Big Data and Information Analytics

BDIA, 9: 350–371.
DOI: 10.3934/bdia.2025017
Received: 16 November 2025
Revised: 02 December 2025
Accepted: 05 December 2025
Published: 11 December 2025

https://www.aimspress.com/journal/BDIA

Review

Revisit the infection attack rate of COVID-19 in the first year of the pandemic

Mengyu Xie*, Li Wen and Daihai He*

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong SAR, China

* **Correspondence:** Email: mengyu.xie@connect.polyu.hk, daihai.he@polyu.edu.hk; Tel: +85298330794, +85227667864.

Abstract: The coronavirus disease 2019 (COVID-19) pandemic has caused a tremendous impact on human society. It is important to predict or estimate the spreading speed of the virus, that is, the proportion of the population being infected in the first year of the pandemic, or the infection attack rate (IAR) in 2020. In this work, we reviewed estimates published in high-profile journals and concluded that the IAR was less than or close to 1/3 in most of the countries/regions. These estimates were built on various data and statistical and mathematical methods. Interestingly, this value was in line with an early prediction by He et al., posted online on 27 March 2020.

Keywords: infection attack rate; cumulative incidence; seroprevalence; SARS-CoV-2; mathematical modeling

1. Introduction

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-COV-2), is the largest global health crisis since the 1968 influenza A/H3N2 pandemic [1].

There are significant differences in the initial stage of the epidemic in different countries which are related to the time of virus transmission, population size, susceptibility, and implementation of prevention and control measures. For example, although European countries and China did not have many deaths in the early stage of the epidemic, they took very early prevention and control measures; South Korea, on the other hand, quickly adopted strict epidemic prevention measures before the large-scale spread of the virus. Comparatively speaking, the United States began to take effective measures

after many deaths [2]. These differences directly affect the estimation of the early infection rate, indicating the profound impact of prevention and control strategies, time nodes, and social structure on the development of the epidemic.

With the continuous development of the epidemic, vaccination is considered the key measure to curb the spread of COVID-19. Pfizer and BioNTech successfully developed the first new coronavirus vaccine in December 2020 [3]. Vaccination plays an important role in improving the immunity of the population and reducing the transmission speed of SARS-CoV-2 in the population. However, due to differences in vaccination rates, virus variants, and prevention and control strategies between countries, the contribution of vaccination to mass immunization varies significantly in different countries and regions. Therefore, the popularity of vaccination not only affects the control of the epidemic but also affects the estimation of the infection rate at a later stage. Consequently, the estimation of cumulative infection rates after vaccine rollout may be biased if pre-vaccination infection levels are overlooked. Although many studies have focused on the data after the vaccine, the infection rate before the vaccine is still critical because it directly affects the prediction of the level of population immunity, the transmission path of the epidemic, and the future evolution of the virus. Especially in the early stage of the epidemic, the transmission characteristics of the virus, the viral load, and the prevention and control measures in different countries have an important impact on the estimation of the infection rate. Therefore, paying special attention to the pre-vaccination infection rate is necessary to guarantee an accurate assessment of the epidemic situation and provide more accurate data for the future pandemic prevention and control strategy.

When estimating the cumulative infection rate of COVID-19, the selected detection method, sample data, and estimation model will affect the accuracy of the results. It is critical to choose the appropriate detection method, because different detection methods may lead to different results, which will affect the estimation of the real infection. At present, the main detection methods of COVID-19 include polymerase chain reaction (PCR) detection and serological detection. The accuracy of PCR detection depends on whether the sample contains enough viral genome and the sensitivity of the reverse transcription-polymerase chain reaction (RT-PCR) [4]. However, in the early stage of infection, especially when the viral load is low, PCR detection may not be able to accurately detect the virus, which leads to the potential infections, and mild patients may be missed, thus underestimating the actual number of infections. Therefore, PCR detection carries a significant risk of false negatives, particularly during the early stage of infection. In contrast, serological tests identify people who have been infected by detecting the immune response (such as antibodies) of the human body to it. Generally, serological tests can detect three types of antibodies: immunoglobulin G (IgG), immunoglobulin M (IgM), and immunoglobulin A (IgA). However, due to the lag of antibodies, individuals may not have produced enough antibodies in the post-infection window (i.e., within the first two weeks after infection), which may lead to the failure of early infected individuals to be detected [5]. Therefore, serological tests may underestimate the early infection rate, especially when there is not enough immune response. To weaken the negative impact of the limitations of these detection methods on the estimation of cumulative infection rate, some studies chose to combine the reported case data and serological data for mathematical modeling and improve the accuracy of estimation by integrating different data sources. These methods provide a more comprehensive estimate of the infection rate and transmission trend, and they are an important basis for global epidemic assessment and the optimization of prevention and control strategies.

In addition, the selection of different types of sample data may significantly affect the estimation

of the cumulative infection rate. For example, antibody detection data may be less sensitive to asymptomatic or slightly infected individuals, resulting in an underestimate of the actual infection rate [6]. The references selected in this study cover various data types, including blood donor serum samples, external antibody detection data, case report data, cross-sectional survey data, case detection and PCR data, model prediction data, and census or large-scale sample data, among others. Each data type may have different bias problems in the process of sample selection and data collection, which will be discussed in detail later.

The methods for estimating the cumulative infection rate can be roughly divided into two categories: the direct estimation methods and indirect estimation methods. Direct estimation methods include multiplying the positive rate of serum by the total population, enzyme-linked immunosorbent assay (ELISA) optical density ratio, Bayesian mixture model, and so forth. The indirect estimation methods reversely calculate the number of infections through death data or estimates the infection by combining Monte Carlo simulation with death data and serum positivity. Some researchers used external data calibration to compare different serological surveys, CDC data, and other official records to verify the accuracy of the estimated results. In addition, because the antibody level gradually declined over time, studies using antibody detection data pay special attention to this factor. In the estimation process, the researchers corrected the effect of antibody decline by adjusting the exposure estimation and serum positive rate. For example, the Rogan-Gladen estimator [7] is widely used in serological research to adjust the sensitivity and specificity of the test to improve the accuracy of serum positive rate estimation. There are significant differences in the statistical test methods. For example, the application of the chi-square test, t-test, Mann-Whitney test, and Bayesian method depends on the focus of the research, and the Bayesian method is often used to adjust the sensitivity and specificity of the test. Following the definition of the primary estimating methodologies, the next step is to illustrate their application in various studies across different nations and how they together enhance our transnational comprehension of cumulative infection rates prior to vaccination.

Based on the previously mentioned methodological approach, this report brings together estimates of the cumulative pre-vaccination infection rate in 23 countries (18 high-income countries and 5 non-high-income countries). High-income countries usually have strong public health infrastructure, such as a higher detection capacity, medical resources, and a better epidemic data collection system, which is more suitable for estimating the pre-vaccination infection rate than for non-high-income countries. The non-high-income countries which we have chosen to study are Afghanistan, Brazil, China, South Africa, and Russia, which can help us fully understand the general situations of Asia, Africa, and Latin America. In particular, China was the initial outbreak location of the COVID-19 epidemic and implemented strict lockdowns along with prevention and control measures during the early stages of the epidemic. The experience of China is of enormous significance globally, especially in terms of the early transmission characteristics of the epidemic and the effect of prevention and control measures [8].

