1
2
3
4   **SARS-CoV-2 seropositivity and subsequent infection risk in healthy young adults:**
5   **a prospective cohort study**
6

7   Andrew G. Letizia*, Yongchao Ge*, Sindhu Vangeti*, Carl Goforth*, Dawn L Weir*, Natalia A.
8   Kuzmina, Hua Wei Chen, Dan Ewing, Alessandra Soares-Schanoski, Mary-Catherine George, William D.
9   Graham, Franca Jones, Preeti Bharaj, Rhonda A. Lizewski, Stephen A. Lizewski, Jan Marayag, Nada
10   Marjanovic, Clare Miller, Sagie Mofsowitz, Venugopalan D. Nair, Edgar Nunez, Danielle M. Parent,
11   Chad K. Porter, Ernesto Santa Ana, Megan Schilling, Daniel Stadlbauer, Victor Sugiharto, Michael
12   Termini, Peifang Sun, Russell. P. Tracy, Florian Krammer, Alexander Bukreyev, Irene Ramos, Stuart C.
13   Sealfon
14
15   *Contributed equally
16
17   **Naval Medical Research Center, Silver Spring, MD USA** (A G Letizia MD, C Goforth PhD, D L Weir
18   PhD, H W Chen PhD, D Ewing BS, W D Graham PhD, F Jones PhD, J Marayag BA, E Nunez AS, C K
19   Porter PhD, E Santa Ana AS, M Schilling PhD, V Sugiharto PhD, P Sun PhD)
20
21   **Department of Neurology. Icahn School of Medicine at Mount Sinai, New York, NY USA** (Y Ge
22   PhD, S Vangeti PhD, M-C George PhD, N Marjanovic MS, C Miller PhD, S Mofsowitz BS, V D Nair
23   PhD, A Soares-Schanoski PhD, I Ramos PhD, S C Sealfon MD)
24
25   **Department of Pathology University of Texas Medical Branch and Galveston National Laboratory,**
26   **Galveston, TX USA** (P Bharaj PhD, N A Kuzmina PhD, A Bukreyev PhD)
27
28   **Naval Medical Research Unit SIX, Lima Peru** (R A Lizewski MD, S A Lizewski PhD)
29
30   **Department of Pathology & Laboratory Medicine, Larner College of Medicine, University of**
31   **Vermont, Burlington, VT USA** (D M Parent BS, R P Tracy PhD)
32
33   **Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY USA** (D
34   Stadlbauer PhD, F Krammer PhD)
35
36   **Naval Medical Readiness and Training Command Beaufort, Beaufort, SC USA** (M Termini MD)
37
38   Correspondence to: Dr. Stuart C. Sealfon Icahn School of Medicine at Mount Sinai, Annenberg 14-44,
39   One Gustave L. Levy Place, New York, NY 10029 USA
40

Summary

Background: The risk of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) subsequent infection among seropositive young adults was studied prospectively.

Methods: The study population comprised 3,249 predominantly male, 18-20-year-old Marine recruits. Upon arrival at a Marine-supervised two-week quarantine, participants were assessed for baseline SARS-CoV-2 IgG seropositivity, defined as a 1:150 dilution or greater on receptor binding domain and full-length spike protein enzyme-linked immunosorbent (ELISA) assays. SARS-CoV-2 infection was assessed by PCR at initiation, middle and end of the quarantine. After appropriate exclusions, including participants with a positive PCR during quarantine, we performed three biweekly PCR tests in both seropositive and in seronegative groups once recruits left quarantine and entered basic training and baseline neutralizing antibody titers on all subsequently infected seropositive and selected seropositive uninfected participants.

Findings: Among 189 seropositive participants, 19 (10.1%) had at least one positive PCR test for SARS-CoV-2 during the six-week follow-up (1.1 cases per person-year). In contrast, 1,079 (48.0%) of the 2,247 seronegative participants tested positive (6.2 cases per person-year). The incidence rate ratio was 0.18 (95% CI 0.11-0.28, $p<0.00001$). Among seropositive recruits, infection was associated with lower baseline full-length spike protein IgG titers ($p<0.0001$). Compared with seronegative recruits, seropositive recruits had about 10-fold lower viral loads (ORF1ab gene, $p<0.005$), and trended towards shorter duration of PCR positivity ($p=0.18$) and more frequent asymptomatic infections ($p=0.13$). Among seropositive participants, baseline neutralizing titers were detected in 45 of 54 (83.3%) uninfected and in 6 of 19 (31.6%) infected participants during the 6 weeks of observation (ID50 difference $p<.0001$).

Interpretation: Seropositive young adults had about one-fifth the risk of subsequent infection compared with seronegative individuals. Although antibodies induced by initial infection are largely protective, they do not guarantee effective SARS-CoV-2 neutralization activity or immunity against subsequent infection. These findings may be relevant for optimization of mass vaccination strategies.

Introduction

As of mid-December 2020, more than 72 million severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections have been diagnosed world-wide[1]. Serological surveys indicate that the actual number of infections has been many times higher than the cumulative incidence of diagnosed cases, with seropositivity rates approaching 10% in some countries and more than 40% in the Brazilian Amazon[2-4]. With the onset of mass SARS-CoV-2 vaccination programs and the increasing proportion of previously infected individuals, the risk of reinfection after natural infection is an important question for modeling the pandemic, estimating herd immunity and guiding vaccination strategies[5,6].

Most individuals mount a sustained serologic response after initial infection[7-11]. Similar to the response to other coronaviruses, the production of IgG antibodies against SARS-CoV-2 peaks several weeks after infection, goes through a decline phase and then stabilizes. The overall humoral response to SARS-CoV-2 is highly variable among individuals[9]. SARS-CoV-2 specific IgG antibodies can be detected in serum from most individuals several months after infection[11]. However, a percentage of infected patients, ranging from 2.5% to 28% in different studies, do not maintain detectable circulating antibodies[9] or neutralizing activity[10-12] at later time points. About 10% of individuals who developed antibodies to SARS-CoV-2, all of whom had full-length spike protein-specific IgG antibody titers lower than 1:320, failed to develop measurable neutralizing activity[11].

Reports have established that SARS-CoV-2 reinfection occurs after previous infection as well as in seropositive individuals[13-22]. Several studies have reported that SARS-CoV-2 IgG antibodies[22,23] and neutralizing antibodies[24] provide protection against subsequent infection. The initial trial results of the adenoviral vector-based SARS-CoV-2 vaccine, ChAdOx1 nCoV-19 (AZD1222), noted that among the 373 participants who were seropositive at baseline, 3 (0.8%) had subsequent positive swab PCR tests; in comparison, among 11,263 baseline seronegative participants in all arms of that study, 218 (1.9%) developed a positive test[21]. A study of SARS-CoV-2 serological status and infection among healthcare workers identified 2 infections among 1265 (0.16%) who were seropositive and 223 among 11,364 (2.0%) who were seronegative[22]. Young adults, of whom a high proportion are asymptomatically infected and become seropositive in the absence of known infection[25,26] can be an important source of transmission

97   to more vulnerable populations[27]. Evaluating the protection against subsequent SARS-CoV-2 infection

98   conferred by seropositivity in young adults is important for determining the need for vaccinating

99   previously infected individuals in this age group.

100   We utilized the COVID-19 Health Action Response for Marines (CHARM) study[26], a

101   longitudinal prospective cohort study, to examine the effect of SARS-CoV-2 seropositivity on the risk of

102   developing SARS-CoV-2 infection in young, healthy adult Marine recruits.

