medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Article Summary Line: Rapid increase in seroprevalence of antibodies against SARS-CoV-2 was seen in Karachi, Pakistan from April to June 2020 with a high conditional risk of infection within the household

Running Title: SARS-CoV-2 serosurvey in Karachi, Pakistan

Keywords: Seroepidemiologic Studies; Coronavirus; Pakistan

Title: Serial population based household survey of antibodies to SARS-CoV-2 in low and high transmission neighborhoods of urban Pakistan

Authors: Muhammad Imran Nisar MBBS, MSc; Nadia Ansari MBBS, MSc; Mashal Amin MBBS, MSc; Aneeta Hotwani MSc; Farah Khalid MSc; Najeeb Rehman BSc; Arjumand Rizvi MPhil; Arslan Memon MBBS, MPH; Zahoor Ahmed MBBS; Ashfaque Ahmed MBBS; Junaid Iqbal PhD; Ali Faisal Saleem MBBS, MSc; Uzma Bashir Aamir MBBS, PhD; Daniel B Larremore PhD; Bailey Fosdick PhD; Fyezah Jehan MBBS, MSc

Affiliations:

The Aga Khan University, Karachi, Sind, Pakistan (MI. Nisar, N. Ansari, M. Amin, A. Hotwani, F. Khalid, N. Rahman, A. Rizvi, J. Iqbal, AF. Saleem, F. Jehan)

Health Department, Government of Sind, Karachi, Pakistan (A. Memon, Z. Ahmed, A. Ahmed)

World Health Organization Country Office, Karachi, Sind, Pakistan (UB. Aamir)

University of Colorado, Boulder, Colorado, USA (DB. Larremore)

Colorado State University, Fort Collins, Colorado, USA (B. Fosdick)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract- 146 words

Serial household antibody sero-surveys informs the pandemic where testing is non-uniform. Young populations with intergenerational co-residence may have different transmission dynamics. We conducted two serial cross-sectional surveys in April and June 2020 in low- and high-transmission neighborhoods of Karachi, Pakistan, using random sampling. Symptoms were assessed and blood tested for antibody using chemiluminescence. Seroprevalence was adjusted using Bayesian regression and post stratification. CRI with 95% confidence intervals was obtained. We enrolled 2004 participants from 406 households. In June 8.7% (95% CI 5.1-13.1) and 15.1% (95% CI 9.4 -21.7) were infected in low- and high-transmission-areas respectively compared with 0.2% (95% CI 0-0.7) and 0.4% (95% CI 0 - 1.3) in April. Conditional risk of infection was 0.31 (95% CI 0.16-0.47) and 0.41(95% CI 0.28-0.52) respectively with only 5.4% symptomatic. Rapid increase in seroprevalence from baseline is seen in Karachi, with a high probability of infection within household.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Text-2541 words

INTRODUCTION

The global COVID-19 pandemic has resulted in more than 15.5 million confirmed cases (until July 24) and more than 633000 deaths, with an estimated case fatality rate (CFR) of 4.1 % *(1)*. Pakistan was among the first of low- and middle- income countries (LMICs) to be affected, and since then, there have been 269000 cases with 5709 deaths (CFR 2.1 %) *(2)*. Karachi, a large metropolitan city became the epicenter of epidemic on 26th Feb 2020. Since then Karachi has seen the largest number of cases (~83000 or 31% of all cases) in Pakistan.

The population demographics of Pakistan are typical of most LMICs with over 65% younger than 30 years of age, nuclear families with intergenerational co-residence and average household sizes of greater than 6 in urban areas. In Karachi, crowded neighborhoods, urban slum dwellings and poor adherence to social distancing measures add to the dynamics of the epidemic. Symptom-based surveillance strategies are the mainstay of prevention but supply (equipment, reagents, and NP-swabs) and demand (stigma and fear) side issues limit its effectiveness along with undetected pre-symptomatic and asymptomatic transmission *(3)*.

In this milieu, population-based sero surveys can be ideal *(4, 5)*. However, high quality reliable population based surveys may not be feasible in an ongoing pandemic because they mandate training on personal protective equipment (PPE) for data collectors and allaying fear or anxiety in potential participants to avoid non-response bias. Sampling entire households can ease procedures of data and blood collection and provide an opportunity to study transmission at household level especially when lockdown and social distancing measures are in place.

Household transmission is a concern in closed congested neighborhoods of metropolitan cities when lockdown measures are in place. It is also of importance because of intergenerational

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

co residence of young and old, especially in light of recent evidence indicating that children can be major drivers of the epidemic *(6)*. Secondary transmission from index cases in households using prospective follow up and active symptom monitoring with NP-PCR has indicated household attack rates of 32.4% (95% CI 22.4%–44.4%) *(7)*. However it is resource intensive and may differ between symptomatic and asymptomatic households, as symptomatic individuals are more likely to transmit the virus *(8-11)*. In a seroprevalence data, household transmission is easily estimable through the conditional risk of infection (CRI), namely the probability that an individual in a household is infected given that another household member is infected.

