Global seroprevalence of SARS-CoV-2 antibodies: a systematic review and meta-analysis

Niklas Bobrovitz*, Rahul Krishan Arora*, Christian Cao, Emily Boucher, Michael Liu, Hannah Rahim, Claire Donnici, Natasha Ilincic, Nathan Duarte, Jordan Van Wyk, Tingting Yan, Lucas Penny, Mitchell Segal, Judy Chen, Mairead Whelan, Austin Atmaja, Simona Rocco, Abel Joseph, David A. Clifton, Tyler Williamson, Cedric P Yansouni, Timothy Grant Evans, Jonathan Chevrier, Jesse Papenburg[†], Matthew P. Cheng[†]

*NB and RKA contributed equally to this paper as co-first authors.
[†]JP and MPC contributed equally to this paper as co-senior authors.

Temerty Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada (N Bobrovitz DPhil, T Yan BHSc[Hons], N Ilincic BScH, M Segal MSc, L Penny);

Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK (RK Arora BHSc [Hons]; DA Clifton DPhil);

Cumming School of Medicine, University of Calgary (H Rahim BHSc [Hons], E Boucher BHSc [Hons], RK Arora BHSc [Hons], C Cao, T Williamson PhD, C Donnici BHSc [Hons], M Whelan BHSc [Hons]);

Department of Social Policy and Intervention, University of Oxford, Oxford, UK (M Liu AB);

Harvard Medical School, Boston, Massachusetts, United States of America (M Liu AB);

Faculty of Engineering, University of Waterloo (A Atmaja; N Duarte; S Rocco; JV Wyk; A Joseph);

Faculty of Medicine and Health Sciences, McGill University (J Chen BHSc[Hons])

School of Population and Global Health, McGill University (TG Evans MD DPhil);

Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University (J Chevrier PhD)

Division of Pediatric Infectious Diseases, Dept. of Pediatrics, McGill University Health Centre (J Papenburg MD MSc)

JD MacLean Centre for Tropical Diseases, McGill University (CP Yansouni MD)

Divisions of Infectious Diseases and Medical Microbiology, McGill University Health Centre, Montreal Qc, Canada (MP Cheng MD, CP Yansouni MD)

Correspondence to:
Dr. Niklas Bobrovitz, Faculty of Medicine, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada, niklas.bobrovitz@mail.utoronto.ca, @nikbobrovitz

Word count: 3169 words
References: 40

Abstract

Background. Studies reporting estimates of the seroprevalence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibodies have rapidly emerged. We aimed to synthesize seroprevalence data to better estimate the burden of SARS-CoV-2 infection, identify high-risk groups, and inform public health decision making.

Methods. In this systematic review and meta-analysis, we searched publication databases, preprint servers, and grey literature sources for seroepidemiological study reports, from January 1, 2020 to August 28, 2020. We included studies that reported a sample size, study date, location, and seroprevalence estimate. Estimates were corrected for imperfect test accuracy with Bayesian measurement error models. We conducted meta-analysis to identify demographic differences in the prevalence of SARS-CoV-2 antibodies, and meta-regression to identify study-level factors associated with seroprevalence. We compared region-specific seroprevalence data to confirmed cumulative incidence. PROSPERO: CRD42020183634.

Findings. We identified 338 seroprevalence studies including 2.3 million participants in 50 countries. Seroprevalence was low in the general population (median 3.2%, IQR 1.0-6.4%) and slightly higher in at-risk populations (median 5.4%, IQR 1.5-18.4%). Median seroprevalence varied by WHO Global Burden of Disease region ($p < 0.01$), from 1.0% in Southeast Asia, East Asia and Oceania to 18.8% in South Asia. National studies had lower seroprevalence estimates than local ($p = 0.02$) studies. Compared to White persons, Black persons (prevalence ratio [RR] 2.34, 95% CI 1.60-3.43) and Asian persons (RR 1.56, 95% CI 1.22-2.01) were more likely to be seropositive. Seroprevalence was higher among people ages 18-64 compared to 65 and over (RR 1.26, 95% CI 1.04-1.52). Health care workers had a 1.74x (95% CI: 1.18-2.58) higher risk compared to the general population. There was no difference in seroprevalence between sexes. There were 123 studies (36%) at low or moderate risk of bias. Seroprevalence estimates from national studies were median 11.9 (IQR 8.0 - 16.6) times higher than the corresponding SARS-CoV-2 cumulative incidence.

Interpretation. Most of the population remains susceptible to SARS-CoV-2 infection. Public health measures must be improved to protect disproportionately affected groups, including non-

White people and adults. Measures taken in SE Asia, E Asia and Oceania, and Latin America and Caribbean may have been more effective in controlling virus transmission than measures taken in other regions.

Funding. Public Health Agency of Canada through the COVID-19 Immunity Task Force.

1. Introduction

As of Nov 17, 2020, there were over 54 million confirmed cases of SARS-CoV-2 infection and 1.3 million deaths worldwide.[1] However, these case counts inevitably underestimate the true cumulative incidence of infection[2] because of limited diagnostic test availability[3], barriers to testing accessibility[4], and asymptomatic infections.[5] The global burden of SARS-CoV-2 infection remains unknown.

Serological assays identify SARS-CoV-2 antibodies, indicating previous infection in unvaccinated persons.[6] Population-based serological testing provides better estimates of the cumulative incidence of infection, complementing diagnostic testing and helping to shape the public health response to COVID-19.

SARS-CoV-2 seroprevalence estimates are reported in published articles and preprints, but also in government and health institute reports, and media.[7] Consequently, few studies have comprehensively synthesized seroprevalence findings. As the world prepares to enter this pandemic's vaccine era, synthesizing seroepidemiology findings is increasingly important to measure the baseline prevalence of SARS-CoV-2 antibodies worldwide, identify disproportionately affected groups, and inform the optimal distribution of COVID-19 vaccines.

To fill this gap, we conducted a systematic review and meta-analysis of SARS-CoV-2 seroprevalence studies globally. We aimed to: (i) describe serosurveys globally and their findings; (ii) identify geographic and study design factors associated with elevated seroprevalence; (iii) identify groups at high risk of previous SARS-CoV-2 infection; and (iv) evaluate how much confirmed infections underestimate the true burden of this pandemic.

2. Methods

2.1 Search strategy and selection criteria

This systematic review and meta-analysis was registered in PROSPERO (CRD42020183634), reported per PRISMA[8] guidelines (Supplementary File 1), and will be regularly updated on an open-access platform (SeroTracker.com).[9]

We searched Medline, EMBASE, Web of Science, and Europe PMC, using a search strategy developed in consultation with a health sciences librarian. Given that many serosurveys are not reported in these databases we also searched public health agency websites and the Google News aggregation platform, invited submissions to our SeroTracker website, and consulted with international experts. We included records published from January 1, 2020 to August 28, 2020. No limits on language were applied. Articles not in English or French were included if they could be extracted in full using machine translation.[10] Full search details are in Supplementary File 2.

We included all SARS-CoV-2 serosurveys in humans. We defined a serosurvey as the serological testing of a defined population over a specified period to estimate the prevalence of SARS-CoV-2 antibodies.[11,12] Included studies had to report a sample size, sampling date and region, and prevalence estimate.

