Nationwide seroprevalence of COVID-19 and identification of risk factors for infection and disease during the first epidemic wave in the Netherlands: towards a longitudinal follow-up

(10)(2e) MSc[1], (10)(2e), PhD[1], (10)(2e), BSc[1], (10)(2e), PhD[1], (10)(2e) (10)(2e), (10)(2e), BSc[1], (10)(2e), MSc[1], (10)(2e) (10)(2e), PhD[1], (10)(2e) MSc[1], (10)(2e), PhD[1], (10)(2e) PhD[1], (10)(2e) PhD[1], (10)(2e) (10)(2e), PhD[1], (10)(2e) Phd[1]

[1]Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3720 MA Bilthoven, the Netherlands.

Corresponding author: (10)(2e) Centre for Infectious Disease Control, National Institute for Health and de Environment (RIVM), the Netherlands, email: (10)(2e) @rivm.nl, telephone: (10)(2e).

ABSTRACT

Background: Reported numbers of COVID-19 cases are an underestimation of the true magnitude of the pandemic. Nationwide serosurveillance studies measuring antibodies against the new coronavirus (SARS-CoV-2) provide insights into the scope of undetected cases, in monitoring the epidemic and guiding interventions. Here, we estimated seroprevalence of SARS-CoV-2 antibodies in the general population of the Netherlands with aim to identify risk factors for infection during the first epidemic wave.

Methods: Participants (n=3,207, aged 2-90 years, from across the country) were enrolled from a previously established representative serosurveillance study (median inclusion date April 3, 2020), and provided a self-collected fingerstick blood sample and completed a questionnaire. IgG antibodies targeted against the spike S1-protein of SARS-CoV-2 were quantified using multiplex-immunoassay. Weighted true seroprevalence was estimated controlling for individual pre-pandemic cross-reactivity and test performance. Logistic regression was used to identify risk factors for seropositivity.

Findings: Seroprevalence in the Netherlands was overall 2·8% [95% CI 2·0-3·7], did not differ between sexes, and ranged between 1·3-4·0% among regions. Estimates were lowest in children (1-3%) and highest in age group 18-39 years (4·9%). Among seropositive participants, 93% reported to have had symptoms, of which anosmia/ageusia (53%) was most outstanding, and antibody concentrations were significantly higher in those with fever or dyspnea. Multivariable analysis further revealed that the odds for seropositivity were significantly higher among persons taking immunosuppressive medication and among Orthodox-Reformed Protestants.

Interpretation: In the midst of the first epidemic wave, we estimate that nearly half a million persons in the Netherlands were infected with SARS-CoV-2, and although this is substantial higher than the cases reported, susceptibility remains high. Risk groups that were identified should be monitored thoroughly during the course of the epidemic. Repeated data collection enables to track

changes in epidemics and immunity, thereby guiding future interventions to control the pandemic, including vaccination once available.

Funding: Ministry of Health, Welfare and Sports (VWS), the Netherlands.

KEYWORDS: COVID-19 pandemic; SARS-CoV-2; seroepidemiology; seroprevalence; risk factors; symptoms; spike S1 IgG antibodies; the Netherlands

RESEARCH IN CONTEXT

Evidence before this study

We searched PubMed for peer-reviewed articles, medRxiv, bioRxiv, arXiv, Research Square, SSRN, Virological and Wellcome Open Research for pre-prints, and online research reports on seroprevalence of and risk factors for SARS-CoV-2 IgG antibodies on a nationwide level, published in English, using the search terms "population-based", "nationwide", "seroprevalence", "SARS-CoV-2 antibodies", "COVID-19", and "risk factors", up to June 23, 2020. We identified two peer-reviewed population-based serosurveys that provided local seroprevalence estimates in outbreak regions, namely Los Angeles County, CA, USA, and the canton of Geneva, Switzerland, which found 4·7% and 4·8%, respectively, to be seropositive at the beginning of April 2020. Studies reporting extensive risk factors analyses were not identified. Three pre-prints reporting seroprevalence estimates on a nationwide level were found, using dissimilar enrollment procedures, designs and/or laboratory methods: one from the Netherlands, using a sample of healthy blood donors aged 18-79 years (seroprevalence: 2·7%); from Brazil, a population-based household serosurvey using rapid serology tests (seroprevalence: 1·4%); and from Luxembourg, enrolling participants from 18 years using a web-panel (seroprevalence: 2·0%).