With this approach, we estimated changes in seroprevalence in low- and high-transmission neighborhoods of Karachi between April and June 2020 and compared the seroprevalence to total reported positive tests in the area to understand the level of under-reporting from gold standard RT-PCR testing. We also determined age - gender stratified estimates of seroprevalence and assessed the role of household transmission through conditional risk of infection (CRI) estimation; using the adapted World Health Organization (WHO) Unity protocol*(12)*.

MATERIALS and METHODS

Study Participants and sample collection

We conducted the study in two areas (Figure 1). Four sub-administrative units or union councils of District East were purposely selected as high-transmission area based on highest number of cases to date and one union council of District Malir which had reported only 4 out of 164 PCR tests was designated as a low-transmission area (Table 1)

Two serial cross-sectional surveys were performed at the household level, between April 15-25 and June 25- July 11, 2020. Four teams, comprising of one data collector and one

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

phlebotomist each conducted the survey. A detailed list of cases was available in District East which allowed households to be selected through systematic random sampling as follows. Reference point was identified randomly from the list of current positive cases and spinning a bottle or pen indicated direction. The reference household with the positive case did not take part in the survey, instead every nth household was sampled based on second to last digit of a bank note until the lane ended in which case the team moved to the next lane or until the day was completed. Same sampling strategy was repeated daily until the overall sample size of the survey was achieved. In case of refusal from a household, the household on the right was sampled. Household level approval and individual participant written informed consent or assent was obtained. A slightly different approach was adapted in District Malir where a list of active cases was not available but a line listing of households from ongoing work *(13)* allowed simple random sampling to be used for household identification daily until the sample size was achieved. In both areas, all household members were eligible to participate irrespective of their case status.

All team members were trained in use of PPE, hand hygiene, disinfection techniques and safe transportation of biological samples. A 3 ml sample of blood from infants and 5 ml from those older than 1 year was collected by a trained phlebotomist and transported to the Infectious Diseases Research Lab for centrifugation, serum separation and storage at -20^{0}C. Demographic and clinical data were collected using the adapted survey instrument from the WHO Unity Protocol for sero-epidemiology survey of COVID-19 *(4)*.

Laboratory analysis

A commercial Elecsys® Anti-SARS-CoV-2 immunoassay (Roche Diagnostics), targeting both IgG and IgM against SARS-CoV-2 was performed at the Nutritional Research Laboratory

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(NRL) at Aga Khan University. The manufacturer reports a specificity greater than 99.8% and sensitivity of 100% for individuals with a positive PCR test at least two weeks prior, and 88.1% sensitivity for those 7-13 days post a PCR positive test *(14)*. Optimization of assay at the NRL was done on a small number of samples; 20 stored sera from before 2019 and 25 newly collected sera from RT-PCR-confirmed COVID-19 cases. All 20 pre-pandemic sera were negative while 20 of 25 sera from cases were positive.

STATISTICAL ANALYSIS

Sample Size

The sample size for each phase of the survey per site was 500, totaling 1000 participants per survey. This allowed us to estimate an age adjusted prevalence for each site from 20-30 % at 95 % confidence with precision of ± 5% and a design effect of 1.5 for household level clustering.

Data Management

All data was double entered on an SQL database and checked for missingness and consistency. Age was categorized and also presented as mean with standard deviations and median with interquartile ranges. Occupation was recoded to create a category of those working outside home. Those reporting fever with or without respiratory symptoms in the previous two months were categorized as symptomatic.

Estimation of overall, age and gender stratified seroprevalence

Age and gender-stratified seroprevalence estimates were computed using a Bayesian hierarchical regression model. This approach, described in detail in the Supplementary Material, accounts for uncertainty due to finite lab validation data,*(15)* and produces estimates with

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

uncertainty across age and gender groups using typical choices of uninformative or weakly informative prior distributions*(16, 17)*.

Seroprevalence estimates were computed for each district and each survey phase independently, and all seroprevalence estimates are expressed as posterior means and 95% equal-tailed credible intervals based on 20000 samples from the Bayesian posterior distribution. All calculations were performed in R and samples from the posterior distributions were obtained using Stan *(18)*.

Household conditional risk of infection (CRI) analysis

CRI, the probability that an individual in a household is infected, given that another household member is infected*(19)* was estimated by a fraction whose numerator is the total number of ordered pairs among infected individuals in the same household and whose denominator is the total number of ordered pairs in the same household in which the first individual in the pair is infected. A 95% confidence interval was estimated via bootstrap for each area by resampling households with replacement.