103

Methods

Study design and participants

106   The prospective study observation period began when Marine recruits arrived at Marine Corps

107   Recruit Depot – Parris Island (MCRDPI) to commence basic training (Figure 1). To mitigate the spread of

108   SARS-CoV-2, just before transferring to MCRDPI, the United States Marine Corps (USMC)

109   implemented two separate quarantine protocols. The first was a two-week home quarantine. After that,

110   the recruits traveled, while masked and socially distanced, to a second USMC-supervised quarantine

111   situated at either a college campus from May until July 2020 or at a hotel from August until October

112   2020. The supervised quarantine employed extensive public health measures, as previously described[26],

113   that were strictly enforced by US Marine instructors at all times. The recruits and staff were forbidden to

114   leave, and no visitors other than deliveries of supplies and food along with local essential workers and the

115   study staff were allowed onto the premises. At the end of this quarantine period, the USMC required all

116   recruits to test negative for SARS-CoV-2 by PCR before proceeding to MCRDPI to initiate basic training.

117   Within 48 hours of arriving at the supervised quarantine location, recruits were offered the

118   opportunity to volunteer for CHARM. Recruits were eligible if they were ≥18 years of age. Since recruits

119   are a vulnerable population and at risk for coercion, special measures were undertaken including study

120   briefers, who are active-duty Navy personnel wore civilian clothes, did not disclose military ranks, did not

121   have members in the recruit's chain of command present, and ensured that participation would not affect

122   a recruit's medical care or influence the grading of a recruit's military performance by superiors.

123   Institutional Review Board approval was obtained from the Naval Medical Research Center (protocol

124   number NMRC.2020.0006) in compliance with all applicable U.S. federal regulations governing the

125   protection of human subjects. All participants provided written informed consent.

126   At enrollment, participants completed a questionnaire consisting of demographic information,

127   risk-factors, reporting of 14 specific COVID-19 related symptoms or any other unspecified symptom, and

128   brief medical-history. At quarantine weeks 0, 1 and 2 a mid-turbinate nares swab for SARS-CoV-2 PCR

129   testing and sera were obtained and questionnaire administered. The follow up questionnaire inquired

130   about the same COVID-19 related symptoms since the last study visit.

131   Participants who had three negative swab PCR results at the beginning, middle and end of

132   quarantine and a baseline serum serology test at the beginning of the supervised quarantine that identified

133   them as seronegative or seropositive for SARS-CoV-2 according to criteria described below were

134   followed prospectively. At weeks 2, 4 and 6 after transfer from the quarantine location to MCRDPI, a

135   mid-turbinate nares swab for SARS-CoV-2 PCR testing and sera were obtained and the symptom

136   questionnaire administered. When clinically indicated due to the development of symptoms, some

137   participants were evaluated at the MCRDPI clinic and diagnosed by rapid testing. If positive, they went to

138   the isolation barracks, where the study team was able to follow up and repeat testing outside of the

139   scheduled longitudinal follow up encounters. Participants without PCR results obtained during the

140   MCRDPI study period were excluded from analysis.

141

142   **Procedures**

143   **SARS-CoV-2 quantitative PCR testing**

144   All swabs in viral transport media (VTM) were kept at 4°C. All assays were performed within 48h

145   of sample collection at high complexity Clinical Laboratory Improvement Amendments (CLIA)-certified

146   laboratories using the US FDA authorized Thermofisher TaqPath™ COVID-19 Combo Kit (Thermo Fisher

147   Scientific, Waltham MA). Lab24Inc (Boca Raton, FL) performed PCR testing from study initiation until

148   the end of July, and the Naval Medical Research Center (Silver Spring, MD) from August until the

149   conclusion of the study.

SARS-CoV2 enzyme-linked immunosorbent assay (ELISA)

150

151   The presence and levels of IgG SARS-CoV-2-specific antibodies in serum were determined using

152   an enzyme-linked immunosorbent assay (ELISA) as previously described[26]. Briefly, 384-well Immulon 4

153   HBX plates (Thermofisher), or 96-well half area Microlon plates (Greiner Bio-One), were coated

154   overnight at 4°C with recombinant his-tagged spike (S) receptor binding domain (S-RBD)

155   (SinoBiological) or full-length spike protein (LakePharma) at a concentration of 2 µg/ml in phosphate-

156   buffered saline (PBS). Plates were washed three times with 0.1% Tween-20 (Fisher Scientific) in PBS

157   (PBS-T) using an automated ELISA plate washer (Aquamax 4000, Molecular devices), and blocked for 1

158   h at room temperature (RT) with 3% milk (BioRad) PBS-T. Blocking solution was removed, and serum

159   samples diluted in 1% milk PBS-T were dispensed in the wells. At least 2 positive controls (sera with

160   known IgG presence), 8 negative controls (sera collected before July 2019) and 4 blanks (no serum) were

161   included in every assay. Plates were incubated for 2 h at room temperature, and then washed 3 times with

162   PBS-T. Next, peroxidase conjugated goat F(ab')2 Anti-Human IgG (Abcam) was added at a dilution

163   1:5,000-1:10,000 dilutions (determined after optimization for each antibody lot) in 1% milk PBS-T, and

164   plates were incubated for 1 h at room temperature. Plates were washed 6 times with PBS-T, developed

165   using o-Phenylenediamine (Sigma), and the reaction was stopped after 10 min with 3M HCl. Optical

166   density (OD) at 492 nm was measured using a microplate reader (SpectramaxM2, Molecular Devices).

167   All serum samples were screened at a 1:50 dilution with S-RBD. Those samples with an OD 492 nm

168   value higher than the average of the negative controls plus 3 times their standard deviation (SD) in the

169   screening assay underwent titration assay (6 serial 1:3 serum dilutions starting at 1:50) using both S-RBD

170   and full-length spike protein. Serum samples were considered positive for each assay when at least 2

171   consecutive dilutions showed higher OD 492 nm than the average of the negative controls plus 3 times

172   their SD at the correspondent dilution or 0.15 OD 492 nM. Specificity was 100% on both S-RBD and

173   full-length spike protein ELISA using 70 control sera obtained before July 2019. Participants were only

174   considered seropositive to SARS-CoV-2 if IgG titrations for both ELISA gave a positive result at a

175   minimum of 1:150 dilution.

176

Neutralization assays

Two-fold serial dilutions of heat-inactivated serum at an initial dilution of 1:20, were prepared in serum free media (Minimum Essential Medium; Thermo fisher Scientific, Cat No. 11095080 containing 25 mM HEPES and 0.05 g/L gentamicin sulfate) and incubated with an equal volume of mNeonGreen SARS-CoV-2[28] for 1 hour at 37°C at a final concentration of 200 plaque forming units in humidified 5% CO_2. Virus-serum mixtures were then added to Vero-E6 monolayers in 96 well optical black plates and incubated at 37°C. Plates were read using the BioTek Cytation 5 plate reader (EX 485 nm, EM 528 nm) at 24 h post-infection. Following background signal correction, neutralization titers at a fluorescent end point of 50% virus reduction (ID_{50}) were determined.

