Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis

Xinhua Chen[1,*], Zhiyuan Chen[1,*], Andrew S. Azman[2,3,4,*], Xiaowei Deng[1], Xinghui Chen[1], Wanying Lu[1], Zeyao Zhao[1], Juan Yang[1], Cecile Viboud[5], Marco Ajelli[6], Daniel T. Leung[7,†], Hongjie Yu[1,†]

[*]These authors are co-first authors contributed equally to this work.

[†]These authors are joint senior authors contributed equally to this work.

Corresponding author: Hongjie Yu, Fudan University, School of Public Health, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai 200032, China; E-mail: yhj@fudan.edu.cn

Affiliations:

1. School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China
2. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
3. Unit of Population Epidemiology, Division of Primary Care Medicine, Geneva University Hospitals, Switzerland
4. Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland
5. Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

6. Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA
7. Division of Infectious Diseases, University of Utah School of Medicine, Salt Lake City, UT, USA

Word count (Abstract): 272

Word count (main text): 3,489

Running Title: Seroevidence of human SARS-CoV-2 infections

Keywords: SARS-CoV-2 infections in humans; Serological Evidence

Disclaimer: The views expressed are those of the authors and do not necessarily represent the institutions with which the authors are affiliated.

Abstract

Background: A rapidly increasing number of serological surveys for anti-SARS-CoV-2 antibodies have been reported worldwide. A synthesis of this large corpus of data is needed.

Purpose: To evaluate the quality of serological studies and provide a global picture of seroprevalence across demographic and occupational groups, and to provide guidance for conducting better serosurveys.

Data sources: PubMed, Embase, Web of Science, medRxiv, bioRxiv, SSRN and Wellcome were searched for English-language papers published from December 1, 2019 to August 28, 2020.

Study selection: Serological studies that evaluated seroprevalence of SARS-CoV-2 infections in humans.

Data extraction: Two investigators independently extracted data from included studies.

Data Synthesis: Most of 178 serological studies, representing tests in >800,000 individuals, identified were of low quality. Close contacts and high-risk healthcare workers had higher seroprevalence of 22.9% (95% CI: 11.1-34.7%) and 14.9% (4.8-25.0%), compared to low-risk healthcare workers and general population of 5.5% (4.6-6.4%) and 6.3% (5.5-7.1%). Generally, young people (0-20 yrs) were less likely to be seropositive compared to the middle-aged (21-55 yrs) populations (RR, 0.8, 95% CI: 0.7-0.8). Seroprevalence correlated with clinical COVID-19 reports with 10 (range: 2 to 34) infections per confirmed

COVID-19 case.

Limitations: Some heterogeneity cannot be well explained quantitatively.

Conclusions: The overall quality of seroprevalence studies examined was low. The relatively low seroprevalence among general populations suggest that in most settings, antibody-mediated herd immunity is far from being reached. Given that ratio of infections to confirmed cases is on the same order of magnitude across different locales, reported case numbers may help provide insights into the proportion of the population infected.

Primary Funding source: National Science Fund for Distinguished Young Scholars (PROSPERO: CRD42020198253).

Introduction

The epidemic of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was first reported in Wuhan, Hubei province in December, 2019 and quickly spread globally (1). As of September 4, 2020, more than 26 million COVID-19 cases, including 869,600 deaths, had been reported in 216 countries/regions (2, 3). The real number of SARS-CoV-2 infections is undoubtedly much higher than the officially reported cases due to various factors, including the occurrence of asymptomatic infections, variable healthcare-seeking for clinically mild cases, varied testing strategies in different countries, false-negative virologic tests, and incomplete case reporting. Therefore, the reported COVID-19 cases based on clinical identification with virologic confirmation only represents the "tip of iceberg", with a large number of asymptomatic and mild infections in the general population who may only be identified by seroepidemiological studies (4).

Serological studies are a useful tool to estimate the proportion of the population previously infected, to quantify the magnitude of transmission of pathogens, estimate the infection fatality rate (5), assess the impact of interventions (6), and when correlates of protection are available, estimate the degree of population immunity (7, 8). Meanwhile, reliable information provided from continuous serological surveillance are valuable for policymakers and health officials when planning public health decision-making.

A large number of serological investigations across the world have been published during the first 8 months of the COVID-19 pandemic, with highly variable estimates of seroprevalence that may largely be due to differences in attack rates, but which also feature heterogeneous sampling strategies and assays used. Two rapid systematic reviews of SARS-CoV-2 seroprevalence were identified (on preprint servers at the time of writing) but both had a limited scope and did not explore important differences between subpopulations, assess study quality or perform standard meta-analyses techniques (9, 10).

Here, we conduct a systematic review and meta-analysis to summarize serologic surveys for SARS-CoV-2 infections in humans, to comprehensively evaluate the study designs, laboratory methods, and outcome interpretations for each included serological study, and to estimate the risk of infections by populations with different presumed levels of exposure to SARS-CoV-2. We aim for these results to help inform decision makers and researchers alike as plans are made for the next phases of the global pandemic.

Methods

Data Sources and Searches

According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline (http://www.prisma-statement.org/) (11), we performed a systematic literature review from three peer-reviewed databases (PubMed, Embase and Web of Science) and four preprint severs (medRxiv,

bioRxiv, SSRN and Wellcome) with predefined search terms **(Appendix Table 1)**.

Study Selection and Extraction

English-language papers published from December 1, 2019 to August 28, 2020 were screened for title and abstract by two independent researchers with the following inclusion criteria: (1) had to be a report of seroprevalence in either the general population or some other well-defined population of non-COVID clinical cases; (2) conducted after the first reported case in the area and (3) reported the specific assays used. We excluded studies that only reported serologic responses among COVID-19 patients, those using samples with known infection status (e.g. validation studies of assays), or reported serological data in animal experiments. Abstracts of congress meetings or conference proceedings, study protocols, media news, commentaries, reviews or case reports were also excluded.

