Original Article

narra j

Original Article

Seroepidemiological investigation of SARS-CoV-2 infection and risk factors in Indonesia before mass COVID-19 vaccination

Tri YM. Wahyono[1*], Renti Mahkota[1,2], Fajaria Nurcandra[3,4], Ansariadi Ansariadi[5], Atik C. Hidajah[6], Helda Helda[1], Fariani Syahrul[6], Indra Dwinata[5] and Nurhayati Kawi[3]

[1]Department of Epidemiology, Faculty of Public Health, Universitas Indonesia, Depok, Indonesia; [2]Doctor of Public Health Program in Epidemiology, Faculty of Public Health, Mahidol University, Bangkok, Thailand; [3]Doctoral Program in Epidemiology, Faculty of Public Health, Universitas Indonesia, Depok, Indonesia; [4]Department of Public Health, Faculty of Health Science, Universitas Pembangunan Nasional Veteran Jakarta, Jakarta, Indonesia; [5]Center for Epidemiology and Population Health Studies, Faculty of Public Health, Universitas Hasanuddin, Makassar, Indonesia; [6]Department of Epidemiology, Biostatistics, Population Studies, and Health Promotion, Faculty of Public Health, Universitas Airlangga, Surabaya, Indonesia

*Corresponding author: triyunis1@ui.ac.id

Abstract

At the onset of the coronavirus disease 2019 (COVID-19) pandemic in Indonesia, surveillance focused on finding and treating symptomatic cases. However, emerging evidence indicated that asymptomatic and pre-symptomatic individuals significantly contributed to viral transmission. This highlights the need for comprehensive surveillance to understand better the actual spread of SARS-CoV-2. Therefore, the aim of this study was to determine the seroprevalence of SARS-CoV-2 antibodies in the general population across Indonesia and identify risk factors associated with infection at the beginning of the pandemic. A cross-sectional survey was conducted across 17 provinces, 69 districts/cities, and 1,020 villages in Indonesia from December 22, 2020, to February 15, 2021. A multistage random sampling technique was employed. Serological testing using enzyme-linked immunosorbent assay (ELISA) was performed to detect anti-SARS-CoV-2. Complex sample analysis, adjusted for weights, was utilized to estimate the national seroprevalence and a generalized linear model with a binomial distribution was applied to identify risk factors. A total of 10,161 individuals were included in the final analysis, with the national seroprevalence being 14.8% (95% confidence interval (CI): 14.2–18.5). The prevalence was higher in females (16.8%; 95%CI: 12.5–22.3), individuals aged 46–59 years (18.6%; 95%CI: 14.2–24.0), and in urban areas (20.1%; 95%CI: 15.0–26.2). The highest prevalence was observed in North Maluku (35.6%; 95%CI: 29.3–42.5). Notably, 54.2% of seropositive individuals were asymptomatic, while 7.5% reported hypertension as a comorbidity. Factors associated with higher seroprevalence were being married (adjusted prevalence ratio (aPR): 1.47; 95%CI: 1.02–2.12), widow (aPR: 1.74; 95%CI: 1.01–3.00), and close contact with confirmed cases (aPR: 2.04; 95%CI: 1.52–2.73). This study revealed a COVID-19 prevalence significantly higher than official estimate in Indonesia, underscoring the need for improved surveillance system to more accurately track disease spread and to inform timely public health responses in the future.

Keywords: COVID-19, infection, prevalence, seroepidemiology, risk factors

Introduction

The emergence and global spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) posed unprecedented challenges to public health systems worldwide [1,2]. The first

Received: December 20, 2024 - Accepted: March 21, 2025 - Published online: March 24, 2025

confirmed cases of coronavirus disease 2019 (COVID-19), caused by SARS-CoV-2, in Indonesia were detected in March 2020, prompting the government to implement large-scale social restrictions to mitigate transmission [3]. However, initial responses largely focused on identifying and managing symptomatic cases, leading to an underestimation of the true burden of infection in the community [4]. Asymptomatic and pre-symptomatic individuals were later recognized as significant contributors to viral transmission, highlighting the necessity for comprehensive surveillance strategies, including seroepidemiological studies, to accurately assess the extent of SARS-CoV-2 exposure in Indonesia [5].

The lack of widespread testing capacity in Indonesia during the early phase of the pandemic further exacerbated the challenge of accurately estimating the true scale of infections. Limited access to diagnostic testing, particularly in rural and remote areas, contributed to the underreporting of cases, leaving large portions of the population unaware of their infection status [4].

To address the limitations of under-detection in COVID-19 surveillance, seroepidemiological investigations were conducted to provide a more accurate measure of SARS-CoV-2 exposure in Indonesia. Previous serosurveys, largely confined to specific cities or provinces, highlighted significant regional variation in infection rates [6-8]. For instance, a seroprevalence study in Jakarta Province reported SARS-CoV-2 antibody positivity of 28.52% (95%CI: 25.44–31.81) and cumulative cases in Tanjung Priok sub-district affecting 2.4% of its population [6]. However, national-scale seroprevalence data remain limited, making it essential to conduct large-scale studies that capture epidemiological patterns across diverse geographic regions.

Furthermore, variations in COVID-19 transmission across Indonesia were likely influenced by differences in population density, mobility patterns, and compliance with public health interventions. Urban areas, characterized by higher population densities and increased mobility, may have experienced greater viral spread compared to rural regions with more limited human interactions. The disparities in infection rates also highlight the need for tailored public health measures that consider local demographic and socioeconomic contexts. Conducting large-scale seroepidemiological studies provides a crucial opportunity to assess these variations and implement evidence-based strategies to mitigate future outbreaks.