We excluded studies conducted only in people with SARS-CoV-2 infection; dashboards that were not associated with a defined serology study; and case reports, case-control studies, and reviews.

2.2 Screening and extraction

Two authors independently screened articles. Data were extracted by one reviewer and verified by a second. We extracted characteristics of the study, sample, antibody test, and seroprevalence. We extracted sub-group data when they were stratified by one variable (e.g., seniors) but not two variables (e.g., female seniors). We contacted study authors to request any missing sub-group seroprevalence data.

2.3 Evaluation of seroprevalence studies and estimates

The intended geographic scope of each estimate was classified as (A) national; (B) regional (e.g., province-level); (C) local (e.g., county-level, city-level); or (D) sublocal (e.g., one hospital department). Countries were classified according to GBD region, and country income status classified by distinguishing the high-income GBD region from other regions.[13,14]

We defined studies of the general population as samples from households, the community, blood donors, or residual sera with the explicit purpose of providing estimates for the population at large and for which the defining features shared by participants were location or age. Special population studies were those sampling from and aiming to provide estimates for populations with additional defining features (e.g., physicians).

We prioritized estimates that tested for IgG antibodies and that used traditional ELISAs, as non-IgG and anti-nucleocapsid antibodies appear to decline over time, while anti-spike IgG antibodies appear to persist for several months after infection.[15–20]

A modified Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Prevalence Studies was used to assess study risk of bias.[21] Studies were classified by overall risk of bias: low, moderate, high, or unclear (detailed criteria in Supplementary File 3).

2.4 Data Analysis

Data processing and descriptive statistics were conducted in Python. *p*-values less than 0.05 were considered statistically significant.

2.4.1 Correcting seroprevalence estimates

To account for imperfect test sensitivity and specificity, seroprevalence estimates were corrected using Bayesian measurement error models, with binomial sensitivity and specificity distributions.[22] We used sensitivity and specificity values from independent evaluations wherever possible[23]; if unavailable, manufacturer-derived values were used (Supplementary File 4).

We presented corrected and uncorrected estimates for all studies. Subsequent analyses were done using corrected seroprevalence estimates. Estimates that could not be corrected by these methods excluded. To assess the impact of correction, we calculated the absolute difference between seroprevalence estimates before and after correction. We also conducted a sensitivity analysis for each analysis with uncorrected data.

2.4.2 Global seroprevalence and associated factors

To examine study-level factors affecting general population seroprevalence estimates, we constructed a multivariable linear meta-regression model, using the *meta* package in R.[24] The outcome variable was the natural logarithm of corrected seroprevalence. Independent predictors were defined *a priori*. Categorical covariates were encoded as indicator variables, and included: study risk of bias (reference: low risk of bias), GBD region (reference: high-income); geographic scope (reference: national); and population sampled (reference: household and community samples). The sole continuous covariate was days since the 100th confirmed case in the country of the study. A quantile-quantile plot and a funnel plot were generated to visually check normality and homoscedasticity.

2.4.3 Population differences in seroprevalence

To quantify population differences in SARS-CoV-2 seroprevalence, we identified subgroup estimates within general population studies that stratified by sex/gender, race/ethnicity, exposure level, occupation, and age groups. We calculated prevalence ratios for each study (e.g., prevalence in males vs. females) and aggregated the ratios across studies using inverse variance-weighted random-effects meta-analysis (Supplementary File 4). Heterogeneity was quantified using the I^2 statistic.[25]

2.4.4 Comparisons of seroprevalence and confirmed SARS-CoV-2 infections

To measure how much confirmed SARS-CoV-2 infections underestimate seroprevalence, we calculated the ratio between seroprevalence estimates of the general population and the cumulative incidence of confirmed SARS-CoV-2 infections. We obtained data on total confirmed SARS-CoV-2 infections[26,27] and population size[28] that geographically matched the

study target populations nine days before the study end date, to reflect the time period between COVID-19 diagnosis and seroconversion (Supplementary File 5).[29–31]

We conducted sensitivity analyses using case data from zero and fourteen days before study end dates and including studies only at low and moderate risk of bias.

3. Results

3.1 Distribution and characteristics of serosurveys

We screened 16,899 titles and abstracts and 1,556 full text articles (Figure 1). We identified 338 unique seroprevalence studies in 281 articles. These studies included 2,305,376 participants and 3,443 seroprevalence estimates.

Of the included studies, 184 (54%) targeted the general population and 154 (46%) targeted special populations (Table 1). Characteristics of individual studies are reported in Supplementary Tables 1 and 2. References are listed at the end of the Supplementary materials.

Fifty countries across all GBD regions were represented among identified serosurveys (Figure 2; Supplementary Figure 1). A minority of studies were conducted in low- and middle-income countries (n = 86, 25%).

Many studies were at high risk of bias (n = 184, 54%), often for not statistically correcting for demographics or for test sensitivity and specificity, using non-probability sampling methods, and using non-representative sample frames (Figure 3, Supplementary Table 3).

3.2 SARS-CoV-2 seroprevalence globally

In studies targeting the general population, median corrected seroprevalence was 3.2% [IQR 1.0-6.4%] (Table 2). These studies included household and community samples (n = 83), residual sera (n = 39), and blood donors (n = 33), with median corrected seroprevalence of 3.5% [IQR 1.2-8.5%], 2.7% [IQR 1.0-4.3%], and 2.8% [IQR 0.9-6.8%], respectively (Supplementary Table 4). The median corrected seroprevalence in studies targeting specific populations was 5.4%, [IQR 1.5-18.4%] (Table 3). Notably, the median corrected seroprevalence was 6.3% [IQR 2.1-18.8%, n = 72 studies] in healthcare workers and caregivers and 6.3% [IQR 2.8-17.8%, n = 21 studies] in specific patient groups (e.g., cancer patients). Essential non-healthcare workers (e.g., first responders) had a median seroprevalence of 10.0% [IQR 1.8-26.3%, n=7 studies] (Supplementary Table 4).

Among high-income countries, the median corrected seroprevalence was 3.4% [IQR 1.3-6.3%]. In the low- and middle-income GBD regions, median corrected seroprevalence ranged from

1.0% [IQR 0.2-2.4%] in Southeast Asia, East Asia, and Oceania to 18.8% [IQR 13.1-35.9%] in South Asia (Table 2).