Data Extraction and Quality Assessment

For eligible studies, we extracted data on the number of participants who provided specimens and the number of these who were seropositive to calculate the seroprevalence (completed variable list were in **Appendix**). The full text of included studies after initial screening were scrutinized to assess the overall eligibility based on the inclusion and exclusion criteria. A third researcher was consulted when the two reviewers disagreed on study assessment. We developed a new scoring system on the basis of a seroepidemiological protocol from

Consortium for the Standardization of Influenza Seroepidemiology (CONSISE), a previously published scoring system for seroprevalence studies of zoonotic influenza viruses, and a seroepidemiological protocol developed by World Health Organization (WHO), to further assess the study quality (12-14). Study design, laboratory assay and outcome adjustment were comprehensively assessed, and each included study were assigned an overall score. Based on their overall score, each study's quality was classified into one of four categories: A, B, C or D. Category A spanned studies with scores ranging from 12 to 10, category B from 9 to 7, category C from 4 to 6, and category D from 0 to 3. We also described the characteristics, laboratory testing method, and primary outcome for each available study in **Appendix Table 2-5**. This review was preregistered with the protocol available in the PROSPERO database (ID: CRD42020198253).

Data Synthesis and Analysis

Seroprevalence was defined as the prevalence of SARS-CoV-2-specific antibodies at or above a designated antibody titer to define a seropositive result in each original study. For serial cross-sectional studies, we calculated the sum of the total number of participants who provided specimens and total number of seropositive individuals during the whole study period, to avoid repeated inclusion of the same study. Similarly, only data from the first blood collection were analyzed for studies with a longitudinal design. For studies that used multiple serological assays, we used the seropositive results from the assay with

the highest sensitivity and specificity validated internally or in previous studies. We also conducted a sensitivity analysis with results from the other assays. If a study used a confirmatory assay (e.g. microneutralization assay) to validate the positive and/or equivocal result from the initial screening, we used the results from the confirmatory assay. For those studies reporting multiple isotypes including IgG, we included only IgG in the main analyses as they remain elevated for a longer period post-infection compared to IgM and IgA (15). While many studies did adjust for various factors, we decided to use the crude (unadjusted) estimates in our analyses to ease interpretation across different studies.

To estimate seroprevalence by different types of exposures, we categorized all study participants into five groups: 1) close contacts, 2) high-risk healthcare workers, 3) low-risk healthcare workers, 4) general populations, and 5) poorly-defined populations **(Appendix Table 5)**. The poorly-defined population classification represents populations with undefined or unknown exposure to laboratory-confirmed or suspected COVID-19 patients (e.g., blood donors, residual blood samples from labs, patients of other diseases, etc.), as well as those participants cannot be exactly categorized as the first four study populations due to limited exposure information [e.g., healthcare workers without reporting usage of PPE (personal protective equipment) or COVID-19-related exposures]. Based on a random-effect meta-analysis model, we used inverse variance method to estimate pooled seroprevalence by different

sub-populations, combined with the use of Clopper-Pearson method to calculate 95% confidence intervals (16, 17).

For seroprevalence estimates from the general population, we further explored potential determinants affecting the seroprevalence, such as sex, age, and the reported cumulative incidence of COVID-19 cases for the location. We calculated cumulative incidence of COVID-19 by taking the total number of confirmed COVID-19 cases in the same target area as a serosurvey as of two weeks before the survey (midpoint) was conducted and dividing by the estimated population size of this target population. The calculation of incidence for the location was based on the following formula: $Incidence = \frac{Cumulative\ number\ of\ reported\ COVID-19\ cases}{Population\ size}$, where the time point for calculating the number of cumulative cases was two weeks before the midpoint of the study period. Furthermore, we meta-analyzed the number of infections per confirmed case across different areas with available epidemiological data collected from the WHO website (18), the COVID-19 dashboard of Johns Hopkins University (2), and local health authorities (19); population size estimates were obtained from the WorldPop (20). The selected studies for estimation of ratio of infections per cases should meet following inclusion criteria: 1) recruited individuals of general population across all ages; 2) had to be representative samples by using random sampling (e.g., multi-stage or stratified random); 3) had local population data and epidemic data for specific study sites.

Variability between studies was determined by the heterogeneity tests with Higgins' I^2 statistic. We explored the reasons for variations among eligible studies and examined whether prevalence of SARS-CoV-2 specific antibodies varied by study location, study quality, peer-review status, and level of exposure by multivariable meta-regression models. All statistical analyses were done using the statistical software R (version 3.6.3), with 'meta' package to conduct the meta-analysis. For all statistical tests, two tailed P value less than 0.05 were considered statistically significant.

Role of the funding source

The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

We identified a total of 6,953 studies after systematically searching multiple data sources with 3,731 coming from peer-reviewed databases, 3,215 from preprint severs and 7 reports from governments or health authorities. After removing duplicates and screening titles and abstracts, 1,389 studies reporting serological evidence of SARS-CoV-2 infections were found. Finally, a total of 178 studies

involving 818,143 participants were included in our meta-analysis after full-text scrutiny **(Figure 1, Appendix Figure 11)**. For the included literature, the majority of studies (155/178, 87%) were cross-sectional studies, with twenty-three (23/178, 13%) studies being follow-up studies. General populations (42/178, 24%) and healthcare workers (both high-risk and low-risk healthcare workers) (67/178, 38%) were sampled with the highest frequency. The age profile of participants varied across studies of different targeted populations **(Appendix Table 2)**. Most studies involved participants aged 18-69 yrs (151/156 studies) and only five studies exclusively recruited young people (< 18 yrs), including two papers focusing on children (< 10 yrs), and one specifically included targeting older persons from nursing home (> 80 yrs) **(Appendix Table 2)**.

