BMJ Open

COVID-19 seroprevalence in Pakistan: a cross-sectional study

Ahsan M Ahmad,[1] Khurram Shahzad,[2] Mariumn Masood ,[2] Maida Umar,[2] Fahad Abbasi,[2] Assad Hafeez[2]

To cite: Ahmad AM, Shahzad K, Masood M, *et al*. COVID-19 seroprevalence in Pakistan: a cross-sectional study. *BMJ Open* 2022;**12**:e055381. doi:10.1136/bmjopen-2021-055381

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-055381).

Received 13 July 2021
Accepted 25 February 2022

[1]Center for Global Public Health-Pakistan, Health Services Academy, Ministry of National Health Services Regulations and Coordination, Opposite National Institute of Health (NIH) Chak Shahzad, Islamabad, Pakistan
[2]Health Services Academy, Ministry of National Health Services Regulations and Coordination, Islamabad, Pakistan

Correspondence to
Dr Mariumn Masood;
msmasood148@gmail.com

ABSTRACT

Objectives This study adapted WHO's 'Unity Study' protocol to estimate the population prevalence of antibodies to SARS CoV-2 and risk factors for developing SARS-CoV-2 infection.

Design This population-based, age-stratified cross-sectional study was conducted at the level of households (HH).

Participants All ages and genders were eligible for the study (exclusion criteria: contraindications to venipuncture-however, no such case was encountered). 4998 HH out of 6599 consented (1 individual per HH). The proportion of male and female study participants was similar.

Primary and secondary outcome measures Following were the measured outcome measures- these were different from the planned indicators (i.e. two out of the three planned indicators were measured) due to operational reasons and time constraints: -
Primary indicators: Seroprevalence (population and age specific).
Secondary indicators: Population groups most at risk for SARS-CoV-2-infection.

Results Overall seroprevalence of SARS-CoV-2 antibodies was 7.1%. 6.3% of individuals were IgG positive while IgM positivity was 1.9%. Seroprevalence in districts ranged from 0% (Ghotki) to 17% (Gilgit). The seroprevalence among different age groups ranged from 3.9% (0–9 years) to 10.1% (40–59 years). There were no significant differences in the overall seroprevalence for males and females. A history of contact with a confirmed COVID-19 case, urban residence and mask use were key risk factors for developing SARS-CoV-2 infection.

Conclusions This survey provides useful estimates for seroprevalence in the general population and information on risk factors for developing SARS-CoV-2 infection in the country. It is premised that similar studies need to be replicated at the population level on a regular basis to monitor the disease and immunity patterns related to COVID-19.

Strengths and limitations of this study

- A sample size of about 5000 individuals was selected, with random selection at each stage i.e. district, union council, villages (for rural union councils), households, and individuals; and there was representation from all four provinces and both regions of the country.
- There was an almost equal representation of males and females; and the rural and urban samples were in accordance with the proportion of urban: rural union councils of each district.
- Seroprevalence was gauged using rapid diagnostic tests, which are not the gold standard (for this purpose) and may have varying sensitivity depending on the time since infection onset.
- Measurement of some of the risk factors for developing SARS-CoV-2 infection was through self-report, and had the potential to introduce recall bias and social desirability bias.
- The age structure of the study population differs from the demographic distribution of the population in the country, which may have caused bias in the study findings.

INTRODUCTION

COVID-19, originating from Wuhan, China was declared a pandemic by the WHO on the 11 March 2020.[1] Pakistan reported its first case on 26 February 2020, witnessing its first peak during July 2020.[2] Among various response measures in Pakistan have been the conduction of research studies to inform response measures to COVID-19 and to enable a better understanding of its epidemiology and spread. This includes partaking in the WHO's global research initiatives such as the 'WHO Unity Studies' through the conduction of national seroprevalence studies.[3] These study protocols have also been adapted around the world at various geographic levels to provide contextual data on the evolving pandemic.