Figure 1. PRISMA flow diagram of study inclusion

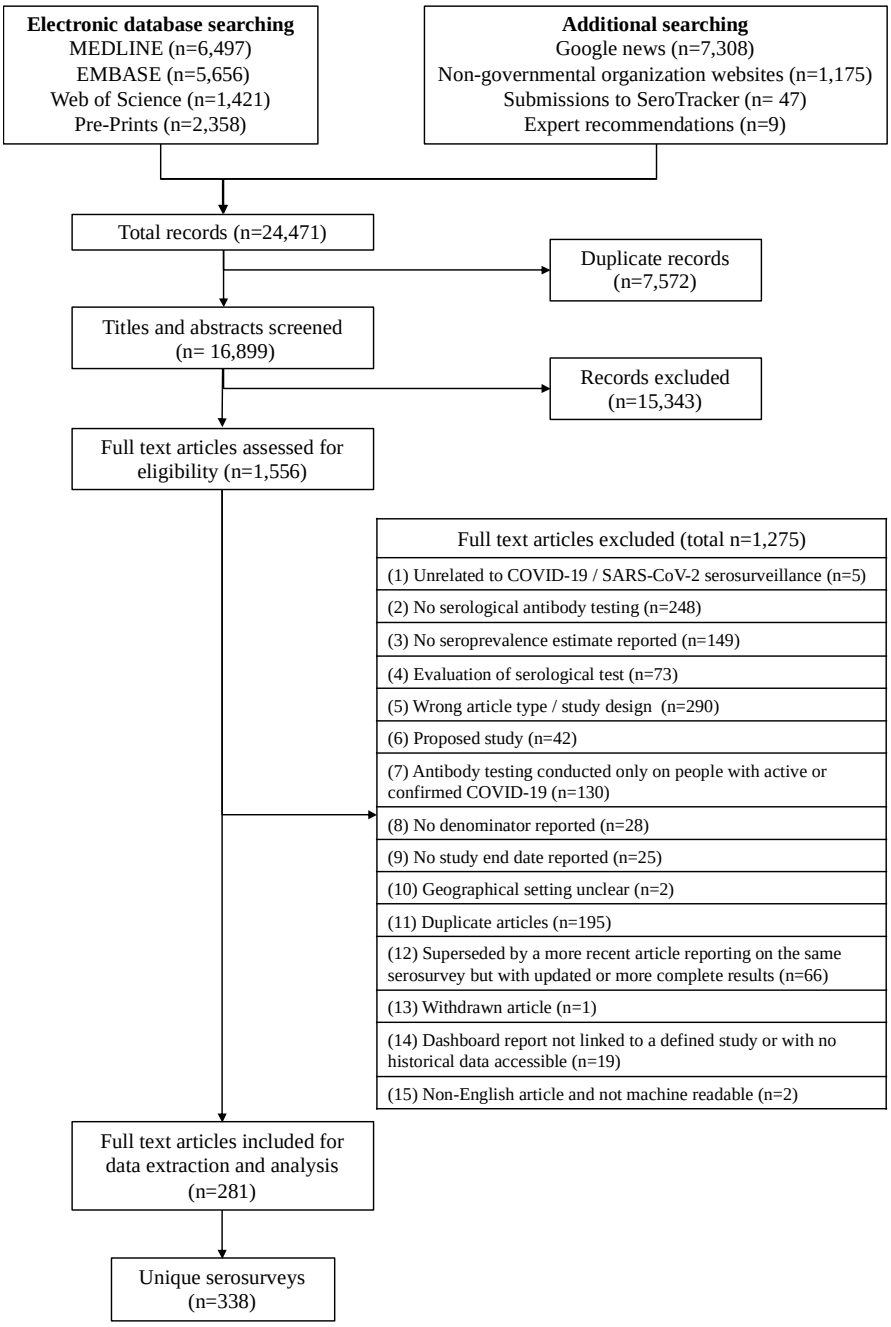
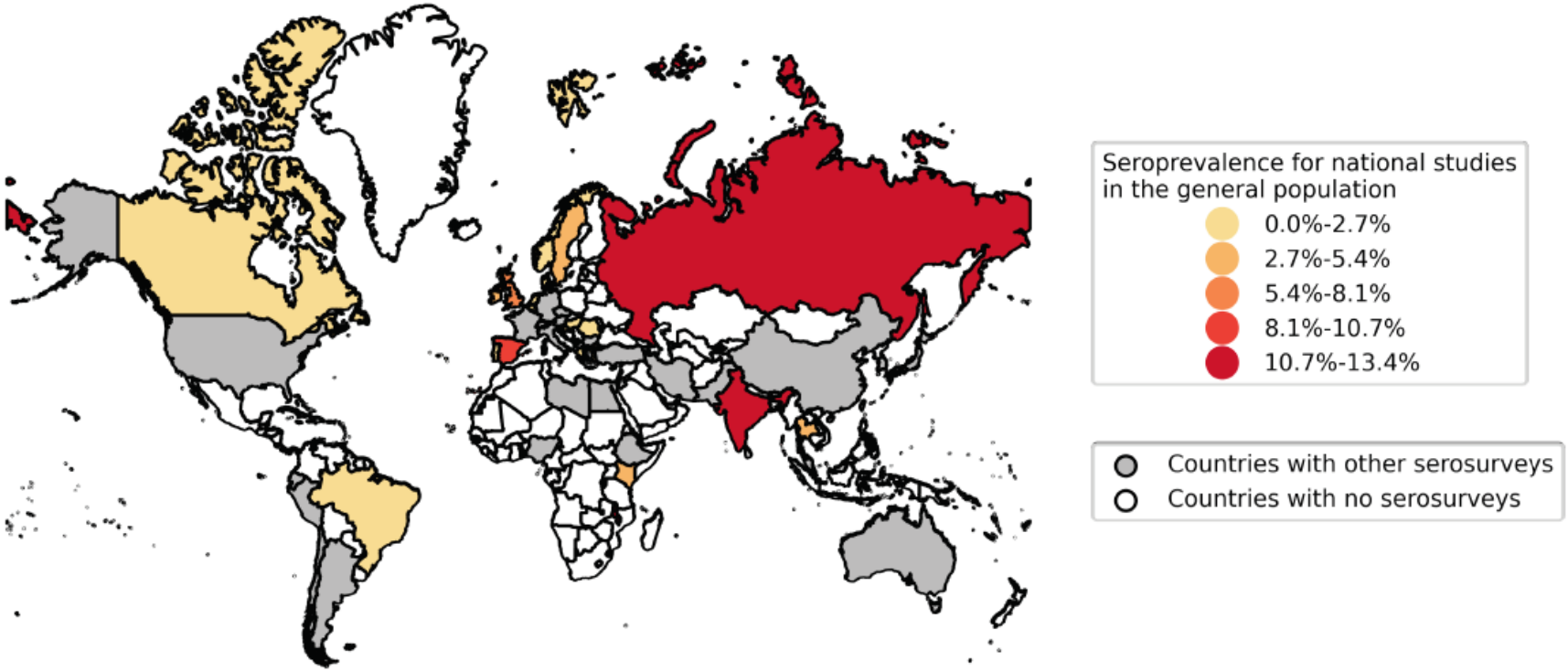


Table 1. Summary characteristics of included articles

Characteristic	Studies n (%)
Geographic scope	
National	58 (17%)
Regional	72 (21%)
Local	129 (38%)
Sublocal	79 (23%)
Age groups*	
Adults (18-64 years)	275 (81%)
Children and Youth (0-17 years)	145 (43%)
Seniors (65+ years)	156 (46%)
Target population	
General population	184 (54%)
Special population†	155 (46%)
County income level‡	
High income	252 (75%)
Low/middle income	86 (25%)
Sampling method	
Probability sampling	104 (31%)
Non-probability sampling	234 (69%)
Antibody tests*	
ELISA	131 (39%)
CLIA	76 (23%)
LFIA	78 (23%)
Other	6 (2%)
Neutralization	4 (1%)
Antibody isotypes reported*	
IgG	279 (83%)
IgM	109 (32%)
IgA	23 (7%)
Risk of bias	
Low	12 (4%)
Moderate	111 (33%)
High	184 (54%)
Unclear	31 (9%)

