to estimate the cumulative incidence of SARS-CoV-2 infection. Because the true infection status was unknown, we derived the following rules: participants with at least one positive ELISA-S, ELISA-NP and SN and no negative test results were assumed to be 'truly infected', those with all three negative results or ELISA-S of <0.7 were assumed to be 'truly non-infected' and the rest of the participants was reclassified according to the MI model using the numerical values from the three serological tests (log-transformed), region, age and gender. Since the specificity was >95% for each serological test independently,[23,26,27] we assumed that the specificity of the MI model for infection status was 100% (see Supplementary Box S1, available as Supplementary data at *IJE* online for details). However, an ELISA-S of <0.7 was sufficient to classify a participant as 'non-infected', which could be biased by the imperfect sensitivity of this method. We estimated the sensitivity of ELISA-S at the threshold of 0.7 in participants with a positive RT–PCR (88%) to adjust our MI estimates and to derive the cumulative incidence of SARS-CoV-2 infection.

Covariates

The association of seropositivity or cumulative incidence was evaluated in relation to age, gender, socio-demographic characteristics, body mass index (BMI), chronic conditions (according to a pre-specified list) and tobacco and alcohol use before the lockdown. Age groups were categorized according to predefined limits (<40, 40–49, 50–59, 60–69 and ≥70 years old) and BMI according to standard cut-offs (<18.5; 18.5 to <25; ≥25 to <30; ≥30 kg/m^2).[28]

The association of seropositivity was also studied in relation to symptoms. COVID-19-like symptoms were defined according to the European Centre for Disease Prevention and Control as at least one of the following: cough, fever, dyspnea and sudden anosmia, ageusia or dysgeusia.[29] COVID-19-like symptoms were based on symptoms reported in the 2 weeks before the questionnaire or as a cough or fever since the beginning of the year. To avoid the risk of misclassification with illnesses caused by other seasonal respiratory pathogens, we restricted our analysis to COVID-19-like symptoms that began on or after 1 March 2020. Participants who did not report any of these symptoms on either questionnaire, did not have a positive COVID-19 diagnosis or did not experience cough or fever since the beginning of the year were classified as 'No symptoms reported'.

Statistical methods

Inverse probability weighting was used to correct for selection and participation bias in the survey (see Supplementary Box S2, available as Supplementary data at *IJE* online, for details). Weighting was performed before the random selection of participants and no further weighting was applied for the participation into the serological study. Using 2016 regional population census data,[30] an initial cohort-specific calibration was performed by generalized raking in relation to the marginal totals of the distribution of age class, gender and socio-professional category in the target population.[31] The weights were rescaled according to the relative sample size of each cohort, then recalibrated according to the same covariates to provide representative estimates of the adult population. This weighting procedure was performed for each region independently. Confidence intervals (CIs) for weighted estimates were computed by bootstrapping. Comparison of

weighted samples by age and gender with the general population showed residual differences between the calibrated age distribution and that of the source population among young adults (Supplementary Figure S2, available as Supplementary data at *IJE* online).

The Markov Chain Monte Carlo method was used for MI. The MI was built from 100 imputed data sets and estimates combined with Rubin's rules.[32]

Chi-square test for trend was used on unweighted data to compare symptoms and healthcare use according to ELISA-S results. Logistic-regression models were used on unweighted data with stratification on the source cohort to identify the determinants of a positive ELISA-S (primary outcome). Indeterminate ELISA-S results were grouped with negative results in the primary analysis. A backward elimination procedure was used to identify independent covariates associated with a positive ELISA-S in multivariable analysis. The initial multivariable model included region, age, gender (forced variables) and all factors associated with seropositivity in univariable models. Elimination of covariates was based on the significance of the Wald chi-square test for parameter estimates at the 0.05 level. Contact with a RT–PCR-positive household member was not considered, to prevent the risk of reverse causation. Multivariable analyses were repeated using secondary outcomes then performed in each region to identify any potential regional-effect modification. Weighting and MI used the survey and the mice package from R software version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria).

Other analyses were performed using SAS 9.4 software (SAS Institute Inc., Cary, North Carolina, USA).

Results

Self-sampling dried-blood spot kits were mailed to 16 000 participants living in IDF, GE or NA who agreed to the serological study and who completed the first and second questionnaires. The dried-blood spot was returned by 15 414 (96%) of these participants and serology was performed on 14 830 (93%) samples: 14 628 (91%) could be interpreted and were included in the analyses. Dried-blood spots were collected between 4 May 2020 and 23 June 2020, with 90% performed before 24 May 2020. The median time between the second questionnaire and the dried-blood spot was 12 days (Q1–Q3: 10–16 days). Participant characteristics are described in Supplementary Table S1, available as Supplementary data at *IJE* online.

Seropositivity and cumulative incidence of SARS-CoV-2

A positive ELISA-S was found in 983 participants, 552 in IDF, 270 in GE and 161 in NA, with weighted seropositivity estimates in the adult population of 10.0% in IDF (95% CI: 9.1%, 10.9%), 9.0% in GE (95% CI: 7.7%, 10.2%) and 3.1% in NA (95% CI: 2.4%, 3.7%) (Table 1). The estimates of positive ELISA-NP and SN were markedly lower. MI classified 941 [standard deviation (SD) = 31] participants as positive, 548 (SD = 23) in IDF, 259

Table 1 Seropositivity to different serological tests and cumulative incidence of SARS-CoV-2 infection

	Ile-de-France (IDF) $n = 6348$	Grand Est (GE) $n = 3434$	Nouvelle-Aquitaine (NA) $n = 4846$
ELISA IgG against the spike protein (ELISA-S)			
Number positive	552	270	161
Unweighted seropositivity	8.7% (8.0%, 9.4%)	7.9% (7.0%, 8.8%)	3.3% (2.8%, 3.9%)
Weighted seropositivity	10.0% (9.1%, 10.9%)	9.0% (7.7%, 10.2%)	3.1% (2.4%, 3.7%)
ELISA IgG against the nucleocapsid protein (ELISA-NP)			
Number positive	315	161	35
Unweighted seropositivity	5.0% (4.5%, 5.5%)	4.7% (4.0%, 5.5%)	0.7% (0.5%, 1.0%)
Weighted seropositivity	5.7% (4.9%, 6.4%)	6.0% (4.9%, 7.0%)	0.6% (0.3%, 0.9%)
Neutralizing anti-SARS-CoV-2 antibodies (SN)			
Number positive	253	120	51
Unweighted seropositivity	4.0% (3.5%, 4.5%)	3.5% (2.9%, 4.2%)	1.1% (0.8%, 1.4%)
Weighted seropositivity	5.0% (4.2%, 5.7%)	4.3% (3.2%, 5.2%)	1.3% (0.8%, 1.7%)
Multiple imputation (MI)			
Number positive	548 ± 23	259 ± 16	134 ± 12
Unweighted cumulative incidence[a]	9.8% (9.0%, 10.6%)	8.6% (7.5%, 9.6%)	3.1% (2.6%, 3.7%)
Weighted cumulative incidence[a]	11.4% (10.1%, 12.8%)	9.8% (8.1%, 11.8%)	2.9% (2.1%, 3.8%)

[a]Adjusted on imperfect test accuracies.