Numerous seroprevalence studies have been conducted around the globe since the onset of the pandemic, with more evidence from large-scale nationwide studies being reported as the situation evolves, and the results showing wide contextual variations.[4] In neighbouring countries, Iran reported a high seroprevalence of 17.1% in a large study conducted across 17 provinces during April to June 2020, although the results were from a much earlier phase of the pandemic.[5] India also conducted a national seroprevalence survey (for IgG) in adults during the same period using ELISA . However, in stark

contrast to Iran, it reported markedly lower seroprevalence of 0.7% in its study population.[6] Studies in the other regions also depict varying patterns of seroprevalence depending on the timelines of the pandemic in their respective countries.[7–9] One of the largest seroprevalence studies around the globe in England has reported a decline in seroprevalence estimates of IgG by 26.5% between June and September 2020, from 6.0% to 4.4% by lateral flow immunoassay tests.[10]

This study in Pakistan is part of the global 'WHO Unity Studies' initiative with the main objective of estimating the seroprevalence of IgG and IgM antibodies to the 'SARS-CoV-2' coronavirus in the general population in the country. This study will not only provide data regarding the exposure of the general population to COVID-19, but would also shed light on some risk factors for developing SARS-CoV-2 infection. Overall, we expect that the estimates would provide us ample evidence to gauge the population-level scenario of COVID-19 in the country as well as provide insights into other epidemiological aspects of the disease, including the risk factors for developing SARS-CoV-2 infection.

METHODS

Patient and public involvement statement: Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

This population-based, age-stratified cross-sectional study was conducted at the level of households (HH) between 21 October 2020 and 8 November 2020.

This was a nationwide study in the four provinces (Punjab, Sindh, Baluchistan and Khyber Pakhtunkhwa) and two regions (Azad Jammu & Kashmir (AJK) and Gilgit Baltistan (GB)) of Pakistan. Individuals of all ages and genders were eligible to participate.

Sample size calculation was done using Open Source Epidemiologic Statistics for Public Health (https://www.openepi.com/Menu/OE_Menu.htm) using multistage sampling design (with expected prevalence of 11% which were preliminary estimates from a previous national level study[11]); difference between upper and lower limit of the interval estimate as 1.25% (0.75% on either side); and a design effect of '2'. The resultant sample size of 4803 was rounded off to 5000 and distributed equally among the ten districts under study as well as among all age brackets.

The study was conducted in 10 districts of the country according to the following criteria: one high and one low prevalence district each was selected from provinces; and the highest-prevalence district each was selected from regions AJK and GB. In Pakistan's context, high prevalence was taken as a cumulative of more than 500 COVID-19 cases (tier 1); and low prevalence as a cumulative of less than 500 cases (tier 2) by the cut-off date of 30 June 2020. The population of district and other characteristics were not used during sampling due to variable testing rates by population in various districts.

After district selection, union council (UC) selection was done with the aim of recruiting 500 HH/participants randomly from each district (25 participants defined as one cluster). UCs were randomly selected from each district; the selection was in accordance with the 'urban UC: rural UC' ratio for each district. Where there were less than 20 UCs in a district, the number of clusters were increased to make the total equal to 500 participants per district. Systematic random sampling was employed for the next stage, that is, HH selection. Thereafter, one individual was randomly selected from each consenting HH in line with the age distribution of the study. Each cluster of 25 had 5 participants from each age group. In this way, 25 HH/individuals were recruited for each cluster, with random selection at different stages (as described) aiming to reduce potential sources of selection bias.

Data collection for the survey included on-the-spot recruitment of HH in the selected localities. After obtaining informed consent, each participant (one individual randomly selected out of all the eligible from each HH) was asked to provide information to the enumerator to fill a pretested questionnaire. The questionnaire was adapted from WHO Unity Studies protocol. It collected information on sociodemographic variables, medical and symptom history, preventive behaviours (note: handwashing was defined as follows—individuals were enquired about number of times they had washed their hands with soap for 20 s; the variable was categorised into those washing hands at least six times and those less than six times), complications, and a history of recent death in family. This was followed by a rapid diagnostic test (RDT) performed by a trained phlebotomist.[3] RDT was performed with 'Bioperfectus' kits for IgG/IgM, and the results were provided to each participant on the spot (within twenty minutes). Prior to the survey, data collectors and field teams underwent training, which included PPE usage for infection prevention and control. Monitoring visits were also conducted during the data collection period to enhance data quality and results. We computed descriptive statistics and univariate logistic regression that examined associations of seropositivity with age, gender, location, tier, gender and symptoms with 95% CI.