The overall quality of studies was extremely low based on the quality scoring system **(Appendix Table 6)**, ranging from 0 points to 11 points for all included serological studies **(Appendix Figure 1)**. The median score for the 42 population-based seroprevalence studies was 5.5 points, with only three studies achieving Grade A, and 13 studies achieving Grade B. Among studies involving populations with a potentially higher risk of infection, the median score for studies of close contacts, high-risk and low-risk healthcare workers was 2.5 points, 2 points and 3 points, respectively (Grade D). None of these studies conducted within specific high-risk populations achieved Grade A and only three

studies achieved Grade B.

Approximately two-third of studies (108/178, 61%) described serological results from convenience samples, while only twenty-two studies (22/178, 12%) used multi-stage/stratified random sampling to select study participants~~ or recruited from multiple centers~~. Thirty-three studies (19%) used more than one serological assay, mainly including enzyme-linked immunosorbent assays (ELISA), lateral flow immunoassays (LFIA), and chemiluminescence immunoassays (CLIA), with immunoglobulin G (IgG) the isotype most frequently tested. Additionally, thirteen studies utilized a microneutralization assay to detect neutralizing antibodies. Among 117 studies that provided the targeted protein for serological assays, 87 (74%) studies used tests targeting the S-protein, and 60 (51%) studies used tests targeting the N-protein. Twenty-seven (27/178, 16%) studies adjusted their results to correct for demographic factors (age and/or sex), and twenty-two (22/178, 13%) adjusted for sensitivity and specificity of the serological assay(s). Only thirteen studies corrected for both socialdemographic factors and test performance. Overall, the quality of the studies was low, and most did not adhere to published recommendations for seroepidemiological study design(12, 14).

The majority of serological studies were conducted in Europe, United States, and China **(Appendix Figure 9)**. Twenty-three studies (23/157, 15%) were

conducted after more than 75% of the total cases had been reported as of Aug 28, 2020, most of which (14/23, 61%) were conducted in China, which may reflect a more accurate cumulative infection attack rate compared to those conducted before or during a rapidly unfolding epidemic **(Appendix Figure 2-7)**. For those twenty-three studies, the median time from the peak of the epidemic to the serosurvey midpoint was 51 days (range: 16-104 days).

Among different study populations, close contacts of confirmed or suspected COVID-19 patients had the highest seroprevalence of 22.9% (95% CI: 11.1-34.7%), followed by high-risk healthcare workers, low-risk healthcare workers and general population, of 14.9% (95% CI: 4.8-25.0%), 5.5% (95% CI: 4.6-6.4%) and 6.3% (95% CI: 5.5-7.1%), respectively **(Figure 3A, Appendix Figure 10)**. The seroprevalence of the populations in studies that did not specify exposure was 4.6% (95% CI: 4.2-5.1%) **(Figure 3A, Appendix Figure 10)**. We found similar qualitative results in sub-analyses limited to either peer-reviewed studies/official reports or pre-prints **(Appendix Table 8)**. No significant differences were found in pooled estimation of seroprevalence by peer-review status, or symptom history **(Appendix Figure 11)**. Sensitivity analysis for studies using multiple assays showed that few changes were observed in seroprevalence by different populations compared to our main analysis **(Appendix Table 9)**.

The seroprevalence increased with age for study participants under 65 yrs, constituting 1.9% (95% CI: 0.5-3.3%), 3.6% (95% CI: 1.9-5.4%) and 5.2% (95% CI: 3.2-7.2%) for study participants younger than 20 yrs, 20-49 yrs, and 50- 64 yrs **(Appendix Figure 12A)**. After further imperfectly merging age groups, the relative risk of seropositivity in the young (<20 yrs) was 20% lower than that of the middle-age group (RR=0.8, 95%CI: 0.7-0.8), suggesting young individuals may be less susceptible to SARS-CoV-2 infections **(Appendix Table 11)**. However, no significant difference was found between sexes, with seroprevalence of 5.9 % (95% CI: 4.8-7.1%) and 5.6% (95% CI: 4.4-6.8%) for male and female, respectively **(Appendix Figure 12B)**.Among 20 studies that reported sex-specific seroprevalence, 60% of studies had higher point estimates of seroprevalence in male than female. Across studies males had no statistically significant risk of being seropositive compared to females (RR=1.104, 95% CI: 0.997-1.223) **(Appendix Table 11, Appendix Figure 12)**.

The relationship between reported COVID-19 incidence and the number of infections identified through serologic surveys can be useful for trying to understand the evolution of the pandemic without serologic surveillance in each and every locale. For studies of the general population, the cumulative incidence of COVID-19 correlated with seroprevalence across locations (Pearson correlation coefficient: 0.81) **(Figure 3B)**. For studies including individuals across all ages, combined with available epidemiological data and population size,

the ratio of infections to confirmed cases varied across locations, with a pooled ratio of 9.7 (95% CI: 6.6-14.5), suggesting that on average only 1 in 10 infections are detected as virologically confirmed COVID-19 cases. While most estimated ratios were smaller than 15, studies from Brazil had estimates as high as 34, and a study in Denmark had an estimate of 2 infections per case **(Figure 4)**.

Discussion

With the increasing availability of serological assays for SARS-CoV-2, a large body of literature describing seroprevalence studies in different populations has also emerged. In this study, we examine the quality and results of over a hundred reports of seroprevalence studies from around the globe, both published, and in preprint form. In general, the quality of existing serological studies was extremely low, involving less rigorous sampling strategies, poorly-validated and non-standardized laboratory methods, and lack of statistical correction in analyses. As expected, we found that close contacts and high-risk healthcare workers had higher prevalence of SARS-CoV-2 specific antibodies than that of low-risk healthcare workers and the general population. Young individuals may be less susceptible to SARS-CoV-2 infections, while the difference of infection risk between sexes was not obvious. Both findings are in agreement with previous independent estimates based on the analysis of contact-tracing data (21-23). Additionally, we find that the ratio of infections per confirmed case is generally on the same order of magnitude across populations with very different

surveillance and healthcare systems.