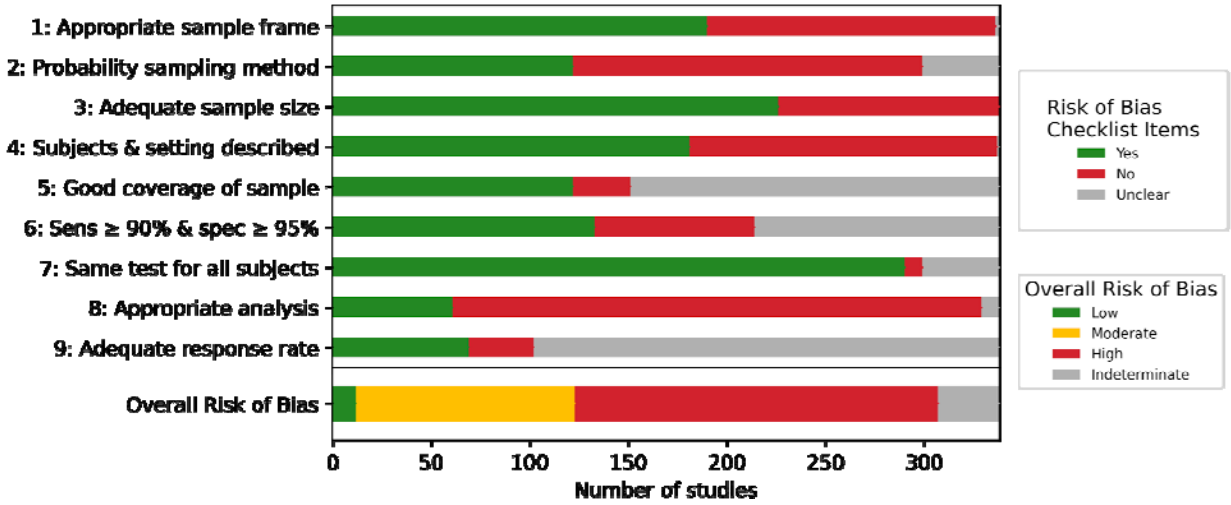
*Studies could have met multiple criteria so the sum of percentages may exceed 100%. †Studies sampling from and aiming to provide estimates for a population with features in common other than geographic location and age (e.g., particular occupation, health status, COVID-19 exposure status). ‡Classified according to the WHO global burden of disease region groupings (high vs other - low/middle).
Abbreviations: ELISA= enzyme-linked immunosorbent assay; CLIA=chemiluminescence immunoassay; LFIA=lateral flow immunoassay.

Figure 2. Map of national seroprevalence studies in the general population



Countries with national-level general population seroprevalence studies are colored on the map, based on the seroprevalence reported in the most recent such study in each country. Countries with no such national serosurveys but with "other serosurveys" are colored in grey; this includes local and regional studies, as well as studies in special populations.

Figure 3. Study risk of bias summary



Item 1: Was the sample frame appropriate to address the target population? Item 2: Were study participants recruited in an appropriate way? Item 3: Was the sample size adequate? Item 4: Were the study subjects and setting described in detail? Item 5: Was data analysis conducted with sufficient coverage of the identified sample? Item 6: Were valid methods used for the identification of the condition? Item 7: Was the condition measured in a standard, reliable way for all participants? Item 8: Was there appropriate statistical analysis? Item 9: Was the response rate adequate, and if not, was the low response rate managed appropriately? Item 10: Overall risk of bias.

Table 2. Summary of seroprevalence data for general populations by global burden of disease region, geographic scope, and risk of bias

Characteristic	No. studies	No. countries	Median sample size [IQR]	Median uncorrected seroprevalence [IQR]	No. studies with adjustable data	Median corrected seroprevalence [IQR]	Risk of bias
All studies	184	36	1200 [750-3873]	3.6% [1.5-6.3%]	155	3.2% [1.0-6.4%]	L: 5%, M: 51%, H: 36%, U: 8%
GBD region							
Central Europe, Eastern Europe, and Central Asia	3	2	90000 [50237-370000]	14.0% [7.3-16.9%]	2	6.3% [3.4-9.1%]	L: 0%, M: 67%, H: 33%, U: 0%
High-income	135	22	1200 [782-3297]	4.0% [1.7-6.2%]	111	3.4% [1.3-6.3%]	L: 4%, M: 52%, H: 36%, U: 9%
Latin America and Caribbean	21	3	900 [900-4500]	2.7% [0.9-5.2%]	20	1.8% [0.6-4.1%]	L: 19%, M: 62%, H: 14%, U: 5%
North Africa and Middle East	2	2	631 [580-682]	10.5% [5.2-15.8%]	2	2.8% [1.5-4.1%]	L: 0%, M: 50%, H: 50%, U: 0%
South Asia	9	2	2702 [1235-21387]	15.0% [3.7-23.5%]	6	18.8% [13.1-35.9%]	L: 11%, M: 22%, H: 44%, U: 22%
Southeast Asia, East Asia, and Oceania	11	2	2199 [516-15667]	0.5% [0.4-2.7%]	11	1.0% [0.2-2.4%]	L: 0%, M: 36%, H: 64%, U: 0%
Sub-Saharan Africa	3	3	185 [142-1642]	4.9% [4.0-15.2%]	3	6.4% [5.8-12.5%]	L: 0%, M: 33%, H: 67%, U: 0%
Scope							
National	51	20	3098 [1200-8317]	4.3% [2.3-5.8%]	48	3.9% [2.1-6.3%]	L: 4%, M: 67%, H: 27%, U: 2%
Regional	57	14	1132 [827-3500]	2.6% [1.0-6.2%]	53	2.4% [0.6-5.1%]	L: 11%, M: 61%, H: 23%, U: 5%
Local	71	19	900 [634-2438]	4.0% [1.5-8.5%]	50	3.4% [1.2-8.7%]	L: 3%, M: 32%, H: 51%, U: 14%
Sub-local	5	4	186 [123-401]	4.1% [3.0-5.2%]	4	3.9% [0.8-22.6%]	L: 0%, M: 20%, H: 60%, U: 20%
Risk of bias							
Low	10	7	4326 [1482-19872]	2.1% [0.3-5.3%]	10	1.6% [0.1-4.4%]	-
Moderate	93	26	1224 [870-4612]	4.0% [1.9-6.9%]	90	3.2% [1.1-6.5%]	-
High	66	21	1200 [492-3169]	3.6% [1.3-5.9%]	54	3.4% [1.1-6.4%]	-
Unclear	15	7	896 [485-3328]	2.2% [1.5-5.5%]	1	2.0% [2.0-2.0%]	-

Abbreviations: No.= number; IQR= interquartile range; L = low; M = moderate; H = high; U = unclear; GBD = global burden of disease region

Table 3. Summary of seroprevalence data for specific populations by GBD region, geographic scope, and risk of bias