Results

A total of 2004 participants were enrolled across two phases from District East and District Malir. Figure 2 Panel A and B describe the flow of participants. Among households who agreed to take part, individual participation rate was 82.3% (1000 out of 1215 household members eligible) in phase 1 and 76.5% (1004 out of 1312 household members eligible) in phase 2. Table 2 describes the baseline sociodemographic and clinical characteristics of the enrolled participants.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

In Phase 1 of the study, only 2 of 500 samples tested positive in District East and 0 of 500 tested positive in District Malir. In Phase 2 of the study, 100 of 500 samples (20.0%) tested positive in District East and 64 of 504 samples (12.7%) tested positive in District Malir. Both districts showed marked and significant increase in seroprevalence between time points. Of the total 166 participants who tested positive, only 9 (5.4%) gave a history of any symptoms.

To measure whether individuals in the same household were more likely to have similar sero- status, we computed the conditional risk of infection (CRI) for phase two. CRI estimates were 0.41 (95% CI 0.28-0.52) in District East and 0.31 (95% CI 0.16-0.4) in District Malir. These estimates are consistent with reported within household transmission, but cannot be completed attributed to within household transmission in the absence of a prospective design*(19).*

Given the correlation among household seropositivity values, we estimated seroprevalence via a Bayesian regression model (see Methods) which took into account household membership, age, and gender for each individual, as well as the lab validation data reported by the test manufacturer*(14).* Seroprevalence estimates by age and gender were then post stratified to adjust for the demographic makeup of the respective district.

In Phase 1, post-stratified seroprevalence was estimated to be 0.4% (95% CI 0%-1.3%) in District East and 0.2% (95% CI 0%-0.7%) in District Malir. In Phase 2, post-stratified seroprevalence was estimated to be 15.1% (95% CI 9.4%-21.7%) in District East, and 8.7% (95% CI 5.1%-13.1%) in District Malir. Seropositivity rates were indistinguishable between men and women within each district. Point estimates of seroprevalence were lower across all age

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

groups in District Malir than in District East, however there was significant overlap between the credible intervals for the two locations (Figure 3)

Discussion

The two serial serosurveys conducted in a metropolitan city of Pakistan in low- and high-transmission neighborhoods indicate a sharp increase in seroprevalence rates between April and June. While increase in the high transmission area is slightly greater and anticipated based on their identification as current "hotspots" of infection, as per the local health authorities, there has been a notable increase in the low transmission areas as well. The survey did not identify any difference in patterns of seroprevalence rates across age and gender categories. In both areas infection was mostly seen in individuals without any symptoms and within household sero-concordance suggests household transmission rates were high.

This is the first report from Pakistan and a low- and middle-income country, of a longitudinal population representative household survey conducted simultaneously in areas of low and high transmission.

The sharp elevation in seroprevalence in an area of presumably low transmission indicates that the virus continues to spread unchecked in populations where test timing is irregular or test numbers are suboptimal. It will be critical to monitor such populations because of the possible risk of resurgence within such areas once lockdown and contact tracing in high transmission areas is eased, as is currently being seen in North America *(20).*

Reasons for heterogeneity in SARS-CoV-2 transmission are largely unclear. Factors such as low socioeconomic area with poor access to clean water and sanitation and recurrent pneumonia and diarrheal illness in low-transmission area may be in play *(21, 22).*Transmission

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

of COVID-19 in areas with poor hygienic conditions is possibly reduced due to environmental microbes and resulting trained immunity seen with other SARS infections *(23)*. However, several of the District East neighborhoods that were part of the study had similar characteristics including household size. This heterogeneity and its underlying causes needs to be further explored.

Seroprevalence in children was not different from that seen in adults and elderly. Recent evidence suggests that children may be more likely to transmit the virus. However data from population-based studies in any countries suggest children may be less likely to be infected *(24).* Our study found that children might be infected at a rate similar to adults. This may be due to the household nature of our survey during school closure, where we were powered better for the pediatric population compared to other studies. However, the margins of errors were broad suggesting the need of better-powered studies to establish rates of infection in children relative to adults. Specifically, such studies could better elucidate the role of children in transmission blocking strategies such as lockdowns, school closures and vaccination measures, once available.

Spread of infection in the household is an important consideration in SARS-CoV2 transmission, especially in areas where there is lockdown and large families are roomed-in in small poorly ventilated spaces, typical of the neighborhoods of Karachi (25). The probability of an individual to have an infection in the presence of another infected household member in our study, as measured by conditional risk of infection (CRI) , was found to be similar to the actual and modelled secondary attack rates reported in other studies.(19)CRI can function as a substitute in situations where comprehensive surveillance and disease notification strategy is absent and secondary attack rates are difficult to calculate.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The strength of our study is knowledge of baseline seropositivity absent in most serosurveys and the serial nature with a two month inter survey interval allowing for a study of change in seroprevalence. A third round of survey is planned in August 2020. About one third or more of our sample includes children less than 18 years of age focusing on an understudied age group in the pandemic. Probability sampling which is community based is less biased for estimating prevalence and results are more generalizable than those from self-selecting or specific groups of individuals. Adaptation of the UNITY protocol allows for pooling of our results with other reports. The Electro Chemi Luminescence (ECL) technology used to test antibodies is sensitive and precise as per manufactures report *(14)*.