Added value of this study

This is the first study providing nationwide representative seroprevalence estimates of SARS-CoV-2 across all ages, as well as identifying risk groups for infection and disease via collection of an extensive set of sociodemographic and clinical characteristics. Building on an existing representative serosurveillance study established in 2016/17, we could respond rapidly to the urgent need of relevant nationwide longitudinal serology studies shortly after first COVID-19 case was confirmed in the Netherlands. Availability of paired sera, i.e., pre-pandemic and present sera, enabled us to

validate seroconversion, and, since we controlled for the nationwide population structure as well as test performance, precise estimates could be provided. Using our expertise in multiplex immunoassays, a high throughput assay quantifying SARS-CoV-2 specific IgG antibodies allowed us to analyze patterns in term of response and severity of reported symptoms. Finally, our study has been designed for repeated collection of sera and information on potential (new appearing) risk factors in the same individuals which will aid in the understanding of antibody kinetics and progression of the epidemic.

Implications of all the available evidence

We estimated an overall seroprevalence in the general population of the Netherlands (with over 17 million inhabitants) of 2·8% [95% CI 2·0-3·7], equivalent to nearly half a million infected persons. This is in striking contrast with the reported number of cases, and underlined the underestimation of the true pandemic size without seroprevalence data. Nonetheless, the proportion of persons susceptible to SARS-CoV-2 is high and disease severity is substantial. Estimates were highest in young adults, and lowest in children. Interestingly, persons taking immunosuppressive drugs as well as those from the Orthodox-Reformed Protestant community appeared to have higher odds for seropositivity than others. Our analyses provide additional information on the proportion of asymptomatic persons, on the correlation between degree of symptoms and antibody responses, and overall severity of disease on a population level – given a substantial overall infection-hospitalization and -fatality ratios estimated – for instance when compared to other respiratory diseases, such as seasonal influenza. Globally, nationwide serology studies will increase our understanding of the size of the pandemic and provide guidance to decisionmakers in taking appropriate control measures, such as vaccination once an effective vaccine becomes available.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causative agent of coronavirus disease (COVID-19), emerged in Wuhan, China, in early December 2019. On March 11, 2020, the World Health Organization declared COVID-19 a pandemic, with millions of confirmed cases as of June, 2020.[1,2] The first patient in the Netherlands was confirmed on February 27, 2020.[3] Cases primarily clustered in the southeastern part of the country, but were also reported in other regions quickly hereafter. Multi-pronged interventions to suppress the spread of the virus, including social distancing and home quarantine were implemented on March 16, 2020. By June 22, 2020, 49,658 cases, 11,851 hospitalizations, and 6,090 related deaths have been reported in the Netherlands.[3]

Reported COVID-19 cases worldwide are unquestionably an underestimation of the true magnitude of the pandemic as the scope of undetected cases remain largely unknown – due to difference in restrictive testing policy and registration in different countries, and proportion of asymptomatic cases in general.[4] Large-scale serosurveillance studies measuring SARS-CoV-2-specific antibodies in serum on a nationwide level, combined with extensive questionnaire data on potential risk factors, are key to better assess the true number of infections – and hence overall disease severity – identify groups at risk, and guide decisionmakers for appropriate control measures[5]. Repeated data collection is paramount in estimating the scale of the epidemic, providing insights into (prolonged) immunity, and guiding future interventions, for instance on targeting susceptible groups eligible for vaccination once an effective vaccine will become available.[6] Unfortunately, such nationwide studies are scarce.

Therefore, we have set-up a nationwide prospective serosurveillance study in the general population of the Netherlands (PIENTER-Corona, PICO) during the first epidemic wave, initiated quickly after the

lockdown was in effect. Our cohort is unique as it comprises data available from a previous serosurvey established in 2016/17 (PIENTER-3) of (1) a representative sample of Dutch citizens, across all ages and throughout the country; and (2) a separate sample enriched for Orthodox-Reformed Protestants whom are suggested to have been exposed to SARS-CoV-2 more frequently due to their socio-geographical clustered lifestyle, e.g., frequent church gatherings.[7,8] The presented serological framework and findings of our first round of inclusion can support public health policy in the Netherlands as well as globally.