(SD = 16) in GE, 134 (SD = 12) in NA, with a weighted cumulative incidence of SARS-CoV-2 infection of 11.4% in IDF (95% CI: 10.1%, 12.8%), 9.8% in GE (95% CI: 8.1%, 11.8%) and 2.9% in NA (95% CI: 2.1%, 3.8%).

Symptoms and healthcare use

Participants with a positive ELISA-S had a higher rate of self-reported symptoms than those with negative tests except for skin lesions (Table 2). COVID-19-like symptoms were reported in 47% of ELISA-S-positive participants a median of 56 days (Q1: 40 days; Q3: 61 days) before collection of blood samples, whereas the rate was 24% at 53 days (46, 60) in those with indeterminate results, and 19% at 53 days (42, 61) in participants with negative results ($P < 0.0001$). The proportion of positive ELISA-S in participants with COVID-19-like symptoms was higher in IDF and GE than in NA (Figure 1). It also varied during lockdown and decreased from 25% in IDF (95% CI: 21%, 29%), 26% in GE (95% CI: 20%, 32%) and 5.3% in NA (95% CI: 2.5%, 9.9%) when the onset of COVID-19-like symptoms was reported during week 12 (16–22 March—the beginning of lockdown) to 2.7% in IDF (95% CI: 0.1%, 14%), 0.0% in GE (95% CI: 0.0%, 23%) and 2.9% in NA (95% CI: 0.1%, 15.3%), when the onset of symptoms was reported during week 18 (Figure 1).

In participants with a positive ELISA-S, a positive ELISA-NP was found in 452/958 (47%) and a positive SN was found in 371/969 (38%). Positivity to either test was associated with reported symptoms: a positive ELISA-NP was found in 29/185 (16%) participants with no symptoms reported, in 335/454 (74%) with COVID-19-like symptoms and in 88/319 (28%) who reported other symptoms ($P < 0.0001$; Supplementary Figure S3, available as Supplementary data at *IJE* online), whereas a positive SN was found in 40/188 (21%), 250/459 (54%) and 81/322 (25%), respectively ($P < 0.0001$).

Factors associated with seropositivity

On univariable analysis, the rate of positive ELISA-S was higher in IDF and GE than in NA (Table 3) as well as in younger adult groups with an observed peak in ages 35–44 years old in each region (Figure 2). The association with age was similar with positive MI, although a higher proportion of positive SN or ELISA-NP was observed in the youngest age groups in the IDF and NA regions (Supplementary Figures S4 and S5, available as Supplementary data at *IJE* online).

Multivariable analysis showed an independent and positive association between positive ELISA-S, IDF and GE compared with NA for younger age and at least one child or adolescent living in the same household (Table 4). A negative association was found with active smoking (vs no smoking). The observed associations were confirmed with MI and were consistent overall with ELISA-NP and SN (Supplementary Tables S2–S4, available as Supplementary data at *IJE* online). When multivariable analysis was performed in each region separately, the associations did not differ between IDF and GE but the pattern in NA was different from the latter two, with a higher odds ratio (OR) in young age groups in the former [OR = 3.30 (95% CI: 1.79, 6.09) in <40 vs 50–60 years old and OR = 3.89 (95% CI: 2.18, 6.95) in 40–50 vs 50–60 years old] than in IDF [OR = 1.80 (95% CI: 1.37, 2.38) and OR = 1.83 (95% CI: 1.39, 2.40), respectively] and GE [OR = 1.45 (95% CI: 0.97, 2.17) and OR = 1.44 (95% CI: 0.99, 2.11), respectively]. Moreover, there was an association with female gender in NA [OR = 2.11 (95% CI: 1.42, 3.14)] but not in IDF [OR= 1.00 (95% CI: 0.83, 1.22)] or GE [OR = 1.02 (95% CI: 0.76, 1.36)] (Figure 3).

Discussion

In May–June 2020 following the first wave of the COVID-19 pandemic and the subsequent lockdown in France, the seropositivity to anti-SARS-CoV-2 antibodies was 10–9% in the adult population in the two regions with the highest rate of disease and 3% in a region with a low rate. The proportions of neutralizing antibody titres of ≥40 or positive ELISA IgG against the NP protein were half that detected by ELISA IgG against the spike protein. Seropositivity was strongly associated with reported symptoms and nearly half of the participants who tested positive experienced symptoms of COVID-19, whereas one in five did not recall having any symptoms. The associations between seropositivity and age, living with at least one child or adolescent and smoking status were consistent across all regions.

Our seropositivity estimates were consistent with seroprevalence estimates reported in other studies conducted in France during the same period, either in a representative sample of the general population (9.2% in IDF, 6.7% in GE and 2.0% in NA)[28] or based on residual sera (8.8%, 8.6% and 3.2%, respectively).[33,34] Estimated cumulative incidences were also close to the cumulative proportions of infection predicted by models at the end of the lockdown period[35] and in the range reported in similar studies in Europe conducted during the same period.[3,5–9] Half of the participants with a positive ELISA-S had an episode corresponding to the definition of a COVID-19 case and the reported symptoms corresponded to those described in similar studies.[6,7] One in five positive participants did not report symptoms during the 2 weeks before the first and second questionnaires or cough or fever from the beginning

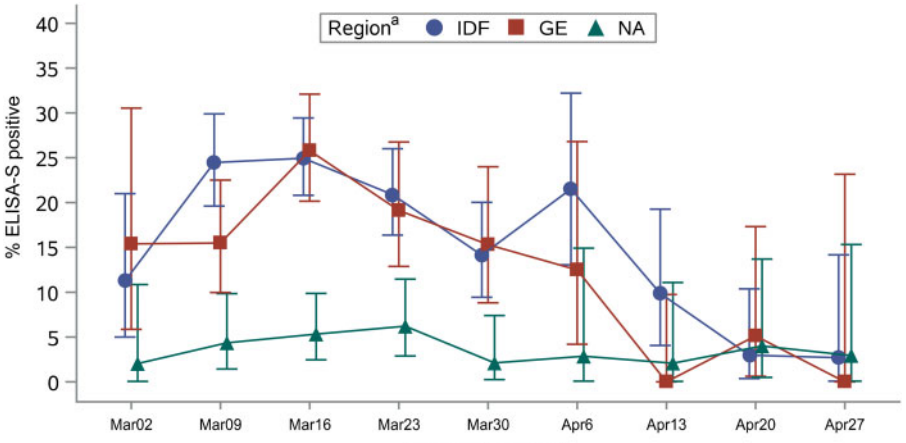
Table 2 Symptoms and healthcare use according to ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) results