Limitations include the limited geographical area of the study and the slight lag in duration of conduct of the second serosurvey that may have led to a slight over-estimation in District Malir. In-house validation on local samples was also not performed due to a limited supply chain of testing kits in Pakistan, however this was somewhat compensated by modeling directly on the data reported by the manufacturer.

Conclusion

There is a large increase in seroprevalence to SARS-CoV-2 infection even in areas where transmission is reportedly low. Most seropositives are asymptomatic and a majority of the population is still seronegative. There is high probability of an individual to be infected given exposure to another infected in the household, irrespective of symptoms. Enhanced surveillance activities of COVID-19 are required especially in low-transmission sites in order to determine the real direction of the pandemic and the risks of household transmission in tightly knit neighborhoods in urban LMIC settings.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgments

The authors would like to acknowledge all the data collectors, phlebotomists and laboratory personnel who made this happen in the most difficult of circumstances.

Author Bio

Dr M Imran Nisar is an Assistant Professor in the Department of Pediatrics and Child Health at The Aga Khan University. He is an epidemiologist by training and his areas of interest include maternal & child health and infectious diseases.

Funding: This study was funded by the Infectious Disease Research Laboratory(IDRL) at the Aga Khan University in Karachi, Pakistan. None of the authors have any financial interest to disclose.

References

1. University JH. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020 [cited 2020 24/7/2020]; Available from: https://coronavirus.jhu.edu/map.html
2. Pakistan Go. COVID-19 Health Advisory Platform by Ministry of National Health Services Regulations and Coordination. 2020 [cited 2020 24/7/2020]; Available from: https://covid.gov.pk/stats/pakistan
3. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases. 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

4. Organization WH. Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection, 26 May 2020: World Health Organization; 2020.

5. Winter AK, Martinez ME, Cutts FT, Moss WJ, Ferrari MJ, McKee A, et al. Benefits and challenges in using seroprevalence data to inform models for measles and rubella elimination. The Journal of infectious diseases. 2018;218(3):355-64.

6. Heald-Sargent T, Muller WJ, Zheng X, Rippe J, Patel AB, Kociolek LK. Age-Related Differences in Nasopharyngeal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Levels in Patients With Mild to Moderate Coronavirus Disease 2019 (COVID-19). JAMA Pediatrics.

7. Wu J, Huang Y, Tu C, Bi C, Chen Z, Luo L, et al. Household Transmission of SARS-CoV-2, Zhuhai, China, 2020. Clinical Infectious Diseases. 2020.

8. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. Nature. 2020:1-.

9. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic population. New England Journal of Medicine. 2020.

10. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science. 2020;368(6491).

11. Streeck H, Schulte B, Kuemmerer B, Richter E, Höller T, Fuhrmann C, et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. medrxiv. 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

12. Organization WH. Coronavirus disease (COVID-19) technical guidance: The Unity Studies: Early Investigations Protocols. 2020 [cited 2020; Available from: Coronavirus disease (COVID-19) technical guidance: The Unity Studies: Early Investigations Protocols

13. Ilyas M, Naeem K, Fatima U, Nisar MI, Kazi AM, Jehan F, et al. Profile: health and demographic surveillance system in peri-urban areas of Karachi, Pakistan. Gates Open Research. 2018;2(2):2.

14. Ltd. RDI. ElectroChemiLuminescence (ECL). 2020 [cited; Available from: https://diagnostics.roche.com/global/en/article-listing/electroChemiLuminescence-unique-immunoassay-technology.html

15. Larremore DB, Fosdick BK, Zhang S, Grad YH. Jointly modeling prevalence, sensitivity and specificity for optimal sample allocation. bioRxiv. 2020.

16. Larremore DB, Fosdick BK, Bubar KM, Zhang S, Kissler SM, Metcalf CJE, et al. Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys. medRxiv. 2020.

17. Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. medRxiv. 2020.

18. Stan - Stan. 2020 [cited 2020 25/7/2020]; Available from: https://mc-stan.org/about/

19. Curmei M, Ilyas A, Evans O, Steinhardt J. Estimating Household Transmission of SARS-CoV-2. medRxiv. 2020.

20. Tuite AR, Greer AL, De Keninck S, Fisman DN. Risk for COVID-19 Resurgence Related to Duration and Effectiveness of Physical Distancing in Ontario, Canada. Annals of Internal Medicine. 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

21. Quadri F, Nasrin D, Khan A, Bokhari T, Tikmani SS, Nisar MI, et al. Health care use patterns for diarrhea in children in low-income periurban communities of Karachi, Pakistan. The American journal of tropical medicine and hygiene. 2013;89(1_Suppl):49-55.