2. Methodology

The authors performed a systematic review of relevant literature and conducted a manual search to examine the seroprevalence of SARS-CoV-2 infection, with a specific focus on studies addressing cumulative infection rates before the widespread vaccination. This review was conducted and reported following the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Figure 1).

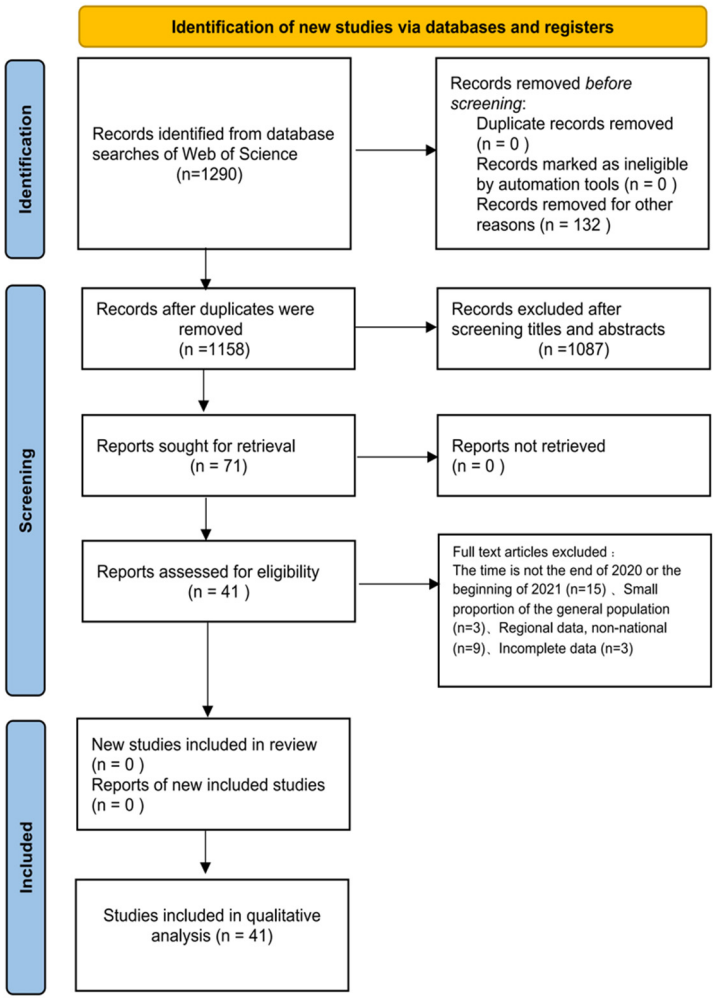

Figure 1. Flowchart of the identification, inclusion, and exclusion of studies. The boxes represent different steps of the process. Arrows indicate the flow from one step to the next.

2.1. Search methodology and article selection

This study systematically reviewed the relevant literature on the estimation of the true cumulative infection rate prior to vaccination. A literature search was conducted using the Web of Science electronic database, using various combinations of keywords including "SARS-CoV-2 cumulative infection rate", "SARS-CoV-2 attack rate", "SARS-CoV-2 cumulative incidence", "SARS-CoV-2 seroprevalence", "COVID-19 cumulative infection rate", "COVID-19 attack rate", "COVID-19 cumulative incidence", and "COVID-19 seroprevalence". A total of 1290 potentially relevant articles were initially retrieved.

To ensure academic rigor and comparability of the included literature, we used the filtering tools of the Web of Science platform to restrict the document type to "Article", yielding 1158 records. CiteSpace was employed for duplicate removal. Based on this dataset, we manually screened the titles and abstracts, excluding studies that (1) only covered localized or regional data, (2) did not estimate infection rates prior to vaccination, (3) did not provide final infection rate estimates, or (4) focused solely on specific subpopulations rather than the general population. After screening, 71 representative studies were retained.

Following a more rigorous selection process, we identified 41 high-quality studies for in-depth analysis. One of these studies provided a global estimate of cumulative infection up to November 2021, while the remaining 40 studies covered 18 high-income countries and 5 non-high-income countries (including China). The selection of countries ensured not only broad geographic representation but also diversity and comprehensiveness in terms of pandemic progression, public health response strategies, and data quality. Some high-income countries—such as the United States, Japan, and France—provided longitudinal data on infection rates spanning from early 2020 to the end of 2021. These datasets offered robust empirical support for the early-stage prediction model proposed by our research team and validated our projection, made on March 27, 2020, that the infection attack rate of COVID-19 in the early phase of the pandemic ranged between 14.8% and 26.8%. For data extraction, the following information was recorded for each included study: first author, year of publication, study dates, country, estimated cumulative infection rate (or attack rate), detection method, number of participants, and population type for each included study. The studies were then categorized alphabetically according to the first letter of each country's English name. For studies that reported estimates at multiple time points, each estimate was recorded individually and organized chronologically, allowing for a more comprehensive analysis of the dynamic evolution of the pandemic across different nations. Comprehensive details are presented in Table 1. The literature search was completed on June 28, 2025, and only English literature was included, ensuring the wide applicability and comparability of the results.

2.2. Selection criteria and thematic interpretation

Among the 41 selected articles, the majority were published in Q1 or Q2 journals as classified by the Journal Citation Reports (JCR), which ensures a high level of methodological rigor and data reliability. The literature spans a broad geographical scope and incorporates diverse estimation approaches, thereby contributing to a more comprehensive and robust analysis of cumulative infection rates.

The comparative analysis of these studies is categorized into three key dimensions:

1) estimation of cumulative infection rates across different regions,
2) diagnostic methods and sample data sources,
3) estimation methodologies.

Specifically, the distribution of studies by country is as follows: one study focused on Australia, three on Canada, one on Croatia, one on the Czech Republic, two on Denmark, two on the United Kingdom (U.K.), one on Finland, two on France, two on Germany, one on Greece, one on Israel, one on Italy, three on Japan, one on South Korea, one on the Netherlands, two on Norway, one on Sweden, and nine on the United States. Additionally, there is one study from each of several other countries, along with a global study that provides a comprehensive overview of global infection patterns.

These studies collectively encompass a wide range of countries and regions, thereby ensuring the geographical diversity and representativeness necessary for a robust and comprehensive assessment. The breadth of coverage provides substantial empirical support for our estimation of pre-vaccination cumulative infection rates across different global contexts.

3. Results

3.1. Cumulative estimated infection rates in different regions

To begin, the study provides an important benchmark for the global cumulative infection rate prior to the emergence of the Omicron (B.1.1.529) variant: as of November 14, 2021, more than 40% of the world's population had been infected with SARS-CoV-2 at least once [9]. Additionally, by August 2021, approximately 50% of the U.S. population had completed vaccination [10].