	ELISA-S test result			
	Negative $n = 13\,369$	Indeterminate $n = 276$	Positive $n = 983$	*P*-value (trend test)
Symptoms reported on first or second questionnaire				
Fever or feverishness	1172 (9)	31 (11)	278 (28)	<0.0001
Cough	1829 (14)	44 (16)	305 (31)	<0.0001
Dyspnea	692 (5)	20 (7)	142 (14)	<0.0001
Anosmia/ageusia	258 (2)	21 (8)	248 (25)	<0.0001
Headaches	3851 (29)	105 (38)	451 (46)	<0.0001
Rhinorrhea	3440 (26)	79 (29)	353 (36)	<0.0001
Fatigue	2375 (18)	71 (26)	399 (41)	<0.0001
Stiffness, myalgia	2138 (16)	53 (19)	300 (31)	<0.0001
Nausea	622 (5)	21 (8)	106 (11)	<0.0001
Diarrhoea	1646 (12)	43 (16)	199 (20)	<0.0001
Chest pain	1086 (8)	31 (11)	166 (17)	<0.0001
Skin lesion (second questionnaire)	384 (3)	13 (5)	39 (4)	0.4131
Cough or fever from the beginning of the year				
< 1 March, 2020	1347 (10)	30 (11)	86 (9)	0.1612
≥1 March 2020	1134 (8)	34 (12)	321 (33)	<0.0001
Did not report any symptoms in first and second questionnaire or cough or fever from the beginning of the year ('No symptoms reported')	4839 (37)	86 (32)	188 (19)	<0.0001
COVID-19-like symptoms[a]	2501 (19)	67 (24)	460 (47)	<0.0001
Delay between onset of symptoms and questionnaire [median (Q1–Q3)]	20 (14–21)	20 (15–22)	20 (15–23)	0.2004
Delay between onset of symptoms and dried-blood spot sample [median (Q1–Q3)]	53 (42–61)	53 (46–60)	56 (49–61)	0.0003
Medical diagnosis of COVID-19	311 (2)	16 (6)	257 (27)	<0.0001
1 March	21 (7)	1 (8)	1 (0)	
1–16 March	67 (22)	4 (25)	50 (20)	
17–29 March	128 (41)	9 (56)	148 (57)	
30 March, 12 April	66 (21)	1 (6)	50 (20)	
> April 12	28 (9)	1 (6)	8 (3)	
Missing	301	10	28	
RT–PCR tested				
Positive/total tested	21/164 (13)	3/7 (43)	68/78 (87)	<0.0001
No date reported	9	1	5	
1–16 March	0	0	6	
17–29 March	5	1	17	
30 March, 12 April	2	1	22	
> April 12	5	0	18	
Positive RT–PCR in another household member	52 (0.4)	2 (0.7)	50 (5)	<0.0001
Sought medical advice for possible COVID-19				
GP visit	413/1859 (22)	16/61 (26)	159/334 (48)	<0.0001
Hospital visit	34/1868 (2)	1/61 (2)	21/336 (6)	<0.0001
Hospitalization	10/1871 (0.5)	0/61 (0)	10/338 (3)	<0.0001

[a]ECDC definition: cough or fever or dyspnea or sudden anosmia, ageusia or dysgeusia, with symptoms onset ≥1 March 2020, $n = 3028$, three missing in ELISA-S-negative group.

of the year. This was lower than the proportion of asymptomatic persons reported in Spain[6] or England,[7] which was ~30%, perhaps due to different methods of data collection of symptoms.

A lower seroprevalence with increasing age was reported in several population-based serological studies.[3,5,17] Although men are known to be at a higher risk of severe COVID-19, hospitalization and deaths than women,[36] we found an association between seropositivity and female gender in NA, which was also reported in a recent Italian study[8] but not in a systematic review of seroprevalence studies.[3] This association was



Number of participants with COVID-19-like symptoms by week of symptoms onset

IDF	71	286	409	292	184	79	71	67	37
GE	39	142	221	136	98	40	36	39	14
NA	49	115	169	145	95	35	48	50	34

Number of participants with ELISA-S positive test

IDF	8	70	102	61	26	17	7	2	1
GE	6	22	57	26	15	5	0	2	0
NA	1	5	9	9	2	1	1	2	1

[a] IDF: Ile-de-France, GE: Grand Est, NA: Nouvelle-Aquitaine

Figure 1 Proportion of participants with COVID-19-like symptoms and a positive ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) according to the date of the onset of symptoms.

only found in the region with a lower prevalence and may be related to the specific dynamics of transmission in this area.

A substantial proportion of participants were potentially infected before lockdown, probably in the workplace or in the community. This could explain why we did not find any specific association with social health inequalities, whereas, conversely, univariable analyses showed associations between a positive serology and being healthy, working adults with higher incomes and educational levels. As in other studies, univariable analysis identified the size of the household and the number of rooms,[33] but only living with at least one child remained associated with seroprevalence on multivariable analysis, indicating that children could play an important role in household-related transmission.[37]

Finally, active smoking was associated with a lower rate of ELISA-S- or SN-positive results—a finding consistently reported in other studies.[7,33,38–40] Smoking status was collected in the source cohort before the peak of the pandemic and thus could not have been affected by preventive behaviours in smokers. Although smoking is a risk factor for severe COVID-19 in infected patients,[41] its role in the risk of infection remains unclear because certain components of the smoke (such as nicotine) regulate ACE2-receptor expression, which is involved in SARS-CoV-2 entry into cells.[42,43]

Our study has several limitations. First, the primary endpoint is based on a test that does not have 100% sensitivity and specificity. With regard to infection by the SARS-CoV-2, some participants were probably misclassified. To overcome this limitation, we used a statistical imputation model combining the three serological results in participants with an ELISA-S of $\geq$0.7 to achieve a 100% specificity and we applied a correction to deal with the imperfect sensitivity of the ELISA-S test. The risk factors identified on multivariable analysis with MI were identical to those obtained with ELISA-S, which supports the robustness of our primary results.

The second potential limitation is that the selected adult population in each region may not be representative. Certain social categories were probably under- or over-represented and a voluntary response bias may have occurred. Although selection and participation biases were accounted for with an appropriate weighting and raking method, our findings cannot be considered to be strictly representative of the general adult population in these regions, particularly in the younger age groups. Nevertheless, the large number of subjects from all social categories makes it possible to draw robust conclusions on the factors associated with a positive serological test.