The aim of this study was to determine the seroprevalence of SARS-CoV-2 antibodies in the general population across Indonesia and to identify sociodemographic and other risk factors associated with infection. Unlike prior studies [6-8], which were restricted to specific provinces or urban centers; this study encompassed 17 provinces, representing both high- and low-burden COVID-19 regions. The study applied a systematic multistage sampling approach to ensure a representative dataset, thereby enhancing the generalizability of findings to the broader Indonesian population. By integrating epidemiological data with serological analysis, this research provided a robust estimate of the true infection burden and offered critical insights into the transmission dynamics of SARS-CoV-2 in Indonesia. Furthermore, this study might contribute to the global effort to understand COVID-19 epidemiology, providing valuable data that can be used for cross-country comparisons and pandemic preparedness planning. The lessons learned from this unprecedented global crisis can serve as a foundation for enhancing public health resilience against emerging infectious diseases. Understanding the seroprevalence and associated risk factors of SARS-CoV-2 infection is crucial for guiding public health responses and preparedness strategies for future outbreaks.

Methods

Study design and setting

A cross-sectional study was conducted to evaluate the seroprevalence of SARS-CoV-2 infection in Indonesia. The data and blood collections were carried out between December 22, 2020, and February 15, 2021, before the commencement of the national mass vaccination program. The protocol used in this study was adapted from the protocol developed by the Consortium for the Standardization of Influenza Seroepidemiology (CONSISE) [9]. CONSISE is a global partnership aiming to develop influenza investigation protocols and standardize seroepidemiology to inform public health policy concerning pandemic, zoonotic, and seasonal influenza. The present study employed a population-based, age-stratified sampling approach across 17 of the 34 provinces in Indonesia, covering 69 districts/cities and 1,020 villages. Further, provinces were classified into high-burden and low-burden categories based on the point prevalence of COVID-19 as of August 20, 2020. National surveillance data informed this classification to ensure representative sampling. High- and low-burden provinces were categorized based on COVID-19 morbidity and mortality. The map of selected provinces with their burden status (high-burden or low-burden) based on national COVID-19 prevalence from Indonesian government is presented in **Figure 1**.

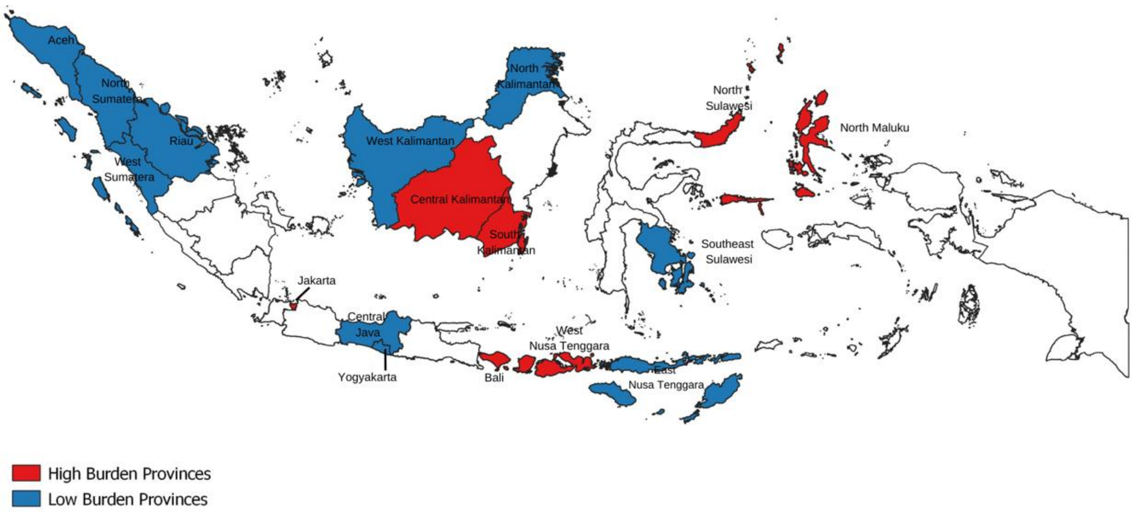

Figure 1. Map of selected provinces in Indonesia categorized as high-burden or low-burden based on coronavirus disease 2019 (COVID-19) prevalence from the Indonesian government. Provinces were classified as high-burden if most of the regions were designated red and orange zones, while provinces were classified as low-burden if most of the region was designated yellow and green zones based on national COVID-19 prevalence data as of August 20, 2020. Geospatial mapping utilized a base layer provided by the Geospatial Information Agency of the Republic of Indonesia (2024), accessible at https://www.indonesia-geospasial.com. The map was prepared using QGIS, a free and open-source geographic information system.

Sampling size and sampling strategy

The sample size was calculated using a formula based on a significance level (α=0.05) and a precision of 0.5 for a finite population. A 95% confidence interval (95%CI) was applied to all variables measured in this study. The target population included all households within the study area, with a minimum calculated sample size of 8,330 participants. To account for invalid test results and non-responses during specimen collection, an additional 22.5% was added to the minimum sample size, resulting in a total sample size of 10,200 participants. The sample size was calculated using the online statistical tool OpenEpi (http://www.openepi.com).

Participants were randomly selected from households across the study sites. Within each selected village, one individual representing their household was chosen. If the selected participant declined to participate, their closest neighbor was randomly selected as a replacement. The sample distribution was stratified by geographical burden, ensuring proportional representation from high- and low-burden provinces, districts/cities, and clusters/villages. This stratification aimed to provide a balanced and representative dataset across different regions.

Participants

This study included participants selected through multistage random sampling from households across all study sites. The inclusion criteria were individuals aged 1–70 years residing in the selected provinces, districts/cities, and villages, regardless of prior or current SARS-CoV-2 infection status. Participants who refused to provide written informed consent or had contraindications to venipuncture were excluded from the study.

Data collection and measurement

Data was collected by a trained team of interviewers and community cadres who utilized the World Health Organization (WHO) globally standardized questionnaire and template. The data was gathered at designated locations, such as primary healthcare facilities or village centers. To maintain physical distancing, team members individually contacted selected households and individuals to prevent large gatherings, following the protocols implemented during the COVID-19 pandemic.