METHODS

Study design

In 2016/17, the National Institute for Public Health and the Environment of the Netherlands (RIVM) initiated a large-scale nationwide representative serosurveillance study (PIENTER-3) (n=7,600; age-range 0-89 years). The primary aim was to obtain insight into the protection against vaccine-preventable diseases offered by the national immunization program in the Netherlands. A comprehensive description of PIENTER-3 has been published previously.[9] Briefly, participants were selected via a two-stage cluster design, comprising 40 municipalities in five regions nationwide (henceforth 'national sample', NS), and nine municipalities in the low vaccination coverage municipalities (LVC), inhabited by a relative large proportion of Orthodox-Reformed Protestants (**Figure1**). Among other materials, sera and questionnaire data had been collected from all participants. Hence, the PIENTER-3 study acted as baseline sample of the Dutch population for the present PICO-study since 6,102 participants (80%) consented to be approached for follow-up (after updating addresses and screening of possible deaths). The study was powered to estimate an overall seroprevalence with a precision of at least 2·5%. The PICO-study protocol was approved by the Medical Ethics Committee (Clinical Trial Registration NTR8473) and available.

Study population and materials

On March 25, 2020, an invitation letter was sent. Invitees (age-range 2-92 years) who were willing to participate registered online. After enrollment, participants received an instruction letter on how to self-collect a fingerstick blood sample in a microtainer (maximum of 0·3mL). Blood samples were returned to the RIVM-laboratory in safety envelopes. Serum sample were stored at -20 °C awaiting analyses. Materials were collected between March 31 and May 11, with the majority (80%) in the first week of April 2020 (median collection date was April 3). Simultaneous with the blood collection,

participants were asked to complete an (online) questionnaire, including questions regarding sociodemographic characteristics and potential determinants for SARS-CoV-2 seropositivity. All participants provided written informed consent.

Laboratory methods

Serum samples were tested for the presence of SARS-CoV-2 specific IgG antibodies using a fluorescent bead-based multiplex-immunoassay using Luminex technology as described.[10] Briefly, recombinant spike S1 (Sino Biologicals) was coupled to carboxylated beads. In lieu of an international reference, 15 PCR-confirmed COVID-19 patient's sera were pooled and an arbitrary unitage (AU/mL) assigned. Sera (1:200) and reference were diluted in SM01 (Surmodics), supplemented with 2% FBS, and incubated with the beads for 45 minutes at room temperature. Following washing steps, 1:400 diluted PE-conjugated Goat Fabs Anti-human IgG was incubated for 30 minutes as above. Plates were washed and acquired on Luminex LX200 or FlexMAP 3D. Mean fluorescence intensity was converted to concentrations (AU/mL) by interpolation from a 5-parameter logistic standard curve. A cutoff concentration for seropositivity was determined by ROC analysis of 400 pre-pandemic control samples (including PIENTER-3 participants who participated in the current PICO-study (n=108)) as well as patients with confirmed influenza-like illnesses caused by coronaviruses and other viruses, and a selection of sera from 115 PCR-confirmed COVID-19 cases with mild, or severe disease symptoms. A specificity-optimized cut-off value (99%) was chosen, resulting in a cutoff value for this specific sample set of 2·37 AU/mL. Sensitivity at this cutoff was 84 ·4%. Seropositive PICO-samples and those with a concentration 25% below the cutoff were retested (n=138), and a geometric mean concentration (GMC) of the first and second measure was calculated and used in further analyses. Pre-pandemic PIENTER-3-samples were available from 129/138 PICO-samples and were tested as described above to correct for false-positive results (for n=26 eventually) (**FigureS1a**).

Statistical analyses

Study population and COVID-19-related symptoms

Data management and analyses were conducted in SAS v.9·4 (SAS Institute Inc., USA) and R v.3·6. P values <0·05 were considered statistically significant. Sociodemographic characteristics and COVID-19-related symptoms (general, respiratory, gastrointestinal) developed since the start of the epidemic were stratified by sample (NS vs. LVC), or sex, respectively, and described for seropositive and seronegative participants. Differences were tested via Pearson's Chi-squared or Fisher's exact test if appropriate. Differences in GMC between reported symptoms in seropositive participants were determined by calculating the difference in log-transformed concentrations of those who developed symptoms at least four weeks prior to the sampling – ensuring a plateaued response – and tested by means of a Mann-Whitney U test.