Statistical Analysis

Analyses, figures and tables were generated using R 3.6.3. Race was categorized as non-Hispanic white, non-Hispanic black, non-Hispanic other, and Hispanic. Cochran-Armitage Chi-square test for trend was used to compare proportions testing positive for SARS-CoV-2 by increasing titers. Cumulative incidence rates computed by Kapan-Meir method were used to estimate the risk of SARS-CoV-2 infection between seropositive and seronegative participants and also between different titers among the seropositive participants. Observational follow-up began upon arrival at MCRDPI and participants were censored at the first observed positive PCR, or at the latest time with valid PCR assay or at the termination of the study (6 weeks of observation). The Cox proportional hazards model controlled for age, sex and race. The p-values from the cumulative incidence curves were determined by the log-rank test and the p-value and CI for Cox proportional hazard model was computed by the R function coxph. The cycle threshold (Ct) values of viral genes were quantile normalized to remove the batch effects between Lab24Inc and Naval Medical Research Center laboratory assays. The CI and p-values for comparing the mean of the Ct values for two groups are computed based on the two sample t-test.

203 **Role of the Funding Source**

204 The funders had no role in the design of the protocol, data collection, data management, data analysis,

205 data interpretation, or writing of the report. SCS, AGL, YG, CP and IR had access to all deidentified data.

206 SCS and AGL had final responsibility for the decision to submit for publication.

207

Results

209   The flow of participants through the study is shown in Figure 1. Two weeks before the initiation

210 of basic training and the start of the prospective cohort study period, a total of 3,249 out of 4,657 eligible

211 new Marine recruits (70%) enrolled in CHARM and underwent a supervised two-week quarantine.

212 Exclusions before the prospective study period included 73 participants who were SARS-CoV-2 PCR

213 positive on at least one of the three PCR tests performed during quarantine, 53 who lacked baseline

214 serology results and 47 who were lost to follow-up.

215   Of the remaining 3,076 participants, 225 were baseline seropositive, having SARS-CoV-2 IgG

216 titers in serum samples obtained at the beginning of quarantine that were greater than 1:150 both with S-

217 RBD and with full-length spike protein ELISA, and 2,851 were baseline seronegative. Among the

218 seropositive participants, 36 (16.0%) were excluded from analysis due either to being lost to follow up

219 (n=34), or to lacking valid PCR results during the study period (n=2). In the seronegative group, 604

220 (21.2%) participants were excluded due to either being lost to follow up (n=532), or lacking valid PCR

221 results (n=72). Participants were lost to follow up for specific reasons unknown to the study team,

222 including dropping out of the study, being separated from the Marines or being removed from the base for

223 medical or administrative reasons. Demographic characteristics of the two groups in the study cohort are

224 shown in Table 1. Most participants were 18-20 years old and male. The two groups were well-balanced

225 with the exception of a higher proportion of participants who self-identified as Hispanic and as black in

226 the seropositive group.

227   A total of 19 out of 189 (10.1%, 1.1 cases per person-year) seropositive participants and 1,079 out

228 of 2,247 (48.0%, 6.2 cases per person-year) seronegative participants had at least one positive SARS-

229 CoV-2 PCR result during the six-week study period, representing an 0.18 (95% CI: 0.11 to 0.28, p<0.001)

230   incidence rate ratio for SARS-CoV-2 infections in the seropositive group (Table 2). The temporal

231   incidence of infection in seropositive and seronegative groups is shown in Figure 2A. After adjusting the

232   effects of race, age and sex on the SARS-CoV-2 infections, the hazard ratio comparing seropositive

233   participants and seronegative participants was 0.16 (95% CI, 0.10-0.25, p<0.001; supplementary Table 1).

234   Within the seropositive group, we assessed the association between the SARS-CoV-2 IgG

235   baseline titers and the risk of infection. As shown in Figure 2B, supplementary Figure 2 and Table 3, we

236   found a strong association between subsequent PCR positive infection and lower titers of IgG antibodies

237   directed to full-length spike protein (p<0.0001) as well as to S-RBD (p=0.0019). The detailed Cox

238   proportional hazard analysis in supplementary Table 1 gives a hazard ratio of 0.45 (95% CI 0.32-0.65,

239   p<0.001) and 0.67 (95% CI 0.47-0.96, p=0.028) respectively for the full-length spike protein titer and S-

240   RBD titer, both log-transformed.

241   We examined baseline SARS-CoV-2 IgG neutralizing antibody activity in all seropositive

242   participants who became PCR positive during the observation period and in the first 54 participants who

243   were seropositive but remained PCR negative. Neutralizing activity was above the limit of detection in 45

244   of 54 (83.3%) seropositive participants who never became PCR positive and in 6 of 19 (31.6%) of

245   participants infected during the 6 weeks of observation. The neutralizing activity assessed as 50%

246   inhibitory dose (ID50) was significantly higher in the participants who did not become PCR positive

247   during the study (p<.0001, Cochran-Armitage test, Table 3, supplementary Figure 2).

248   We also compared virus loads estimated by PCR Ct values between the seronegative and

249   seropositive PCR infected groups and found that seronegative individuals had on average 4, 2.6 and 3.3

250   lower cycle values for ORF1ab gene (p=0.004, 95% CI 1.23 to 6.67), S gene (p=0.11, 95% CI -0.58 to

251   5.77) and N gene (p=0.033, 95% CI 0.27 to 6.33), respectively, than seropositive individuals. The lower

252   Ct values suggest an approximately 10-fold higher virus load in the samples from seronegative

253   participants (Table 4). In addition, seronegative participants tended to have a longer duration of PCR

254   positivity than seropositive individuals (p=0.18). The proportion of asymptomatic participants was 84.2%

255   (16 of 19) and 67.8% (732 of 1079) in seropositive and seronegative infected individuals, respectively, a

256   difference that did not reach significance (p=0.13).

257

Discussion

259   This study of primarily young, male Marine recruits found that the presence of antibodies to

260   SARS-CoV-2 conferred an 82% reduced incidence rate of SARS-CoV-2 infection. The percentage of

261   symptomatic infections was comparable in seropositive and seronegative participants. Among the

262   seropositive group, the participants that became infected had lower antibody titers than those that were

263   uninfected, and were more likely to lack detectable baseline neutralizing antibody activity. Our results

264   indicate that although antibodies induced by infection are largely protective, they do not guarantee

265   effective immunity against subsequent infection.

266   This study leveraged a two-week USMC-mandated quarantine period during which baseline

267   SARS-CoV-2 antibody status was established on arrival. Only baseline seronegative or seropositive

268   participants who had multiple negative SARS-CoV-2 PCR tests during the supervised quarantine were

269   included in the prospective study. The three negative PCR tests during quarantine helped ensure that

270   infections diagnosed during basic training were not persistent infections but incident infection occurring

271   during the prospective period. The two-week home quarantine preceding the supervised quarantine, as

272   well as the relatively low frequency of infections diagnosed on arrival and during quarantine, further

273   support that only incident infections were included in our analyses. The aggregate infection rate in both

274   groups during the six weeks of observation at MCRDPI was 1,098 out of 2,436 (45.1%) participants. In

275   contrast, only 28 of 3,249 (0.9%) participants were PCR positive on arrival at the supervised quarantine

276   and less than 2% of participants became PCR positive during the two-week quarantine period. In view of

277   the consecutive two periods of quarantine, the relatively low rate of infection during quarantine, and the

278   three consecutive negative PCR tests, it is unlikely that any participants with persistent infection

279   preceding their arrival at MCRDPI were entered into the prospective study. Even if this occurred, such

280   cases would be unlikely to affect the relative risk calculations comparing the seropositive and

281   seronegative groups. This methodology allowed for the creation of two well defined groups who entered

282   basic training without active infection and differed primarily by baseline serology.