Personnel involved in the study underwent training in infection prevention and control (IPC) procedures in line with national and local guidelines. IPC measures included hand hygiene and the correct use of surgical masks, particularly during close contact with participants, to reduce the risk of infection transmission. The relevant WHO technical guidance for COVID-19 IPC was adhered to throughout the data collection process [10].

Eligible participants provided informed consent prior to participation. After consent, each participant completed a demographic, clinical, and exposure data questionnaire (COVID-19 infection history, comorbid, and risk factors). Blood samples (2 mL) were then collected by trained nurses using vacutainer tubes, with samples processed in compliance with study protocols. The blood samples were stored in cool boxes with ice packs and transported to the laboratory on the same day for antibody analysis via enzyme-linked immunosorbent assay (ELISA).

Each blood sample was uniquely labeled to correspond with the participant's questionnaire data for tracking purposes. The time of collection, transport conditions, and laboratory arrival times were recorded for each specimen. Samples were promptly sent to six provincial laboratories, known as the Center for Technical Laboratory in Environmental Health (*Balai Teknik Kesehatan Lingkungan dan Pengendalian Penyakit* (BTKL-PP)) for serology testing. The six provincial laboratories were: BBTKL-PP Palembang (South Sumatra), BBTKL-PP Jakarta (DKI Jakarta), BBTKL-PP Yogyakarta (DI Yogyakarta), BBTKL-PP Surabaya (East Java), BBTKL-PP Banjar Baru (South Kalimantan), and BBTKL-PP Makassar (South Sulawesi). National Institute of Health Research and Development laboratories of the Indonesian Ministry of Health carried out quality control and validation of results.

Serum samples were tested for COVID-19-specific antibodies using the Wantai SARS-CoV-2 Ab ELISA test (Wantai Biological Pharmacy Enterprise, Beijing, China), a highly sensitive and specific assay (94.3% sensitivity and 100% specificity) [11]. The test detects total antibodies (IgG and IgM) and is intended to identify individuals with an immune response to SARS-CoV-2, indicating prior or recent infection. The Wantai SARS-CoV-2 Ab ELISA test was not used to diagnose acute SARS-CoV-2 infection. Testing was limited to laboratories certified under the Clinical Laboratory Improvement Amendments of 1988 (CLIA), 42 U.S.C 263a, which meet requirements to perform high-complexity tests [12-14]. The test was performed in line with WHO guidance.

Each team comprised one specimen collector, one interviewer from the primary health care facility (*Puskesmas*), and a community cadre. A team leader and assistant from the District Health Office coordinated activities within each district/city, ensuring adherence to guidelines and protocols. The team managed the daily data collection process, ensuring all questionnaires were properly filled, collated, and recorded.

Data storage

Data was handled following the national guidelines. The questionnaire was structured in a written format, and responses were immediately recorded. To maintain data integrity, the team leader conducted daily checks on the collected data to identify and address discrepancies, ensuring errors were minimized. Any inconsistencies were promptly addressed by revisiting the field for re-validation. The primary goal was to guarantee the accuracy and completeness of data by the end of fieldwork. Data from each district was securely submitted to the central data manager for analysis and storage.

Data analysis

Statistical analysis was conducted using Stata version 15.1 (StataCorp, College Station, USA), licensed to the Faculty of Public Health, Universitas Indonesia. To minimize bias and enhance the precision of the estimates, a comprehensive weighting procedure of the seroprevalence data was implemented across five stages: (a) data exploration, initial examination of data from both the questionnaire and laboratory results; (b) data merging, integration of the questionnaire and laboratory data; (c) weight calculation, derivation of base weights according to the sampling design employed in the survey, followed by weight calibration using the 2020 National Socio-Economic Survey reference population; (d) prevalence calculation, estimation of prevalence using both weighted and unweighted data; and (e) prevalence correction, adjustment of prevalence estimates to account for potential misclassification errors, including quality control results and the accuracy of diagnostic tests used.

Descriptive statistics were generated to cover participant numbers, weighted prevalence estimates, and corresponding 95%CIs. Moreover, a complex sample analysis, incorporating adjusted weights, was used to estimate the prevalence of SARS-CoV-2 infection across various subgroups, including sex, age, region of residence, provincial burden, symptoms, and comorbidities.

For the multivariate analysis, a generalized linear model (GLM) with a binomial distribution was employed to identify sociodemographic risk factors associated with SARS-CoV-2 infection. Variables with a $p<0.25$ were retained for inclusion in the final multivariate GLM, which accounted for sampling design and weight adjustments. This model aimed to control for sampling frame errors and potential biases, ultimately providing a final risk factor model. Prevalence ratios (PRs) and adjusted prevalence ratios (aPRs) and their 95%CIs were reported to quantify the association between sociodemographic variables and SARS-CoV-2 infection prevalence.

Results

Characteristics of the participants

A total of 10,200 individuals were approached to participate in this study, and 99.6% (10,161) of them completed the survey and had ELISA results. The characteristics of the included participants are presented in **Table 1**. The sample consisted of 40.12% males and 59.88% females. By age group, 7.5% were aged 1–18 years, 18.25% were 19–30 years, 37.24% were 31–45 years, 27.52% were 46–59 years, and 9.5% were 60 years or older. A total of 59.32% of participants resided in rural areas, while 40.68% lived in urban areas. Based on province burden, 41.23% of participants were from high-burden provinces, while 58.77% were from low-burden provinces. Regarding symptoms, 45.33% of participants reported having symptoms, while 54.67% were asymptomatic. Among participants with comorbidities, 10.21% had hypertension, 3.24% had diabetes mellitus, and 2.02% had asthma, while lower percentages were recorded for ischemic heart disease (0.79%), tuberculosis (0.56%), thyroid disorders (0.3%), immune deficiency (0.12%), chronic renal disease (0.11%), cancer (0.11%), and chronic liver failure (0.06%) (**Table 1**).