Characteristic	No. studies	No. countries	Median sample size [IQR]	Median uncorrected seroprevalence [IQR]	No. studies with adjustable data	Median corrected seroprevalence [IQR]	Risk of bias
All studies	153	34	516 [150-1282]	5.5% [1.6-14.4%]	129	5.4% [1.5-18.4%]	L: 1%, M: 12%, H: 77%, U: 10%
GBD region							
Central Europe, Eastern Europe, and Central Asia	6	5	483 [373-590]	2.8% [2.3-4.3%]	6	1.6% [1.1-5.2%]	L: 0%, M: 0%, H: 67%, U: 33%
High-income	117	19	498 [146-1247]	5.9% [1.7-14.5%]	96	5.7% [2.0-22.2%]	L: 1%, M: 11%, H: 79%, U: 9%
Latin America and Caribbean	0	0	-	-	0	-	-
North Africa and Middle East	4	3	110 [76-310]	3.5% [2.0-6.2%]	3	9.0% [7.2-19.3%]	L: 0%, M: 25%, H: 75%, U: 0%
South Asia	5	2	1000 [212-4202]	17.6% [15.6-19.8%]	3	39.4% [30.3-67.9%]	L: 20%, M: 0%, H: 40%, U: 40%
Southeast Asia, East Asia, and Oceania	19	2	1027 [465-2975]	4.0% [0.7-8.9%]	18	3.0% [0.4-9.1%]	L: 0%, M: 21%, H: 74%, U: 5%
Sub-Saharan Africa	3	3	500 [305-728]	16.8% [8.8-21.4%]	3	11.1% [5.7-18.2%]	L: 0%, M: 0%, H: 100%, U: 0%
Scope							
National	7	6	857 [546-6261]	2.7% [1.6-4.2%]	6	1.2% [0.6-2.9%]	L: 0%, M: 14%, H: 57%, U: 29%
Regional	15	9	3609 [444-9349]	2.5% [1.1-5.3%]	14	3.2% [1.6-8.4%]	L: 13%, M: 27%, H: 60%, U: 0%
Local	58	21	688 [204-1492]	5.7% [1.3-13.8%]	54	5.4% [1.1-18.1%]	L: 0%, M: 19%, H: 72%, U: 9%
Sub-local	74	17	276 [110-944]	8.0% [2.2-18.7%]	55	8.8% [3.0-25.4%]	L: 0%, M: 3%, H: 85%, U: 12%
Risk of bias							
Low	2	2	16497 [10350-22644]	29.1% [16.6-41.6%]	2	50.2% [27.1-73.3%]	-
Moderate	18	13	1556 [904-6668]	5.9% [3.0-9.5%]	18	5.4% [2.8-11.5%]	-
High	118	26	308 [132-954]	5.7% [1.6-14.6%]	105	6.3% [1.3-21.2%]	-
Unclear	16	9	1148 [773-3651]	3.7% [0.9-22.0%]	4	2.2% [1.7-15.4%]	-

Abbreviations: No.= number; IQR= interquartile range; L = low; M = moderate; H = high; U = unclear; GBD = global burden of disease region.

3.3 Differences in seroprevalence by demographic characteristics

There were significant within-study differences in seroprevalence based on age, race/ethnicity, status as healthcare worker, and contact exposure (Table 4). There was no difference in the risk of infection based on sex/gender. Results for uncorrected prevalence estimates are reported in Supplementary Table 5.

Table 4: Differences in seroprevalence by demographic characteristics

Factor	Reference Group	Comparison Group	Number of Studies	Risk Ratio (95% CI)*	Heterogeneity (I^2)
Age	Adults (18-64)	Seniors (65+)	52	1.26 [1.04-1.52]	92.9%
	Adults (18-64)	Youth (0-18)	38	1.23 [0.99-1.52]	75.1%
	Adults 18-64	-	-	Reference	-
Sex/Gender	Male	Female	56	1.05 [0.95-1.17]	85.1%
Race	White	Black	14	2.34 [1.60-3.43]	96.6%
	White	Asian	9	1.56 [1.22-2.01]	85.4%
	White	Indigenous	2	4.32 [0.79-23.72]	95.3%
	White	-	-	Reference	-
Close contact with COVID-19 patients	Individuals with no close contact	Individuals with close contact	8	2.74 [1.58-4.76]	99.4%
	Health care workers with no close contact	Health care workers with close contact	12	1.40 [1.15-1.71]	91.9%
Health care worker status	Non-health care workers and caregivers	Health care workers and caregivers	8	1.74 [1.18-2.58]	96.2%

*Using corrected seroprevalence estimates. Abbreviations: CI= confidence interval.

3.4 Impact of serology assay sensitivity and specificity on seroprevalence findings

Tests that have been independently evaluated were used in 145 studies (42.9%; Supplementary Table 6). Test sensitivity and specificity were reported in 185 studies (54.7%), with sensitivity ranging from 37-100% and specificity from 85-100%. Only 69 studies (20.4%) corrected seroprevalence estimates for test sensitivity and specificity.

We corrected seroprevalence estimates from 254 studies (75.1%) for imperfect sensitivity and specificity, and used author-corrected estimates in 30 (8.9%) studies where uncorrected estimates were unavailable. Data were insufficient to correct estimates from 48 studies (14.2%). The

median absolute difference between corrected and uncorrected seroprevalence estimates was 1.2% [IQR 0.4-3.1%].