283   The two groups had similar demographic profiles, with the exception of a slightly higher

284 prevalence in self-identified Hispanic and non-Hispanic black participants in the seropositive group. The

285 seropositive group was almost 50% Hispanic and 22% non-Hispanic black participants compared to 22%

286 and 12%, respectively in the seronegative group. This is likely due to minority populations having higher

287 seroprevalence rates during the COVID-19 pandemic in general and among young adults specifically[27].

288   Rates of infection and the risk reduction provided by seropositivity are important for

289 understanding transmission dynamics for COVID-19, for epidemiologic modeling, and for estimating and

290 achieving herd immunity levels, a major goal of mass vaccination strategies. Herd immunity is difficult to

291 predict if the infection risk after natural and vaccine-induced immunity is unknown. Since SARS-CoV-2

292 vaccines may not provide sterile immunity, it is possible that both previously infected and vaccinated

293 individuals may later become infected. It is not known if either can contribute to transmission events. We

294 found only a modest, approximately 10-fold decrease in nares virus load as estimated by swab PCR Ct

295 levels in the seropositive compared with the seronegative infected participants. This finding suggests that

296 some reinfected individuals could have a similar capacity to transmit infection as those who are infected

297 for the first time. The rate at which reinfection occurs after vaccines and natural immunity are important

298 for estimating the percentage of the population that needs to be vaccinated to suppress the pandemic.

299   The crowded living conditions, demanding regimen and requirement for personal contact during

300 basic training despite the pandemic leads not only to an increased risk for respiratory epidemics[29], but also

301 potentially to higher exposure levels. The close quarters and constant contact among recruits that are

302 needed for team building allows a viral infection to rapidly proliferate within a unit. The physically and

303 mentally demanding training environment may also suppress immunity. These conditions may contribute

304 to the high infection rate we observed during the six-week study period. These factors are not typically

305 present in the civilian community. Therefore, the study setting limits the generalizability of our findings

306 to other settings where the frequency and intensity of exposure and the susceptibility of the host might

307 differ.

308   The clinical outcomes between seropositive and seronegative groups were similar, with the

309 majority (84.2% and 67.8%, respectively) being asymptomatic. The two groups did not show a significant

310   difference in the duration of PCR positivity. No participants in either group needed inpatient care.

311   Although our findings are limited to healthy young adults, studying this population does have the

312   advantage of reducing the confounding factors of age and co-morbid illnesses[30]. Infection in seropositive

313   participants was associated with lower SARS-CoV-2 IgG titers and absent or lower levels of neutralizing

314   antibody activity. Young adults have high rates of asymptomatic and pauci-symptomatic infection, which

315   has been associated with lower levels of antibodies and potentially a less robust immune memory

316   response. This could lead to higher overall rates of re-infection among this population compared to other

317   populations. We did not examine the role of cell mediated immunity or host, environment and virus

318   factors leading to reinfection.

319   Since the study population is a fairly accurate representation of the races and ethnicities in the US

320   population among 18-20-year-olds, the results are most applicable to young male adults. The relative risk

321   of infection may be different in seropositive females and in adults of other ages or health status. Other

322   limitations in our study include not being able to investigate the exposure event during a seropositive

323   participant's initial infection prior to arrival at quarantine, the inability to confirm initial SARS-CoV-2

324   infection by PCR in the seropositive group, and potentially missing detectable infections that occurred

325   between biweekly sampling.

326   Our investigation is likely to underestimate the risk of SARS-CoV-2 infection in previously

327   infected individuals because the seronegative group includes an unknown number of previously infected

328   participants who did not have significant IgG titers in their baseline serum sample. Despite this

329   underestimation, we found that previously infected participants identified by seropositivity are susceptible

330   to repeat infection, with nearly one-fifth the incidence rate of those without evidence of previous

331   infection. This suggests that COVID-19 vaccination may be necessary for control of the pandemic in

332   previously infected young adults.

333

334   **Contributor**s

335   AGL, IR, and SCS provided overall leadership and guidance to the investigation. YG performed

336   statistical analysis and data curation. SV, CKP, CG, DLW, RL, and IR contributed to data analysis. RL,

337   JM, EN, ESA, SM, NM, contributed to sample collection and management. CG, DLW, WDG contributed

338   to data collection. SV, NAK, DE, AS, PB, NM, CM, DMP, SM, DS, PS performed serological assays.

339   HWC, and VS performed PCR assays. AGL, SV, DLW, NAK, AS, RL, CM, CKP, MT, and IR

340   contributed to data interpretation. AGL, CG, FJ, RL, EN supervised sample collection. AGL, CG, RL

341   supervised data collection. RPT, FK, AB, and IR supervised serology assays. MS supervised PCR assays.

342   MG and MT provided project administration and coordination. YG, SV, IR, CKP and VDN prepared

343   figures. AGL, YG, SV, IR and SCS wrote and revised the original draft. All authors reviewed and

344   approved the final manuscript.

345

346   **Declaration of Interests**

347   AGL, YG, SV, CG, DLW, HWC, DE, MCG, WDG, FJ, RL, JM, NM, CM, SM, VDN, EN, CKP, ESA,

348   ASS, MS, VS, MT, PS, RPT, IR and SCS declare no competing interests. DP owned stock in Co-

349   Diagnostics Inc during the conduct of the study. DP held stock in Co-Diagnostics, Inc. DS and FK have

350   filed a patent regarding serological assays and for SARS-CoV-2 and Icahn School of Medicine at Mount

351   Sinai has founded a company to commercialize serological assays they developed. AGL, CG, DLW,

352   HWC, DEW, DG, FJ, JM, EN, CKP, ESA, MS, VS, PS, RAL, SAL, and MT are military Service

353   members or GS employees. This work was prepared as part of their official duties. Title 17, U.S.C., §105

354   provides that copyright protection under this title is not available for any work of the U.S. Government.

355   Title 17, U.S.C., §101 defines a U.S. Government work as a work prepared by a military Service member

356   or employee of the U.S. Government as part of that person's official duties. The views expressed in the

357   article are those of the authors and do not necessarily express the official policy and position of the US

358   Navy, the Department of Defense, the US government or the institutions affiliated with the authors.

359
360   **Acknowledgements**
361

372   **Table and Figure legends**

373   **Table 1: Demographic characteristics of participants studied prospectively in the seropositive and**

374   **seronegative groups.** A total of 120 participants are excluded due to being lost to follow up, lacking any

375   baseline valid IgG or becoming PCR positive during the quarantine period. The table includes all 3,076

376   participants who entered training and were followed prospectively, including the 640 participants who

377   were later excluded for further analysis

378

379   **Table 2: Comparison of SARS-Cov-2 infection (PCR positive) at MCRDPI between the seropositive**

380   **and seronegative groups.** The analysis is based on the 2,436 participants who had valid PCR data

381   obtained during the prospective follow-up period.