The survey was based on self-administered questionnaires, which may be subject to reporting bias, despite validation by participant representatives. Except for cough or

Table 3 Factors associated with a positive ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) (vs negative or indeterminate)

	ELISA-S positive/total	% (exact 95% CI)	Weighted seropositivity estimates (%)	Odds ratio[a]	95% CI	*P*-value
Regions						<0.0001
Ile-de-France	552/6348	8.7 (8.0, 9.4)	10.0 (9.1, 10.9)	2.64	2.20, 3.17	<0.0001
Grand Est	270/3434	7.9 (7.0, 8.8)	9.0 (7.7, 10.2)	2.44	1.99, 2.98	<0.0001
Nouvelle-Aquitaine	161/4846	3.3 (2.8, 3.9)	3.1 (2.4, 3.7)	Reference		
Age group (years)						<0.0001
<40	245/2262	10.8 (9.6, 12.2)	10.9 (9.2, 12.4)	2.06	1.68, 2.53	<0.0001
40–50	332/2897	11.5 (10.3, 12.7)	12.9 (11.2, 14.5)	2.14	1.77, 2.59	<0.0001
50–60	176/3019	5.8 (5.0, 6.7)	5.6 (4.3, 6.8)	Reference		
60–70	133/3272	4.1 (3.4, 4.8)	5.2 (4.0, 6.4)	0.68	0.54, 0.86	<0.0001
≥70	97/3175	3.1 (2.5, 3.7)	3.8 (2.9, 4.6)	0.50	0.38, 0.65	<0.0001
Gender						0.0012
Male	327/5809	5.6 (5.1, 6.3)	7.2 (6.2, 8.2)	Reference		
Female	656/8818	7.4 (6.9, 8.0)	8.6 (7.7, 9.4)	1.27	1.10, 1.46	
Living area						0.0006
Rural	118/2176	5.4 (4.5, 6.5)	7.1 (5.6, 8.5)	Reference		
<20 000 inhabitants	129/1863	6.9 (5.8, 8.2)	8.9 (7.1, 10.5)	1.23	0.95, 1.60	0.1124
20 000 to 100 000 inhabitants	211/2797	7.5 (6.6, 8.6)	8.0 (6.6, 9.3)	1.34	1.06, 1.70	0.0135
>100 000 inhabitants	524/7769	6.7 (6.2, 7.3)	8.0 (7.2, 8.8)	1.59	1.27, 1.99	<0.0001
Missing	23					
Household size and composition						
Number of adults (including participant)						<0.0001
1	176/2851	6.2 (5.3, 7.1)	7.2 (5.8, 8.4)	Reference		
2	329/6533	5.0 (4.5, 5.6)	6.6 (5.7, 7.3)	0.84	0.69, 1.01	0.0685
3+	478/5244	9.1 (8.4, 9.9)	10.3 (9.1, 11.4)	1.55	1.29, 1.86	<0.0001
Number of children (<18 years old)						<0.0001
0	574/10 848	5.3 (4.9, 5.7)	6.7 (6.0, 7.4)	Reference		
1+	409/3780	10.8 (9.9, 11.9)	11.8 (10.3, 13.2)	2.16	1.89, 2.48	
Number of rooms						0.0069
1–2	131/1696	7.7 (6.5, 9.1)	9.5 (7.8, 11.0)	1.06	0.86, 1.29	0.6076
3–4	421/5715	7.4 (6.7, 8.1)	8.5 (7.5, 9.4)	Reference		
5–6	323/5366	6.0 (5.4, 6.7)	6.8 (5.8, 7.8)	0.81	0.70, 0.98	0.0063
7+	100/1700	5.9 (4.8, 7.1)	7.2 (5.2, 9.0)	0.78	0.62, 0.98	0.0317
Missing	151					
Total household monthly income						<0.0001
<1000€	10/201	5.0 (2.4, 9.0)	5.4 (1.6, 8.4)	0.6	0.32, 1.15	0.1234
1000–1499	18/447	4.0 (2.4, 6.3)	4.8 (2.2, 7.0)	0.48	0.30, 0.77	0.0026
1500–1999	61/1000	6.1 (4.7, 7.8)	8.2 (5.8, 10.3)	0.76	0.57, 0.99	0.0411
2000–2999	138/2500	5.5 (4.7, 6.5)	7.1 (5.6, 8.4)	0.67	0.55, 0.82	<0.0001
3000–3999	207/3426	6.0 (5.3, 6.9)	7.5 (6.2, 8.7)	0.77	0.65, 0.92	0.0034
>4000	477/6045	7.9 (7.2, 8.6)	9.5 (8.4, 10.5)	Reference		
Missing	1009					
Educational level						<0.0001
<High-school degree	59/1629	3.6 (2.8, 4.7)	5.0 (3.5, 6.3)	Reference		
High-school degree or undergraduate	349/6032	5.8 (5.2, 6.4)	7.8 (6.9, 8.7)	1.69	1.27, 2.24	0.0003
Graduate degree or doctorate	464/5646	8.2 (7.5, 9.0)	9.3 (8.3, 10.2)	2.43	1.84, 3.21	<0.0001
Missing	1321					
Professional activity before lockdown						<0.0001
Student	mai-81	6.2 (2.0, 13.8)	7.2 (0.1, 12.6)	0.68	0.27, 1.68	0.4023
Working	741/8309	8.9 (8.3, 9.6)	10.5 (9.5, 11.4)	Reference		
Looking for a job	30/402	7.5 (5.1, 10.5)	7.8 (4.7, 10.4)	0.83	0.57, 1.21	0.3305
Retired	182/5381	3.4 (2.9, 3.9)	4.3 (3.5, 5.0)	0.35	0.30, 0.42	<0.0001

(Continued)