Table 1. Participant characteristics included in anti-severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) seroprevalence study in Indonesia (n=10,161)

Characteristics	Number of participants tested	Percentage (%)
Sex		
Male	4,077	40.12
Female	6,084	59.88
Age group (years old)		
1–18	762	7.50
19–30	1,854	18.25
31–45	3,784	37.24
46–59	2,796	27.52
≥60	965	9.50
Living area		
Rural	6,028	59.32
Urban	4,133	40.68
Province burden category		
High	4,189	41.23
Low	5,972	58.77
Provinces		

Characteristics	Number of participants tested	Percentage (%)
Aceh	600	5.90
Bali	600	5.90
Yogyakarta	600	5.90
Jakarta	598	5.89
Central Java	599	5.90
West Kalimantan	600	5.90
South Kalimantan	600	5.90
Central Kalimantan	600	5.90
North Kalimantan	600	5.90
North Maluku	595	5.86
West Nusa Tenggara	600	5.90
East Nusa Tenggara	597	5.88
Riau	599	5.90
Southeast Sulawesi	577	5.68
North Sulawesi	597	5.88
West Sumatera	600	5.90
North Sumatera	599	5.90
Symptom		
Yes	4,606	45.33
No	5,555	54.67
Comorbidity		
Diabetes mellitus	329	3.24
Hypertension	1,037	10.21
Ischemic heart disease	80	0.79
Tuberculosis	57	0.56
Thyroid	30	0.30
Asthma bronchial	205	2.02
Immune deficiency	12	0.12
Chronic liver failure	6	0.06
Chronic renal disease	11	0.11
Cancer	11	0.11
Chronic obstructive pulmonary disease	21	0.21
Others	638	6.28

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection seroprevalence

The unweighted cumulative seroprevalence of SARS-CoV-2 infection in Indonesia from December 22, 2020, to February 15, 2021, was 20.5% (95%CI: 19.7–21.3). After adjusting for the sampling frame, population distribution from the National Socio-Economic Survey (SUSENAS) [15], and ELISA test kit validity, the weighted cumulative seroprevalence was 14.8% (95%CI: 11.6–18.8). This result suggested approximately 14.8% of Indonesia's population, or 40 million individuals, were infected with SARS-CoV-2. The weighted seroprevalence of SARS-CoV-2 in Indonesia varied across different characteristics (**Table 2**).

Table 2. Seroprevalence of anti-severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in Indonesia (n=10,161)

Characteristics	Number of seropositive participants	Weighted prevalence estimate (%)	95%CI weighted prevalence estimate
Sex			
Male	665	12.85	9.93–16.48
Female	1,417	16.81	12.45–22.33
Age group (years old)			
1–18	111	9.13	4.81–16.64
19–30	373	15.07	11.49–19.52
31–45	791	17.6	14.87–20.70
46–59	611	18.59	14.15–24.02
≥60	196	17.51	13.05–23.08
Living area			
Rural	874	12.53	8.87–17.42
Urban	1,208	20.05	15.03–26.23
Province burden category			
High	1,058	16.83	11.8–23.44
Low	1,024	13.54	9.48–18.98
Provinces			

Characteristics	Number of seropositive participants	Weighted prevalence estimate (%)	95%CI weighted prevalence estimate
Aceh	149	21.06	18.95–23.33
Bali	131	18.59	15.14–22.62
Yogyakarta	24	3.73	1.81–7.54
Jakarta	185	22.83	18.95–27.24
Central Java	104	13.22	7.309–22.75
West Kalimantan	98	15.92	10.29–23.81
South Kalimantan	171	27.22	21.96–33.21
Central Kalimantan	51	7.50	4.42–12.44
North Kalimantan	54	6.20	3.46–10.89
North Maluku	165	35.64	29.29–42.54
West Nusa Tenggara	107	12.34	6.28–22.82
East Nusa Tenggara	49	3.07	1.19–7.71
Riau	81	14.22	12.53–16.10
Southeast Sulawesi	214	32.94	28.30–37.94
North Sulawesi	195	30.94	28.96–33.00
West Sumatera	149	23.55	20.30–27.16
North Sumatera	155	16.57	15.14–18.10
Symptom			
Yes	1,015	45.84	34.72–57.38
No	1,067	54.16	42.62–65.28
Comorbidity			
Diabetes mellitus	85	1.90	1.21–2.95
Hypertension	238	7.53	5.62–10.02
Ischemic heart disease	24	0.59	0.26–1.33
Tuberculosis	11	0.23	0.09–0.60
Thyroid	7	0.14	0.03–0.51
Asthma bronchial	44	2.40	1.08–5.25
Immune deficiency	2	0.01	0.001–0.05
Chronic liver failure	1	0.00	0.001–0.04
Chronic renal disease	1	0.08	0.01–0.63
Cancer	2	0.07	0.01–0.33
Chronic obstructive pulmonary disease	21	0.09	0.02–0.39
Others	161	7.21	4.75–10.79

Females had a higher seroprevalence (16.81%; 95%CI: 12.45–22.33) compared to males (12.85%; 95%CI: 9.93–16.48). Age-wise, the lowest seroprevalence was observed in individuals aged 1–18 years (9.13%; 95%CI: 4.81–16.64), while the highest was in the 46–59 age group (18.59%; 95%CI: 14.15–24.02). Participants living in urban areas had a higher seroprevalence (20.05%; 95%CI: 15.03–26.23) compared to those in rural areas (12.53%; 95%CI: 8.87–17.42). Seroprevalence was slightly higher in high-burden provinces (16.83%; 95%CI: 11.8–23.44) than in low-burden provinces (13.54%; 95%CI: 9.48–18.98). Asymptomatic individuals had a seroprevalence of 54.16% (95%CI: 42.62–65.28), underscoring the role of subclinical infections in transmission (**Table 2**).