22. Kerai S, Nisar I, Muhammad I, Qaisar S, Feroz K, Raza A, et al. A Community-Based Survey on Health-Care Utilization for Pneumonia in Children in Peri-Urban Slums of Karachi, Pakistan. The American journal of tropical medicine and hygiene. 2019;101(5):1034-41.

23. Netea MG, Giamarellos-Bourboulis EJ, Domínguez-Andrés J, Curtis N, van Crevel R, van de Veerdonk FL, et al. Trained Immunity: a Tool for Reducing Susceptibility to and the Severity of SARS-CoV-2 Infection. Cell. 2020.

24. Li X, Xu W, Dozier M, He Y, Kirolos A, Theodoratou E. The role of children in transmission of SARS-CoV-2: A rapid review. Journal of global health. 2020;10(1).

25. Statistics GoPBo. Housing units by number of rooms and type [cited; Available from: www.pbs.gov.pk

Address for correspondence: Fyezah Jehan, Associate Professor, Department of Paediatrics and Child Health, Aga Khan University, Stadium Road, Karachi 74800, Pakistan; Email: Fyezah.jehan@aku.edu; Tel: +92 21 3486 4981

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1- District size, population density and population of the selected union councils

District	Union councils	Population
East	Dalmia/ Shanti Nagar	50,692
	Pehlwan Goth	136,229
	Safoora Goth	223,771
	Faisal cantonment	69,576
Malir	Ibrahim Hyderi	85,000

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. General characteristics of the study participants

Characteristics	District East Phase 1 (n= 500)	District East Phase 2 (n= 500)	District Malir Phase 1 (n= 500)	District Malir Phase 2 (n= 504)
Sex, male n (%)	256 (51.2%)	225 (45.0%)	207 (41.4%)	199 (39.5%)
Age in years , mean(SD)	26.2 (17.9)	25.9(16.7)	28.5 (17.9)	24.32 (16.7)
Age, categories, n (%)				
0-4 year	35 (7.0%)	22 (4.4%)	26 (5.2%)	33 (6.6%)
5-9 year	57 (11.4%)	54 (10.8%)	52 (10.4%)	74 (14.7%)
10-18 year	107 (21.4%)	139 (27.8%)	86 (17.2%)	120 (23.9%)
19-39 year	185 (37.0%)	170 (34.0%)	203 (40.6%)	186 (37.0%)
40-59 year	83 (16.6%)	97 (19.4%)	103 (20.6%)	65 (12.9%)
60 and above	33 (6.6%)	18 (3.6%)	30 (6.0%)	25 (5.0%)
Household size, mean (SD)	6.1(4.17)	6.6(3.6)	6(2.8)	5.89(3.2)
Working outside home, n (%)	155 (31.0%)	145 (29.0%)	143 (28.6%)	124 (24.6%)
Comorbids, n (%)				
None	453 (92.3%)	463 (93.0%)	473 (94.6%)	479 (95.0%)
Diabetes	20 (4.0%)	14 (2.8%)	6 (1.2%)	10 (2.0%)
Hypertension	14 (2.8%)	13 (2.6%)	16 (3.2%)	11 (2.2%)
Asthma or Allergy	3 (0.6%)	3 (0.6%)	1 (0.2%)	0 (0.0%)
Chronic Hepatitis	2 (0.4%)	3 (0.6%)	3 (0.6%)	3 (0.6%)
Chronic Heart Disease	0 (0.0%)	3 (0.6%)	1 (0.2%)	1 (0.2%)
Asymptomatic	435 (87.0%)	453 (90.6%)	472 (94.4%)	487 (96.6%)

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163451; this version posted August 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