The multivariate binary logistic regression model

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

where p is the probability that an individual is seropositive, β_0 is the intercept, β_i are the coefficients and X_i represents the independent variables. The outcome variable represents the seropositivity which is a binary categorical variable whereas location, age, mask use and contact with COVID-19 positive person are the predictor variables. A forward-stepwise process was used for the selection of significant variables in the final model. During the analysis, some interaction terms were considered but not included in the final model because they were not statistically significant. The selection of multivariate logistic model was based on Homer-Lemeshow goodness of fit,

Table 1 Seroprevalence at district level

District	Reported prevalence based on RT(Real Time)-PCR	Positive (%)	Negative (%)
Quetta	High	16 (3.2)	484 (96.8)
Mardan		31 (6.2)	470 (93.8)
Rawalpindi		34 (6.8)	467 (93.2)
Ghotki		0 (0.0)	498(100)
Muzaffarabad	Low	39 (8.6)	460 (92.2)
Gilgit		85 (17.0)	415 (83.0)
Sibbi		24 (4.8)	474 (95.2)
Abbottabad		42 (8.4)	460 (91.6)
Lodhran		43 (8.6)	457 (91.4)
Jacobabad		37 (7.4)	462 (92.6)

biological interpretability and statistical significance. The significance level or alpha was 0.05. Data analysis for the survey was performed using SPSS V.23.

RESULTS

During field work, recruitment was continued till the target sample was achieved. A total of 6599 HH were reached for on-the-spot participation, of which 1601 (24%) did not provide consent to participate. A total of 4998 HH (with one individual per HH) consented to participate in the study across the 10 selected districts of Pakistan. The proportion of males in the recruited participants was 51%. The mean age of males (31.7 years) was similar to that of females (32.8 years). Almost two-thirds (62%) of individuals were recruited from rural areas. Male to female ratio of sampled individuals was similar for rural and urban areas.

Information on seroprevalence was available for all 4998 participants with no missing values. The overall seroprevalence of SARS-CoV-2 antibodies was 7.1%. Almost 6.3% of individuals were IgG positive while IgM positivity was 1.9%. Seroprevalence in districts ranged from 0.0% (Ghotki) to 17.0% (Gilgit). Most of the districts reported a range of 6% to 9% (table 1). The seroprevalence among different age groups ranged from 4% (0–9 years) to 10% (40–59 years) (table 2). A total of 4% reported to have had contact with a COVID-19 positive individual.

The use of preventative behaviours was also studied. The use of face masks while going out in public was reported to be 63%. Mask use was similar in urban and rural areas (63%). It was highest in 20–59 years age group (68%), while 60+ group reported relatively less use (49%). Mask use increased incrementally with education, from 36% in those non-educated to 82% in individuals above matric (ten years of education). Handwashing (washing hands at least six times with soap and water for 20 s in last 24 hours) was reported relatively less compared to mask use (39%). Handwashing was higher in urban (44%) areas; among females (43%); and increased with education, being highest in individuals above matric (table 3).

Symptoms during past 3 months were enquired to look for possible association with COVID-19 seropositivity. Total of 23 symptoms were enquired relating to multiple systems. The symptoms shown in table 2 were significantly higher in seropositive individuals. Sore throat, fatigue and joint aches were strongly associated with seropositivity. Among COVID-19 seropositive individuals, 68% had at least one symptom in last 2 months, while 32% reported to be completely asymptomatic during this period.

Almost 24% of individuals reported having at least one comorbidity. Hypertension was reported the most (18%), followed by diabetes (5%) and chronic kidney disease (2%). The reported occurrence of heart disease (1%) and asthma (2%) was relatively lower. Reported prevalence of at least one comorbidity increased with age, with maximum being reported for 60+ age group (54%), followed by 40–59 year age group (40%).

Multivariate logistic regression analysis was performed to identify factors associated with seropositivity (table 4). Urban residents were more likely to test positive for SARS CoV-2 antibodies than rural residents (OR 1.29, 95% CI 1.04 to 1.61). Individuals aged 20 and above were about twice as likely to be seropositive than those who were 0–9 years old. Odds of seropositivity were also high among individuals who did not wear face mask (OR 1.54, 95% CI 1.20 to 1.975) and in those who reported contact with a COVID-19-positive person (OR 1.81, 95% CI 1.16 to 2.83).