Regionally, the highest SARS-CoV-2 seroprevalence was recorded in North Maluku (35.64%; 95%CI: 29.29–42.54), followed by Southeast Sulawesi (32.94%; 95%CI: 28.30–37.94) and North Sulawesi (30.94%; 95%CI: 28.96–33.00), while the lowest was in East Nusa Tenggara (3.07%; 95%CI: 1.19–7.71) and Yogyakarta (3.73%; 95%CI: 1.81–7.54). The results of the identification of SARS-CoV-2 seropositive in **Table 2** were mapped to provide an overview of its spread and seroprevalence. It can be seen which provinces with SARS-CoV-2 are included in the low (0–10%), moderate (11–20%), and high (>20%). A detailed description of the SARS-CoV-2 seroprevalence based on the provinces is presented in **Figure 2**.

Factors associated with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection seroprevalence

The univariate analysis identified several significant variables based on the crude prevalence ratio (**Table 3**). Sex was significantly associated with seroprevalence, with females having a higher likelihood of being seropositive compared to males (prevalence ratio (PR): 1.31; 95%CI: 1.01–1.70; p=0.046). Marital status was also significant, with married individuals (PR: 1.54; 95%CI: 1.07–2.22; p=0.024) and widowed individuals (PR: 1.91; 95%CI: 1.12–3.26; p=0.021) having higher seroprevalence compared to those who had never married.

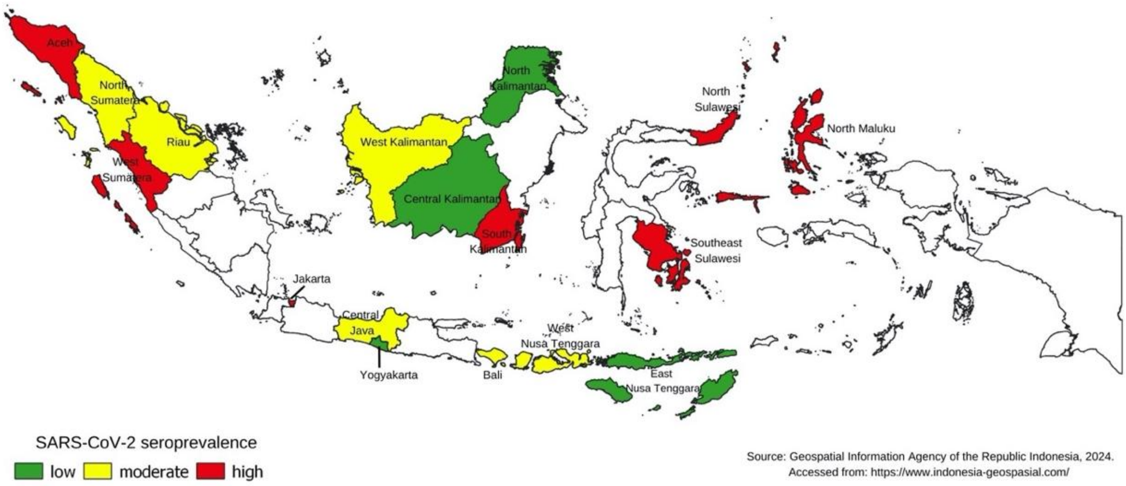

Figure 2. Seroprevalence of SARS-CoV-2 in Indonesia between December 22, 2020, and February 15, 2021. The selected provinces were categorized into low (0–10%), medium (11–20%), and high (>20%) based on the seroprevalence results obtained. The geospatial mapping used base layers provided by the Geospatial Information Agency of the Republic of Indonesia (2024), which can be accessed at https://www.indonesia-geospasial.com. The map was prepared using QGIS, a free and open geographic information system.

Close contact history played a crucial role, as individuals who had known close contact with a COVID-19 case had significantly higher odds of being seropositive (PR: 2.17; 95%CI: 1.58–2.99; p<0.001). Among sources of close contact, exposure within the family (PR: 3.33; 95%CI: 1.89–5.85; p<0.001), neighbors (PR: 2.41; 95%CI: 1.36–4.25; p=0.005), and friends (PR: 2.67; 95%CI: 1.10–6.48; p=0.032) significantly increased the likelihood of seropositivity compared to workplace exposure (**Table 3**). These findings highlight the role of close contact, particularly within households and social circles, in the transmission of SARS-CoV-2. Age group, education level, occupation, and divorced marital status were not significantly associated with SARS-CoV-2 seroprevalence (**Table 3**). Additionally, having an unknown source of close contact or exposure outside of family, neighbors, and friends, such as in the workplace, did not show a significant association with seroprevalence.

Determinants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) seroprevalence in Indonesia

In muti-variate analysis, sex, education level, occupation, marital status, and closed contact were included in the analysis. Then, those variables were selected using the forward selection method. Of the total participants, only 609 participants admitted to having close contact with COVID-19 patients, so the source of closed contact was not included in the multivariate modeling.

However, only marital status and contact with confirmed cases emerged as the primary factors influencing COVID-19 seroprevalence (**Table 4**). Marital status influenced seroprevalence, with married individuals (aPR: 1.47; 95%CI: 1.02–2.12) and widows (aPR: 1.74; 95%CI: 1.01–3.00) having higher adjusted aPR than individuals who were never married. Individuals who had contact with confirmed COVID-19 cases were significantly more likely to be seropositive for SARS-CoV-2 antibodies. Those who had close contact with case had a 2.04-fold increased likelihood of seropositivity (95%CI: 1.52–2.73) compared to those who never had contact with COVID-19 cases (**Table 4**).