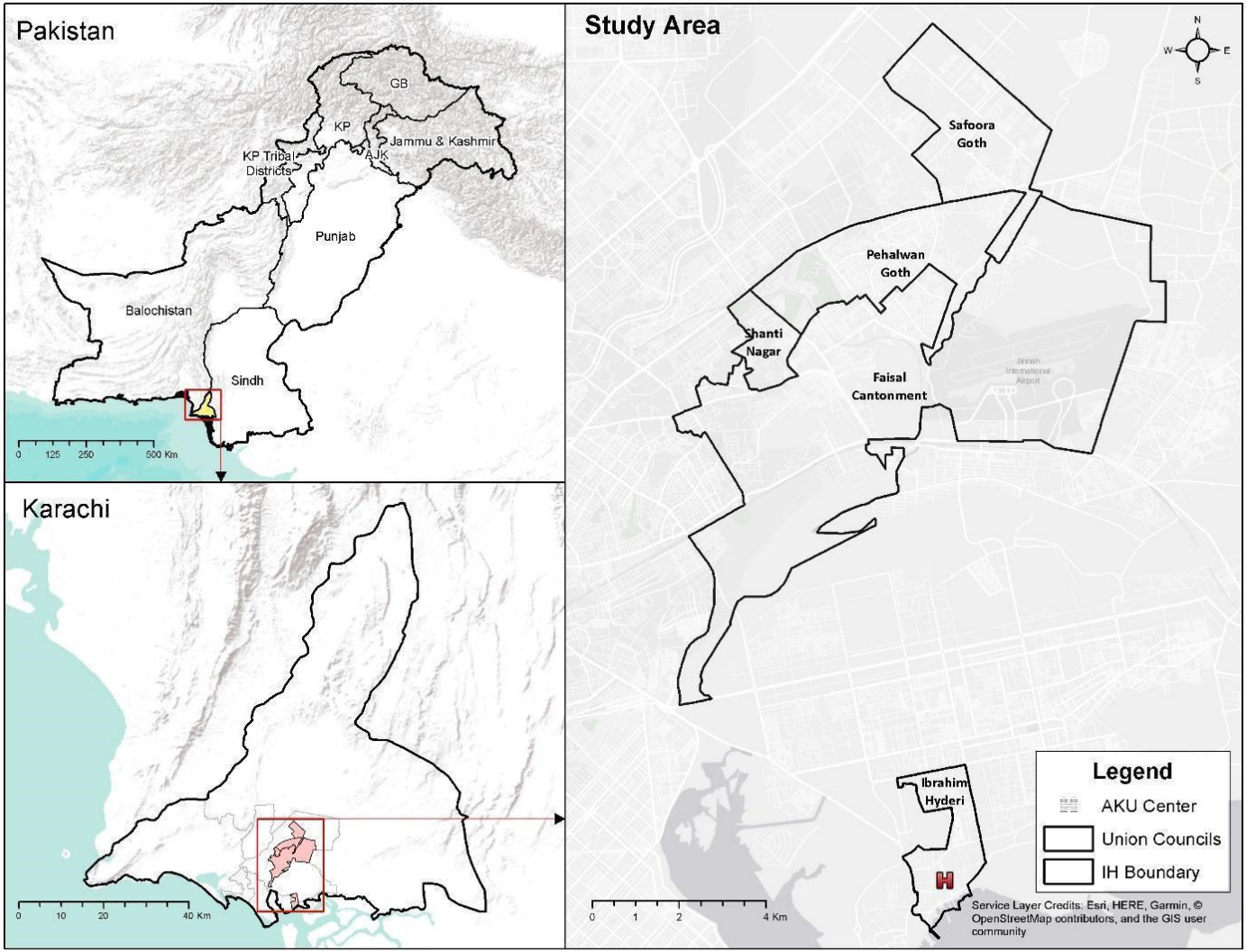
Sought care	9 (1.8%)	9 (1.8%)	3 (0.6%)	4 (0.8%)
Hospitalization	1 (0.2%)	5 (1.0%)	1 (0.2%)	0 (0.0%)
History of Travel	2 (0.4%)	11 (2.2%)	8 (1.6%)	1 (0.2%)
Contact with suspected or confirmed COVID-19 case	1 (0.2%)	1 (0.2%)	0 (0.0%)	1 (0.2%)