For any discrepancy and to ensure completeness of data, field teams were contacted and cross-checked. All cases are included in the analysis and missing data up to 1% were considered acceptable because of low occurrence.

DISCUSSION

The survey through the use of Unity Studies' age-stratified approach, estimated the national seroprevalence of IgG and IgM antibodies for COVID-19 to be 7.1% in Pakistan based on RDT testing. Among the included districts, the highest prevalence was observed for Gilgit, followed by Lodhran and Muzaffarabad (table 1). This study was initiated in the last week of October 2020 and field activities were completed by the second week of November 2020. This was a time frame, when the first wave of the pandemic was considered to have largely subsided, the new number of cases per day was markedly lower and there was a threat of a second wave of the pandemic in the forthcoming winter months of November onwards. During this time frame <1000 confirmed cases per day were being recorded, in the backdrop of the highest daily number of cases (6825) that had been reported on 13 June 2020.[2]

Although other tests such as ELISA offer greater accuracy than RDTs in terms of antibody detection,[12] due to practical and operational issues, RDTs were opted for this large-scale population-based study to estimate the

Table 2 Seroprevalence by age, place, district tier, gender and (significant) symptoms

Variables		Positive (%)	Negative (%)	Odds ratio (95% CI)	P value
Seroprevalence for COVID-19		351 (7.1)	4647 (93.0)		
Gender (n=4998)	Male	178 (6.9)	2394 (93.1)	0.97 (0.78 to 1.20)	0.879
	Female	173 (7.1)	2253 (92.9)		
Age (years) (n=4997)	0–9	38 (3.9)	935 (96.1)		
	10–19	46 (4.5)	973 (95.5)	0.43 (0.29 to 0.63)	0.000
	20–39	80 (7.8)	950 (92.2)	0.49 (0.34 to 0.71)	0.000
	40–59	101 (10.1)	899 (89.9)	0.88 (0.64 to 1.21)	0.390
	60+	85 (8.7)	890 (91.3)	1.18 (0.87 to 1.59)	0.309
Location (n=4998)	Urban	154 (8.1)	1741 (91.9)	1.30 (1.04 to 1.62)	0.014
	Rural	197 (6.3)	2906 (93.7)		
District prevalence (reported cases) (n=4998)	Low prevalence (tier-2)	270 (9.0)	1919 (96.0)	2.34 (1.81 to 3.02)	0.000
	High prevalence (tier-1)	81 (4.1)	2728 (91.0)		
Contact with COVID-19 positive case (n=4971)	Yes	25 (13.9)	155 (86.1)	2.34 (1.81 to 3.03)	0.001
	No	292 (6.7)	4041 (93.3)		
Sore throat (n=4987)	Yes	118 (10.6)	998 (89.4)	1.86 (1.46 to 2.32)	0.000
	No	233 (6.0)	3638(94)		
Fatigue (n=4991)	Yes	43 (10.9)	351 (89.1)	1.706 (1.22 to 2.39)	0.002
	No	308 (6.7)	4289 (93.3)		
Joint ache (n=4988)	Yes	65 (10.7)	541 (89.3)	1.734 (1.31 to 2.30)	0.000
	No	284 (6.5)	4.98 (93.5)		
High-grade fever (n=4986)	Yes	123 (8.6)	1313 (91.4)	1.365 (1.09 to 1.72)	0.008
	No	228 (6.4)	3322 (93.6)		
Cough (n=4993)	Yes	98 (8.5)	1055 (91.5)	1.317 (1.03 to 1.67)	0.026
	No	253 (6.6)	3587 (93.4)		
Runny nose (n=4987)	Yes	97 (8.5)	1044 (91.5)	1.32 (1.03 to 1.7)	0.026
	No	253 (6.6)	3593 (93.4)		

seroprevalence of antibodies to COVID-19, with the results for on-the-spot testing in the field available within twenty minutes. Some researchers using similar kits to those used in this survey reported the sensitivity (to detect IgG/IgM) to range between 41% (at 1–5 days since symptom onset in patients positive by RT-PCR) to 100% (at >20 days since symptom onset in patients positive by RT-PCR); while the reported specificity when compared with PCR was 95%[13] This time-dependent sensitivity of RDTs to detect SARS-Cov-2 antibodies has also been noted elsewhere.[14] Due to the widely varying sensitivity of the testing method, it is likely that the actual seroprevalence may have been much higher than this study's estimates.