382

383   **Table 3. SARS-CoV-2 S-RBD and full-length spike IgG titers and neutralizing antibody activity in**

384   **PCR positive and PCR negative seropositive participants.** p-values were assessed by the Cochran-

385   Armitage test. ID50 is the titer at which a 50% reduction in virus infection was observed.

386

387   **Table 4: Comparison of symptoms and cycle threshold values between SARS-CoV-2 infected (PCR**

388   **positive) seropositive and seronegative groups.**

389

390   **Figure 1: Flow chart of the study design and outcomes.** The baseline analyses were performed within

391   two days of enrollment and two weeks before the prospective study period. Nearly all participants were

392   tested at scheduled biweekly intervals during the prospective study period. A few participants were

393   diagnosed by the MCRDPI clinic, and were tested by the study team at times not corresponding to the

394   regularly scheduled longitudinal follow ups. Participants lost to follow up either dropped out of the study,

395   were separated from the Marine Corps or were removed from the base for medical or administrative

396   reasons. The study team did not know the reason for participants missing study visits.

397

398   **Figure 2: SARS-CoV-2 PCR positive incidence curves during the six-week follow-up period**. A.

399   Overall cumulative incidence for testing PCR positive in the baseline seropositive and seronegative

400   groups. B. Cumulative incidence for testing PCR positive in the seropositive group at different baseline

401   full-length spike protein IgG titers which ranged from 1:150 to 1:12150. The charts are derived from

402   Kaplan-Meier survival analyses and the cumulated censored participants are listed.

403

404

405

406

407

References

408

409   1.   Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in

410   real time. *Lancet Infect Dis* 2020; **20**(5): 533-4.

411   2.   Buss LF, Prete CA, Jr., Abrahim CMM, et al. Three-quarters attack rate of SARS-CoV-2

412   in the Brazilian Amazon during a largely unmitigated epidemic. *Science* 2020.

413   3.   Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain

414   (ENE-COVID): a nationwide, population-based seroepidemiological study. *The Lancet* 2020;

415   **396**(10250): 535-44.

416   4.   Stadlbauer D, Tan J, Jiang K, et al. Repeated cross-sectional sero-monitoring of SARS-

417   CoV-2 in New York City. *Nature* 2020.

418   5.   Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-

419   19 protective immunity. *Lancet* 2020; **395**(10236): 1527-9.

420   6.   Overbaugh J. Understanding protection from SARS-CoV-2 by studying reinfection. *Nat*

421   *Med* 2020; **26**(11): 1678-9.

422   7.   Beaudoin-Bussieres G, Laumaea A, Anand SP, et al. Decline of Humoral Responses

423   against SARS-CoV-2 Spike in Convalescent Individuals. *mBio* 2020; **11**(5).

424   8.   Gudbjartsson DF, Norddahl GL, Melsted P, et al. Humoral Immune Response to SARS-

425   CoV-2 in Iceland. *N Engl J Med* 2020; **383**(18): 1724-34.

426   9.   Self WH, Tenforde MW, Stubblefield WB, et al. Decline in SARS-CoV-2 Antibodies

427   After Mild Infection Among Frontline Health Care Personnel in a Multistate Hospital Network -

428   12 States, April-August 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**(47): 1762–6.

429   10.   Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing

430   antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat*

431   *Microbiol* 2020; **5**(12): 1598-607.

432   11.   Wajnberg A, Amanat F, Firpo A, et al. Robust neutralizing antibodies to SARS-CoV-2

433   infection persist for months. *Science* 2020; **370**(6521): 1227-30.

434   12.   Crawford KHD, Dingens AS, Eguia R, et al. Dynamics of neutralizing antibody titers in

435   the months after SARS-CoV-2 infection. *J Infect Dis* 2020.

436   13.   Chan PKS, Lui G, Hachim A, et al. Serologic Responses in Healthy Adult with SARS-

437   CoV-2 Reinfection, Hong Kong, August 2020. *Emerg Infect Dis* 2020; **26**(12): 3076-8.

438   14.   Huang J, Zheng L, Li Z, et al. Kinetics of SARS-CoV-2 positivity of infected and

439   recovered patients from a single center. *Sci Rep* 2020; **10**(1): 18629.

440   15.   Lee JS, Kim SY, Kim TS, et al. Evidence of Severe Acute Respiratory Syndrome

441   Coronavirus 2 Reinfection After Recovery from Mild Coronavirus Disease 2019. *Clin Infect Dis*

442   2020.

443   16.   Lu J, Peng J, Xiong Q, et al. Clinical, immunological and virological characterization of

444   COVID-19 patients that test re-positive for SARS-CoV-2 by RT-PCR. *EBioMedicine* 2020; **59**:

445   102960.

446   17.   Mulder M, van der Vegt D, Oude Munnink BB, et al. Reinfection of SARS-CoV-2 in an

447   immunocompromised patient: a case report. *Clin Infect Dis* 2020.

448   18.   Munoz Mendoza J, Alcaide ML. COVID-19 in a patient with end-stage renal disease on

449   chronic in-center hemodialysis after evidence of SARS-CoV-2 IgG antibodies. Reinfection or

450   inaccuracy of antibody testing. *IDCases* 2020; **22**: e00943.

451   19.   Selhorst P, Van Ierssel S, Michiels J, et al. Symptomatic SARS-CoV-2 reinfection of a

452   health care worker in a Belgian nosocomial outbreak despite primary neutralizing antibody

453   response. *Clin Infect Dis* 2020.

454   20.   Van Elslande J, Vermeersch P, Vandervoort K, et al. Symptomatic SARS-CoV-2

455   reinfection by a phylogenetically distinct strain. *Clin Infect Dis* 2020.

456   21.   Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-

457   19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled

458   trials in Brazil, South Africa, and the UK. *The Lancet* 2020.

459   22.   Lumley SF, O'Donnell D, Stoesser NE, et al. Antibody Status and Incidence of SARS-

460   CoV-2 Infection in Health Care Workers. *New England Journal of Medicine* 2020.

461   23.   Harvey RA, Rassen JA, Kabelac CA, et al. Real-world data suggest antibody positivity to

462   SARS-CoV-2 is associated with a decreased risk of future infection. *medRxiv* 2020:

463   2020.12.18.20248336.

464   24.   Addetia A, Crawford KHD, Dingens A, et al. Neutralizing Antibodies Correlate with

465   Protection from SARS-CoV-2 in Humans during a Fishery Vessel Outbreak with a High Attack

466   Rate. *J Clin Microbiol* 2020; **58**(11): e02107-20.

467   25.   Letizia AG, Ge Y, Goforth CW, et al. SARS-CoV-2 seropositivity among US Marine

468   recruits *Emerging Infectious Diseases* 2021; **revision awaiting decision**.

469   26.   Letizia AG, Ramos I, Obla A, et al. SARS-CoV-2 Transmission among Marine Recruits

470   during Quarantine. *N Engl J Med* 2020; **383**(25): 2407-16.

471   27.   De Sanctis V, Ruggiero L, Soliman AT, Daar S, Di Maio S, Kattamis C. Coronavirus

472   Disease 2019 (COVID-19) in adolescents: An update on current clinical and diagnostic

473   characteristics. *Acta Biomed* 2020; **91**(2): 184-94.

474   28.   Xie X, Muruato A, Lokugamage KG, et al. An Infectious cDNA Clone of SARS-CoV-2.

475   *Cell Host Microbe* 2020; **27**(5): 841-8 e3.

476   29.   Gray GC, Callahan JD, Hawksworth AW, Fisher CA, Gaydos JC. Respiratory diseases

477   among U.S. military personnel: countering emerging threats. *Emerg Infect Dis* 1999; **5**(3): 379-

478   85.

479   30.   Frasca D, Blomberg BB. Aging induces B cell defects and decreased antibody responses

480   to influenza infection and vaccination. *Immun Ageing* 2020; **17**(1): 37.