Table 3 Continued

	ELISA-S positive/total	% (exact 95% CI)	Weighted seropositivity estimates (%)	Odds ratio[a]	95% CI	*P*-value
Not working due to health conditions	7/125	5.6 (2.3, 11.2)	3.8 (0.5, 6.3)	0.58	0.27, 1.25	0.1621
No professional activity (housewife or husband)	16/306	5.2 (3.0, 8.4)	7.2 (2.3, 11.1)	0.53	0.32, 0.88	0.0144
Missing	24					
Essential job position						
Healthcare worker, Yes vs No	60/568	10.6 (8.2, 13.4)	11.6 (8.3, 14.4)	1.62	1.23, 2.13	0.0007
Other essential job, Yes vs No	122/1425	8.6 (7.2, 10.1)	11.9 (9.4, 14.0)	1.29	1.06, 1.58	0.0114
Professional activity during lockdown						<0.0001
Not working	240/6295	3.8 (3.4, 4.3)	4.9 (4.1, 5.6)	0.39	0.33, 0.46	<0.0001
Stopped working	127/1457	8.7 (7.3, 10.3)	8.3 (6.4, 10.0)	0.91	0.74, 1.12	0.3717
Working from home, remote working	410/4444	9.2 (8.4, 10.1)	11.6 (10.1, 13.0)	Reference		
Partially working from home	75/759	9.9 (7.9, 12.2)	12.6 (9.4, 15.1)	1.06	0.82, 1.38	0.6516
Working outside home	96/1134	8.5 (6.9, 10.2)	11.1 (8.7, 13.2)	0.89	0.70, 1.12	0.3062
Other	12/242	5.0 (2.6, 8.5)	7.8 (1.0, 12.6)	0.57	0.32, 1.04	0.065
Missing	297					
Smoking status before lockdown						<0.0001
Active smoker	98/1750	5.6 (4.6, 6.8)	7.1 (5.4, 8.7)	0.74	0.59, 0.92	0.0079
Ex-smoker	353/5973	5.9 (5.3, 6.5)	7.1 (6.1, 8.0)	0.73	0.63, 0.84	<0.0001
Non-smoker	516/6670	7.7 (7.1, 8.4)	8.9 (7.9, 9.9)	Reference		
Missing	235					
Alcohol use before lockdown (in g/dy)						0.0821
<5	426/5803	7.3 (6.7, 8.0)	8.5 (7.6, 9.3)	Reference		
(5, 10)	176/2641	6.7 (5.7, 7.7)	8.0 (6.5, 9.3)	0.94	0.78, 1.13	0.4971
(10, 20)	205/2963	6.9 (6.0, 7.9)	8.4 (6.9, 9.7)	1.03	0.86, 1.23	0.738
(20, 30)	63/1359	4.6 (3.6, 5.9)	5.8 (3.6, 7.7)	0.7	0.53, 0.92	0.0201
≥30	64/1128	5.7 (4.4, 7.2)	6.9 (4.5, 8.9)	0.87	0.66, 1.15	0.3276
Missing	734					
Body mass index (kg/m^2)						0.039
<18.5	33/499	6.6 (4.6, 9.2)	5.6 (3.0, 7.7)	0.87	0.61, 1.25	0.4578
(18.5–25)	619/8521	7.3 (6.7, 7.8)	8.2 (7.5, 9.0)	Reference		
(25–30) (overweight)	239/3995	6.0 (5.3, 6.8)	8.2 (7.0, 9.3)	0.83	0.71, 0.97	0.0191
>=30 (obese)	82/1409	5.8 (4.7, 7.2)	6.8 (5.0, 8.5)	0.78	0.61, 0.99	0.04
Missing	204					
Chronic diseases						<0.0001
Yes	259/4756	5.5 (4.8, 6.1)	6.7 (5.7, 7.7)	0.72	0.62, 0.83	<0.0001
No	715/9767	7.3 (6.8, 7.9)	8.7 (7.8, 9.4)	Reference		
Don't know	8/80	10.0 (4.4, 18.8)	9.9 (0.0, 16.2)	1.46	0.70, 3.06	0.3138
Missing	25					
Chronic diseases (Yes vs No)						
Asthma, chronic obstructive pulmonary diseases, other respiratory diseases	91/1534	5.9 (4.8, 7.2)	6.5 (4.9, 7.9)	0.81	0.64, 1.02	0.0695
Diabetes	19/481	4.0 (2.4, 6.1)	5.8 (2.5, 8.5)	0.63	0.39, 1.00	0.0519
Hypertension	67/1553	4.3 (3.4, 5.5)	4.9 (3.4, 6.3)	0.59	0.46, 0.77	<0.0001
Other cardiovascular diseases	19/451	4.2 (2.6, 6.5)	6.7 (3.3, 9.4)	0.64	0.40, 1.01	0.0569
Cancer	53/830	6.4 (4.8, 8.3)	7.6 (5.3, 9.7)	0.83	0.62, 1.11	0.1963
Anxiety, depression	30/404	7.4 (5.1, 10.4)	7.8 (3.8, 11.0)	1.08	0.74, 1.58	0.6955
Other	106/1826	5.8 (4.8, 7.0)	7.3 (5.6, 8.8)	0.78	0.63, 0.97	0.0228
Missing	25					

[a]With stratification on the source cohort.

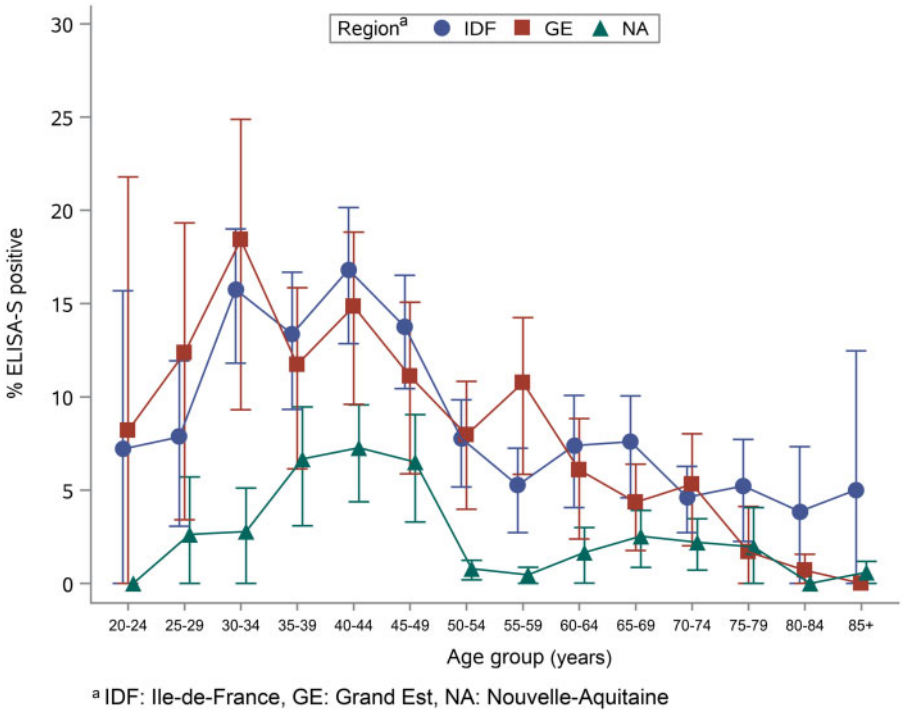


Figure 2 Proportion of participants with a positive ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) by age (weighted estimates)

Table 4. Multivariable analysis of factors associated with a positive ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) (vs negative or indeterminate)