In this review, we included nine studies that estimated the cumulative infection rate in the United States between July 2020 and June 2021 [11–19]. These estimates ranged from 3.5% to 36%. Particularly, studies based on serological detection methods showed a relatively stable upward trend, from 3.5% in July 2020 to approximately 20% in January 2021. Among them, the work is especially critical, as it reported estimates at two key time points: 3.5% (95% CI: 3.2%–3.8%) in July 2020 and 20.2% (95% CI: 19.9%–20.6%) in May 2021 [13].

For other high-income countries, two representative studies were selected to enable cross-national comparison. In France, estimates from July 2020 to January 2021 ranged from 7.6% (95% CI: 7.3%–7.8%) to 14.9% (95% CI: 13.2%–16.9%) [20, 21]. In contrast, Japan reported significantly lower infection rates over the same period, ranging from 0.12% (95% CI: 0.06%–0.24%) to 3.8% [22–24]. In Germany, the infection rate remained low in 2020 (below 2%) [25,26]. In Denmark, the rate increased from 4.1% (95% CI: 3.1%–4.9%) in late 2020 to 7.2% (95% CI: 6.3%–7.9%) by February 2021 [27,28]. In the United Kingdom, the infection rate rose from 6% in July 2020 to 9.1% (95% CI: 8.7%–9.8%).

Regarding Canada, available evidence indicates that the cumulative infection rate remained below 5% up to April 2021, followed by a gradual increase to approximately 9% by November 2021. Therefore, it can be reasonably inferred that the pre-vaccination infection rate was below 5% [29–31]. Other high-income countries also reported relatively low cumulative infection rates (below 10%) around or prior to January 2021, including Australia (<0.50%), Finland (<7.00%), Greece (9.09%), Israel (6.30%), Italy (6.90%), Norway (<3.20%), and South Korea (0.62%) [32–39].

For most non-high-income countries, the pre-vaccination cumulative infection rates were generally higher: Afghanistan (29.80%), Brazil (11.00%), Russia (19.20%, IQR: 19.0%–19.6%), and South Africa (14.50%, 95% CI: 13.90%–15.20%) [40–43]. For China, based on data from January 2020 to January 2022, the cumulative infection rate by the end of 2020 can be indirectly inferred to range between 3.2% and 37.2% [44].

3.2. Test method and sample data

The detection of active SARS-CoV-2 infection primarily relies on RT-PCR, whereas serological assays such as ELISA are used to identify past infection by detecting antibodies [45]. Antibody testing assesses the immune response by measuring the presence and concentration of immunoglobulins IgG and IgM in blood, serum, or plasma samples. The most commonly used antibody detection methods include lateral flow assay (LFA) and ELISA [5].

Among the 41 articles included in this review, 33 employed only serological detection methods, one utilized PCR-based approaches, one study combined PCR and serological (ELISA) techniques to detect S1-specific IgG antibodies, and the remaining 6 articles employed mathematical modeling approaches to estimate the attack rate, utilizing data on reported deaths, confirmed cases, and other relevant epidemiological indicators.

It is evident that most studies estimating cumulative infection rates relied on serological detection methods. Although serological assays may be influenced by cross-reactivity and timing of antibody

response, and potential confounders such as non-specific IgM may introduce significant variability in COVID-19 patients—thereby leading to potential inaccuracies—serological approaches offer important advantages [6]. Seroepidemiological methods are better suited for capturing medium- to long-term antibody responses following SARS-CoV-2 infection, especially in contexts where the virus induces a stable immune response. Under such conditions, the seropositivity rate can serve as a reasonable proxy for cumulative incidence over a defined period.

In addition, because antibodies can persist for several months following infection, serological investigations offer a reliable basis for estimating cumulative infections over a recent period. Beyond estimating infection rates, serological detection can also be employed to assess residual susceptibility and monitor the development of population-level (herd) immunity. Importantly, even when mortality is not considered, serological methods remain of significant public health relevance. Furthermore, if mortality can be reasonably estimated, the value of serological investigations in understanding the true burden of infection is further enhanced [46].

For studies employing serological detection, particular attention should be given to the effects of antibody waning. This phenomenon refers to the gradual decline in immune protection over time, which may result from factors such as decreasing antibody concentrations, loss of immune memory, or the emergence of vaccine-resistant variants [47]. Antibody waning can increase susceptibility to infection and, consequently, impact the accurate estimation of the infection attack rate. Accordingly, most serology-based studies included in this review accounted for antibody waning by correcting for the sensitivity and specificity of test results and adjusting for the potential impact of declining antibody levels.

Blood donor samples are commonly used due to their accessibility and large sample size, not because donors had antibodies for therapeutic purposes. For instance, Jones et al. (2021) analyzed over 1.5 million blood donor samples [13]. Similarly, Reedman et al. (2022) utilized data provided by Canadian Blood Services (CBS), which included samples from donors aged 17 and older across all provinces except Quebec [29]. Other studies used population-based or commercial residual serum samples; for example, García-Carreras et al. (2023) analyzed anti-N antibody data derived from commercial laboratory samples provided by the CDC [18].

For studies employing PCR detection methods, data sources include U.S. case reports derived from electronic health records (IBM Watson) and surveillance platforms such as the National Notifiable Diseases Surveillance System (NNDSS). In the United Kingdom, PCR testing data are obtained from the COVID-19 dashboard (confirmed case data) and the Office for National Statistics (ONS), which relies on random population sampling. Common challenges associated with PCR-based approaches include limitations in testing capacity, lack of representativeness among participants, and under-detection of asymptomatic infections.

In addition, when neither antibody testing nor PCR detection is feasible, mathematical modeling offers a valuable alternative for inferring infection rates. The primary challenges associated with using testing data include the potential undercounting of asymptomatic infections, limited testing resources, and the failure to account for recovered individuals. Although mortality data may be criticized for potentially including deaths unrelated to COVID-19, it is generally considered more reliable than confirmed case counts, which often fail to capture most infections [48].

In this review, two studies employed mathematical models based on reported case data and incorporated serological findings to improve accuracy. For instance, Pei et al. (2021) used CDC-reported case counts alongside serological data from ten CDC surveillance sites for model calibration [15]. Two

additional studies utilized mathematical modeling based on both reported cases and mortality data. For example, Impouma et al. (2021) analyzed cumulative case and death data from 41 countries and regions across Africa [40].

4. Estimation method

4.1. Direct estimation method

The direct estimation methods calculate the infection rate using either serological data or mathematical models without relying on indirect assumptions. A straightforward example involves multiplying the seroprevalence by the estimated total population to infer the number of infections [12]. Additionally, probability-based models have been used to estimate total SARS-CoV-2 infections, including symptomatic cases and hospitalizations, such as in the study by Reese et al. (2020) [11]. Several statistical models have been applied in this context. A Bayesian mixed-effects model was used to estimate the historical infection probability of SARS-CoV-2 [21], while a Bernoulli model was applied to estimate seroprevalence [42]. Moreover, the optical density (OD) ratio from ELISA testing has also been used to estimate infection probability, particularly in identifying recent infections [23]. The main advantage of these methods is that they rely directly on empirical serological data or comprehensive modeling frameworks, making them especially suitable for large-scale epidemiological studies.