481

Label	Levels	Seropositive	Seronegative	Total
Total N (%)		225 (7.3)*	2851 (92.7)	3076
Age	Mean (SD)	19.0 (1.8)†	19.1 (1.9)	19.1 (1.9)
Sex	F	22 (9.8)	229 (8.0)	251 (8.2)
	M	203 (90.2)	2622 (92.0)	2825 (91.8)
Race	Non-Hispanic White	56 (24.9)	1698 (59.6)	1754 (57.0)
	Non-Hispanic Black	50 (22.2)	349 (12.2)	399 (13.0)
	Non-Hispanic Other	7 (3.1)	183 (6.4)	190 (6.2)
	Hispanic	112 (49.8)	621 (21.8)	733 (23.8)
Non-US residence	No	216 (96.0)	2749 (96.4)	2965 (96.4)
	Yes	4 (1.8)	19 (0.7)	23 (0.7)
	N/A‡	5 (2.2)	83 (2.9)	88 (2.9)
Non-US birth	No	189 (84.0)	2626 (92.1)	2815 (91.5)
	Yes	32 (14.2)	192 (6.7)	224 (7.3)
	N/A	4 (1.8)	33 (1.2)	37 (1.2)

482 **Table 1: Demographic characteristics of participants studied prospectively in the seropositive and**
483 **seronegative groups.** A total of 120 participants were excluded due to being lost to follow up, lacking
484 any baseline valid IgG or becoming PCR positive during the quarantine period. The table includes all
485 3,076 participants who entered training and were followed prospectively, including the 640 participants
486 who were later excluded for further analysis (see Figure 1).
487 * The continuous variable age is summarized with mean (SD)
488 † Categorical variables are summarized with counts (percentages)
489 ‡ If a participant answered unknown or left a question blank, then the value is grouped into "N/A".
490

491

Group	Seropositive	Seronegative
Total	189	2247
PCR positive	19	1079
Incidence proportion (%)	10.1	48
Proportion difference (%)	-38 (-45 to -31, p<0.001)	
Observed person year	17.1	175.2
Incidence rate	1.11	6.16
Incidence rate ratio	0.18 (0.11 to 0.28, p<0.001)	

492
493 **Table 2: Comparison of SARS-Cov-2 infection (PCR positive) at MCRDPI between the seropositive**
494 **and seronegative groups.** The analysis is based on the 2,436 participants who had valid PCR data
495 obtained during the prospective follow-up period.
496
497

Assay	Group	N	No. (%) of people at the titer or ID50 interval					mean±sem	p-value
titer			150	450	1350	4050	12150		
S-RBD	PCR+	19	5(26.3)	5(26.3)	6(31.6)	3(15.8)	0(0)	674.5±181.1	0.017
	PCR-	170	6(3.5)	49(28.8)	81(47.6)	27(15.9)	7(4.1)	1186.3±86.2	
Full-length spike protein	PCR+	19	6(31.6)	2(10.5)	2(10.5)	6(31.6)	3(15.8)	1202.6±472.7	1.20E-05
	PCR-	170	3(1.8)	2(1.2)	41(24.1)	83(48.8)	41(24.1)	3723.7±259.9	
ID50 range			<20	[20,40)	[40,80)	[80,160)	[160,320)		
Neutralization	PCR+	19	13(68.4)	2(10.5)	3(15.8)	1(5.3)	0(0)	16.8±3.1*	9.80E-05
	PCR-	54	9(16.7)	10(18.5)	19(35.2)	15(27.8)	1(1.9)	48.2±5.7	

498
499 **Table 3. SARS-CoV-2 S-RBD and full-length spike IgG titers and neutralizing antibody activity in**
500 **PCR positive and PCR negative seropositive participants.** p-values were assessed by the Cochran-
501 Armitage test. ID50 is the titer at which a 50% reduction in virus infection was observed.
502
503 *The mean and SEM is computed after converting undetectable titer <20 to be 10.
504

	Seropositive	Seronegative	Difference comparison
Total N	**19**	**1079**	
Infection duration>7 days	6 (31.6%)*	510 (47.3%)	-0.16 (-0.38 to 0.07, p=0.175)†
Symptomatic	3 (15.8%)	347 (32.2%)	-0.16 (-0.38 to 0.05, p=0.129)
N gene Ct	27.7 (7.6) ‡	24.4 (5.5)	3.30 (0.27 to 6.33, p=0.033) ¶
S gene Ct	26.9 (7.1)	24.3 (5.3)	2.60 (-0.58 to 5.77, p=0.109)
ORF1ab Ct	28.0 (7.0)	24.0 (5.3)	3.95 (1.23 to 6.67, p=0.004)

505
506 **Table 4: Comparison of symptoms and cycle threshold values between SARS-CoVCov-2 infected**
507 **(PCR positive) seropositive and seronegative groups.**
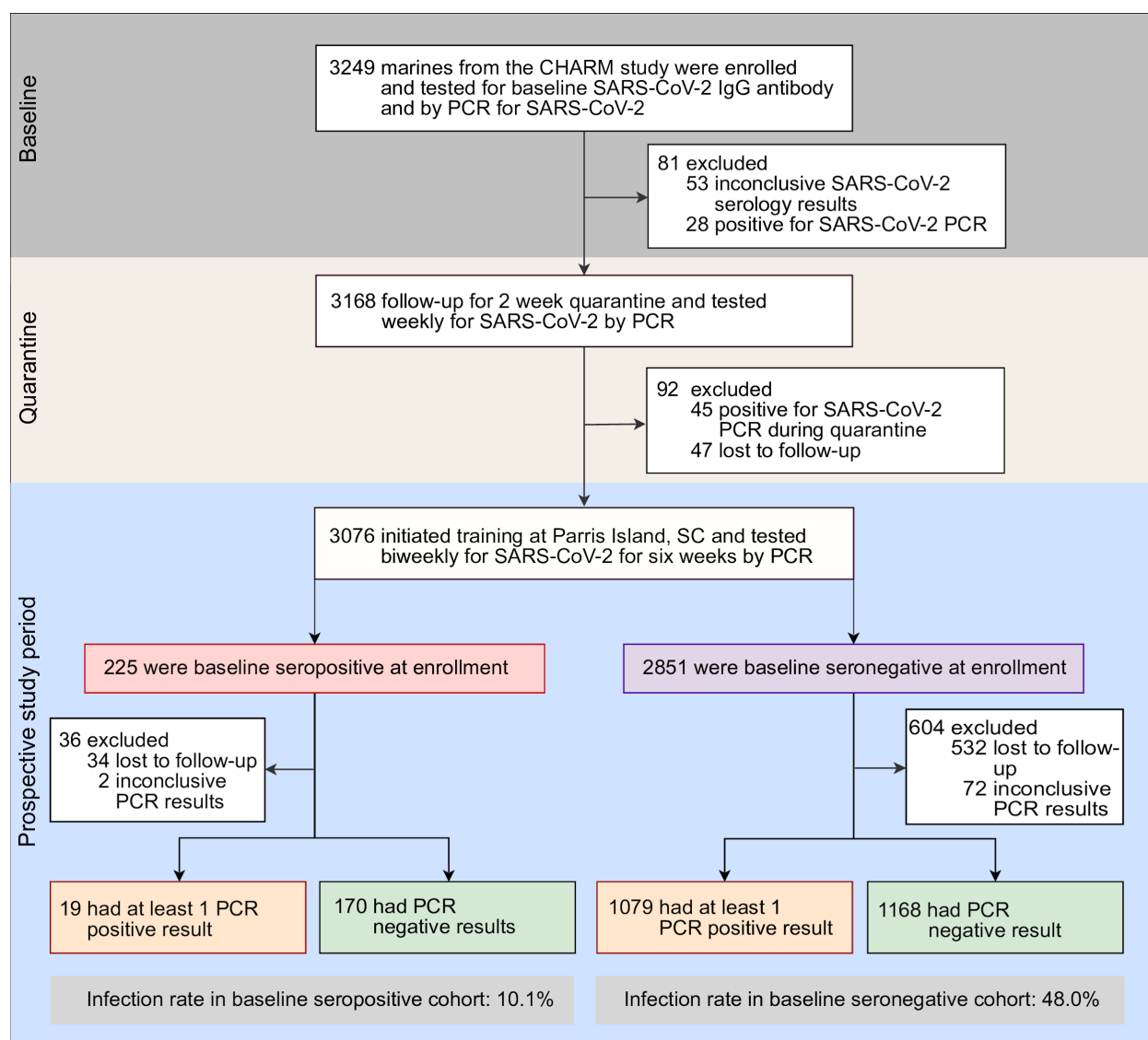508 * Binary variables are summarized with counts (percentages)
509 † Difference in proportion for binary variables
510 ‡ The cycle threshold (Ct) values of three viral genes are summarized by mean (SD)
511 ¶ Difference in mean for the Ct values
512
513