	Odds ratio[a]	95% CI	*P*-value
Regions			
Ile-de-France	2.43	2.02, 2.93	<0.0001
Grand Est	2.24	1.83, 2.75	<0.0001
Nouvelle-Aquitaine	Reference		
Age group (years)			
<40	1.84	1.49, 2.28	<0.0001
40–50	1.92	1.57, 2.36	<0.0001
50–60	Reference		
60–70	0.77	0.60, 0.97	0.0299
0.56	0.42, 0.74	<0.0001	
Gender			
Male	Reference		
Female	1.14	0.99, 1.32	0.0792
Household size and composition—number of children (<18 years old)			
0	Reference		
1+	1.3	1.11, 1.53	0.0014
Smoking status before lockdown			
Active smoker	0.71	0.57, 0.89	0.0033
Ex-smoker	0.96	0.83, 1.11	0.5607
Non-smoker	Reference		

[a]With stratification on the source cohort. 235 participants—16 with an ELISA-S-positive result were excluded from the multivariable model due to missing smoking status.

fever, which were to be reported from the beginning of the year, we limited the questionnaires to a detailed description of symptoms present in the past 14 days to avoid recall bias. Thus, we may have missed symptoms related to SARS-CoV-2 infection other than cough or fever that occurred in early March. On the other hand, we cannot formally associate the self-reported COVID-19 symptoms with a positive serological result and, once again, misclassification may have occurred. Also, we cannot exclude that some (37% of those contacted) participants in the survey may have been more likely to complete the questionnaires because they experienced symptoms, which could bias certain estimates of rates of symptoms or seropositivity and be incompletely corrected by weighting. However, the rate of ELISA-S-positives in participants with COVID-19-like symptoms [15% (95% CI: 14%, 17%)] were consistent with the rates reported in other studies (e.g. 16.9%[6], 8.9–12.7%[15]).

Finally, we were surprised by the low proportion of ELISA-NP- or SN-positive tests in participants with an ELISA-S-positive result. Interestingly, ELISA-NP and SN seropositivity was strongly associated with reported symptoms and participants with COVID-19-like symptoms were 74% positive with ELISA-NP and 54% positive with SN. This dropped to 16% and 21%, respectively, in participants with no symptoms reported. It is therefore likely that factors such as symptoms[10,44] or age[4] affect the intensity and heterogeneity of the antibody response and

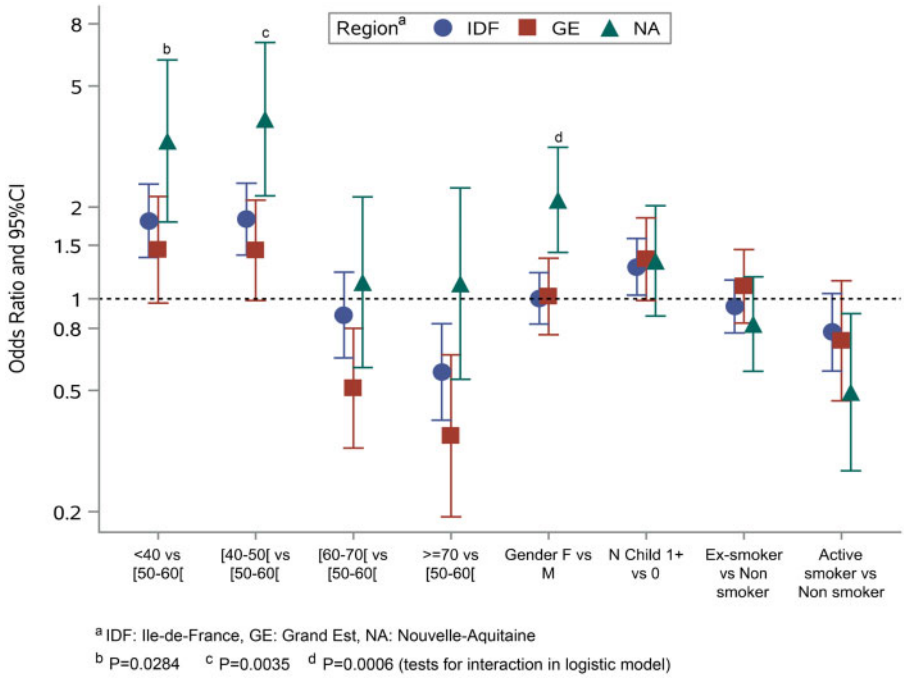


Figure 3 Risk factors of positive ELISA IgG against the SARS-CoV-2 spike protein (ELISA-S) by French region

explain, in part, these discrepancies. In addition, the results of analyses of these secondary criteria are globally consistent with those of the primary analysis.

This study has several strengths. In particular, it is based on well-characterized general-population cohorts. Moreover, serological samples were collected within 1–3 months after the period of intense circulation of SARS-CoV-2 and all serological tests were centralized and performed blinded to participants' characteristics or clinical history. Several serological methods were combined, including neutralization, to improve the interpretation of test results and to derive estimates of the cumulative incidence of SARS-CoV-2 infection.

In conclusion, our study shows that the level of seropositivity to anti-SARS-CoV-2 antibodies remained low in the French regions most affected by the first wave of SARS-CoV-2. Longer-term clinical and serological follow-up is needed to evaluate the duration of the humoral response, the risk of infection or reinfection and to establish the correlates of protection—a key element in preparing for the evaluation of vaccines against SARS-CoV-2.

Supplementary data

Supplementary data are available at *IJE* online.

Ethics approval

Ethical approval and written or electronic informed consent were obtained from each participant before enrolment in the original cohort. The SAPRIS survey was approved by the Inserm ethics committee (approval #20–672 dated 30 March 2020). The SAPRIS-SERO study was approved by the Sud-Mediterranée III ethics committee (approval #20.04.22.74247) and electronic informed consent was obtained from all participants for dried-blood spot testing.

Funding

This study

ANR (Agence Nationale de la Recherche) [#ANR-20-COVI-000, #ANR-10-COHO-06], Fondation pour la Recherche Médicale [#20RR052-00], Inserm (Institut National de la Santé et de la Recherche Médicale) [#C20–26]. The sponsor and funders facilitated data acquisition but did not participate in the study design, analysis, interpretation or drafting.

Cohorts funding

The CONSTANCES Cohort Study is supported by the Caisse Nationale d'Assurance Maladie (CNAM), the French Ministry of Health, the Ministry of Research and the Institut national de la Santé et de la Recherche Médicale. CONSTANCES benefits from a grant from the French National Research Agency [grant number ANR-11-INBS-0002] and is also partly funded by MSD, AstraZeneca, Lundbeck and L'Oreal. The E3N-E4N cohort is supported by the following institutions: Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, Inserm, University Paris-Saclay, Gustave Roussy, the MGEN and the French League Against Cancer. The NutriNet-Santé study is supported by the following public institutions: Ministère de la Santé, Santé Publique France, Institut National de la Santé et de la Recherche Médicale

(Inserm), Institut National de la Recherche Agronomique (INRAE), Conservatoire National des Arts et Métiers (CNAM) and Sorbonne Paris Nord. The CEPH-Biobank is supported by the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

Data availability

With regard to data availability, the data of the study are protected under the protection of health data regulation set by the French National Commission on Informatics and Liberty (Commission Nationale de l'Informatique et des Libertés, CNIL). The data can be available upon reasonable request to the corresponding author (fabrice.carrat@iplesp.upmc.fr), after a consultation with the steering committee of the SAPRIS-SERO study. French law forbids us from providing free access to SAPRIS-SERO data; access could however be given by the steering committee after legal verification of the use of the data. Please feel free to come back to us should you have any additional question.