Table 3. Univariate analysis showing the factors associated with anti-severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) seroprevalence in Indonesia before coronavirus disease 2019 (COVID-19) mass vaccination

Variables	SARS-CoV-2 seroprevalence				Crude prevalence ratio	95%CI	*p*-value
	Positive		Negative	95%CI			
	Estimate (%)	95%CI	Estimate (%)				
Age group (n=10,161)							
1–5	11.76	1.54–53.11	88.24	46.89–98.46	*Ref*		
6–18	8.99	4.71–16.51	91.00	83.49–95.29	0.76	0.12–5.04	0.766
19–30	15.07	11.49–19.52	84.93	80.48–88.51	1.28	0.19–8.46	0.784
31–45	17.6	14.87–20.70	82.40	79.30–85.13	1.5	0.22–10.20	0.661
46–59	18.59	14.15–24.02	81.41	75.98–85.85	1.58	0.25–10.11	0.607
≥60	17.51	13.05–23.08	82.49	76.92–86.95	1.49	0.22–10.02	0.663
Sex (n=10,161)							
Male	12.85	9.97–16.48	87.15	83.52–90.07	*Ref*		
Female	16.81	12.45–22.33	83.19	77.67–87.55	1.31	1.01–1.70	0.046*
Education level (n=10,161)							
No school	9.73	4.68–19.13	90.27	80.87–95.32	*Ref*		
Elementary	11.84	7.85–17.46	88.16	82.54–92.15	1.22	0.57–2.61	0.591
Junior high school	12.82	7.98–19.96	87.18	80.04–92.02	1.32	0.64–2.70	0.425
Senior high school	17.85	14.98–21.12	82.15	78.88–85.02	1.83	0.87–3.85	0.102
College/university	18.8	14.13–24.57	81.20	75.43–85.87	1.93	0.92– 4.06	0.078
Occupation (n=10,161)							
Unemployment	10.60	5.35–19.90	89.40	80.10–94.65	*Ref*		
Housewife	21.45	16.09–28.01	78.55	71.99–83.91	2.02	0.97–4.21	0.058
Government employee	19.93	10.81–33.82	80.07	66.18–89.19	1.88	0.71–4.94	0.184
Army	17.72	3.79–54.03	82.28	45.97–96.21	1.67	0.37–7.60	0.481
Private employee	16.63	13.11–20.87	83.37	79.13–86.89	1.57	0.82–2.99	0.157
Entrepreneur	16.71	11.33–23.94	83.29	76.06–88.67	1.58	0.80–3.10	0.172
Farmer	10.74	6.95–16.25	89.26	83.75–93.05	1.01	0.43–2.39	0.974
Fisherman	9.91	4.60–20.05	90.09	79.95–95.4	0.93	0.31–2.79	0.897
Student	10.14	5.77–17.23	89.86	82.77–94.23	0.96	0.51–1.80	0.884
Others	15.26	9.14–24.38	84.74	75.62–90.86	1.44	0.63–3.29	0.362
Marital status (n=10,161)							
Never married	11.20	7.31–16.78	88.80	83.22–92.69	*Ref*		
Married	17.24	14.06–20.95	82.76	79.05–85.94	1.54	1.07–2.22	0.024*
Widow	21.39	12.15–34.89	78.61	65.11–87.85	1.91	1.12–3.26	0.021*
Divorced	15.46	10.42–22.32	84.54	77.68–89.58	1.38	0.87–2.19	0.156
Closed contact (n=10,161)							
Yes	29.52	23.23–36.7	70.48	63.3–76.77	2.17	1.58–2.99	<0.001**
No	13.60	10.30–17.74	86.40	82.26–89.70	*Ref*		
Unknown	16.95	11.41–24.44	83.05	75.56–88.59	1.25	0.85–1.83	0.239
Source of closed contact (n=609)							
Family	41.92	28.87–56.22	58.08	43.78–71.13	3.33	1.89–5.85	<0.001**
Neighbors	30.35	18.61–45.38	69.65	54.62–81.39	2.41	1.36–4.25	0.005**
Friends	33.72	12.87–63.67	66.28	36.33–87.13	2.67	1.10–6.48	0.032*
Worker	12.61	7.99–19.32	87.39	80.68–92.01	*Ref*		
Others	26.67	6.96–63.88	73.33	36.12–93.04	2.12	0.61–7.37	0.220

*Statistically significant at $p<0.05$
**Statistically significant at $p<0.01$

Table 4. Multivariate model of determinant factors influencing coronavirus disease 2019 (COVID-19) seroprevalence in Indonesia

Variables[a]	Standard error (SE)	*t*-value	Adjusted prevalence ratio (aPR)	95%CI	*p*-value
Marital status (*single*)					
Married	0.25	2.26	1.47	1.02–2.12	0.039*
Widow	0.44	2.18	1.74	1.01–3.00	0.045*
Divorced	0.29	1.55	1.39	0.88–2.17	0.142
Close contact with cases (*no*)					
Yes	0.28	5.20	2.04	1.52–2.73	<0.001**
Unknown	0.20	1.08	1.20	0.84–1.71	0.297

[a]Group in italic is a reference group
*Statistically significant at $p<0.05$
**Statistically significant at $p<0.01$

Discussion

This study found an overall seroprevalence of 14.8% and two key determinants included marital status and close contact with confirmed COVID-19 cases, indicating significant disparities in exposure and infection risks across the population. This study determined the seroprevalence of COVID-19 prior to the mass vaccination campaign in Indonesia to be 14.8%, consistent with routine surveillance data from Bali Province, which reported a similar prevalence of 16.73% [16]. In comparison, seroprevalence in Wuhan was lower at 6.92% from August 18 to September 20, 2020 [16]. While in Jakarta, Indonesia's pandemic epicenter, from November 23, 2020, to February 19, 2021, reported a higher antibody prevalence of 28.52% [6]. Notably, our seroprevalence is lower than India's weighted community-based serosurvey in July 2020, which reported 18.4% [17]. Our seroprevalence was also lower than a household-based serosurvey from August 10 to September 11, 2020, in South Sudan, which found a prevalence of 22.3% [18]. These differences reflect disparities in transmission dynamics, demographic factors, and regional public health interventions.

The pattern of COVID-19 spread in Indonesia before mass vaccination may have been the same as in India [16]. During the first wave of the COVID-19 pandemic, Indonesia implemented self-isolation orders. Unfortunately, this situation created stigma among COVID-19 patients for several weeks [19]. Public knowledge and understanding regarding COVID-19 were still limited when this serosurvey was conducted. This situation likely accounted for the proportion of seroprevalence of SARS-CoV-2 infections in Indonesia.