Representative serosurveys can provide useful snapshots of the infection history of a population, although the timing and type of sampling can have large influences on their findings. We found that the majority of earlier studies of SARS-CoV-2 seroprevalence were cross-sectional studies employing convenience sampling. We found a limited number of studies (22/178, 12%) whose design satisfied criteria for representative sampling. An optimal study design for estimating seroprevalence includes a detailed sampling framework, rigorous sampling methods (i.e. multi-stage, stratified sampling), and adjustments for social-demographical features or other potential confounders (12).

Various detection assays were used for determination of seropositivity. We found large variations in test performance, targeted antigens and immunoglobulin isotypes, and threshold used. Furthermore, more than half of the studies lacked independent validations of the sensitivity and specificity of the diagnostic kits prior to assessment of serosurvey samples to verify their initial results (24-26). Notably, although the WHO established a population-based serological study protocol, there is a lack of standardized guidelines and procedures for laboratory testing, which may contribute towards such heterogeneity in both performance and reporting of results. We strongly call on national and international governance bodies to develop standardized antibody testing protocols and

reporting practices and create biobanks of reference standards (e.g. monoclonal antibodies), to reduce lab-to-lab variations, thus facilitating the comparability and interpretability amongst seroprevalence studies. Despite the WHO recommendations, the estimates described by many of the population-based serosurveys did not adjust for demographic structure of the target population (14), nor the testing performance (sensitivity and specificity) of the assay, which made the comparison among studies difficult.

The highest seroprevalence was observed in close contacts, reflective of unprotected contact and highly-intensive exposures to confirmed or suspected cases. Similarly, we found high seroprevalence among high-risk healthcare workers, defined as those who provided routine medical care to COVID-19 patients without PPE. On the contrary, low-risk healthcare workers, defined as those wearing adequate PPE (e.g. protective suits, mask, gloves, goggles, face shields, and gowns) or those who provided care for non-COVID-19 patients, had seroprevalence estimates closer to (and with overlapping confidence intervals) that of the general population(27). We found a pooled seroprevalence of 6.3% in the general population, suggesting that globally, the number of persons infected thus far is unlikely to approach that satisfying estimates of what it would take to achieve antibody-mediated herd immunity (28-30). We also found notable differences in seroprevalence between age groups, with the seroprevalence increasing with age among participants younger than 65 yrs. We found that the

young were less likely to be infected by SARS-CoV-2 compared to middle-aged individuals, which was consistent with the lower numbers of virally-confirmed cases in young children compared to elder age groups (31). However, the prevalence of SARS-CoV-2 specific antibodies among elder persons were low, which may be explained by poorer serologic responses after infection (32), lower rates of infection as a result of biological differences or, perhaps more likely, due to behavior changes leading to reduced frequency of infectious contacts.

.

High heterogeneity was observed for all study populations. Although meta regression could partly explain such varied seroprevalence between different study populations with different type of exposures, the internal heterogeneity for the same population should be interpreted carefully. The seroprevalence of the general population is likely a reflection of the duration and intensity of community transmissions. We showed using a regression analysis that higher cumulative incidence of reported cases is associated with higher seroprevalence. For locations where data on the number of confirmed cases were available, we found that the number of infections per confirmed case varied (2.0-34.2) but the order of magnitude was quite consistent across locales. Such evidence indicates the existence of a large number of undetected cases in the community and provides a range of scaling factors for translating the observed confirmed cases

into unobserved infections in the community.

Our study has several limitations. First, although we have performed meta-regression and subgroup analysis to explore heterogeneity of varied seroprevalence for different populations, there are still some factors that we have not taken into account, so that some heterogeneity cannot be well explained quantitatively. Second, misclassification bias may occur due to the limited information on exposures for the study populations, especially for the "Poorly-defined" populations. For some key information (e.g., the use of PPE for healthcare workers) that cannot be extracted from the original articles, we have tried to contact the authors, but the response rate was low. Third, our findings involving cumulative incidence of confirmed cases are limited by the inherently heterogeneous incidence across locations, due to the diversity of testing strategy, case tracing and identification, and public health interventions. Lastly, only raw data without correcting test characteristic were used in the analysis to ensure comparability within studies. Lastly, our included study only involved a total of 40 (40/197, 20%) countries around the world, with large geographic gaps in the sero-data of Africa, South Asia and South America.

In conclusion, the overall quality of the existing seroprevalence studies of SARS-CoV-2 is low and international efforts to standardize study design and assays are urgently needed (14). Pooled estimates of SARS-CoV-2 seroprevalence

based on currently available data demonstrate a higher infection risk among close contacts and healthcare workers lacking PPE, while the relatively low prevalence of SARS-CoV-2 specific antibodies among general populations suggests that the majority of examined populations have not been infected. Therefore, antibody-mediated herd immunity is likely far from being reached in most settings, and continuous serological monitoring is necessary to inform public health decision making.

Contributors

H.Y. designed and supervised the study. X.C. and Z.C. did the literature search, set up the database and did all statistical analyses. X.C., Z.C., A.S.A., and D.T.L. co-drafted the first version of the article. X.C, Z.C., X.D., W.L, Z.Z, and X.C. helped with checking data and did the figures. D.T.L., A.S.A., J.Y, M.A., C.V. and H.Y. commented on the data and its interpretation, revised the content critically. All authors contributed to review and revision and approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Declaration of interests

H.Y. has received investigator-initiated research funding from Sanofi Pasteur, GlaxoSmithKline, and Yichang HEC Changjiang Pharmaceutical Company; M.A. has received research funding from Seqirus; D.T.L. and A.S.A. has received research funding from the US National Institutes of Health. None of those

research funding is related to COVID-19. All other authors report no competing interests.