Figure 1: Flow chart of the study design and outcomes. The baseline analyses were performed within two days of enrollment and two weeks before the prospective study period. Nearly all participants were tested at scheduled biweekly intervals during the prospective study period. A few participants were diagnosed by the Marine Corps Recruit Depot Parris Island clinic, and were tested by the study team at times not corresponding to the regularly scheduled longitudinal follow ups. Participants lost to follow up either dropped out of the study, were separated from the Marine Corps or were removed from the base for medical or administrative reasons. The study team did not know the reason for participants missing study visits.
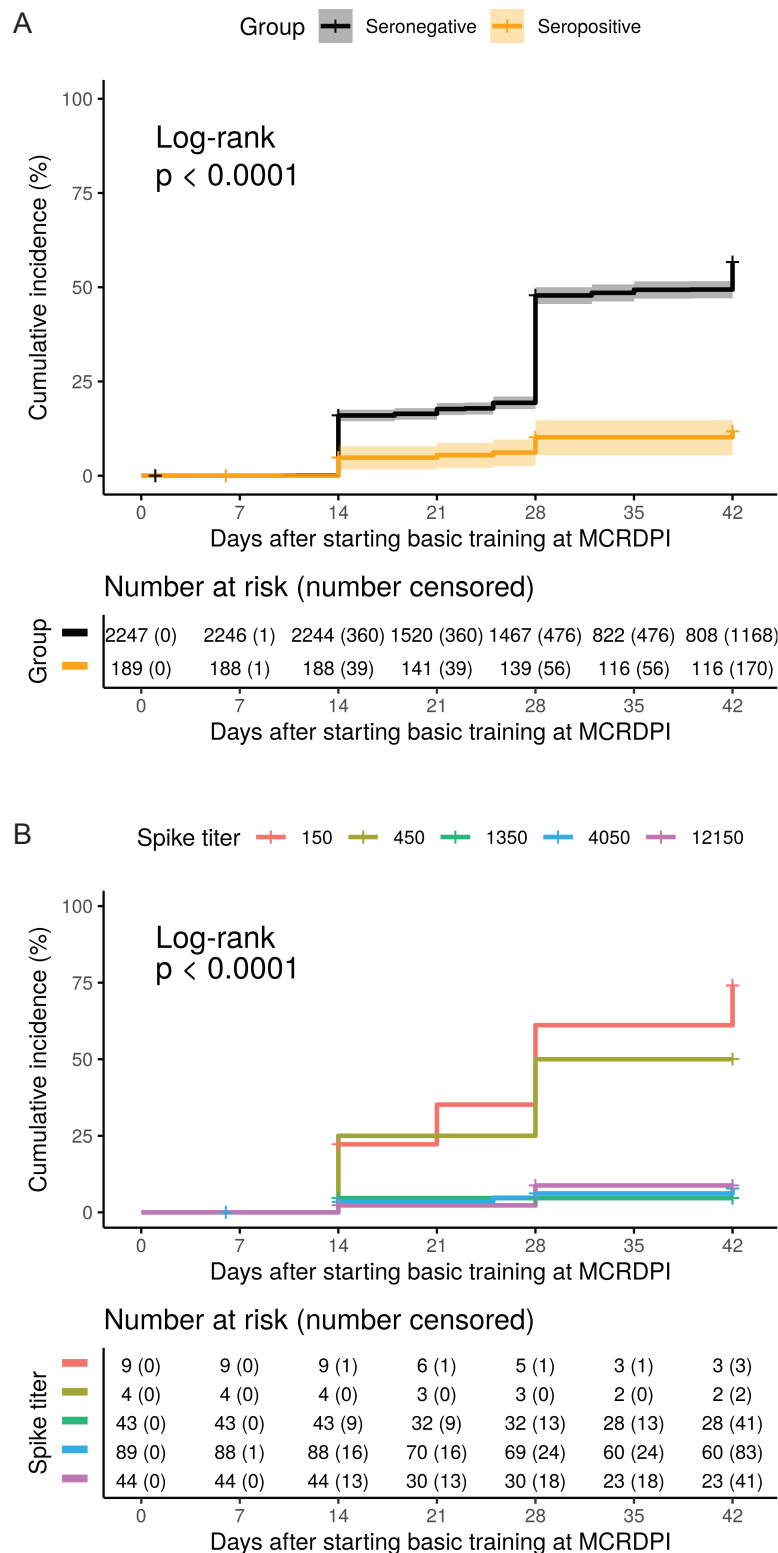
Figure 2: SARS-CoV-2 PCR positive incidence curves during the six-week follow-up period. A. Overall cumulative incidence for testing PCR positive in the baseline seropositive and seronegative groups. B. Cumulative incidence for testing PCR positive in the seropositive group at different baseline full-length spike protein IgG titers which ranged from 1:150 to 1:12150. The cumulative incidence rate is computed by the Kaplan-Meier method, and the cumulated censored participants are listed.