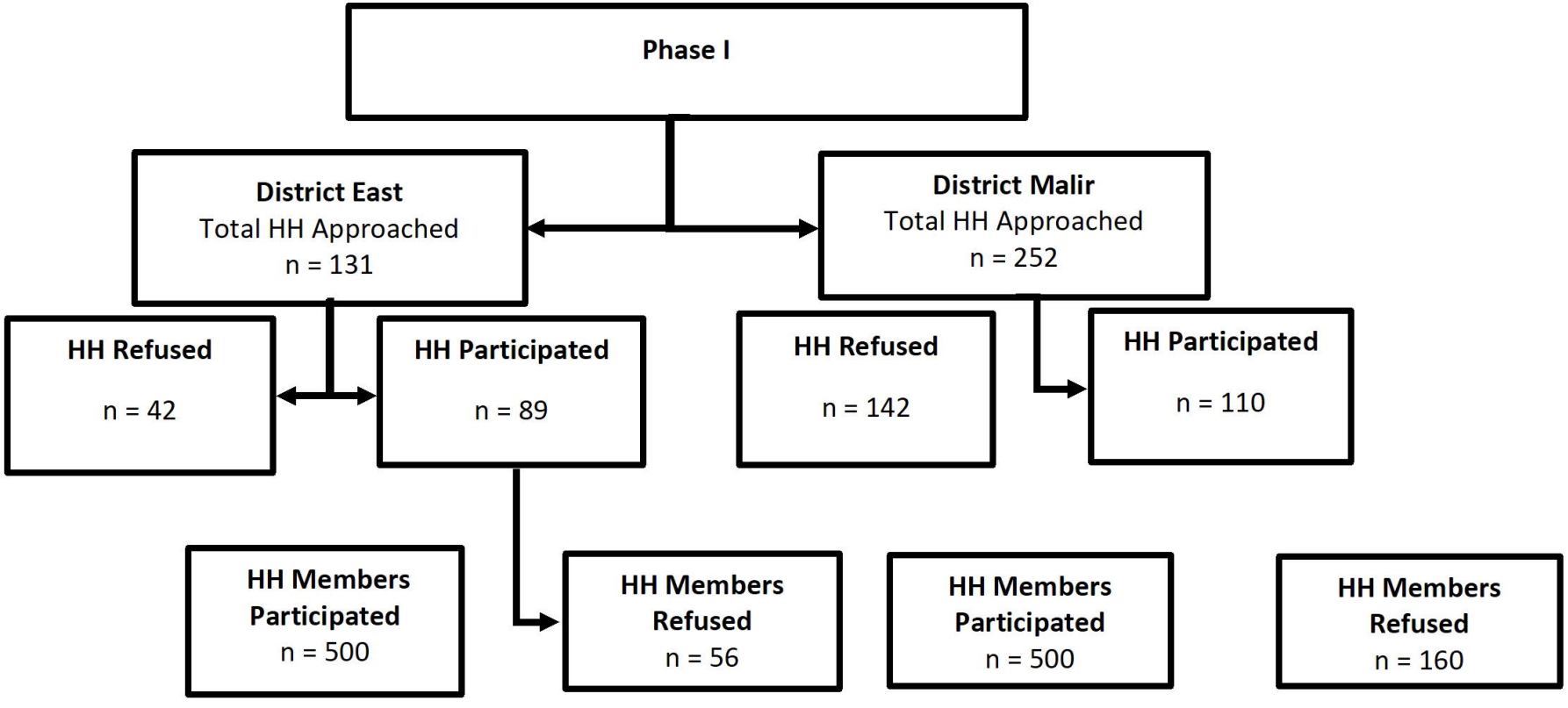
Figure 1. Study sites.

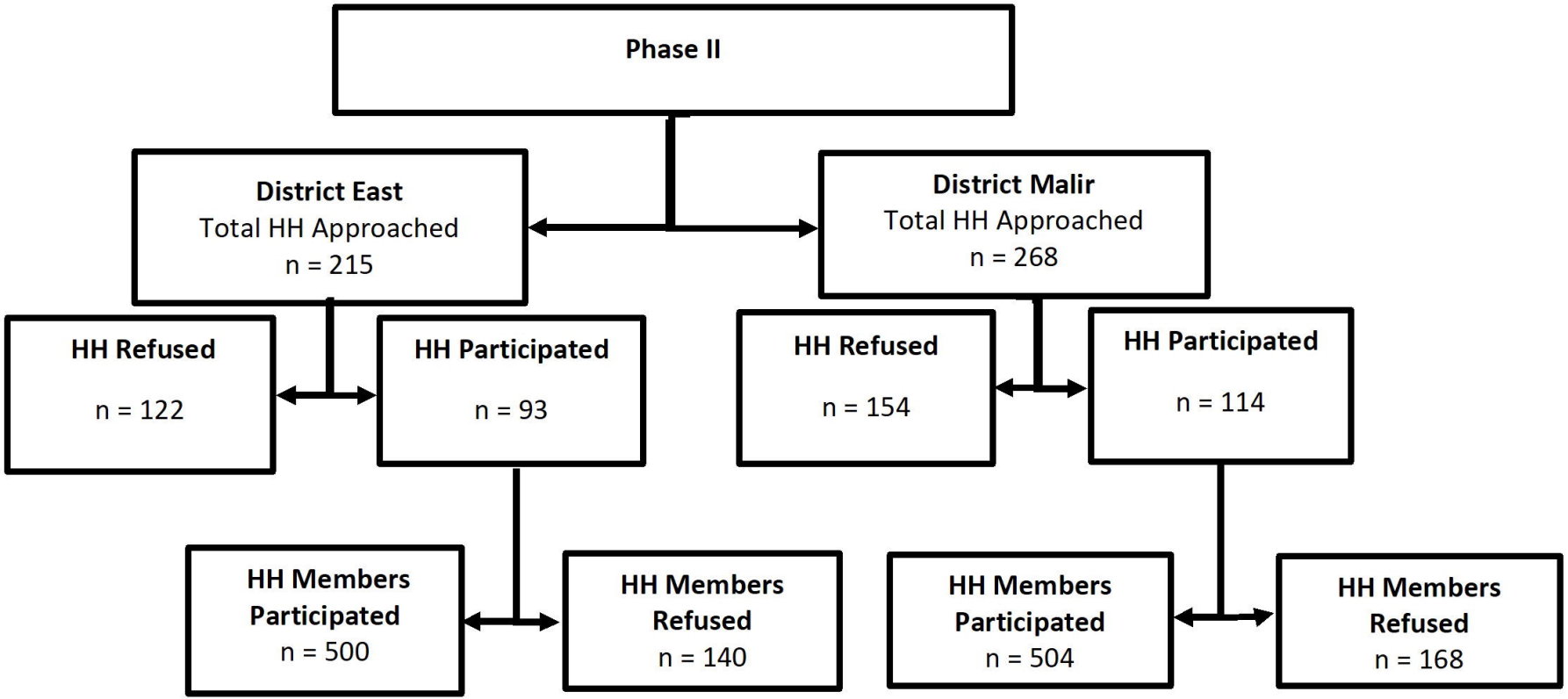
Figure 2 Panel A. Flow chart of participants in the phase 1 of study