Acknowledgments

We thank Wei Wang, Qianhui Wu, and Junbo Chen from the Fudan University. This study was funded by the National Science Fund for Distinguished Young Scholars (grant no. 81525023), Program of Shanghai Academic/Technology Research Leader (grant no. 18XD1400300), National Science and Technology Major project of China (grant no. 2017ZX10103009-005, 2018ZX10201001-010), the US National Institutes of Health (R01 AI135115 to D.T.L. and A.S.A.)

References

1. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020.
2. CSSE J 2020 Accessed at https://coronavirus.jhu.edu/map.html on 17 Aug 2020.
3. WHO 2020 Accessed at https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
4. Munster VJ, Koopmans M, van Doremalen N, et al. A Novel Coronavirus Emerging in China - Key Questions for Impact Assessment. N Engl J Med. 2020;382(8):692-4.
5. Erikstrup C, Hother CE, Pedersen OBV, et al. Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. Clin Infect Dis. 2020.

6. Sughayer MA, Mansour A, Al Nuirat A, et al. Covid-19 Seroprevalence rate in healthy blood donors from a community under strict lockdown measures. 2020:2020.06.06.20123919.
7. Griffin S. Covid-19: Herd immunity is "unethical and unachievable," say experts after report of 5% seroprevalence in Spain. Bmj. 2020;370:m2728.
8. Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis. 2020.
9. Bobrovitz N, Arora RK, Yan T, et al. Lessons from a rapid systematic review of early SARS-CoV-2 serosurveys. 2020:2020.05.10.20097451.
10. Arora RK, Joseph A, Van Wyk J, et al. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. Lancet Infect Dis. 2020.
11. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
12. Horby PW, Laurie KL, Cowling BJ, et al. CONSISE statement on the reporting of Seroepidemiologic Studies for influenza (ROSES-I statement): an extension of the STROBE statement. Influenza Other Respir Viruses. 2017;11(1):2-14.
13. Sikkema RS, Freidl GS, de Bruin E, et al. Weighing serological evidence of human exposure to animal influenza viruses - a literature review. Euro Surveill. 2016;21(44).
14. WHO. Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection. 2020.

15. Ma H, Zeng W, He H, et al. Serum IgA, IgM, and IgG responses in COVID-19. Cell Mol Immunol. 2020;17(7):773-5.

16. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934;26:404-13.

17. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998;17(8):857-72.

18. WHO 2020 Accessed at https://www.who.int/news-room/feature-stories/detail/who-updates-covid-19-dashboard-with-better-data-visualization on 28 Aug 2020.

19. FOPH 2020 Accessed at https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html#-1680104524 on 11 Aug 2020.

20. WorldPop 2020 Accessed at https://www.worldpop.org/ on 28 Aug 2020.

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

22. Hu S, Wang W, Wang Y, et al. Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. 2020:2020.07.23.20160317.

23. Viner RM, Mytton OT, Bonell C, et al. Susceptibility to and transmission of COVID-19 amongst children and adolescents compared with adults: a systematic review and meta-analysis. 2020:2020.05.20.20108126.

24. Chen Y, Tong X, Wang J, et al. High SARS-CoV-2 antibody prevalence among healthcare workers exposed to COVID-19 patients. J Infect. 2020.

25. Daniel C. Payne PSES-J, MPH1; Gosia Nowak, MPH2; Uzo Chukwuma,

MPH2; Jesse R. Geibe, MD2; Robert J. Hawkins, PhD, DNP2; Jeffrey A. Johnson, PhD1; Natalie J. Thornburg, PhD1; Jarad Schiffer, MS1; Zachary Weiner, PhD1; Bettina Bankamp, PhD1; Michael D. Bowen, PhD1; Adam MacNeil, PhD1; Monita R. Patel, PhD1; Eric Deussing, MD2; CDC COVID-19 Surge Laboratory Group; Bruce L. Gillingham, MD2. SARS-CoV-2 Infections and Serologic Responses from a Sample of U.S. Navy Service Members — USS Theodore Roosevelt, April 2020. MMWR Morb Mortal Wkly Rep. 2020.

26. Kelvin Kai-Wang To* VC-CC, Jian-Piao Cai*, Kwok-Hung Chan*, Lin-Lei Chen, Lok-Hin Wong, Charlotte Yee-Ki Choi,, Carol Ho-Yan Fong AC-KN, Lu Lu, Cui-Ting Luo, Jianwen Situ, Tom Wai-Hin Chung, Shuk-Ching Wong, Grace See-Wai Kwan,, Siddharth Sridhar JF-WC, Cecilia Yuen-Man Fan, Vivien W M Chuang, Kin-Hang Kok, Ivan Fan-Ngai Hung, Kwok-Yung Yuen. Seroprevalence of SARS-CoV-2 in Hong Kong and in residents evacuated from Hubei province, China: a multicohort study. Lancet Microbe. 2020.

27. Liu M, Cheng SZ, Xu KW, et al. Use of personal protective equipment against coronavirus disease 2019 by healthcare professionals in Wuhan, China: cross sectional study. Bmj. 2020;369:m2195.

28. 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. The Lancet. 2020.

29. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet. 2020.

30. Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose

Novel Coronavirus Disease (COVID-19). Clin Infect Dis. 2020.

31. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. N Engl J Med. 2020;382(24):2302-15.

32. Pawelec G, Weng NP. Can an effective SARS-CoV-2 vaccine be developed for the older population? Immun Ageing. 2020;17:8.