This seroprevalence survey provided an unbiased estimate of serologic COVID-19 prevalence, reflecting the prevalence of persons infected with SARS-CoV-2 before mass vaccination. Bias correction was performed using indirect standardization, whereas direct or indirect standardization depended on the values used in the correction, which could also introduce bias. Furthermore, for the uncertainty of the seroprevalence in Indonesia, a correction was made to estimate bias using the true values of Indonesian data (Center Bureau of Statistics for demographic data and valid laboratory data). Therefore, our weighted seroprevalence estimate reflected the true value of SARS-CoV-2 infection seroprevalence in Indonesia.

The weighted seroprevalence survey in Indonesia before mass vaccination was 14.8% (95%CI: 14.2–18.5), which, when extrapolated to Indonesia's total population of 277 million, suggests there were approximately 40,996,000 infections before mass vaccination was conducted. Conversely, the number of reported cases in Indonesia from routine COVID-19 surveillance systems was approximately 1 million COVID-19 cases [3]. This indicates that the detection capacity of surveillance in Indonesia remained low. Notably, several countries face difficulties in real-time detection of morbidity and mortality from community transmission of SARS-CoV-2 [20].

GLM logistic regression was used to identify factors associated with the occurrence of COVID-19 in Indonesia. The analysis revealed variables associated with SARS-CoV-2 seroprevalence: marital status and contact with COVID-19 cases. Married and widowed individuals had a higher risk of COVID-19 infection than those unmarried. Married people have higher stress than unmarried people due to the burden of living as married or ever-married persons [21,22]. Furthermore, male or female married individuals are more likely to be infected with SARS-CoV-2 because they have financial responsibilities to their families and, therefore, they have to work and meet many people [23], then they may get infected from workplace and infect their family at home. Marital status was found to cause negative lifestyle changes during the COVID-19 pandemic. Changing these negative habits might have impacted health during the COVID-19 pandemic [24]. Persons in contact with COVID-19 cases had a higher seroprevalence than those without contact. The frequency of contact, whether within or outside the family, occurred in clusters. These clusters could be family units, shared communities, work communities, or schools, with the highest number of clusters in Indonesia composed of family units.

This serosurvey had some limitations as well as strengths. A 2% disease prevalence was assumed in the sample size calculation based on the consideration that the survey targeted a rare disease. However, the observed prevalence was not rare, exceeding 14.8% (94%CI: 14.2–18.5). This assumption was made because, at the time of the study, COVID-19 prevalence remained low in provinces outside Java. Nevertheless, the findings indicate that COVID-19 had already spread across Indonesia but remained undetected by the surveillance system due to asymptomatic cases and underreporting. To minimize recall bias, efforts were made to measure and collect data on all relevant variables using a standardized questionnaire. This questionnaire covered all risk factors and COVID-19 signs and symptoms. Training sessions for enumerators and field coordinators were conducted both virtually and in person to enhance their understanding and improve their skills in questionnaire administration and interview techniques. However, the issue of temporality remains unavoidable in cross-sectional surveys. Selection bias might have occurred in this study; however, it could have been minimized by randomly selecting all sampling units. Provinces, districts, and cities in this study were selected using stratified random sampling. Villages were then chosen through cluster random sampling, applying proportionate probability to size based on population size. Furthermore, households and individuals were selected using simple random sampling. However, due to the complexity of the sampling frame, bias might still have occurred as a result of the sampling design effect. To address this, bias was accounted for at each stage of the sample selection process. Additionally, all biases across parameters were measured and mitigated through a multivariate analysis.

Conclusion

This seroepidemiological study on SARS-CoV-2 infection in Indonesia included 10,161 individuals from 17 provinces, 69 districts/cities, and 1,020 villages. The weighted overall seroprevalence of anti-SARS-CoV-2 was 14.8% (95%CI: 11.6–18.8). Seroprevalence was higher among females than males, individuals aged 46–59 compared to other age groups, those graduated from college/university compared to other educational attainment groups, urban than rural residents, those working as housewife and government employee, those who were married or widow than unmarried individuals, and among those who had close contact with COVID-19 cases. However, two key determinants of seropositive were marital status and close contact with confirmed COVID-19 cases. Notably, the seroprevalence of anti-SARS-CoV-2 in the community was significantly higher than reported in the surveillance system, indicating low detection capacity of surveillance system in the country. This study highlights that strengthening the surveillance system's capacity is essential to improve early detection and response for future pandemics. Enhancing active case finding, integrating seroprevalence studies into routine surveillance, and improving data integration between laboratory testing and epidemiological monitoring will help identify future outbreaks more accurately in the future.

Ethics approval

The Ministry of Health - National Institute of Health Research and Development gave this study ethical approval (No. LB.02.01/2/KE.642/2020).

Acknowledgments

We would like to thank the Consortium for the Standardization for Influenza Seroepidemiology (CONSISE) for allowing us to adapt the CONSISE protocol. The initial protocol used for this study was developed by Maria Van Kerkhove, Isabel Bergeri, Rebecca Grant, Lorenzo Subissi (from the Health Emergencies Program, WHE, World Health Organization), Joseph Wu, Malik Peiris (Hong Kong University), John Watson (US CDC), Marta Valenciano, and Anthony Nardone (Epiconcept). The present study was conducted by the Faculty of Public Health, Universitas Indonesia, in collaboration with Faculty of Public Health of Universitas Airlangga and Universitas Hasanuddin, Ministry of Health (Directorate General of Diseases Prevention and Control and the National Institute for Health Research Development), Provincial Health Offices in 17 selected provinces, District Health Offices in 69 districts/cities as well as Public Health Centers (*Puskesmas*) that covered 1,020 villages all over Indonesia. Achmad Naufal Azhari from the WHO Country Office in Indonesia provided technical assistance in finalizing the study protocol. All the researchers appreciate the support from six laboratories that assayed the samples for anti-SARS-CoV-2 antibodies using the ELISA test and all respondents who contributed to this study. We thank all study participants, enumerators, and supervisors involved in data collection and Muhammad Noor Farid for his assistance in data analysis.