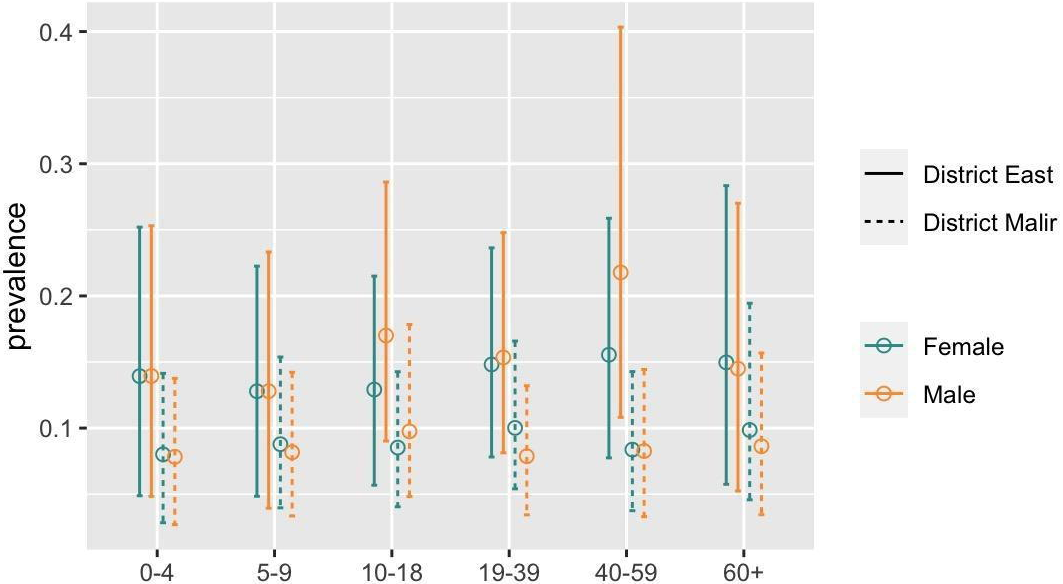
Figure 2 Panel B. Flow chart of participants in the phase two of study

Figure 3: Prevalence estimates by age and gender based on the data from the second survey. The circle represents the posterior mean seroprevalence and the bar represents the 95% equal-tailed credible interval. Posterior mean estimates for District East are consistently greater than those for District Malir, although there is significant overlap in the credible intervals for all age and gender subpopulations. No consistent patterns exist between the prevalence rates for males and females.


Pakistan
GB
KP
Jammu & Kashmir
KP Tribal Districts
AJK
Punjab
Balochistan
Sindh
0 125 250 500 Km
Karachi
0 10 20 40 Km
Study Area
N
W
E
S
Safoora Goth
Pehalwan Goth
Shanti Nagar
Faisal Cantonment
Ibrahim Hyderi
0 1 2 4 Km
Legend
AKU Center
Union Councils
IH Boundary
Service Layer Credits: Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS user community


Phase I
District East
Total HH Approached
n = 131
District Malir
Total HH Approached
n = 252
HH Refused
n = 42
HH Participated
n = 89
HH Refused
n = 142
HH Participated
n = 110
HH Members Participated
n = 500
HH Members Refused
n = 56
HH Members Participated
n = 500
HH Members Refused
n = 160


Phase II
District East
Total HH Approached
n = 215
District Malir
Total HH Approached
n = 268
HH Refused
n = 122
HH Participated
n = 93
HH Refused
n = 154
HH Participated
n = 114
HH Members Participated
n = 500
HH Members Refused
n = 140
HH Members Participated
n = 504
HH Members Refused
n = 168


prevalence
0.4
0.3
0.2
0.1
0-4
5-9
10-18
19-39
40-59
60+
District East
District Malir
Female
Male