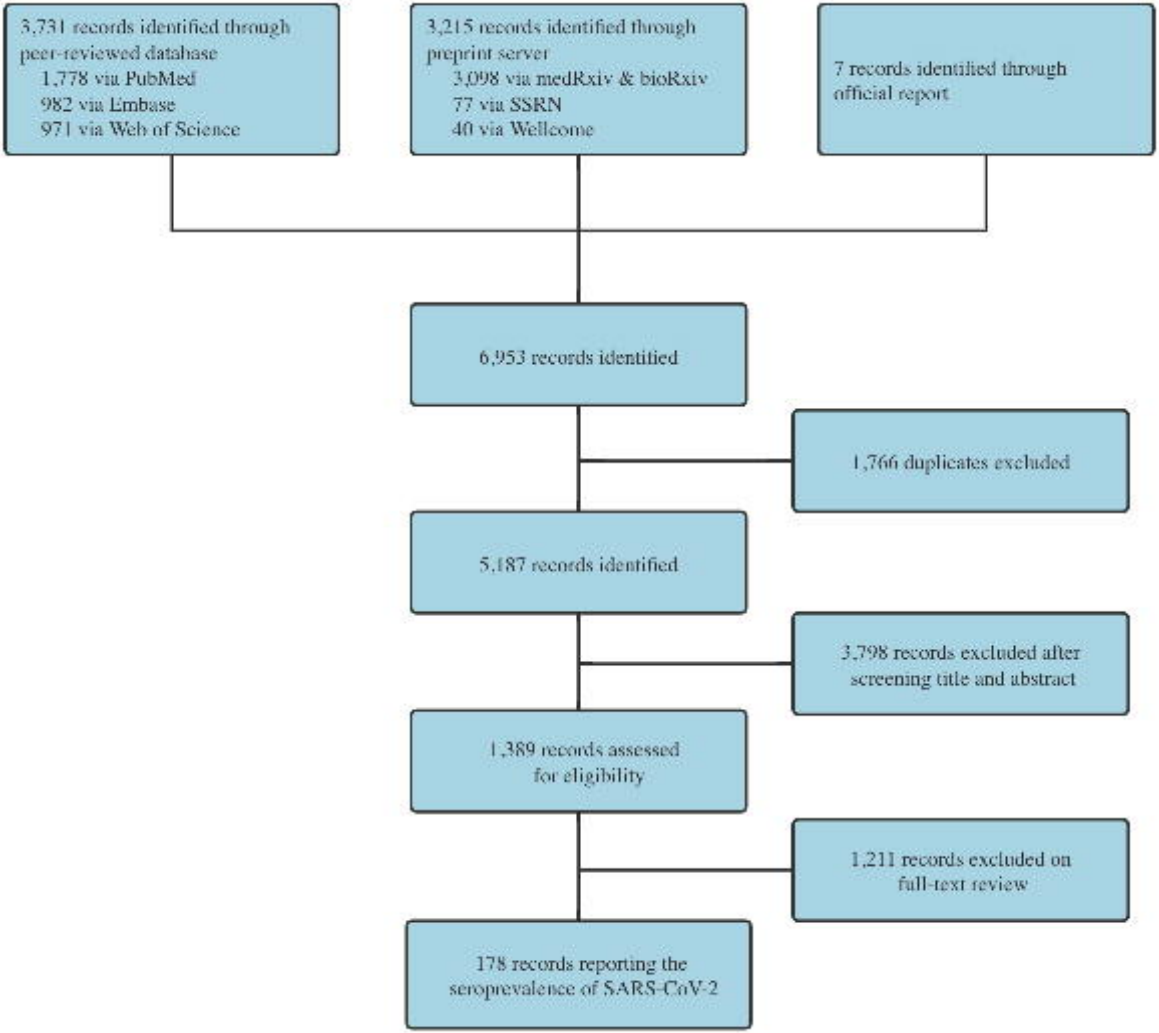
Figure Legends

Figure 1. Flowchart of the selection of serological studies of SARS-CoV-2 infection, December 2019-August 2020