Competing interests

All the authors declare no conflicts of interest.

Funding

The WHO Country Office in Indonesia funded this work.

Underlying data

Derived data supporting the findings of this study are available from the corresponding author upon request.

Declaration of artificial intelligence use

We hereby confirm that no artificial intelligence (AI) tools or methodologies were utilized at any stage of this study, including data collection, analysis, visualization, or manuscript preparation. The authors entirely conducted this study manually.

How to cite

Wahyono TYM, Mahkota R, Nurcandra F, *et al*. Seroepidemiological investigation of SARS-CoV-2 infection and risk factors in Indonesia before mass COVID-19 vaccination. Narra J 2025; 5 (1): e1957 - http://doi.org/10.52225/narra.v5i1.1957.

References

1. World Health Organization. Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection. Geneva: World Health Organization; 2020.
2. World Health Organization. Assessment of risk factors for coronavirus disease 2019 (COVID-19) in health workers: Protocol for a case-control study. Geneva: World Health Organization; 2020.
3. Kementerian Kesehatan Republik Indonesia. Current situation of the novel coronavirus (COVID-19). Available from: https://infeksiemerging.kemkes.go.id/dashboard/covid-19. Accessed: 22 April 2021.
4. Wu SL, Mertens AN, Crider YS, *et al.* Substantial underestimation of SARS-CoV-2 infection in the United States. Nat Commun 2020;11(1):4507.
5. Chimeddorj B, Nyamadawa, Davaalkham, *et al.* Novel coronavirus (COVID-19) population-based age-stratified sero-epidemiological study in Mongolia (a prospective, nationwide study). Available from: https://www.thelancet.com/cms/10.1016/j.lanwpc.2021.100317/attachment/aca67146-d00c-4a76-9ffc-9701826958f9/mmc1.pdf. Accessed: 22 April 2021.
6. Herlinda O, Bella A, Kusnadi G, *et al.* Seroprevalence of antibodies against SARSCov-2 in the high impacted sub-district in Jakarta, Indonesia. PLoS One 2021;16(12):1-13.
7. Sawitri AAS, Yuliyatni PCD, Astuti PAS, *et al.* Seroprevalence of SARS-CoV-2 antibodies in Bali Province: Indonesia shows underdetection of COVID-19 cases by routine surveillance. PLOS Glob Public Health 2022;2(8):e0000727.
8. Ahmad RA, Indriani C, Arisanti RR, *et al.* Seroprevalence of SARS-CoV-2 and risk factors in Bantul Regency in March-April 2021, Yogyakarta, Indonesia. PLOS Glob Public Health 2023;3(6):e0000698.

9. Van Kerkhove MD, Broberg E, Engelhardt OG, *et al.* The consortium for the standardization of influenza seroepidemiology (CONSISE): A global partnership to standardize influenza seroepidemiology and develop influenza investigation protocols to inform public health policy. Influenza Other Respir Viruses 2013;7(3):231-234.
10. World Health Organization. Coronavirus disease (COVID-19) technical guidance: Infection prevention and control/WASH. Geneva: World Health Organization; 2020.
11. Beijing Wantai Biological Pharmacy Enterprise. WANTAI Total Ab ELISA for antibody levels product. Beijing: Beijing Wantai Biological Pharmacy Enterprise Co., Ltd; 2020.
12. World Health Organization. Diagnostic testing for SARS-CoV-2: Interim guidance. Geneva: World Health Organization; 2020.
13. Gha A, Meurant R. COVID-19 serological tests: How well do they actually perform?. Diagnostics 2020;10(7):453.
14. Jarvis LM. Biotech start-ups hit by coronavirus work stoppages. C&EN Glob Enterp 2020;98(7):13.
15. Badan Pusat Statistik. Profil statistik kesehatan 2021. Jakarta: Badan Pusat Statistik; 2021.
16. Murhekar MV, Bhatnagar T, Selvaraju S, *et al.* SARS-CoV-2 antibody seroprevalence in India, August–September, 2020: Findings from the second nationwide household serosurvey. Lancet Glob Health 2021;9(3):e257-e266.
17. Selvaraju S, Kumar MS, Thangaraj JWV, *et al.* Population-based serosurvey for severe acute respiratory syndrome coronavirus 2 transmission, Chennai, India. Emerg Infect Dis 2021;27(2):586-589.
18. Wiens KE, Mawien PN, Rumunu J, *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Juba, South Sudan: A population-based study. MedRxiv Prepr Serv Health Sci 2021:1-28.
19. Rochwulaningsih Y, Sulistiyono ST, Utama MP, *et al.* Integrating socio-cultural value system into health services in response to COVID-19 patients' self-isolation in Indonesia. Humanit Soc Sci Commun 2023;10(1).
20. Richter E, Liebl D, Schulte B, *et al.* Analysis of fatality impact and seroprevalence surveys in a community sustaining a SARS-CoV-2 superspreading event. Sci Rep 2023;13(1):1-13.
21. Thomeer MB. Relationship status-based health disparities during the COVID-19 pandemic. Soc Curr 2023;10(1):17-40.
22. Nkire N, Nwachukwu I, Shalaby R, *et al.* COVID-19 pandemic: Influence of relationship status on stress, anxiety, and depression in Canada. Ir J Psychol Med 2022;39(4):351-362.
23. Salah R, Al-Doghmi N. The impact of the COVID-19 pandemic and lockdown on female university employees: A sociolinguistic case study. Int J Soc Cult Lang 2023;11(2):12-27.
24. Abe M, Arima H, Satoh A, *et al.* Marital status, household size, and lifestyle changes during the first COVID-19 pandemic: Nippon data2010. PLoS One 2023;18(3):e0283430.