532 **Supplementary Tables and Figures**

			N	Hazard ratio (univariable)	Hazard ratio (multivariable)
Group Analysis*	Group	Seronegative	2247	-‡	-
		Seropositive	189	0.16 (0.10-0.25, p<0.001)	0.16 (0.10-0.25, p<0.001)
	Age	Mean (SD)	19.1 (1.9)	0.97 (0.94-1.00, p=0.088)	0.97 (0.94-1.00, p=0.070)
	Sex	F	217	-	-
		M	2219	1.03 (0.85-1.25, p=0.746)	1.06 (0.87-1.29, p=0.541)
	Race	Non-Hispanic White	1352	-	-
		Non-Hispanic Black	324	0.75 (0.62-0.91, p=0.004)	0.86 (0.70-1.04, p=0.120)
		Non-Hispanic Other	162	0.97 (0.77-1.22, p=0.793)	0.96 (0.76-1.22, p=0.763)
		Hispanic	598	0.94 (0.81-1.08, p=0.360)	1.08 (0.93-1.24, p=0.324)
S-RBD Titer Analysis†	Titer	Mean (SD)	10.1 (1.4)	0.67 (0.48-0.94, p=0.022)	0.67 (0.47-0.96, p=0.028)
	Age	Mean (SD)	19.0 (1.8)	1.00 (0.78-1.27, p=0.970)	1.00 (0.79-1.27, p=0.975)
	Sex	F	17	-	-
		M	172	0.95 (0.22-4.10, p=0.941)	1.19 (0.25-5.55, p=0.829)
	Race	Non-Hispanic White	46	-	-
		Non-Hispanic Black	45	0.68 (0.11-4.08, p=0.674)	0.59 (0.10-3.60, p=0.571)
		Non-Hispanic Other	5	3.41 (0.35-32.83, p=0.288)	2.47 (0.23-26.67, p=0.456)
		Hispanic	93	2.28 (0.65-7.98, p=0.199)	2.05 (0.58-7.25, p=0.268)
Full-Length Spike Protein Titer Analysis†	Titer	Mean (SD)	8.1 (1.1)	0.49 (0.35-0.68, p<0.001)	0.45 (0.32-0.65, p<0.001)
	Age	Mean (SD)	19.0 (1.8)	1.00 (0.78-1.27, p=0.970)	1.04 (0.79-1.35, p=0.794)
	Sex	F	17	-	-
		M	172	0.95 (0.22-4.10, p=0.941)	1.07 (0.19-6.22, p=0.936)
	Race	Non-Hispanic White	46	-	-
		Non-Hispanic Black	45	0.68 (0.11-4.08, p=0.674)	0.68 (0.11-4.19, p=0.678)
		Non-Hispanic Other	5	3.41 (0.35-32.83, p=0.288)	1.62 (0.11-23.89, p=0.724)
		Hispanic	93	2.28 (0.65-7.98, p=0.199)	4.82 0.79-10.11, p=0.112)

533
534 **Supplementary Table 1: Cox proportional hazard regression analysis of group and titer association**
535 **with SARS-CoV-2 infection.**
536 * The predictor in multivariable analysis includes predictor variables group, age, sex and race using all
537   2,436 participants with valid PCR data to detect SARS-CoV-2 infection during the six-week
538   prospective study period.
539 † The predictor in multivariable analysis includes predictor variables titer, age, sex and race using 189
540   participants in the seropositive group. The titers (S-RBD and full-length spike protein) have been log
541   transformed. The univariate analysis for both S-RBD titer and full-length spike protein titer gives the
542   same results of age, sex and race due to use the same set of participants and were kept for both to make
543   it is easier to compare the univariable and multivariable analysis.
544 ‡ The dashed symbol indicates the reference level against which other levels were compared.
545

546
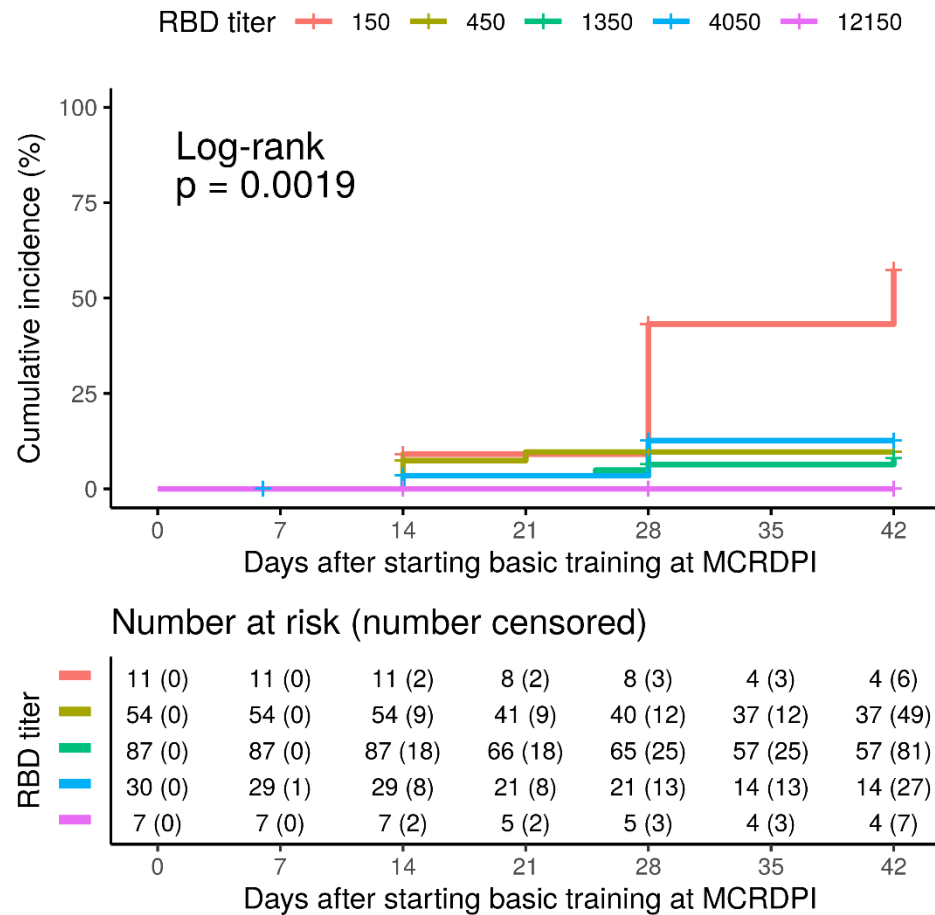
Supplementary Figure 1: SARS-CoV-2 PCR positive incidence curves at different baseline RBD IgG titers during the six-week follow-up period. Cumulative incidence for testing PCR positive in the seropositive group at different baseline RBD IgG titers, which ranged from 1:150 to 1:12150. The numbers at risk and cumulated censored participants are listed below the table.
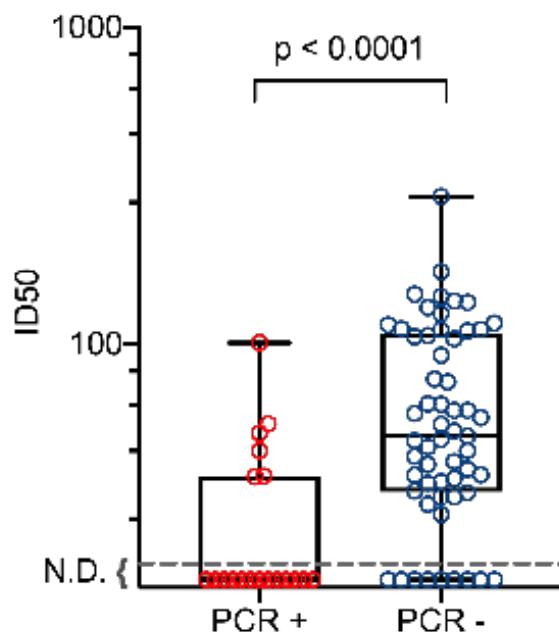
557
558 **Supplementary Figure 2: Baseline neutralizing antibodies in PCR+ and PCR- seropositive**
559 **participants.** The baseline 50% inhibitory dose (ID50) was determined for all seropositive participants
560 who became SARS-CoV-2 PCR positive during the subsequent six-week observation period and in the
561 first 54 seropositive participants enrolled who did not test PCR positive. N.D. indicates no inhibition
562 detected at a 1:20 dilution which was the lowest titer evaluated. P value was assessed by the Cochran-
563 Armitage test.
564
565

566
567