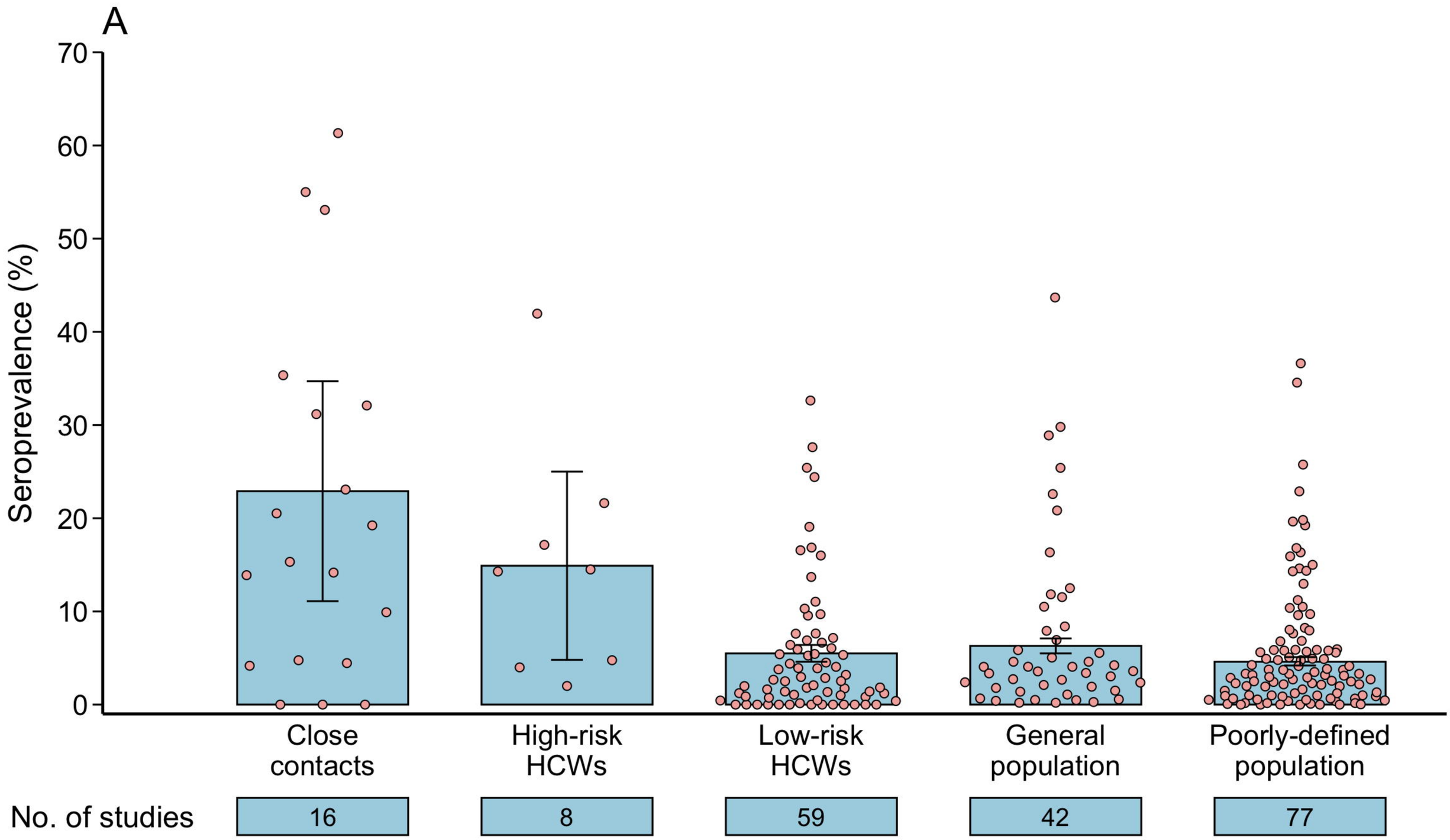
Figure 2. Geographical distribution of SARS-CoV-2 serosurveys in humans by study populations, December 2019-August 2020. The color of the map indicates the cumulative incidence of reported cases (18) with darker colors representing higher values. **(A)** Close contact. **(B)** Healthcare workers. **(C)** General population. **(D)** Poorly-defined population

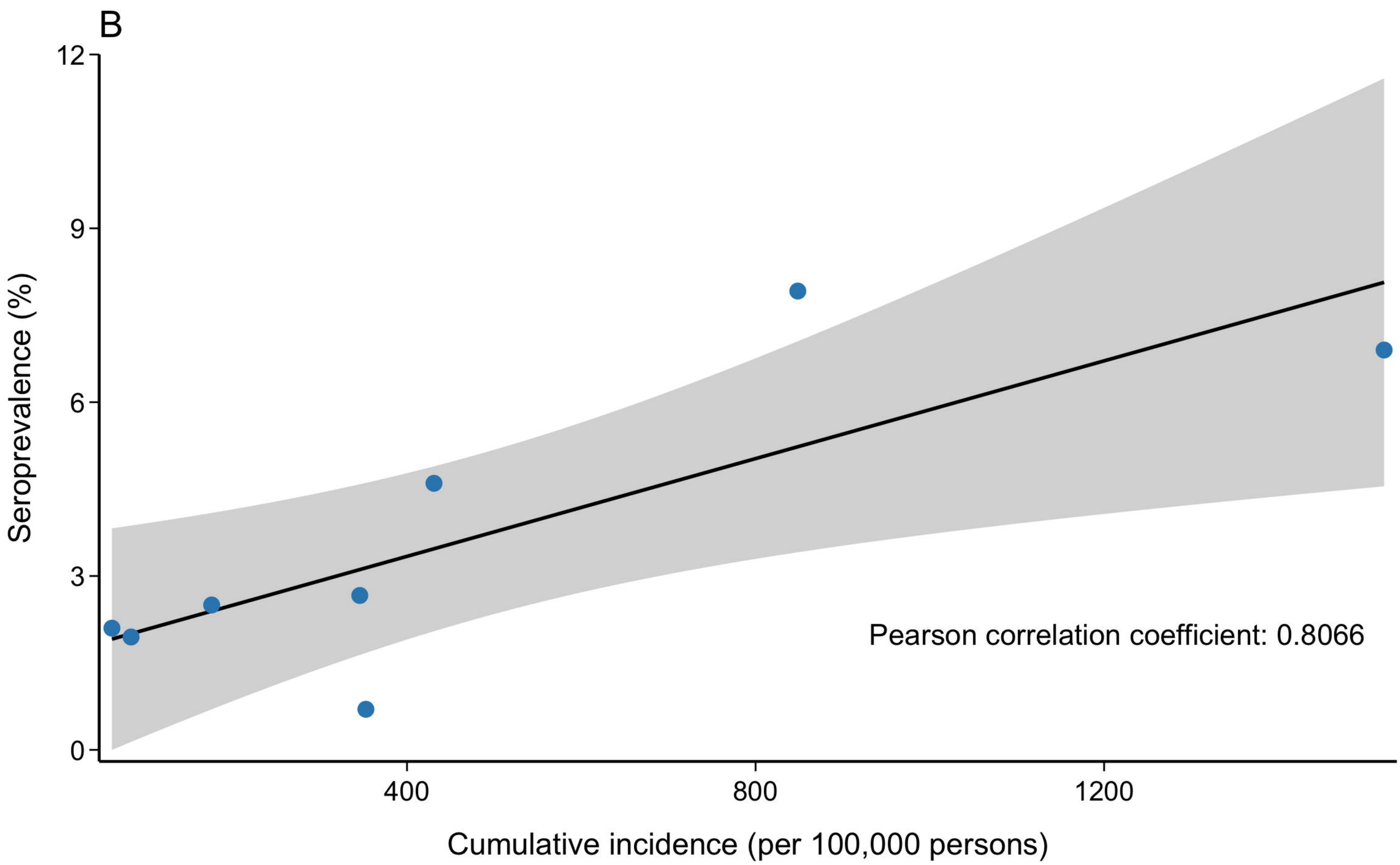
Figure 3. Estimated seroprevalence by study populations and regression analysis between seroprevalence and cumulative incidence in the local among general population. **(A)** Seroprevalence by study populations. The bar represents the pooled estimates. The vertical line represents the 95%CI. Each dot represents the result of one single study. **(B)** Regression analysis. Studies involved general population across all ages were included in this analysis.