It should be noted that the districts' population size and the age structure of the districts' population were not taken into account during district selection or selection of individuals of different age groups from within HH, this may have potentially introduced bias in the seroprevalence estimates.

Additionally, the high and low prevalence definition ideally should have been based on percent of cases reported by population in a district. During the initial days of the epidemic, the number of cases reported were quite low. Only 24 districts (out of 136) had reported more than 500 cases. Less than 100 cases were reported by 41 districts. Thus, a strategic decision was made to consider districts reporting more than 500 cases as high prevalence—which has its limitations.

The estimated seroprevalence was 62 times that of the cases reported by 30th October 2020 in the sampled districts. This points towards a general lack of testing in sampled districts. In Pakistan, testing mostly had been done in symptomatic cases and their contacts. Thus, a large pool of subclinical infections remained undetected. The variation and low diagnostic testing are likely to be attributable to the gap between seroprevalence and reported cases.

About one-third of seropositive individuals had reported to have experienced symptoms during the past 2 months. Six out of the studied 23 symptoms experienced during the past 2 months were found to be significantly associated with seropositivity in the univariate analysis, most of which were respiratory/pharyngeal symptoms including sore throat, shortness of breath, cough, and

Table 3 Mask use and handwashing practices by sociodemographic characteristics

Variables		Mask use (%)			Handwashing (%)		
		Yes	No	Total (%)	Yes	No	Total (%)
Overall		3128 (62.6)	1844 (37.1)		1946 (38.9)	3052 (61.1)	
Location	Urban	1185 (62.9)	699 (37.1)	1884 (37.8)	839 (44.3)	1056 (55.7)	1895 (37.9)
	Rural	1943 (62.9)	1145 (37.1)	3088 (62.1)	1107 (35.7)	1996 (64.3)	3103 (62.1)
District Tiers	Tier-1	1121 (56.5)	864 (43.5)	1985 (39.9)	762 (38.1)	1238 (61.9)	2000 (40.0)
	Tier-2	2007 (67.2)	980 (32.8)	2987 (60.0)	1184 (39.5)	1814 (60.5)	2992(60)
Gender	Male	1654 (64.6)	907 (35.4)	2561 (51.5)	905 (35.2)	1667 (33.4)	2572 (51.5)
	Female	1474 (61.1)	937 (38.9)	2411 (48.5)	1041 (42.9)	1385 (27.7)	2426 (48.5)
Age	0–19	576 (59.5)	392 (40.5)	968 (24.5)	264 (27.1)	709 (72.9)	973 (24.5)
	20–59	1364 (67.5)	657 (32.5)	2021 (51.5)	934 (46.0)	1096 (54.0)	2030 (51.0)
	60+	476 (49.1)	50.9 (50.9)	970 (24.5)	382 (39.2)	593 (2398)	975 (24.5)
Education	No education	412 (36.0)	731 (39.0)	1143 (36.9)	326 (28.4)	822 (71.6)	1148 (36.9)
	Primary	194 (61.0)	124 (24.8)	318 (10.3)	137 (42.9)	182 (57.1)	319 (10.3)
	Matric	491 (75.2)	162 (18.4)	653 (21.2)	307 (46.5)	353 (53.5)	660 (21.2)
	Above matric	797 (81.6)	180 (64.0)	977 (36.9)	573 (58.4)	408 (13.1)	981 (31.6)

runny nose. High grade fever, joint aches and fatigue were the three generalised symptoms associated with seropositivity. Although the symptoms were self-reported with a possibility for recall bias, information on symptoms was obtained before the testing was done hence it is likely that any misclassification of symptoms may have been non-differential.