Seroprevalence estimates

Apparent seroprevalence estimates (with 95% Wilson confidence intervals (CI)) for SARS-CoV-2-specific antibodies were calculated taking into account the survey design (i.e., controlling for region and municipality) and were weighted by sex, age, ethnic background and degree of urbanization to match the distribution of the general Dutch population in both NS and LVC sample. True seroprevalence estimates were derived by correcting the apparent estimates via the Rogan & Gladen bias correction for the test performance (with sensitivity of 84.4% and assuming a specificity of 100% after cross-validation with pre-sera).[11] Smooth age-specific true seroprevalence estimates were obtained with a logistic regression in a Generalized Additive Model using penalized splines, as implemented in the R package mgcv.

Risk factors for SARS-CoV-2 seropositivity

A logistic regression model was used to identify risk factors for SARS-CoV-2 seropositivity, applying a full case analysis (n=3,100). Potential risk factors included sociodemographic characteristics (sex, age

group, region, ethnic background, Orthodox-Reformed Protestants, (maternal) educational level, household size, (parent with a) contact profession, healthcare worker), and COVID-19 related factors (contact with a COVID-19 patient, number of persons contacted yesterday, working from home (normally and in the last week (during lockdown)), comorbidities (combining chronic lung disease, diabetes, history of malignancy, immunodeficiency, cardio-vascular disease, kidney disease), and use of blood pressure medication, immunosuppressive medication, statins and antivirals/antibiotics in the last month). Crude odds ratios (OR) in univariable analyses were a priori adjusted for sex, age and region to account for the survey design. Variables with $p<0{\cdot}10$ were entered in the multivariable analysis, and backward selection – manually dropping variables one-by-one – was performed to identify significant risk factors based on $p<0{\cdot}05$. Adjusted ORs and corresponding 95% CIs were provided.

Role of the funding source

The funders had no role in study design, data collection, data analysis, data interpretation, writing of the report, and in submission of the paper. FvdK and RV had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Study population

Of 6,102 invitees, 3,487 persons registered. In total, 3,207 (53%) donated a serum sample and filled-out the questionnaire, of which 2,657 persons from NS and 570 from LVC. Participants from across the country participated (**Figure1**) and age ranged from 2-90 years (**Table1**). In NS, slightly more women (56%) participated, most (88%) were of Dutch descent, nearly half had a high educational level, and 45% was religious. Twenty percent of persons between age 25-66 years were healthcare workers and 56% of the (parents of) participants reported to have had daily contact with patients, clients and/or children in their profession/volunteer work normally. Over half of the participants lived in a ≥2-person household, and 78% reported to have had contact with less than five people outside their own household yesterday (during lockdown), of which more than half with nobody. Comorbidities that were reported most frequently included chronic lung and cardiovascular disease (both 13%), and a history of malignancy (5%). In line with the population distribution, the LVC sample was characterized by a relative high proportion of Orthodox-Reformed Protestants from Dutch descent (**Table1**).

Seroprevalence

Overall weighted true seroprevalence in NS was 2·8% [2·0-3·7] and did not differ between sexes or between NS and LVC (**Table1**). Seroprevalence was lowest in the Northern region (1·3%) and highest in the Mid-West (4·0%). Estimates slightly differed between people from different ethnic backgrounds (Dutch: 2·8%, Western: 2·0%, non-Western: 3·4%), although not statistically significant. Estimated seroprevalence was lowest in children – gradually increasing from 1% at 2 years to 3% at 17 years – was highest in age group 18-39 years (4·9%) and ranged between 2-3·5% up to 90 years (**Figure2**). In both samples, seroprevalence was highest in Orthodox-Reformed Protestants (>7%).

FigureS1b displays the distribution of IgG concentrations for all participants by age, and **FigureS2** shows the seroprevalence smoothed by age in LVC.

Symptoms and antibody response

In total, 63% of all participants reported to have had one or more COVID-19-related symptom(s) since the start of the epidemic, with runny nose (37%), headache (33%), and cough (30%) being most common (**Table2**). Over 70% of both the seropositive and seronegative participants had symptoms at least two weeks prior to the blood sampling. All reported symptoms were significantly higher in seropositive compared to seronegative persons, except for stomach ache. Interestingly, among those seropositive, some reported to have had symptoms already in mid-February – of which the first at February 12 by a participant from Noord-Brabant – 15 days prior to the official first notification. The majority of the seropositive participants (93%) reported to have had symptoms (90% of men vs. 95% of women). Among them, median duration of illness was 8·5 days (interquartile range: 4·0-12·5), 16% (n=12) visited a general practitioner and one was admitted to the hospital.