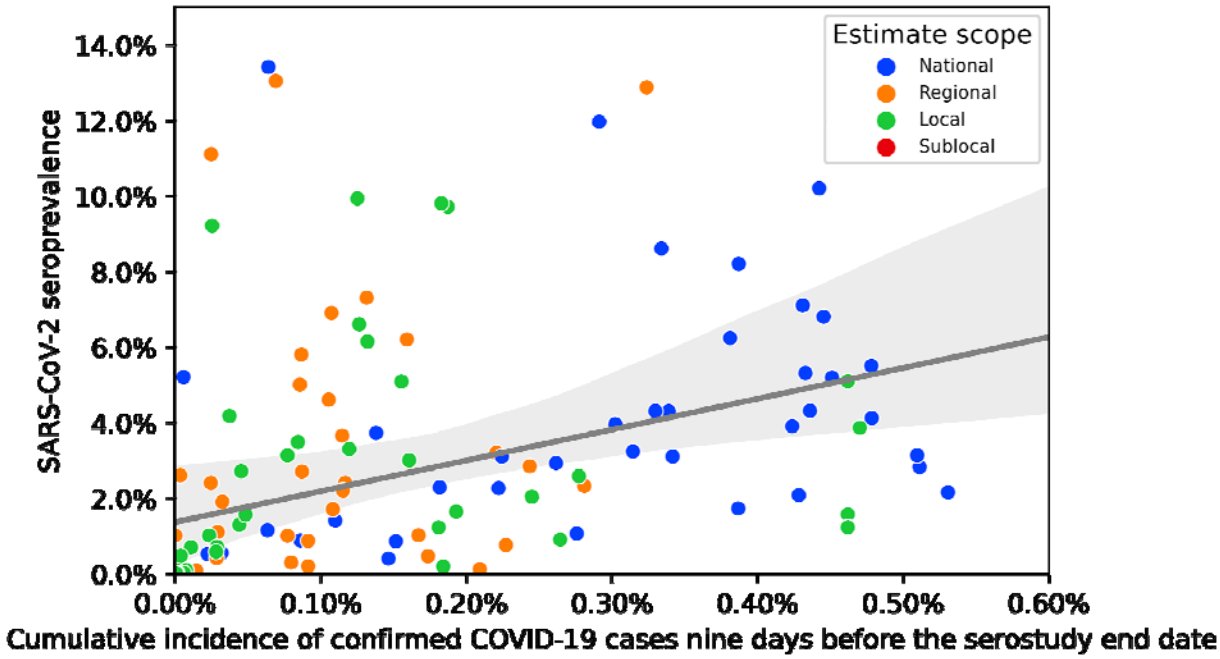
3.5 Factors affecting seroprevalence

On multivariable meta-regression, studies at low risk of bias reported lower corrected seroprevalence estimates relative to moderate risk of bias studies (prevalence ratio 0.38x, 95% CI 0.19-0.73) and high risk of bias studies (0.44x, 95% CI 0.21-0.89) times estimates from studies at moderate and high risk of bias, respectively (Supplementary Table 7). Blood donors and residual sera groups, both used as proxies for the general population, reported significantly lower corrected seroprevalence estimates compared to household and community samples (blood donors: 0.64x, 95% CI 0.42-0.98; residual sera: 0.63x, 95% CI 0.41-0.96). National studies reported similar seroprevalence estimates to regional studies (0.98x, 95% CI 0.66-1.46), but lower estimates than local (0.59x, 95% CI 0.38-0.91) and sublocal studies (0.45x, 95% CI 0.15-1.33). Finally, compared to high-income countries, countries in Sub-Saharan Africa (2.92x, 95% CI 1.00-8.50) reported higher seroprevalence estimates, while countries in Southeast Asia, East Asia, and Oceania (0.49x, 95% CI 0.11-0.39) and Latin America and Caribbean (0.50x, 95% CI 0.30-0.81) reported lower seroprevalence estimates. Visual checks confirmed that model assumptions of normality and homoscedasticity were met.

3.6 Seroprevalence to cumulative incidence ratio

The median ratio between corrected seroprevalence estimates and the corresponding cumulative incidence of SARS-CoV-2 infection was 14.5 (IQR 8.2 - 39.7, n = 125 studies; Figure 4). This ratio was higher for estimates from local studies (median 24.0, IQR 8.4 - 47.9, n=44 studies) than national studies (median 11.9, IQR 8.0 - 16.6, n=40 studies) and regional studies (median 15.7, 7.9-55.5, n=41 studies). Using the cumulative incidence on the same day as the serosurvey end date (11.9 [IQR 6.0 – 24.2]) and 14 days (16.9 IQR 9.2 – 56.7] prior yielded similar results (Supplementary Figures 2, 3).

Figure 4. Seroprevalence to cumulative case incidence ratios using cumulative incidence nine days prior to the serosurvey end date



4. Discussion

This systematic review and meta-analysis provides an overview of global SARS-CoV-2 seroprevalence based on data from 2,305,376 participants in 338 serosurveys from 281 reports. Seroprevalence remains low in the general population (median 3.2%, IQR 1.0-6.4%), with slightly higher seroprevalence in at-risk populations (e.g., health care workers, specific patient groups, and essential non-healthcare workers; median 5.4%, IQR 1.5-18.4%).

Seroprevalence varied considerably between GBD regions after correcting for study characteristics and test sensitivity and specificity. Given the limited evidence for altitude or climate effects on SARS-CoV-2 transmission[32], variations likely reflect differences in community transmission and public health responses. Stakeholders should carefully review the infection control measures implemented in Southeast Asia, East Asia, and Oceania, as well as Latin America and the Caribbean, as they appear to have been effective at limiting SARS-CoV-2 transmission.

Our results suggest clear population differences in SARS-CoV-2 burden, with marginalized and high-risk groups disproportionately affected. Differences in infection risk based on race might be attributed to crowding, higher-risk occupation roles (e.g., front-line service jobs) and other systemic inequities. Our review further found that health care workers and individuals who had close contact with confirmed COVID-19 cases had a higher risk of seropositivity, consistent with previous reports.[33] Some of these groups (Black, Asian, and minority ethnic) are also known to have higher infection fatality rates.[34,35] Such differences may inform enrolment in vaccine clinical trials and policy on vaccine distribution.

Disproportionately few studies (25%) have been conducted in low- and middle-income countries. Results from the ongoing WHO Unity studies will help to bridge this knowledge gap and inform an equitable plan for global vaccine distribution. Similarly, even in high income countries, only a handful of studies have targeted people experiencing homelessness or continuing care facility residents and staff, despite their heightened risk for SARS-CoV-2 transmission and poor health outcomes.[36,37]

Nearly half (n = 77; 42%) of studies examining SARS-CoV-2 seroprevalence in the general population used blood from donors and residual sera as a proxy. Our results showed that these studies report seroprevalence estimates that are 40% lower than studies of household and community-based samples. It has previously been shown that these groups contain disproportionate numbers of people that are young, White, college graduates, employed, physically active, and never-smokers.[38,39] Investigators using these proxy sampling frames should ensure robust correction for demographic differences or consider incorporating a correction factor (1.67x) to yield more representative estimates.

Systematic reviews of SARS-CoV-2 serological test accuracy have found that many tests have poor sensitivity and specificity.[15,16] Of the studies included here, only 69 (20.4%) corrected for test sensitivity and specificity - fewer than the 84 (24.9%) serosurveys which failed altogether to report identifying information for test used. Our study corrected seroprevalence estimates for test sensitivity and specificity. The median absolute difference between corrected and uncorrected estimates was 1.2% — a substantial change, given that the median corrected seroprevalence reported in general population studies was 3.2%. This difference emphasizes the importance of conducting such corrections to minimize bias in serosurvey data.

Seroprevalence estimates were 14.5 times higher than the corresponding cumulative incidence of COVID-19 infections. Within countries, there are substantial differences in seroprevalence between national studies and local studies. This study reports the new finding that there is more pronounced under-ascertainment when data from local seroprevalence studies are used (24.0 local vs. 11.9 national vs. 15.7 regional). This may be because many local studies have been conducted in hot-spot regions, where transmission overwhelmed diagnostic testing capacity. This level of under-ascertainment suggests that confirmed SARS-CoV-2 infections are an especially poor indicator of the true extent of infection burden in these hot-spot areas, and emphasizes the importance of interpreting seroprevalence findings in context and accounting for geographic scope.