Acknowledgements

The authors warmly thank all the volunteers of the CONSTANCES, E3N-E4N and NutriNet-Santé cohorts. We thank the staff of the CONSTANCES, E3N-E4N and NutriNet-Santé cohorts who have worked with dedication and engagement to collect and manage the data used for this study and to ensure continuing communication with the cohort participants. We thank the CEPH-Biobank staff for their adaptability and the quality of their work. In the virology department, we thank Dr Nadège Brisbarre and the technical staff for impeccable management of samples and serological assays. We thank Mrs Dale Roche-Lebrec for her help in editing the manuscript.

Author contributions

Conceptualization: Carrat, de Lamballerie, Rahib, Lapidus, Meyer, Charles, Ancel, Jusot, Rouquette, Bajos, Severi, Touvier, Zins. *Data curation:* Rahib, Blanché, Lusivika-Nzinga, Nicol, Legot, Druesne-Pecollo, Esseddik. *Formal analysis:* Carrat, Lapidus, Artaud, Kab, Renuy, Szabo de Edelenyi. *Funding acquisition:* Carrat, de Lamballerie, Severi, Touvier, Zins, Bajos. *Investigation:* de Lamballerie, Blanché, Priet, Saba Villarroel, Fourié, Nicol, Legot, Druesne-Pecollo, Esseddik. *Methodology:* Carrat, de Lamballerie, Lapidus. *Project administration:* Lai, Gagliolo. *Supervision:* Carrat, de Lamballerie, Rahib, Lydié, Charles, Ancel, Deleuze, Severi, Touvier, Zins, Bajos. *Writing—original draft:* Carrat. *Writing—review and editing:* All authors.

The SAPRIS study group

Nathalie Bajos (co-Principal investigator), Fabrice Carrat (co-Principal investigator), Pierre-Yves Ancel, Marie-Aline Charles, Florence Jusot, Claude Martin, Laurence Meyer, Ariane Pailhé, Gianluca Severi, Alexis Spire, Mathilde Touvier, Marie Zins.

The SAPRIS-SERO study group

Fabrice Carrat (Principal investigator), Pierre-Yves Ancel, Marie-Aline Charles, Gianluca Severi, Mathilde Touvier, Marie Zins. Sofiane Kab, Adeline Renuy, Stephane Legot, Celine Ribet, Emmanuel Wiernik, Marcel Goldberg, Marie Zins (CONSTANCES cohort). Fanny Artaud, Pascale Gerbouin-Rérolle, Mélody Enguix, Camille Laplanche, Roselyn Gomes-Rima, Lyan Hoang, Emmanuelle Correia, Alpha Amadou Barry, Nadège Senina, Gianluca Severi (E3N-E4N cohort). Julien Allegre, Fabien Szabo de Edelenyi, Nathalie Druesne-Pecollo, Younes Esseddik, Serge Hercberg, Mathilde Touvier (NutriNet-Santé cohort). Marie-Aline Charles, Pierre-Yves Ancel, Valérie Benhammou, Anass Ritmi, Laetitia Marchand, Cecile Zaros, Elodie Lordmi, Adriana Candea, Sophie de Visme, Thierry Simeon, Xavier Thierry, Bertrand Geay, Marie-Noelle Dufourg, Karen Milcent (Epipage2 and Elfe child cohorts). Clovis Lusivika-Nzinga, Gregory Pannetier, Nathanael Lapidus, Isabelle Goderel, Céline Dorival, Jérôme Nicol, Fabrice Carrat (IPLESP—methodology and coordinating data centre). Cindy Lai, Hélène Esperou, Sandrine Couffin-Cadiergues (Inserm). Jean-Marie Gagliolo (Institut de Santé Publique). Hélène Blanché, Jean-Marc Sébaoun, Jean-Christophe Beaudoin, Laetitia Gressin, Valérie Morel, Ouissam Ouili, Jean-François Deleuze (CEPH-Biobank). Stéphane Priet, Paola Mariela Saba Villarroel, Toscane Fourié, Souand Mohamed Ali, Abdenour Amroun, Morgan Seston, Nazli Ayhan, Boris Pastorino, Xavier de Lamballerie (Unité des Virus Emergents).

Conflict of interest

None declared.

References

1. Koopmans M, Haagmans B. Assessing the extent of SARS-CoV-2 circulation through serological studies. *Nat Med* 2020;**26**: 1171–72.
2. Arora RK, Joseph A, Van Wyk J *et al.* SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. *Lancet Infect Dis* 2021; **21**:e75–e76.
3. Bobrovitz N, Aro ra RK, Cao C *et al.* Global seroprevalence of SARS-SoV-2 antibodies: a systematic review and meta-analysis. *medRxiv* 2020; doi:10.1101/2020.11.17.20233460, 18 November preprint: not peer-reviewed.
4. Gudbjartsson DF, Norddahl GL, Melsted P *et al.* Humoral immune response to SARS-CoV-2 in Iceland. *N Engl J Med* 2020; **383**:1724–34.
5. Stringhini S, Wisniak A, Piumatti G *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet* 2020;**396**: 313–19.
6. Pollan M, Perez-Gomez B, Pastor-Barriuso R *et al.* Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* 2020;**396**:535–44.
7. Ward H, Atchinson C, Whitaker M *et al.* Antibody prevalence for SARS-CoV-2 following the peak of the pandemic in England: REACT2 study in 100,000 adults. *medRxivk* 2020; doi: 10.1101/2020.08.12.20173690, 21 August preprint: not peer-reviewed.