Most seropositive participants reported to have had at least one respiratory symptom(s) (86%), of which runny nose and cough (both 61%), and general (84%) symptoms, with headache (65%), arthralgia (57%) and anosmia/ageusia (53%), were the most frequent (**Table2**). Nearly all symptoms were more common in women, except for anosmia/ageusia, cough and irritable/confusion, which were more dominant in men. Almost 75% of the seropositive participants met the COVID-19 case definition of fever and/or cough and/or dyspnea, which improved to 80% when anosmia/ageusia was included – while remaining 36% in those seronegative. GMC was generally higher in seropositive participants with any symptom(s) (with onset at least four weeks before blood sampling) as compared to those without (20·2 vs 8·7 AU/mL, p=0.16,), albeit not statically significantly. Significant dissimilarities were however observed between seropositive persons with fever vs. without (48·2 vs. 11·6 AU/mL, p=0.01), and with dyspnea vs. without (78·6 vs. 13·5 AU/mL, p=0·04).

Risk factors

Logistic regression analysis was performed for the total cohort (n=3,100) to identify risk factors associated with SARS-CoV-2 seropositivity (**Table3**). Variables that were associated with seropositivity ($p<0{\cdot}1$) in univariable analyses included age group, Orthodox-Reformed Protestant, contact with a COVID-19 case, and use of immunosuppressive and antibiotic/antiviral medication in the last month. Multivariable analysis revealed that substantial higher odds were observed for those who took immunosuppressive medication, had been in contact with a COVID-19 case, are Orthodox-Reformed Protestant (vs. others), and those from age groups 18-24 and 25-39 years (as compared to age group 2-12 years).

DISCUSSION

Here, we have estimated the seroprevalence of SARS-CoV-2-specific antibodies and identified risk factors for infection and disease in the general population of the Netherlands during the first epidemic wave in April, 2020. To our knowledge, this is the first representative nationwide cohort – covering ages 2-90 years – that links SARS-CoV-2 seropositivity to an extensive set of sociodemographic and clinical parameters, including pre-pandemic sera to confirm specific seroconversion. Given the high fraction of susceptibility throughout the population as of yet, these data can guide future interventions, including strategies for vaccination as that might be one of the most realistic solutions to overcome this pandemic.

The first sampling of this PICO-study revealed that 2·8% [2·0-3·7] of the Dutch population had detectable SARS-CoV-2-specific serum IgG antibodies. Seropositive participants reported to have had COVID-19-related symptoms back in mid-February, suggesting that the virus circulated in our country in the beginning of February already. Our overall estimate is in line with preliminary results from another study conducted in the Netherlands in the beginning of April which found 2·7% to be seropositive, albeit this study was performed in healthy blood donors aged 18-79 years[12]. Worldwide, various seroprevalence studies are ongoing. However, the few current studies in literature mostly cover specific regions or COVID-19 hotspots – with possibly bias in selection of participants and/or smaller age-ranges – with rates ranging between 2-5% in April (e.g., in Los Angeles County (CA, USA)[13], Geneva (Switzerland)[14] and Luxembourg[15]). Estimates also very much depend on test performances. Particularly, when seroprevalence is relatively low, specificity of the assay should approach near 100% to diminish false-positive results and minimize overestimation.[16] Although we cannot rule-out false-positive samples completely, our assay was validated using a broad range of positive and negative (including other coronaviruses) SARS-CoV-2 samples, the PICO-

samples were cross-linked to pre-pandemic sera, and bias correction for test performances was applied to represent the estimates most accurate.