4.2. Indirect estimation method

The indirect estimation methods rely on the auxiliary data sources and inference models to estimate infection rates. A common approach involves back-calculating the time of infection based on reported deaths and the assumed time lag between infection and death. Monte Carlo simulations have also been employed to generate infection rate estimates, which can be validated using serological data [22]. In addition, the COVID-19 Prevalence Calculator developed by Resolve to Save Lives was used to estimate four key epidemiological indicators: the total number of infections, the number of underreported cases, the detection rate, and the cumulative infection rate [40]. These indirect approaches often improve estimation accuracy through statistical modeling techniques and data calibration.

4.3. Methods to improve estimation accuracy

Among the above methods, external data calibration is a common and effective strategy to improve estimation accuracy. For example, comparing CDC data with community seroprevalence surveys provides a reliable benchmark for model calibration. For the estimation of seroprevalence, the Rogan-Gladen estimator [7] is widely used to adjust for test sensitivity and specificity, thereby improving the accuracy of prevalence estimates:

$$\text{Estimated true prevalence} = \frac{\text{Observed prevalence} + (\text{Specificity} - 1)}{\text{Sensitivity} + (\text{Specificity} - 1)} \tag{1}$$

4.4. Statistical test methods

In these estimation methods, the choice of statistical test typically depends on the research

objective and the characteristics of the data. For example, the chi-square test is commonly employed to compare seroprevalence across different population groups, while the t-test is suitable for paired comparisons [13]. Logistic regression has been used to identify infection risk, estimate odds ratios, and model the probability of infection or seropositivity [49]. Bayesian methods are often applied to adjust for the sensitivity and specificity of diagnostic tests [33]. The Mann-Whitney U test is appropriate for comparisons between groups when data are not normally distributed [44], and the Clopper-Pearson method is frequently used to calculate exact confidence intervals for binomial proportions [39].

5. Discussion

A key contribution of this review lies in its focus on the proportion of infections among the populations of high-income countries during 2020 and early 2021. Based on a comprehensive synthesis of the existing literature, we estimate that the average cumulative infection rate in high-income countries was approximately one-third or lower during the initial phase of the pandemic. Specifically, the early cumulative infection rates in several representative high-income countries were as follows: the United States (3.50%–31.0%, July–December 2020), France (7.60%–14.90%, July 2020–January 2021), Germany (less than 2.00%, throughout 2020 and early 2021), Denmark (4.10%–7.20%, late 2020–February 2021), and the United Kingdom (6.00%–9.10%, July–October 2020).

In contrast, other high-income countries such as Australia, Canada, Finland, Greece, Israel, Italy, South Korea, Norway, and Japan reported cumulative infection rates below 10.00% during the same period. These differences may be attributed to variations in public health policies, intervention strategies, and the timing of local outbreaks. Nevertheless, several countries exhibited relatively high cumulative infection rates by the end of 2020 or early 2021, including Croatia (25.10%), the Czech Republic (28.00%), the Netherlands (19.50%), and Sweden (20.10%).

For example, South Korea implemented proactive containment measures immediately after the first confirmed case on January 20, 2020. The Korea Centers for Disease Control and Prevention (KCDC) introduced rapid and accessible diagnostic testing alongside efficient contact tracing technologies, both of which were critical in curbing the spread of the virus [50]. Japan's response strategy centered on three core pillars: early detection of cases, ensuring intensive care and medical service capacity, and promoting behavioral modifications among citizens [51]. Canada's relatively low infection rate may be attributed to geographic and demographic factors such as its vast land area and low population density. Furthermore, the coordinated division of responsibilities between the federal and provincial governments—wherein the federal level oversees economic policy and border control, and provinces manage public health—also contributed to the effective pandemic response [52]. Similarly, for other high-income countries, the literature generally indicates that pre-vaccination cumulative infection rates were approximately one-third or lower.

In analyzing the selection of detection methods and sample data, we found that the majority of studies estimating cumulative infection rates employed serological testing. Compared to PCR-based diagnostics, serological methods offer distinct advantages in identifying individuals who have previously been infected, particularly those with mild or asymptomatic cases. However, the limitations of serological detection—most notably, antibody waning and issues of sample representativeness—must be carefully considered. To address these concerns, many studies applied correction techniques such as sensitivity and specificity adjustments as well as corrections for antibody attenuation. For

instance, the Rogan-Gladen estimator [7] is commonly used to refine estimates of true seroprevalence. In terms of statistical analysis, the selection of specific test methods varied depending on the research objectives and data characteristics. The application of diverse statistical methodologies reflects considerable methodological heterogeneity across studies, which may influence the robustness of cumulative infection rate estimates.

Although this review provides a comprehensive perspective on the early cumulative infection rates of COVID-19 by analyzing 41 relevant studies, several limitations should be acknowledged. First, like all systematic reviews, the results of this study may also be influenced by publication bias. Especially in the early stages of the epidemic outbreak, the available data is extremely limited, research resources are tight, and studies with more significant results, higher positivity rates, or stronger policy significance are often published quickly, while studies with smaller sample sizes or insignificant results may be less likely to be published promptly. This tendency towards publication may lead to a combined estimate of the overall infection rate deviating from the true level. In addition, the quality of data varies greatly among countries in the early stages of the epidemic, and there are significant differences in case reporting standards, with common issues of underreporting and delayed reporting. This may not only lead to systematic bias in the estimation of infection rates but also further weaken the comparability between different studies.

Second, the limitations of serological testing methods themselves have a significant impact on the accuracy and comparability of the included studies. A core issue is the phenomenon of antibody decline over time (antibody waning). The sensitivity and specificity of different assays, as well as their ability to detect low antibody concentrations, vary considerably. As a result, studies conducted later in the pandemic are more likely to systematically underestimate infection rates. In addition, the sampling time points of various studies were after different epidemic waves, which further amplified the bias caused by antibody attenuation, and existing calibration methods are still insufficient to fully compensate for this difference. Closely related to this, the high heterogeneity of the criteria for determining serum positivity also significantly affects comparability between studies. The included studies used reagents from different manufacturers, followed different operating procedures, and employed different positive thresholds (cutoffs). Samples that are considered positive in one study may be classified as negative in another study due to different threshold settings. This fundamental lack of standardization not only introduces structural biases between studies but also directly weakens the comparability of serum positivity rates in this review. Finally, this review is restricted to data from high-income and select non-high-income countries. Future research should expand to include data from a broader range of low- and middle-income countries to better understand transmission dynamics under diverse economic conditions and public health systems.

6. Conclusions

Based on a comprehensive review of 41 relevant studies, the early prediction by He et al. (2020) [53]—that the infection attack rate of COVID-19 would range between 14.8% and 26.8% in the worst-case scenario during the first year of the pandemic—has been largely validated. This forecast addressed a critical gap in estimating the early cumulative infection rate of COVID-19 prior to the availability of widespread vaccination and provides an essential reference point for understanding the true infection burden during the initial phase of the outbreak.