Similar to what has been reported in some other national studies (including a large-scale national-level HH study by Pollán *et al* in Spain), gender was not found to be significantly associated with seropositivity for SARS CoV-2 antibodies.[8 9 15] The association of age with seropositivity increased with age until 59 years, and declined slightly in those above this age bracket. As observed previously, nasal gene expression of ACE2 has been postulated to be responsible for this age-related pattern, and children have been reported to be less susceptible to contracting COVID-19 than adults potentially due to the role of innate immunity[16], and the protective effect of lower levels of this enzyme in their nasal epithelium.[17 18]

The risk of seropositivity doubled in those with a history of exposure to a diagnosed COVID-19 patient and was found to be statistically significant. A study in Italy reported an even higher OR of 2.5 in those who had previous contact with a case.[19] Association of COVID-19 positivity with HH contact with known cases of COVID-19 has also found to be significant in other prevalence studies.[20]

Table 4 Factors associated with seropositivity

Variables		B	SE	Adjusted OR (95% CI)*	P value
Location	Urban	0.256	0.113	1.29 (1.035 to 1.612)	0.024
	Rural†				
Age (years)	0–9†				
	10–19	0.112	0.225	1.11 (0.7 to 1.737)	0.618
	20–39	0.650	0.204	1.92 (1.285 to 2.854)	0.004
	40–59	0.975	0.197	2.65 (1.803 to 3.899)	0.000
	60+	0.894	0.202	2.45 (1.646 to 3.630)	0.000
Mask use	Yes	0.434	0.126	1.54 (1.205 to 1.975)	0.001
	No†				
Contact with COVID-19 positive case	Yes	0.596	0.227	1.81 (1.163 to 2.831)	0.009
	No†				
Constant		–3.599			

*Hosmer-Lemshow goodness-of-fit test: χ^2 value=8.322, p=0.403.
†Reference category.

Differences in seroprevalence between tier 1 (districts considered to be high transmission areas by the end of June 2020) and tier 2 districts (districts considered to be low transmission areas at the end of June 2020), were significant at about 4.9%. However, it is important to note that the seroprevalence was lower in tier 1 districts in this study. This pattern suggests that the transmission scenarios would have evolved in 3–4 months since the chosen cut-off date, and the areas earlier considered to be higher transmission at the end of June (i.e. tier 1 districts) may have now become areas of relatively lower transmission and vice versa. The said difference is understood to be possibly due to a higher proportion of population in previously high-risk districts, having possibly experienced the exposures, with possibility of reduction in IgM levels across a time span of 90 or more days (i.e. the time span between June and October 2020).

High reported seroprevalence in tier 1 vs tier 2 districts could also be due to other confounding factors. The average population size of the districts in tier 1 was 2.9 million compared with 0.9 million in tier 2. Similarly, the average population density in tier 1 districts was more than double of that in tier 2 districts (852 vs 348 persons per square kilometre).

Urban and rural areas were sampled from each district in accordance with the urban to rural ratio for that particular district. Overall, urban areas reported a higher prevalence (8.1%) than rural areas (6.3%), and urban residents were more likely to be seropositive for COVID-19 than rural residents (table 4). While some researchers have argued that larger city sizes tend to have higher attack rates,[21] other studies have gone further so as to report that while urban areas do have a propensity for earlier outbreaks than rural areas, population density is not significantly associated with COVID-19 cases.[22] Our results seem to align with the former, that is, showing a significant association of urban residence with COVID-19 seropositivity.

Among studied behaviours, use of masks was reported by about two-thirds of the study participants with similar values in rural and urban areas. Mask use was found to be significantly linked with seropositivity in logistic regression in our study (table 4), in line with wide-ranging evidence.[1 23 24] On the other hand, hand-washing (at least six times per day) had a lower prevalence (39%) in the study population and was not found to be a significant risk factor in this study, although other researchers have reported protective benefits of hand hygiene.[25] Since these have both been recommended preventive behaviours during the pandemic, the results may have been affected by Social Desirability Bias[26] in the study population, causing over-reporting of these behaviours. In such a case, any true association of these behaviours with seropositivity for COVID-19 may have been masked, if they do indeed influence seropositivity for SARS-CoV-2.

Various studies to estimate the seroprevalence of COVID-19 have been conducted in the country at different scales. For example, preliminary results from a national seroprevalence study that was conducted during July 2020 reflected a prevalence of 11.2%.[11] Seroprevalence studies at a smaller scale in the country have also been conducted in Karachi city and Islamabad Capital Territory (ICT), with the former reporting figures of 9.7% and 15.1% in 'low-transmission' and 'high-transmission' areas in the city, respectively,[27] and the latter reporting an overall seroprevalence of 14.5%.[28 29] Results from the aforementioned sub-national surveys differ from those of the current study in which the seroprevalence was found to 4.1% in tier 1 districts and 9.2% in tier 2 districts, respectively. This may have been due the differing time frames of the studies in Karachi and ICT, which were conducted during the first wave of COVID-19 in the country; while the current study was conducted during October–November 2020, when the first wave of the pandemic had largely subsided.