Reports indicate that the incubation time of COVID-19 is approximately 5 days[17] and median seroconversion time of IgG post-symptom onset is 14 days.[18] The seroprevalence estimated here reflects therefore the epidemiological situation of virus spread in the Netherlands in mid-March, directly after the lockdown was proclaimed. Extrapolating our estimate of 2·8% to the Dutch population (17,423,980 inhabitants, March 2020[19]) suggests that almost half a million inhabitants were infected (487,871 [365,904-644,687]). Dutch surveillance data reveal that 11,029 cases, 5,230 hospitalizations and 1,570 deaths were confirmed for persons with a first day of illness up to and including March 20 (i.e., thereby accounting for the median incubation time since mid-March). Applying the total number of estimated infections based on seroprevalence, a rough approximation of the overall infection-hospitalization ratio would be 1 in 93 [70-123] persons (or 1·07% [0·81-1·43]) and an infection-fatality ratio of 1 in 311 [233-411] persons (or 0·32% [0·24-0·43]), which is significantly lower than based on reported cases. Importantly, as the risk of hospitalization – as well as mortality – for COVID-19 increases strongly with age, this ratio is substantially higher for elderly, especially since a proportion of elderly might not have been documented as a confirmed COVID-19 case given data on excess mortality.[3] Overall disease severity might be affected by availability of medical resources and could thus be different per country. Still, these rough estimates are in line with those based on early estimates from Chinese data, by Verity and colleagues[4], or those calculated from Gangelt (Germany)[20], and reveal that overall severity of COVID-19 is substantially higher than e.g., the recent H1N1 influenza pandemic in 2019 (with 0·1%), but lower than SARS-CoV-1.[21]

Seroprevalence was highest in adults aged 18-39 years, which is in line with the serosurvey among blood donors in the Netherlands.[12] The elevation in these younger adults may be explained by

increased social contacts typical for this age group, in addition to specific social activities in February, such as skiing holidays in the Alps (from where the virus disseminated quickly across Europe), or carnival festivities in the Netherlands (multiple superspreading events). Estimates were lowest in children, who seem to be at decreased risk for developing (severe) COVID-19. [22,23] Further, significantly higher odds for seropositivity were seen in Orthodox-Reformed Protestants. This community lives socio-geographically clustered in the Netherlands, i.e., work, school, leisure and church are intertwined heavily. As observed in other countries, particularly frequent attendance of church with close distance to others, including singing activities, might have fueled the spread of SARS-CoV-2 within this community in the beginning of the epidemic[7,8]. Repeated data collection will offer additional insights for policymakers worldwide on whether such gatherings and/or groups need specific attention. Whereas comorbidities were not associated with seropositivity in this study, immunosuppressive drug use did display substantial higher odds (note: we did not have information of specific drugs). These agents are related to increased risk of infection generally as they have the potential to cause lymphopenia and/or impair lymphocyte function.[24] Recent data suggest that immunosuppressive treatment is not associated with worse COVID-19 outcomes[25], yet continued surveillance is warranted as these patients might be more prone to (future) infection, for instance due to a possible attenuated humoral immune response.

The majority of seropositive participants exhibited one or more symptoms, mostly general and respiratory. Asymptomatic proportions reported by clinical studies ranges tremendously (5-80%); the observed overall fraction in the present serology study (7%) might be a conservative estimate as the self-reported symptoms could have been due to other reasons or circulating pathogens along the recalled period (i.e., 62% of the seronegative participants reported symptoms too). Also, attributable proportion of COVID-19 on displaying symptoms might be different across ages and should be explored further. Besides clear evidence of pre-symptomatic transmission[26], the debate on the role of asymptomatic persons in transmission is ongoing and needs to be elucidated via well-

designed contact-tracing studies. Interestingly, clinical studies have observed anosmia/ageusia to be associated with SARS-CoV-2 infection, especially as an early indicator.[27] This notion is also supported here at a population-based level. In the pandemic context, sudden onset of anosmia/ageusia seems to be a useful clinical screening tool besides the existing case definition, which can contribute to early disease recognition and hence minimize transmission by rapid self-isolating.

Questions regarding the duration of protective immunity against SARS-CoV-2, especially in asymptomatic/mild cases that develop lower levels of antibodies as observed here and in previous studies, remain unanswered.[5,28] Nevertheless, higher antibody levels itself do not necessarily mean protection against disease, as other features, such as avidity and neutralizing capacity, but also cell-mediated immunity are expected to contribute to immune protection. To illustrate, for SARS-CoV-1 it was shown that despite the absence of neutralizing antibodies, virus-specific memory T-cells could be identified several years after recovery[29], and strikingly, SARS-CoV-2 reactive $CD4^+$ T cells were detected in 40-60% of pre-pandemic samples recently, signifying potential cross-reactive T-cell recognition from circulating human coronaviruses.[30]