By comparing various detection methods, estimation techniques, and statistical approaches, this review highlights the methodological complexity and heterogeneity inherent in estimating the actual

infection rate of COVID-19. Although serological testing is subject to limitations such as antibody waning and sample representativeness, adjustments for test sensitivity and specificity have enhanced its utility. When properly corrected, serological data offer valuable insights into the cumulative infection burden. Both direct and indirect estimation approaches—including Bayesian models, mathematical simulations, and data calibration methods—have contributed substantially to refining our understanding of the transmission dynamics and severity of the pandemic.

Furthermore, the findings synthesized in this review have significant implications for future research, particularly in forecasting long-term epidemic trends and preparing for emerging infectious disease threats. The experience gained from estimating early cumulative infection rates serves as a foundational reference for developing public health strategies, optimizing predictive models, and strengthening preparedness efforts. The methodologies discussed herein not only support ongoing research and epidemiological investigations but also inform policy-making and public health interventions. These insights are not only relevant to COVID-19 but also offer strategic guidance for addressing future pandemics.

Credit author statement

Mengyu Xie: writing—Original Draft, Data Curation, Visualization; Li Wen: writing—Review & Editing, Project administration; Daihai He: conceptualization, writing—Review & Editing, Supervision.

Data availability

This is a systematic review article. All data utilized in this study are available within the cited published literature.

Use of AI tools declaration

No generative AI tools were used in the creation of this manuscript.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. Shi Y, Wang G, Cai XP, Deng JW, Zheng L, Zhu HH, et al. (2020) An overview of COVID-19. *J Zhejiang Univ Sci B* 21: 343–360. https://doi.org/10.1631/jzus.B2000083
2. Middelburg RA, Rosendaal FR, (2020) COVID-19: How to make between-country comparisons. *Int J Infect Dis* 96: 477–481. https://doi.org/10.1016/j.ijid.2020.05.066
3. Saleh A, Qamar S, Tekin A, Singh R, Kashyap R, (2021) Vaccine development throughout history. *Cureus* 13: e16635. https://doi.org/10.7759/cureus.16635
4. Esakandari H, Nabi-Afjadi M, Fakkari-Afjadi J, Farahmandian N, Miresmaeili SM, Bahreini E, (2020) A comprehensive review of COVID-19 characteristics. *Biol Proced Online* 22: 1–10. https://doi.org/10.1186/s12575-020-00128-2

5. Yüce M, Filiztekin E, Özkaya KG, (2021) COVID-19 diagnosis—A review of current methods. *Biosens Bioelectron* 172: 112752. https://doi.org/10.1016/j.bios.2020.112752
6. Giri B, Pandey S, Shrestha R, Pokharel K, Ligler FS, Neupane BB, (2021) Review of analytical performance of COVID-19 detection methods. *Anal Bioanal Chem* 413: 35–48. https://doi.org/10.1007/s00216-020-03019-6
7. Rogan WJ, Gladen B, (1978) Estimating prevalence from the results of a screening test. *Am J Epidemiol* 107: 71–76. https://doi.org/10.1093/oxfordjournals.aje.a112450
8. Peng F, Tu L, Yang Y, Hu P, Wang R, Hu Q, et al. (2020) Management and treatment of COVID-19: The Chinese experience. *Can J Cardiol* 36: 915. https://doi.org/10.1016/j.cjca.2020.05.022
9. Barber RM, Sorensen RJ, Pigott DM, Bisignano C, Carter A, Amlag JO, et al. (2022) Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: A statistical analysis. *Lancet* 399: 2351–2380. https://doi.org/10.1016/S0140-6736(22)01519-1
10. Liu C, Lee J, Ta C, Soroush A, Rogers JR, Kim JH, et al. (2021) A retrospective analysis of COVID-19 mRNA vaccine breakthrough infections-risk factors and vaccine effectiveness, preprint, medRxiv: 2021.10.05.21264583. https://doi.org/10.1101/2021.10.05.21264583
11. Reese H, Iuliano AD, Patel NN, Garg S, Kim L, Silk BJ, et al. (2021) Estimated incidence of coronavirus disease 2019 (COVID-19) illness and hospitalization—United States, February–September 2020. *Clin Infect Dis* 72: e1010–e1017. https://doi.org/10.1093/cid/ciaa1780
12. Angulo FJ, Finelli L, Swerdlow DL, (2021) Estimation of US SARS-CoV-2 infections, symptomatic infections, hospitalizations, and deaths using seroprevalence surveys. *JAMA Network Open* 4: e2033706. https://doi.org/10.1001/jamanetworkopen.2020.33706
13. Jones JM, Stone M, Sulaeman H, Fink RV, Dave H, Levy ME, et al. (2021) Estimated US infection-and vaccine-induced SARS-CoV-2 seroprevalence based on blood donations, July 2020-May 2021. *JAMA* 326: 1400–1409. https://doi.org/10.1001/jama.2021.12814
14. Louca S, (2021) SARS-CoV-2 infections in 165 countries over time. *Int J Infect Dis* 111: 336–346. https://doi.org/10.1016/j.ijid.2021.05.012
15. Pei S, Yamana TK, Kandula S, Galanti M, Shaman J, (2021) Burden and characteristics of COVID-19 in the united states during 2020. *Nature* 598: 338–341. https://doi.org/10.1038/s41586-021-03578-8
16. Chitwood MH, Russi M, Gunasekera K, Havumaki J, Klaassen F, Pitzer VE, et al (2022) Reconstructing the course of the COVID-19 epidemic over 2020 for US states and counties: Results of a Bayesian evidence synthesis model. *PLoS Comput Biol* 18: e1010465. https://doi.org/10.1371/journal.pcbi.1010465
17. Sullivan PS, Siegler AJ, Shioda K, Hall EW, Bradley H, Sanchez T, et al. (2022) Severe acute respiratory syndrome coronavirus 2 cumulative incidence, United States, August 2020–December 2020. *Clin Infect Dis* 74: 1141–1150. https://doi.org/10.1093/cid/ciab610
18. García-Carreras B, Hitchings MD, Johansson MA, Biggerstaff M, Slayton RB, Healy JM, et al. (2023) Accounting for assay performance when estimating the temporal dynamics in SARS-CoV-2 seroprevalence in the US. *Nat Commun* 14: 2235. https://doi.org/10.1038/s41467-023-37950-3
19. Wiegand RE, Deng Y, Deng X, Lee A, Meyer WA, Letovsky S, et al. (2023) Estimated SARS-CoV-2 antibody seroprevalence trends and relationship to reported case prevalence from a repeated, cross-sectional study in the 50 states and the District of Columbia, United States—October 25, 2020–February 26, 2022. *Lancet Regional Health Am* 18: 100578. https://doi.org/10.1016/j.lana.2023.100578