Overall, it is likely that the seroprevalence estimates may have been affected by low sensitivity of the testing methods (causing underestimation of the values); and the estimates for association of seropositivity with risk factors for developing SARS-CoV-2 infection may have been affected by biases including recall and social desirability bias among the study population. The seroprevalence estimates provided by this study may be interpreted with caution as an estimated value for the general population, particularly since the age structure of the study population differs from the demographic distribution of the population in the country.

CONCLUSIONS

Notwithstanding the limitations of the study, this survey provides useful prevalence estimates as well as information on risk factors for developing SARS-CoV-2 infection. The results also show that the youngest age groups have the lowest proportion of seropositivity as compared with those aged 40 years and above. Interestingly, tier 1 districts (considered to be high risk based on the number of PCR test based confirmed cases by the end of June 2020) reflected lower prevalence as compared with tier 2 districts which may perhaps be depicting reversing patterns at the population level over time. A history of contact with a confirmed COVID-19 case, being an urban resident, and mask use were key risk factors for developing SARS-CoV-2 infection. Keeping view of these findings, it is premised that similar studies need to be replicated at the population level on a regular basis to monitor the disease and immunity patterns related to COVID-19.

Acknowledgements Muhammad Siddique supported various survey activities, including provision of population data, field activities, and training. We are thankful for the support and contribution of everyone who has been a part of the research process.

Contributors AH and AMA supervised the overall process; AMA contributed to the analysis, manuscript writing and monitoring of data collection; KS led the coordination of all survey activities including design, training, administrative aspects and data analysis; led the process of preparing tools for data collection; contributed to training of data collectors and monitoring of the data collection process, and to the analysis and reporting of Results; MM supported preparation of data collection tools, contributed to training of data collectors, drafted the

household sampling methodology, conducted household refusal data analysis, contributed to monitoring of data collection, conducted literature reviews to analyse epidemiological data from the survey, and prepared the initial manuscript draft; MU contributed to data management, analysis and reporting. FA supported preparation of data collection tools, contributed to monitoring of data collection and training of data collectors. AMA is the guarantor of the research. All authors reviewed and approved the manuscript.

Funding This study was jointly funded by WHO, and Health Services Academy (HSA), Islamabad, Ministry of National Health Services Regulations and Coordination, Government of Pakistan.The funding organizations were involved in all aspects of the survey including study design/technical supervision; collection, analysis and interpretation of the data; in the writing of the manuscript; and in the decision to submit the paper for publication. Funding/grant award number: F-11-2020-WHO/DAi/HSA.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Institutional Ethical Review Committee of Health Services Academy, Ministry of National Health Services, Islamabad, Pakistan No. 7-82/IERC-HSA/2020-33. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. These are deidentified participant data, available from Health Services Academy (HSA) (Dr Shahzad Ali Khan, Vice Chancellor, HSA. Email: shahzad@hsa.edu.pk). Data may be reused for future meta-analysis upon reasonable justification of benefit at regional/global level after permission. Additional information is also available (protocol inclusive of study tool and statistical analysis plan).