This study has some limitations. First, although our response rate was high (>50%) and inhabitants across half of the total municipalities in the Netherlands were included, some COVID-19 hotspots might be missed – as these predominantly clustered in this epidemic – due to the study design. Estimates on an lower regional level was thus not conceivable. To improve geographical coverage, our second round of inclusion has been extended with an additional sample of persons throughout all municipalities. Second, our study population consisted of more Dutch (88%) than non-Dutch persons as well as relative more healthcare workers (20%) when compared to the general population (76% and 14%, respectively).[19] Although both groups did not display significant higher estimates, and additional selectivity in response was further minimized by weighting our study sample on a set of sociodemographic characters to match the Dutch population, seroprevalence

could be slightly influenced. Third, some potential determinants for seropositivity could have been missed as we might have been underpowered to detect small differences given the low prevalence in this phase or questioned had not been addressed.

To conclude, we estimate that nearly half a million Dutch inhabitants were infected with SARS-CoV-2 amidst the first epidemic wave in the beginning of April, 2020. This is in striking contrast with the reported number of cases, and underlined the underestimation of the true pandemic size without seroprevalence data. Nonetheless, the proportion of persons susceptible to SARS-CoV-2 is high and disease severity is substantial. Globally, nationwide seroepidemiological studies are urgently needed for better understanding of COVID-19 spread, related risk factors, and measures applied to mitigate dissemination.[6] Additionally, the prospective nature of our study will enable us to gain key insights on the duration and quality of antibody responses in infected persons, and hence possible protection of disease by antibodies.[5] Serosurveys will thus play a major role in guiding future interventions, such as strategies for vaccination (of risk groups), since initial vaccine availability will be limiting.

CONTRIBUTORS

(10)(2e) wroth the manuscript. (10)(2e) conceived the study. (10)(2e) conducted the epidemiological data analyses, (10)(2e) provided code for the GAM model and statistical advise, and HdM supervised the analyses. (10)(2e) and (10)(2e) supervised the logistics and performed – among other technicians mentioned in the acknowledgments – the laboratory analyses. (10)(2e) and (10)(2e) designed the laboratory assay and supervised the laboratory analyses. (10)(2e) (10)(2e), and (10)(2e) processed the data. (10)(2e) was principal investigator of the study. All authors contributed to the interpretation of data, and read, edited and approved the final manuscript.

DECLARATION OF INTEREST

We declare no competing interests.

ACKNOWLEDGEMENTS

First of all, we gratefully acknowledge the participants of the PICO-study. Secondly, this study would not have been possible without the instrumental contribution of colleagues from the National Institute of Public Health and Environment (RIVM), Bilthoven, the Netherlands, more specially the department of Immunology of Infectious Disease and Vaccines regarding logistics and/or laboratory analyses ((10)(2e), (10)(2e), (10)(2e), (10)(2e) (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e), and (10)(2e)), the Epidemiology and Surveillance department concerning logistics ((10)(2e), (10)(2e), (10)(2e), (10)(2e), (10)(2e) (10)(2e), (10)(2e), (10)(2e), (10)(2e), and (10)(2e)), methodological insights ((10)(2e), (10)(2e)) (10)(2e), (10)(2e), (10)(2e)

(10)(2e), (10)(2e), and (10)(2e)) and manuscript reviewing ((10)(2e), and (10)(2e)), department of IT and Communication for help with the invitations ((10)(2e) (10)(2e), (10)(2e), and (10)(2e)), student interns for additional support ((10)(2e) for creating FigureS1; (10)(2e), (10)(2e) and (10)(2e) for logistics; and (10)(2e) (10)(2e) for data entry of paper questionnaires), and Sidekick-IT, Breda, the Netherlands, regarding data flow ((10)(2e)). This study was funded by the ministry of Health, Welfare and Sports (VWS), the Netherlands.

DATA SHARING

Our data are accessible to researchers upon reasonable request for data sharing to the corresponding author.

REFERENCES

1. John Hopkins University. COVID-19 Dashboard by the Center for Systems Science and Engineering. 2020. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6 (accessed June 22 2020).

2. World Health Organization (WHO). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed May 12 2020).

3. National Institute for Public Health and the Environment (RIVM). Daily Situational Report of COVID-19 in the Netherlands - June 22, 2020 (in Dutch). Bilthoven: RIVM, 2020.

4. Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis* 2020; **20**(6): 669-77.

5. Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-19 protective immunity. *Lancet* 2020; **395**(10236): 1527-9.

6. Goudsmit J. The paramount importance of serological surveys of SARS-CoV-2 infection and immunity. *Eur J Epidemiol* 2020; **35**(4): 331-3.

7. James A, Eagle L, Phillips C, et al. High COVID-19 Attack Rate Among Attendees at Events at a Church - Arkansas, March 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**(20): 632-5.

8. Yong SEF, Anderson DE, Wei WE, et al. Connecting clusters of COVID-19: an epidemiological and serological investigation. *Lancet Infect Dis* 2020.

9. Verberk JDM, Vos RA, Mollema L, et al. Third national biobank for population-based seroprevalence studies in the Netherlands, including the Caribbean Netherlands. *BMC Infect Dis* 2019; **19**(1): 470.

10. den Hartog G, Schepp RM, Kuijer M, et al. SARS-CoV-2-specific antibody detection for sero-epidemiology: a multiplex analysis approach accounting for accurate seroprevalence. *medRxiv* 2020: 2020.06.18.20133660.

11. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol* 1978; **107**(1): 71-6.

12. Slot E, Hogema BM, Reusken CBEM, et al. Herd immunity is not a realistic exit strategy during a COVID-19 outbreak. *Research Square* 2020 (preprint).

13. 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**(23): 2425-7.

14. 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.

15. Snoeck CJ, Vaillant M, Abdelrahman T, et al. Prevalence of SARS-CoV-2 infection in the Luxembourgish population: the CON-VINCE study. *medRxiv* 2020 (preprint): 2020.05.11.20092916.

16. Farnsworth CW, Anderson NW. SARS-CoV-2 Serology: Much Hype, Little Data. *Clinical Chemistry* 2020.

17. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020; **382**(13): 1199-207.

18. Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clin Infect Dis* 2020.

19. Statistics Netherlands (CBS). Statline database. 2020. https://opendata.cbs.nl/statline/#/CBS/en/ (accessed May 20 2020).

20. Streeck H, Schulte B, Kuemmerer B, et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. *medRxiv* 2020: 2020.05.04.20090076.

21. Riley S, Kwok KO, Wu KM, et al. Epidemiological characteristics of 2009 (H1N1) pandemic influenza based on paired sera from a longitudinal community cohort study. *PLoS Med* 2011; **8**(6): e1000442.

22. She J, Liu L, Liu W. COVID-19 epidemic: Disease characteristics in children. *J Med Virol* 2020; **92**(7): 747-54.

23. Zhang J, Litvinova M, Liang Y, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 2020.

24. Kronbichler A, Gauckler P, Windpessl M, et al. COVID-19: implications for immunosuppression in kidney disease and transplantation. *Nature Reviews Nephrology* 2020; **16**(7): 365-7.

25. Fredi M, Cavazzana I, Moschetti L, et al. COVID-19 in patients with rheumatic diseases in northern Italy: a single-centre observational and case–control study. *The Lancet Rheumatology*.

26. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* 2020; **26**(5): 672-5.

27. Tostmann A, Bradley J, Bousema T, et al. Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020. *Euro Surveill* 2020; **25**(16).

28. Yongchen Z, Shen H, Wang X, et al. Different longitudinal patterns of nucleic acid and serology testing results based on disease severity of COVID-19 patients. *Emerg Microbes Infect* 2020; **9**(1): 833-6.

29. Channappanavar R, Zhao J, Perlman S. T cell-mediated immune response to respiratory coronaviruses. *Immunol Res* 2014; **59**(1-3): 118-28.

30. Grifoni A, Weiskopf D, Ramirez SI, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* 2020.

FIGURES LEGENDS

Figure1. Geographical representation of absolute number of participants in the PICO-study, the Netherlands, first round of inclusion, per municipality. The size of the dots reflect the absolute number of participants. Thicker grey and smaller light grey boundaries represent provinces and municipalities, respectively, and orange and blue boundaries characterize municipalities from the national and low vaccination coverage sample, respectively.

Figure2. Smooth age-specific SARS-CoV-2 seroprevalence in the general population of the Netherlands, beginning of April 2020.