20. Hozé N, Paireau J, Lapidus N, Kiem CT, Salje H, Severi G, et al. (2021) Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: A modelling study. *Lancet Public Health* 6: e408–e415. https://doi.org/10.1016/S2468-2667(21)00111-5
21. Glemain B, de Lamballerie X, Zins M, Severi G, Touvier M, Deleuze JF, et al. (2024) Estimating SARS-CoV-2 infection probabilities with serological data and a Bayesian mixture model. *Sci Rep* 14: 9503. https://doi.org/10.1038/s41598-024-59991-4
22. McKenzie L, Shoukat A, Wong KO, Itahashi K, Yasuda E, Demarsh A, et al. (2022) Inferring the true number of SARS-CoV-2 infections in Japan. *J Infect Chemother* 28: 1519–1522. https://doi.org/10.1016/j.jiac.2022.07.010
23. Yamayoshi S, Iwatsuki-Horimoto K, Okuda M, Ujie M, Yasuhara A, Murakami J, et al. (2022) Age-stratified seroprevalence of SARS-CoV-2 antibodies before and during the vaccination era, Japan, February 2020–March 2022. *Emerging Infect Dis* 28: 2198. https://doi.org/10.3201/eid2811.220424
24. Arashiro T, Arai S, Kinoshita R, Otani K, Miyamoto S, Yoneoka D, et al. (2023) National seroepidemiological study of COVID-19 after the initial rollout of vaccines: Before and at the peak of the Omicron-dominant period in Japan. *Influenza Other Respir Viruses* 17: e13094. https://doi.org/10.1111/irv.13094
25. Neuhauser H, Rosario AS, Butschalowsky H, Haller S, Hoebel J, Michel J, et al. (2022) Nationally representative results on SARS-CoV-2 seroprevalence and testing in Germany at the end of 2020. *Sci Rep* 12: 19492. https://doi.org/10.1038/s41598-022-22621-1
26. Offergeld R, Preussel K, Zeiler T, Aurich K, Baumann-Baretti BI, Ciesek S, et al. (2023) Monitoring the SARS-CoV-2 pandemic: Prevalence of antibodies in a large, repetitive cross-sectional study of blood donors in Germany—results from the sebluco study 2020–2022. *Pathogens* 12: 551. https://doi.org/10.3390/pathogens12040551
27. Espenhain L, Tribler S, Sværke Jørgensen C, Holm Hansen C, Wolff Sönksen U, Ethelberg S, (2021) Prevalence of SARS-CoV-2 antibodies in Denmark: Nationwide, population-based seroepidemiological study. *Eur J Epidemiol* 36: 715–725. https://doi.org/10.1007/s10654-021-00758-5
28. Krogsgaard LW, Espenhain L, Tribler S, Sværke Jørgensen C, Hansen CH, Møller FT, et al. (2023) Seroprevalence of SARS-CoV-2 antibodies in Denmark: Results of two nationwide population-based surveys, February and May 2021. *Infect Drug Resist* 16: 301–312. https://doi.org/10.2147/IDR.S395312
29. Reedman CN, Drews SJ, Yi QL, Pambrun C, O'Brien SF, (2022) Changing patterns of SARS-CoV-2 seroprevalence among Canadian blood donors during the vaccine era. *Microbiol Spectrum* 10: e00339-22. https://doi.org/10.1128/spectrum.00339-22
30. Tuite AR, Fisman D, Abe KT, Rathod B, Pasculescu A, Colwill K, et al. (2022) Estimating SARS-CoV-2 seroprevalence in Canadian blood donors, April 2020 to March 2021: Improving accuracy with multiple assays. *Microbiol Spectrum* 10: e02563-21. https://doi.org/10.1128/spectrum.02563-21
31. Murphy TJ, Swail H, Jain J, Anderson M, Awadalla P, Behl L, et al. (2023) The evolution of SARS-CoV-2 seroprevalence in Canada: A time-series study, 2020–2023. *CMAJ* 195: E1030–E1037. https://doi.org/10.1503/cmaj.221559
32. Pota M, Pota A, Sirico ML, Esposito M, (2020) SARS-CoV-2 infections and COVID-19 fatality: Estimation of infection fatality ratio and current prevalence. *Int J Environ Res Public Health* 17: 9290. https://doi.org/10.3390/ijerph17249290

33. Lee K, Jo S, Lee J, (2021) Seroprevalence of SARS-CoV-2 antibodies in South Korea. *J Korean Stat Soc* 50: 891–904. https://doi.org/10.1007/s42952-021-00143-0
34. Tunheim G, Rø GØI, Tran T, Kran AMB, Andersen JT, Vaage EB, et al. (2022) Trends in seroprevalence of SARS-CoV-2 and infection fatality rate in the Norwegian population through the first year of the COVID-19 pandemic. *Influenza Other Respir Viruses* 16: 204–212. https://doi.org/10.1111/irv.13009
35. Bassal R, Keinan-Boker L, Cohen D, Mendelson E, Lustig Y, Indenbaum V, (2022) Estimated infection and vaccine induced SARS-CoV-2 seroprevalence in Israel among adults, January 2020–July 2021. *Vaccines* 10: 1663. https://doi.org/10.3390/vaccines10101663
36. Anda EE, Braaten T, Borch KB, Nøst TH, Chen SL, Lukic M, et al. (2022) Seroprevalence of antibodies against SARS-CoV-2 in the adult population during the pre-vaccination period, Norway, winter 2020/21. *Eurosurveillance* 27: 2100376. https://doi.org/10.2807/1560-7917.ES.2022.27.13.2100376
37. Koureas M, Bogogiannidou Z, Vontas A, Kyritsi MA, Mouchtouri VA, Dadouli K, et al. (2022) SARS-CoV-2 sero-surveillance in Greece: Evolution over time and epidemiological attributes during the pre-vaccination pandemic era. *Diagnostics* 12: 295. https://doi.org/10.3390/diagnostics12020295
38. Vette KM, Machalek DA, Gidding HF, Nicholson S, O'Sullivan MV, Carlin JB, et al. (2022) Seroprevalence of severe acute respiratory syndrome coronavirus 2-specific antibodies in Australia after the first epidemic wave in 2020: A national survey. *Open Forum Infect Dis* 9: ofac002. https://doi.org/10.1093/ofid/ofac002
39. Solastie A, Nieminen T, Ekström N, Nohynek H, Lehtonen L, Palmú AA, et al. (2023) Changes in SARS-CoV-2 seroprevalence and population immunity in Finland, 2020–2022. *Emerg Microbes Infect* 12: 2222849. https://doi.org/10.1080/22221751.2023.2222849
40. Impouma B, Mboussou F, Shahpar C, Wolfe CM, Farham B, Williams GS, et al. (2021) Estimating the SARS-CoV2 infections detection rate and cumulative incidence in the world health organization African region 10 months into the pandemic. *Epidemiol Infect* 149: e264. https://doi.org/10.1017/S0950268821001988
41. Popova AY, Smirnov VS, Andreeva EE, Babura EA, Balakhonov SV, Bashketova NS, et al. (2021) SARS-CoV-2 seroprevalence structure of the Russian population during the COVID-19 pandemic. *Viruses* 13: 1648. https://doi.org/10.3390/v13081648
42. Saeedzai SA, Sahak MN, Arifi F, Aly EA, van Gurp M, White LJ, et al. (2022) COVID-19 morbidity in Afghanistan: A nationwide, population-based seroepidemiological study. *BMJ Open* 12: e060739. https://doi.org/10.1136/bmjopen-2022-060739
43. Figueiredo GM, Tengan FM, Campos SR, Luna EJ, (2023) Seroprevalence of SARS-CoV-2 in Brazil: A systematic review and meta-analysis. *Clinics* 78: 100233. https://doi.org/10.1016/j.clinics.2023.03.003
44. Guang Y, Lina L, Hui L, (2025) Seroprevalence of SARS-CoV-2 antibody before and after both vaccination and natural infection in China. *Immunity Inflammation Dis* 13: e70184. https://doi.org/10.1002/iid3.70184
45. Filchakova O, Dossym D, Ilyas A, Kuanysheva T, Abdizhamil A, Bukasov R, (2022) Review of COVID-19 testing and diagnostic methods. *Talanta* 244: 123409. https://doi.org/10.1016/j.talanta.2022.123409
46. Meyer MJ, Yan S, Schlageter S, Kraemer JD, Rosenberg ES, Stoto MA, (2022) Adjusting COVID-19 seroprevalence survey results to account for test sensitivity and specificity. *Am J Epidemiol* 191: 681–688. https://doi.org/10.1093/aje/kwab222