Seroprevalence to cumulative case ratios provide a roadmap for public health authorities by identifying regions, countries, and locales that may be receiving potentially insufficient levels of testing. These ratios are also valuable for estimating true infection rates from test-confirmed

infection counts in intervals between seroprevalence studies. Applying the 11.9x ratio for national studies to the number of confirmed infections suggests that SARS-CoV-2 may have already infected 643 million people globally, rather than the 54 million reported as of November 17, 2020 — and that the recent global surge may involve 7.1 million new infections each day, as opposed to the 600,000 test-confirmed infections being reported.[1]

Our study has limitations. Firstly, some asymptomatic individuals may not seroconvert and some individuals may have been tested prior to seroconversion, so the data in this study may underestimate the true number of SARS-CoV-2 infections.[40] To ameliorate this, we prioritized estimates that tested for IgG antibodies, which show better persistence in serum compared to non-IgG and anti-nucleocapsid antibodies.[15–20] Secondly, to account for measurement error in seroprevalence estimates resulting from poorly performing tests, it was necessary to use sensitivity and specificity information from multiple sources of varying quality. While we prioritized independent evaluations, these were not available for all tests. Thirdly, the residual heterogeneity in our meta-regression indicates that not all relevant explanatory variables have been accounted for. There may be other factors that confound the associations we identified in our analysis. However, a key driving factor may simply be true differences in spread of infection and impact of the pandemic. Finally, we were only able to incorporate cumulative case incidence published on national, regional, or local government dashboards. This may have systematically excluded areas too under-resourced to conduct or report mass diagnostic testing.

Our systematic review is the largest synthesis of SARS-CoV-2 serosurveillance data to date. Our search was rigorous and comprehensive: we included non-English articles, government reports and unpublished data, and serosurveillance reports obtained via expert recommendations and our SeroTracker website. This comprehensive search is important because many serosurveys — especially in LMICs — have not been published or released as preprints. This is the first systematic review and meta-analysis to correct prevalence estimates for test sensitivity and specificity, revealing that imperfect sensitivity and specificity have major effects on seroprevalence findings. Furthermore, our synthesis accounts for study scope, enabling us to identify gaps between case incidence and serologic testing at different geographic levels and suggesting a need to increase testing capacity in many jurisdictions. To our knowledge, this is

the first study to systematically compare seroprevalence estimates from blood donors, residual sera, and household and community-based general population samples. Finally, this study is part of a regularly-updated systematic review, and summary results will continue to be disseminated throughout the pandemic on a publicly available website (SeroTracker.com).[9]

Serosurveillance efforts so far have mostly taken the form of formal studies led by academic institutions. This approach makes sense for the current role that serosurveys play - to monitor the true burden of infection and identify high-risk groups. However, as vaccines are deployed, the value of serosurveys will likely shift towards measuring population antibody titres as a correlate of protection, and evaluating vaccine effectiveness in the real world. Going forward, serosurveillance efforts may better serve end-users if they take the form of real-time monitoring programs housed in public health units. Leaders who can pair vaccine distribution data with live serosurveys will be well-equipped to track the outcomes of vaccination efforts in their communities in real time.

Our review shows that SARS-CoV-2 seroprevalence remains low in the general population, indicating that many people remain susceptible to infection and suggesting that naturally-derived herd immunity is not achievable without substantial morbidity, mortality, and strain on health services. These findings also highlight the importance of remaining vigilant until effective vaccines are broadly available. There are clear population differences in SARS-CoV-2 burden, with certain marginalized (Black and Asian persons) and at-risk populations (health care workers, essential non-health care workers, specific patient groups, close contacts) disproportionately affected. Policy and decision makers need to better protect these groups to reduce inequity in the distribution and impact of COVID-19. Such differences may inform policy on vaccine distribution.

As the COVID-19 pandemic progresses and serology data accumulate, ongoing evidence synthesis is needed to inform public health policy. We will continue to update our systematic review and seroprevalence dashboard to help address this need.

Contributors
The study was conceived by RKA, NB, TY, TGE, JP, and MPC. The protocol and data collection methods were designed by NB, RKA, CC, EB, ML, ND, JVW, CY, JP, and MPC. Analysis methods were designed by RKA, ML, NB, JoC, JP, and MPC. Article screening, data extraction, and critical appraisal were conducted by NB, RA, CC, EB, ML, HR, CD, NI, ND, JVW, TY, LP, MS, JuC, and MW. Additional data was collected by CC, EB, ML, and NB. The data for this manuscript and the companion dashboard was managed by NB, CC, JVW, ND, AA, SR, and AJ. Data was analyzed by RKA, ML, AA, SR, and AJ. Data was interpreted by NB, RKA, CC, EB, ML, DC, CPY, TW, TGE, JoC, JP, and MPC. The first draft was written by NB, RKA, CC, EB, ML, and MPC. NB, RKA, CC, EB, and ML verified the underlying data. All authors debated, agreed to the findings, and provided critical revisions to the paper.

Declaration of interests
DAC reports personal fees from Oxford University Innovation, Biobeats, and Sensyne Health. MPC reports grants from McGill Interdisciplinary Initiative in Infection and Immunity and grants from Canadian Institutes of Health Research during the conduct of the study; personal fees from GEn1E Lifesciences (as a member of the scientific advisory board) and personal fees from nplex biosciences (as a member of the scientific advisory board), both outside the submitted work. JP reports grants and personal fees from BD Diagnostics, Seegene, Janssen Pharmaceutical and AbbVie, grants from MedImmune and Sanofi Pasteur, outside the submitted work.

Acknowledgments
We would like to thank Dr. Diane Lorenzetti for her assistance in developing the search strategies. We would also like to thank all serosurvey authors who contributed data and enhanced the quality of this review. CPY holds a “Chercheur-boursier clinicien” career award from the Fonds de recherche du Québec – Santé (FRQS). JC holds a Canada Research Chair in Global Environmental Health and Epidemiology.

Role of the funding source
This research was funded by the Public Health Agency of Canada through Canada’s COVID-19 Immunity Task Force. Our funding source had no role in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. We have not been paid to write this article by a pharmaceutical company or other agency. The corresponding author (NB) confirms that all authors had full access to the full data in the study and accepts responsibility to submit for publication.

References

1. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard [Internet]. [cited 2020 Nov 11]. Available from: https://covid19.who.int/

2. Cheng MP, Papenburg J, Desjardins M, Kanjilal S, Quach C, Libman M, et al. Diagnostic Testing for Severe Acute Respiratory Syndrome–Related Coronavirus 2: A Narrative Review. Ann Intern Med. 2020 Jun 2;172(11):726–34.

3. American Society for Microbiology. Supply Shortages Impacting COVID-19 and Non-COVID Testing [Internet]. 2020 [cited 2020 Nov 11]. Available from: https://asm.org/Articles/2020/September/Clinical-Microbiology-Supply-Shortage-Collecti-1

4. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 Testing and Positivity in New York City. Am J Prev Med. 2020/06/25 ed. 2020 Sep;59(3):326–32.

5. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. PLoS Med. 2020 Sep;17(9):e1003346.

6. Cheng MP, Yansouni CP, Basta NE, Desjardins M, Kanjilal S, Paquette K, et al. Serodiagnostics for Severe Acute Respiratory Syndrome-Related Coronavirus 2□: A Narrative Review. Ann Intern Med. 2020 Sep 15;173(6):450–60.

7. Bobrovitz N, Arora RK, Yan T, Rahim H, Duarte N, Boucher E, et al. Lessons from a rapid systematic review of early SARS-CoV-2 serosurveys. medRxiv. 2020 May 14;2020.05.10.20097451.

8. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009 Jul 21;339:b2535.

9. Arora RK, Joseph A, Van Wyk J, Rocco S, Atmaja A, May E, et al. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. Lancet Infect Dis [Internet]. [cited 2020 Nov 11]; Available from: https://doi.org/10.1016/S1473-3099(20)30631-9

10. Jackson JL, Kuriyama A, Anton A, Choi A, Fournier J-P, Geier A-K, et al. The Accuracy of Google Translate for Abstracting Data From Non-English-Language Trials for Systematic Reviews. Ann Intern Med. 2019 Nov 5;171(9):677–9.

11. Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection [Internet]. World Health Organization; 2020 Mar [cited 2020 May 5]. Report No.: WHO/2019-nCoV/Seroepidemiology/2020.1. Available from: https://www.who.int/publications-detail/population-based-age-stratified-seroepidemiological-investigation-protocol-for-covid-19-virus-infection

12. Clapham H, Hay J, Routledge I, Takahashi S, Choisy M, Cummings D, et al. Seroepidemiologic Study Designs for Determining SARS-COV-2 Transmission and Immunity. Emerg Infect Dis. 2020 Sep;26(9):1978–86.

13. Institute for Health Metrics and Evaluation. Institute for Health Metrics and Evaluation Frequently Asked Questions [Internet]. [cited 2020 Nov 11]. Available from: http://www.healthdata.org/gbd/faq

14. Institute for Health Metrics and Evaluation. Global Burden of Disease Study 2019 (GBD 2019) Data Resources [Internet]. [cited 2020 Nov 11]. Available from: http://ghdx.healthdata.org/gbd-2019;

15. Bastos ML, Tavaziva G, Abidi SK, Campbell JR, Haraoui L-P, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ-Br Med J. 2020 Jul 1;370.

16. Deeks JJ, Dinnes, J, Takwoingi, Y, Davenport, C, Spijker, R, Taylor-Phillips, S, Adriano, A, Beese, S, Dretzke, J, Ferrante di Ruffano, L, Harris, IM, Price, MJ, Dittrich, S, Emperador, D, Hooft, L, Leeflang, MMG, Van den Bruel A. Antibody tests for identification of current and past infection with SARS□CoV□2. Cochrane Database Syst Rev [Internet]. 2020;(6). Available from: https://doi.org//10.1002/14651858.CD013652

17. Ward H, Cooke G, Atchison C, Whitaker M, Elliott J, Moshe M, et al. Declining prevalence of antibody positivity to SARS-CoV-2: a community study of 365,000 adults. medRxiv. 2020 Oct 27;2020.10.26.20219725.

18. Bolotin S, Tran V, Osman S, Brown KA, Buchan SA, Joh E, et al. SARS-CoV-2 seroprevalence survey estimates are affected by anti-nucleocapsid antibody decline. medRxiv. 2020 Sep 29;2020.09.28.20200915.

19. Wajnberg A, Amanat F, Firpo A, Altman DR, Bailey MJ, Mansour M, et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. Science. 2020 Oct 28;eabd7728.

20. Isho B, Abe KT, Zuo M, Jamal AJ, Rathod B, Wang JH, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. Sci Immunol. 2020 Oct 8;5(52):eabe5511.

21. Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. Int J Evid Based Healthc. 2015 Sep;13(3):147–53.

22. Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. J R Stat Soc Ser C-Appl Stat.

23. FIND. SARS-CoV-2 diagnostics: performance data [Internet]. 2020. Available from: https://www.finddx.org/covid-19/dx-data/

24. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019 Nov;22(4):153–60.

25. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 5.1.0. The Cochrane Collaboration; 2011. Available from: www.handbook.cochrane.org

26. European Centre for Disease Prevention and Control. Sources – EU/EEA and UK regional data on COVID-19 [Internet]. 2020 [cited 2020 Nov 11]. Available from: https://www.ecdc.europa.eu/en/publications-data/sources-eueea-and-uk-regional-data-covid-19

27. European Centre for Disease Prevention and Control. Sources - Worldwide data on COVID-19 [Internet]. 2020 [cited 2020 Nov 11]. Available from: https://www.ecdc.europa.eu/en/publications-data/sources-worldwide-data-covid-19

28. United Nations, Department of Economics and Social Affairs, Population Division. 2019 Revision of World Population Prospects [Internet]. [cited 2020 Nov 11]. Available from: https://population.un.org/wpp/

29. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA. 2020 Aug 25;324(8):782–93.

30. Wilson N, Kvalsvig A, Barnard LT, Baker M. Case-Fatality Risk Estimates for COVID-19 Calculated by Using a Lag Time for Fatality. Emerg Infect Dis J. 2020;26(6):1339.

31. Vanella P, Wiessner C, Holz A, Krause G, Moehl A, Wiegel S, et al. The role of age distribution, time lag between reporting and death and healthcare system capacity on case fatality estimates of COVID-19. medRxiv. 2020 Jan 1;2020.05.16.20104117.

32. Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, et al. No Association of COVID-19 transmission with temperature or UV radiation in Chinese cities. Eur Respir J. 2020 Jan 1;2000517.

33. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect Dis. 2020 Aug 1;20(8):911–9.

34. Pan D, Sze S, Minhas JS, Bangash MN, Pareek N, Divall P, et al. The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. EClinicalMedicine [Internet]. 2020 Jun 1 [cited 2020 Nov 12];23. Available from: https://doi.org/10.1016/j.eclinm.2020.100404

35. Barroso B., Valverde-Monge M., Canas J.A., Rodrigo-Munoz J.M., Gonzalez-Cano B., Villalobos-Violan V., et al. Presenting prevalence, characteristics and outcome of asthmatic patients with T2 diseases in hospitalized subjects with COVID-19 in Madrid, Spain. J Investig Allergol Clin Immunol. 2020;0.

36. Lewer D, Braithwaite I, Bullock M, Eyre MT, White PJ, Aldridge RW, et al. COVID-19 among people experiencing homelessness in England: a modelling study. Lancet Respir Med [Internet]. [cited 2020 Nov 12]; Available from: https://doi.org/10.1016/S2213-2600(20)30396-9

37. European Centre for Disease Prevention and Control. Surveillance of COVID-19 in long-term care facilities in the EU/EEA [Internet]. Stockholm; 2020 May. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-long-term-care-facilities-surveillance-guidance.pdf

38. Goldman M, Steele WR, Di Angelantonio E, van den Hurk K, Vassallo RR, Germain M, et al. Comparison of donor and general population demographics over time: a BEST Collaborative group study. Transfusion (Paris). 2017 Oct;57(10):2469–76.

39. Patel EU, Bloch EM, Grabowski MK, Goel R, Lokhandwala PM, Brunker PAR, et al. Sociodemographic and behavioral characteristics associated with blood donation in the United States: a population-based study. Transfusion (Paris). 2019/06/20 ed. 2019 Sep;59(9):2899–907.

40. Gallais F, Velay A, Wendling M-J, Nazon C, Partisani M, Sibilia J, et al. Intrafamilial Exposure to SARS-CoV-2 Induces Cellular Immune Response without Seroconversion. medRxiv. 2020 Jan 1;2020.06.21.20132449.