8. Vena A, Berruti M, Adessi A *et al.* Prevalence of antibodies to SARS-CoV-2 in Italian adults and associated risk factors. *JCM* 2020;**9**:2780.
9. Herzog S, De BJ, Abrams S *et al.* Seroprevalence of IgG antibodies against SARS coronavirus 2 in Belgium: a prospective cross-sectional nationwide study of residual samples. *medRxiv* 2020; doi:10.1101/2020.06.08.20125179, 1 October preprint: not peer-reviewed.
10. Azziz NH, Corman VM, Echterhoff AKC *et al.* Seroprevalence and correlates of SARS-CoV-2 neutralizing antibodies: results from a population-based study in Bonn, Germany. *medRxiv* 2020; doi:10.1101/2020.08.24.20181206, 29 August preprint: not peer-reviewed.
11. Xu X, Sun J, Nie S *et al.* Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. *Nat Med* 2020; **26**:1193–95.
12. Silveira MF, Barros AJD, Horta BL *et al.* Population-based surveys of antibodies against SARS-CoV-2 in Southern Brazil. *Nat Med* 2020;**26**:1196–99.
13. Skowronski DM, Sekirov I, Sabaiduc S *et al.* Low SARS-CoV-2 sero-prevalence based on anonymized residual sero-survey before and after first wave measures in British Columbia, Canada, March-May 2020. *medRxiv* 2020; doi:10.1101/2020.07.13.20153148, 15 July preprint: not peer-reviewed.
14. Havers FP, Reed C, Lim T *et al.* Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States. *JAMA Intern Med* 2020; 23 March–12 May (online ahead of print).
15. Sood N, Simon P, Ebner P *et al.* Seroprevalence of SARS-CoV-2-specific antibodies among adults in Los Angeles County, California, on April 10-11, 2020. *JAMA* 2020;**323**:2425–27.
16. Menachemi N, Yiannoutsos CT, Dixon BE *et al.* Population point prevalence of SARS-CoV-2 infection based on a statewide random sample—Indiana, April 25-29, 2020. *MMWR Morb Mortal Wkly Rep* 2020;**69**:960–64.
17. UK Biobank SARS-CoV-2 Serology Study. Weekly Report—21 July 2020. https://www.ukbiobank.ac.uk/2020/07/uk-biobank-covid-19-antibody-study-latest-updates/ (28 August 2020, date last accessed).
18. Carrat F, Touvier M, Severi G *et al.*; for the SAPRIS study group. Incidence and risk factors of COVID-19-like symptoms in the French general population during the lockdown period: a multicohort study. *BMC Infect Dis* 2021;**21**. In press.
19. Zins M, Goldberg M, Team C. The French CONSTANCES population-based cohort: design, inclusion and follow-up. *Eur J Epidemiol* 2015;**30**:1317–28.
20. Clavel-Chapelon F, Group ENS. Cohort Profile: the French E3N cohort study. *Int J Epidemiol* 2015;**44**:801–09.
21. Hercberg S, Castetbon K, Czernichow S *et al.* The NutriNet-Sante Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 2010;**10**:242.
22. Santé Publique France. COVID-19: point épidémiologique du 14 mai 2020. 2020. https://www.santepubliquefrance.fr/content/download/252588/2603686 (24 August 2020, date last accessed).
23. Gallian P, Pastorino B, Morel P, Chiaroni J, Ninove L, de Lamballerie X. Lower prevalence of antibodies neutralizing SARS-CoV-2 in group O French blood donors. *Antiviral Res* 2020;**181**:104880.
24. Morley GL, Taylor S, Jossi S *et al.* Sensitive detection of SARS-CoV-2-specific antibodies in dried blood spot samples. *Emerg Infect Dis* 2020;**26**:2970–73.
25. Zava TT, Zava DT. Validation of dried blood spot sample modifications to two commercially available COVID-19 IgG antibody immunoassays. *Bioanalysis* 2021;**13**:13–28.
26. Patel EU, Bloch EM, Clarke W *et al.* Comparative performance of five commercially available serologic assays to detect antibodies to SARS-CoV-2 and identify individuals with high neutralizing titers. *J Clin Microbiol* 2021;**59**: e02257-20.
27. Rikhtegaran Tehrani Z, Saadat S, Saleh E *et al.* Performance of nucleocapsid and spike-based SARS-CoV-2 serologic assays. *PLoS One* 2020;**15**:e0237828.
28. *Obesity: Peventing and Managing the Global Epidemic*. Report on a WHO consultation. Geneva: WHO, 1997.
29. European Center for Diseases Control. Case definition for coronavirus disease 2019 (COVID-19), as of 29 May 2020. https://www.ecdc.europa.eu/en/covid-19/surveillance/case-definition (15 June 2020, date last accessed).
30. Institut National de la Statistique et des Etudes Economiques. Individus localisés à la région en 2016. Recensement de la population—Fichiers détail, 2016. https://www.insee.fr/fr/statistiques/4171523 (3 February 2021, date last accessed).
31. Deville JC, Sarndal CE, Sautory O. Generalized raking procedures in survey sampling. *J Am Stat Assoc* 1993;**88**:1013–20.
32. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, 1987.
33. Warszawski J, Bajos N, Meyer L *et al*. In May 2020, 4.5% of the population of metropolitan France had developed antibodies against SARS-CoV-2. *Etudes et Resultats (Drees)* 2020;1167. https://drees.solidarites-sante.gouv.fr/IMG/pdf/er1167-en.pdf (1 December 2020, date last accessed).
34. Le VS, Jones G, Anna F *et al.* Prevalence of SARS-CoV-2 antibodies in France: results from nationwide serological surveillance. *medRxiv* 2020; doi:10.1101/2020.10.20.20213116, 21 October preprint: not peer-reviewed.
35. Salje H, Tran Kiem C, Lefrancq N *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* 2020;**369**:208–11.
36. Williamson EJ, Walker AJ, Bhaskaran K *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;**584**:430–36.
37. Jing QL, Liu MJ, Zhang ZB *et al.* Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *Lancet Infect Dis* 2020;**20**:1141–50.
38. Fontanet A, Tondeur L, Madec Y *et al.* Cluster of COVID-19 in northern France: a retrospective closed cohort study. *medRxiv* 2020; doi:10.1101/2020.04.18.20071134, 23 April preprint: not peer-reviewed.
39. Rentsch CT, Kidwai-Khan F, Tate JP *et al.* Covid-19 testing, hospital admission, and intensive care among 2,026,227 United States Veterans aged 54-75 years. *medRxiv* 2020; doi:10.1101/2020.04.09.20059964, 14 April preprint: not peer-reviewed.
40. Simons D, Shahab L, Brown J, Perski O. The association of smoking status with SARS-CoV-2 infection, hospitalisation and mortality from

COVID-19: a living rapid evidence review with Bayesian meta-analyses (version 7). *Addiction* 2020 (online ahead of print).

41. Guan WJ, Ni ZY, Hu Y *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;**382**:1708–20.
42. Oakes JM, Fuchs RM, Gardner JD, Lazartigues E, Yue X. Nicotine and the renin-angiotensin system. *Am J Physiol Regul Integr Comp Physiol* 2018;**315**:R895–906.
43. Smith JC, Sausville EL, Girish V *et al.* Cigarette smoke exposure and inflammatory signaling increase the expression of the SARS-CoV-2 receptor ACE2 in the respiratory tract. *Dev Cell* 2020;**53**: 514–29.e3.
44. Long QX, Tang XJ, Shi QL *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* 2020;**26**:1200–04.