47. Lee B, Song H, Apio C, Han K, Park J, Liu Z, et al. (2023) An analysis of the waning effect of COVID-19 vaccinations. *Genomics Inf* 21: e50. https://doi.org/10.5808/GI.2023.21.4.e50
48. McCulloh I, Kiernan K, Kent T, (2020) Inferring true COVID19 infection rates from deaths. *Front Big Data* 3: 565589. https://doi.org/10.3389/fdata.2020.565589
49. Beser J, Galanis I, Enkirch T, Kühlmann Berenzon S, van Straten E, Duracz J, et al. (2022) Seroprevalence of SARS-CoV-2 in Sweden, April 26 to May 9, 2021. *Sci Rep* 12: 10816. https://doi.org/10.1038/s41598-022-14871-9
50. Lee D, Choi B, (2020) Policies and innovations to battle COVID-19—A case study of South Korea. *Health Policy Technol* 9: 587–597. https://doi.org/10.1016/j.hlpt.2020.07.002
51. Yan B, Zhang X, Wu L, Zhu H, Chen B, (2020) Why do countries respond differently to COVID-19? A comparative study of Sweden, China, France, and Japan. *Am Rev Public Adm* 50: 762–769. https://doi.org/10.1177/0275074020950521
52. Desson Z, Weller E, McMeekin P, Ammi M, (2020) An analysis of the policy responses to the COVID-19 pandemic in France, Belgium, and Canada. *Health Policy Technol* 9: 430–446. https://doi.org/10.1016/j.hlpt.2020.05.006
53. He D, Gao D, Zhuang Z, Cao P, Lou Y, Yang L, (2020) The attack rate of the COVID-19 in a year. *SSRN Electron J* 2020. https://dx.doi.org/10.2139/ssrn.3562044
54. Chen S, Flegg JA, Lythgoe KA, White LJ, (2025) Reconstructing the first COVID-19 pandemic wave with minimal data in England. *Epidemics* 50: 100814. https://doi.org/10.1016/j.epidem.2024.100814
55. Pagen DM, Brinkhues S, Dukers-Mujrers NH, den Heijer CD, Bouwmeester-Vincken N, Hanssen DA, et al. (2022) Exposure factors associated with SARS-CoV-2 seroprevalence during the first eight months of the COVID-19 pandemic in the Netherlands: A cross-sectional study. *PLoS One* 17: e0268057. https://doi.org/10.1371/journal.pone.0268057
56. Piler P, Thon V, Andryšková L, Doležel K, Kostka D, Pavlík T, et al. (2022) Nationwide increases in anti-SARS-CoV-2 IgG antibodies between October 2020 and March 2021 in the unvaccinated Czech population. *Commun Med* 2: 19. https://doi.org/10.1038/s43856-022-00182-1
57. Vilibic-Cavlek T, Stevanovic V, Ilic M, Barbic L, Capak K, Tabain I, et al. (2021) SARS-CoV-2 seroprevalence and neutralizing antibody response after the first and second COVID-19 pandemic wave in Croatia. *Pathogens* 10: 774. https://doi.org/10.3390/pathogens10060774
58. Ward H, Atchison C, Whitaker M, Ainslie KE, Elliott J, Okell L, et al. (2021) SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. *Nat Commun* 12: 905. https://doi.org/10.1038/s41467-021-21200-1

Appendix

Table A.1. Estimated pre-vaccine SARS-CoV-2 attack rates in studies from 5 non-high-income and 18 high-income countries.

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Vette et al. [38]	2022	2020.8	Australia	<0.5%	Serological testing	5132; 2972; 3213; NA	General population (covering all ages); Females (aged 20–39 years old); Blood donors (aged 20–69 years old); Adjusted based on age and geographical distribution of the Australian population
Reedman et al. [29]	2022	2021.1.1	Canada	2.24% (95% CI: 2.08–2.41%)	Serological testing	149,522 donations	Blood donors aged 17 years and older from all provinces except Quebec
Tuite et al. [30]	2022	2021.3	Canada	6.3% (95% CI: 5.1–7.6%)	Serological testing	1500 samples per month (for 12 months)	Retention samples; Blood donors (randomly selected from blood donor clinics)
Murphy et al. [31]	2023	2020.7	Canada	<0.3% (95% CI: 0.02–0.9%)	Serological testing	640,315 blood donors	Blood donors (Canadian Blood Services)
Vilibic-Cavlek et al. [57]	2021	2021.2	Croatia	25.1% (95% CI: 22.8–27.4%)	Serological testing	1436 participants	People aged under 10 years to 70 years and older
Piler et al. [56]	2022	2020.11	Czech Republic	27.9% (95% CI: 26.1%–29.7%)	Serological testing	3626 persons	Individuals who provided blood samples (October 2020–November 2020)

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Espenhain et al. [27]	2021	2020.12	Denmark	4.1% (95% CI: 3.1–4.9%)	Serological testing	10,358 participants	Random sample of individuals aged 12 years and older (including self-reported symptoms and antibody serological tests)
Krogsgaard et al. [28]	2023	2021.2	Denmark	7.2% (95% CI: 6.3–7.9%)	Serological testing	11,275 people	Residents of Denmark aged 12 years and older
Ward et al. [58]	2021	2020.7	England	6.0% (95% CI: 5.8–6.1%)	Serological testing	105,651 people	Adults (18 years and older, from England, registered in the NHS patient registration database)
Chen et al. [54]	2025	2020.10	England	9.1% (95% CI: 8.7–9.8%)	PCR	NA	Individuals aged 2 years and older (randomly selected from households, providing nose and throat swabs for PCR testing)
Solastie et al. [39]	2023	2021.4	Finland	<7%	Serological testing	34,619 subjects	Subjects (participating in random population surveys via regular mail)
Glemain et al. [21]	2024	2020.5	France	7.6% (95% CI: 7.3–7.8%)	Serological testing	82,467 people	People (France, including Paris, Sapris-sero cohort)
Hozé et al. [20]	2021	2021.1	France	14·9% (95% CI: 13·2–16·9%)	Serological testing	9782 adults	Adults (from the two regions with the most serious epidemic in France)
Neuhauser et al. [25]	2022	2020.11	Germany	1.7% (95% CI: 1.2–2.3%)	Serological testing + PCR	31,675 adults	Adults (from the German RKI-SOEP joint survey, covering all 401 administrative regions in Germany)