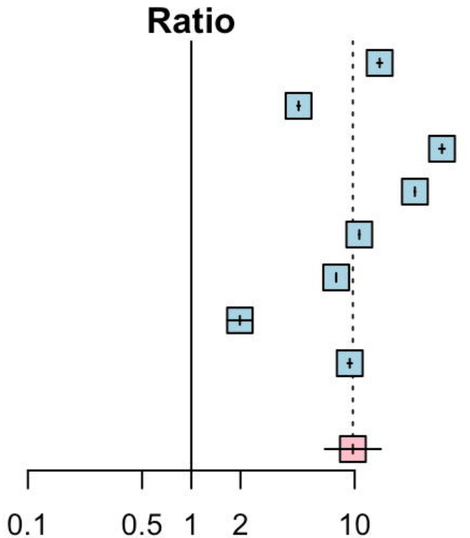
Figure 4. Estimated ratio of human infections with SARS-CoV-2 per COVID-19 confirmed cases.


3,731 records identified through peer-reviewed database
1,778 via PubMed
982 via Embase
971 via Web of Science
3,215 records identified through preprint server
3,098 via medRxiv & bioRxiv
77 via SSRN
40 via Wellcome
7 records identified through official report
6,953 records identified
1,766 duplicates excluded
5,187 records identified
3,798 records excluded after screening title and abstract
1,389 records assessed for eligibility
1,211 records excluded on full-text review
178 records reporting the seroprevalence of SARS-CoV-2

A. Close contact
United States
Japan
China
Thailand
Chile
Norway
Germany
Switzerland
France
Italy
Spain
Incidence of COVID−19 cases (per 100,000 persons)
0−40
41−568
>568
Data unavailable
Seroprevalence (%)
0−4.8
4.9−18
18.1−32.1
32.2−61.3
No. of study participants
1000−10000
<1000
B. Healthcare worker
Canada
United States
Libya
Egypt
Japan
China
Thailand
Sweden
United Kingdom
Denmark
Belgium
Germany
France
Spain
Italy
Greece
Incidence of COVID−19 cases (per 100,000 persons)
0−40
41−568
>568
Data unavailable
Seroprevalence (%)
0−1.6
1.7−4.0
4.1−9.7
9.8−45.3
No. of study participants
>10000
1000−10000
<1000
C. General population
United States
Iran
Libya
Jordan
Pakistan
China
Ethiopia
Nigeria
Ecuador
Brazil
Sweden
United Kingdom
Denmark
Germany
Luxembourg
France
Switzerland
Hungary
Spain
Italy
Greece
Incidence of COVID−19 cases (per 100,000 persons)
0−40
41−568
>568
Data unavailable
Seroprevalence (%)
0−1.7
1.8−3.4
3.6−8.3
8.4−43.7
No. of study participants
>10000
1000−10000
<1000
D. Poorly-defined population
Canada
United States
China
Japan
Korea
India
Thailand
Malaysia
Togo
Nigeria
Kenya
Malawi
Ecuador
Sweden
United Kingdom
Denmark
Netherlands
Germany
Belgium
Czechia
France
Romania
Croatia
Bulgaria
Spain
Italy
Turkey
Incidence of COVID−19 cases (per 100,000 persons)
0−40
41−568
>568
Data unavailable
Seroprevalence (%)
0−1.5
1.6−3.9
4.0−9.5
9.6−36.6
No. of study participants
>10000
1000−10000
<1000


A
Seroprevalence (%)
70
60
50
40
30
20
10
0
Close contacts
High-risk HCWs
Low-risk HCWs
General population
Poorly-defined population
No. of studies
16
8
59
42
77


B
Seroprevalence (%)
12
9
6
3
0
400
800
1200
Cumulative incidence (per 100,000 persons)
Pearson correlation coefficient: 0.8066

Author (Location)	Infections	Cases	Ratio	95%CI
Biggs et al. (Georgia, USA)	45166	3176	14.2	[13.7; 14.7]
Feehan et al. (Orleans and Jefferson Parish, Louisiana, USA)	56761	12515	4.5	[4.4; 4.6]
Gomes et al. (Espirito Santo, Brazil)	84391	2465	34.2	[32.9; 35.6]
Hallal et al. (Brazil)	4165000	178214	23.4	[23.3; 23.5]
Marina et al. (Spain)	2137063	200210	10.7	[10.6; 10.7]
Menezes et al. (Brazil)	5697377	739503	7.7	[7.7; 7.7]
Petersen et al. (Faroe Islands, Denmark)	365	184	2.0	[1.7; 2.4]
Stringhini et al. (Geneva, Switzerland)	39546	4238	9.3	[9.0; 9.6]
Overall			9.7	[6.6; 14.5]

Heterogeneity: $Tau^2 = 0.324$; $Chi^2 = 179493.8$, df = 7 (P = 0); $I^2 = 100.0\%$