ORCID iD

Mariumn Masood http://orcid.org/0000-0002-4041-3200

REFERENCES

1 World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020, 2020. Available: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19 [Accessed 22 December 2020].
2 Government of Pakistan. COVID-19 Dashboard 2020. Available: http://covid.gov.pk/stats/pakistan?locale=en [Accessed 22 December 2020].
3 World Health Organization. Coronavirus disease (COVID-19) technical guidance: the unity studies: early investigation protocols 2020. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/early-investigations [Accessed 22 December 2020].
4 Arora RK, Joseph A, Van Wyk J, *et al*. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. *Lancet Infect Dis* 2021;21:e75-e76.
5 Poustchi H, Darvishian M, Mohammadi Z, *et al*. SARS-CoV-2 antibody seroprevalence in the general population and high-risk occupational groups across 18 cities in Iran: a population-based cross-sectional study. *Lancet Infect Dis* 2021;21:473-481.
6 Murhekar MV, Bhatnagar T, Selvaraju S, *et al*. Prevalence of SARS-CoV-2 infection in India: findings from the National serosurvey, May-June 2020. *Indian J Med Res* 2020;152:48.
7 Xu X, Sun J, Nie S, *et al*. Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. *Nat Med* 2020;26:1193–5.
8 Pollán M, Pérez-Gómez B, Pastor-Barriuso R, *et al*. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* 2020;396:535–44.
9 Herzog S, De Bie J, Abrams S, *et al*. Seroprevalence of IgG antibodies against SARS coronavirus 2 in Belgium: a prospective cross-sectional study of residual samples. *medRxiv* 2020.
10 Ward H, Cooke G, Atchison CJ, *et al*. Declining prevalence of antibody positivity to SARS-CoV-2: a community study of 365,000 adults. *MedRxiv* 2020.
11 Dawn News. 11pc Pakistanis have developed protective immunity: study 2020. Available: https://www.dawn.com/news/1575607/11pc-pakistanis-have-developed-protective-immunity-study [Accessed 22 December 2020].
12 Kontou PI, Braliou GG, Dimou NL, *et al*. Antibody tests in detecting SARS-CoV-2 infection: a meta-analysis. *Diagnostics* 2020;10:319.
13 Whitman JD, Hiatt J, Mowery CT, *et al*. Evaluation of SARS-CoV-2 serology assays reveals a range of test performance. *Nat Biotechnol* 2020;38:1174–83.
14 Zainol Rashid Z, Othman SN, Abdul Samat MN, *et al*. Diagnostic performance of COVID-19 serology assays. *Malays J Pathol* 2020;42:13–21.
15 Lai C-C, Wang J-H, Hsueh P-R. Population-Based seroprevalence surveys of anti-SARS-CoV-2 antibody: an up-to-date review. *Int J Infect Dis* 2020;101:314–22.
16 Yang HS, Costa V, Racine-Brzostek SE, *et al*. Association of age with SARS-CoV-2 antibody response. *JAMA Netw Open* 2021;4:e214302–e.
17 Patel AB, Verma A. Nasal ACE2 levels and COVID-19 in children. *JAMA* 2020;323:2386–7.
18 Bunyavanich S, Do A, Vicencio A. Nasal gene expression of angiotensin-converting enzyme 2 in children and adults. *JAMA* 2020;323:2427.
19 Vena A, Berruti M, Adessi A, *et al*. Prevalence of antibodies to SARS-CoV-2 in Italian adults and associated risk factors. *J Clin Med* 2020;9:2780.
20 Menachemi N, Yiannoutsos CT, Dixon BE, *et al*. Population Point Prevalence of SARS-CoV-2 Infection Based on a Statewide Random Sample - Indiana, April 25-29, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:960–4.
21 Stier A, Berman M, Bettencourt L. *COVID-19 attack rate increases with City size. Mansueto Institute for Urban Innovation Research Paper Forthcoming*, 2020.
22 Carozzi F. Urban density and Covid-19. *SSRN Journal* 2020.
23 Chu DK, Akl EA, Duda S, *et al*. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* 2020;395:1973–87.
24 Gandhi M, Beyrer C, Goosby E. Masks do more than protect others during COVID-19: reducing the inoculum of SARS-CoV-2 to protect the wearer. *J Gen Intern Med* 2020;35:3063–6.
25 Ma Q-X, Shan H, Zhang H-L, *et al*. Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2. *J Med Virol* 2020;92:1567–71.
26 Grimm P. Social desirability bias. In: *Wiley international encyclopedia of marketing*, 2010.
27 Nisar MI, Ansari N, Amin M, *et al*. Serial population based serosurvey of antibodies to SARS-CoV-2 in a low and high transmission area of Karachi, Pakistan. *medRxiv* 2020.
28 Dawn News. 300,000 persons infected with Covid-19 in Islamabad: survey 2020 [Accessed 22 December 2020].
29 Pakistan Go. COVID-19 Tracker (press release, July 2020) 2020. Available: http://covid.gov.pk/press-release-details/2 [Accessed 22 December 2020].