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Offergeld et al. [26]	2023	2021.1	Germany	1.8% (95% CI: 1.4–2.2%)	Serological testing	134,510 anonymous blood donation samples	Blood donors (from 28 regions in Germany)
Koureas et al. [37]	2022	2020.12	Greece	9.09% (95% CI: 7.58–11.41%)	Serological testing	55,947 serum samples	Serum samples
Bassal et al. [35]	2022	2020.9	Israel	6.3% (95% CI: 4.4–8.8%)	Serological testing	NA	Residual serum samples (all ages or not limited to 16+)
Bassal et al. [35]	2022	2020.12	Israel	8.1%	Serological testing	9520 serum samples	Residual serum samples from adults aged 16 and above
Pota et al. [32]	2020	2020.10	Italy	About 6.9%	Mathematical modeling using reported case and death data	NA	NA
McKenzie et al. [22]	2022	2020.6	Japan	0.12% (95% CI: 0.06%–0.24%)	Mathematical modeling based on reported cases and corrected serological data	NA	NA
McKenzie et al. [22]	2022	2020.12	Japan	0.83% (95% CI: 0.68%–1.00%)	Mathematical modeling based on reported cases and corrected serological data	NA	NA
Yamayoshi et al. [23]	2022	2021.1	Japan	About 3.8%	Serological testing	More than 60,000 serum/plasma samples	Serum/plasma samples (from 9 medical institutions)
Arashiro et al. [24]	2023	2021.12	Japan	2.2% (95% CI: 1.9–2.5%)	Serological testing	16,296 participants	Adults aged 20 and older, randomly selected from five counties

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
McKenzie et al. [22]	2022	2021.12	Japan	2.34% (95% CI: 2.06%–2.76%)	Mathematical modeling based on reported cases and corrected serological data	NA	NA
Lee et al. [33]	2021	2020.10	Korea	Approximately 0.62%	Serological testing	July 9: 1500 samples; September 11: 1440 samples; November 23: 1379 samples	Samples from the general population in Korea
Pagen et al. [55]	2022	2020.12	Netherlands	19.5% (95% CI: 18.7–20.3%)	Serological testing	10,001 adults	Adults (convenient sampling, Limburg, Netherlands)
Anda et al. [36]	2022	2020.12	Norway	0.9% (95% CI: 0.7–1.0%)	Serological testing	27,700 people (completed); 31,458 people (respondents)	People aged 16 or above (randomly selected for questionnaire and DBS samples)
Tunheim et al. [34]	2022	2021.1	Norway	3.2% (95% CI: 2.3–4.2%)	Serological testing	900 serum samples (spring 2020); 1812 serum samples (late summer 2020); 1912 serum samples (January 2021); 216 serum samples (August 2019)	Residual serum samples; General population
Beser et al. [49]	2022	2021.5	Sweden	20.1%	Serological testing	2860 valid samples	Children and adults (aged 3–90 years old, randomly selected nationwide)
Jones et al. [13]	2021	2020.7	US	3.5% (95% CI: 3.2–3.8%)	Serological testing	6000 serum samples per month	Blood donors

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Reese et al. [11]	2021	2020.9	US	Approximately 16%	PCR	NA	NA (Surveillance data only)
Sullivan et al. [17]	2021	2020.10	US	11.9% (95% CI: 10.5–13.5%)	Serological testing	4654 adult samples	One individual aged 18 years or older from each selected US household
Wiegand et al. [19]	2023	2020.11	US	8.0% (95% CI: 7.9–8.1%)	Serological testing	1,469,792 serum specimens	Specimens of patients who sought routine screening
Angulo et al. [12]	2021	2020.11	US	14.3% (IQR, 11.6%–18.5%)	Serological testing	16,596 (April); 40,817 (May, June, July); 38,355 (August)	General population from CDC seroprevalence surveys; General population (randomly selected for community serosurvey)
Pei et al. [15]	2021	2020.12	US	About 31% (It is derived from the population susceptibility of 69.0% (95% CI: 63.6–75.4%))	Mathematical modeling using reported case + Serological data correction	NA	NA
García-Carreras [18]	2023	2021.1	US	About 20% (in Figure 3)	Serological testing	NA	NA
Chitwood et al. [16]	2022	2021.1	US	28% (in Figure 5. Percentage of the population ever-infected with SARS-CoV-2 as of January 1, 2021)	Mathematical modeling using reported case and death data	NA	NA
Jones et al. [13]	2021	2021.5	US	20.2% (95% CI: 19.9–20.6%)	Serological testing	6000 serum samples per month (involved in the study)	Blood donors

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Louca [14]	2021	2021.6	US	36% (95% CI: 23–61%)	Mathematical modeling using death data + Serological data correction	NA	Adults aged 20 years or older (in 165 countries); Participants in previous nationwide seroprevalence surveys
Saeedzai et al. [42]	2022	2020.7	Afghanistan	29.8% (95% CI:28.8–30.7%)	Serological testing	9514 people	People from nine regions, including participants aged ⩾ 18 years and 5–17 years.
Figueiredo et al. [43]	2023	2021.11	Brazil	11% (95% CI: 11.0–12.0%)	Serological testing	336,620 participants	Participants (from 135 surveys included in the review and meta-analysis)
Guang et al. [44]	2025	2020.1	China	3.2% (in Figure 1)	Serological testing	More than 10,000 participants	Healthy population (non-high-risk occupation) of confirmed cases; Travelers undergoing mandatory testing
Popova et al. [41]	2021	2020.12	Russia	19.2% (95% CI: 19.0–19.6%)	Serological testing	74,158 volunteers; 14,275 seropositive individuals	Volunteers (surveyed across 26 model regions of Russia); Seropositive individuals (identified with antibodies to the SARS-CoV-2 nucleocapsid antigen)
Impouma et al. [40]	2021	2020.12	South Africa	14.5% (95% CI: 13.9%–15.2%)	Mathematical modeling using reported case and death data	NA	NA

Continued on next page

The first author (reference)	Publication Year	Study dates	Country	Attack rate	Identification method	Sample Size	Participants Type
Barber et al. [9]	2022	2021.11	Global	43.9% (95% CI: 39.9–46.9%)	Serological testing combined with other data sources, such as reported cases	NA	NA

Notes: For entries marked as NA, estimates were not based on participant-level sampling but instead derived from alternative sources such as serological surveys, confirmed cases and deaths, excess mortality, and diagnostic testing. Consequently, information on study participants and experimental design was not reported. Some studies did not report cumulative infection rates with confidence intervals. This was primarily due to two reasons: (1) the attack rate was calculated directly as the ratio of the estimated number of cumulative infections to the total population, and (2) the cumulative infection rate was visually extracted from graphical representations.