Title Page

Insights from a Pan India Sero-Epidemiological survey (Phenome-India Cohort) for SARS-CoV2

Salwa Naushin, MSc[1,32#]; Viren Sardana, M.Tech[1,32#]; Rajat Ujjainiya, MSc[1,32]; Nitin Bhatheja, BTech[1]; Rintu Kutum, MSc[1,32]; Akash Kumar Bhaskar, MSc[1,32]; Shalini Pradhan, MSc[1]; Satyartha Prakash, MTech[1]; Raju Khan, PhD[2,32]; Birendra Singh Rawat, PhD[3,32]; Karthik Bharadwaj Tallapaka, MD[4]; Mahesh Anumalla, MSc[4]; Giriraj Ratan Chandak, PhD[4,32] ; Amit Lahiri, PhD[5,32]; Susanta Kar, PhD[5,32]; Shrikant Ramesh Mulay,PhD[5,32]; Madhav Nilakanth Mugale, PhD[5,32]; Mrigank Srivastava, PhD[5,32]; Shaziya Khan, MSc[5,32]; Anjali Srivastava, MSc[5,32]; Bhawna Tomar, MSc[5,32]; Murugan Veerapandian, PhD[6,32]; Ganesh Venkatachalam, PhD[6,32]; Selvamani Raja Vijayakumar, MBBS[6]; Ajay Agarwal, PhD[7,32]; Dinesh Gupta, MD[7]; Prakash M Halami, PhD[8,32]; Muthukumar Serva Peddha, PhD[8,32]; Gopinath M Sundaram, PhD[8,32]; Ravindra P Veeranna, PhD[8,32]; Anirban Pal, PhD[9,32]; Vinay Kumar Agarwal, MD[9]; Anil Ku Maurya, PhD[9]; Ran Vijay Kumar Singh, PhD[10,32]; Ashok Kumar Raman, MBBS[10]; Suresh Kumar Anandasadagopan, PhD[11,32]; Parimala Karuppanan, MBBS[11]; Subramanian Venkatesan, PhD[11,32]; Harish Kumar Sardana, PhD[12]; Anamika Kothari, MBBS[12]; Rishabh Jain, MTech[12,32]; Anupma Thakur, MTech[12,32]; Devendra Singh Parihar, MTech[12,32]; Anas Saifi, MTech[12,32]; Jasleen Kaur, MTech[12,32]; Virendra Kumar, MCA[12]; Avinash Mishra, PhD[13,32]; Iranna Gogeri, MTech[14,32], Geethavani Rayasam, PhD[15,32]; Praveen Singh, MSc[1,32]; Rahul Chakraborty, MSc[1,32]; Gaura Chaturvedi, MSc[1,32]; Pinreddy Karunakar, MSc[1,32]; Rohit Yadav, MSc[1,32]; Sunanda Singhmar, MSc[1,32]; Dayanidhi Singh, MSc[1,32]; Sharmistha Sarkar, MSc[1,32]; Purbasha Bhattacharya, MSc[1,32]; Sundaram Acharya, MSc[1,32]; Vandana Singh, MSc[1,32]; Shweta Verma, MSc[1,32]; Drishti Soni, MSc[1,32]; Surabhi Seth, MSc[1,32]; Shakshi Vashisht, MSc[1,32]; Sarita Thakran, MSc[1,32]; Firdaus Fatima, MSc[1,32]; Akash Pratap Singh, MSc[1,32]; Akanksha Sharma, MSc[1,32]; Babita Sharma, MSc[1,32]; Manikandan Subramanian, PhD[1]; Yogendra Padwad, PhD[16,32]; Vipin Hallan, PhD[16,32]; Vikram Patial, PhD[16,32]; Damanpreet Singh, PhD[16,32]; Narendra Vijay Tirpude, MVSc[16,32]; Partha Chakrabarti, MD[17,32]; Sujay Krishna Maity, MS(Pharm)[17]; Dipyaman Ganguly, PhD[17,32]; Jit Sarkar, MBBS[17,32]; Sistla Ramakrishna, PhD[18,32]; Balthu Narender Kumar, MBBS[18]; A Kiran Kumar, PhD[18]; Sumit G. Gandhi, PhD[19,32]; Piyush Singh Jamwal, MSc[19]; Rekha Chouhan, MSc[19]; Vijay Lakshmi Jamwal, MSc[19,32]; Nitika Kapoor, MSc[19,32]; Debashish Ghosh, PhD[20,32]; Ghanshyam Thakkar, MBBS[20]; Umakanta Subudhi, PhD[21,32]; Pradip Sen, PhD[22,32]; Saumya Ray Chaudhury, PhD[22,32]; Rashmi Kumar, PhD[22,32]; Pawan Gupta, PhD[22,32]; Amit Tuli, PhD[22,32]; Deepak Sharma, PhD[22,32]; Rajesh P. Ringe, PhD[22]; Amarnarayan D, MBBS[23], Mahesh Kulkarni, PhD[24,32]; Dhanasekaran Shanmugam, PhD[24,32]; Mahesh S Dharne, PhD[24,32]; Syed G. Dastager, PhD[24,32]; Rakesh Joshi, PhD[24,32]; Amita P. Patil, MBBS[24]; Sachin N. Mahajan, MBBS[24]; Abu Junaid Khan, MSc[24,32]; Vasudev Wagh, MSc[24,32]; Rakeshkumar Yadav, MSc[24,32]; Ajinkya Khilari, MSc[24,32]; Mayuri Bhadange, MSc[24,32]; Arvindkumar H. Chaurasiya, MSc[24,32]; Shabda E Kulsange, MSc[24,32]; Krishna Khairnar, PhD[25,32]; Shilpa Paranjape, MD[25]; Jatin Kalita, PhD[26]; G. Narahari Sastry, PhD[26]; Tridip Phukan, PhD[26]; Prasenjit Manna, PhD[26]; Wahengbam Romi, PhD[26]; Pankaj Bharali, PhD[26]; Dibyajyoti Ozah, PhD[26]; Ravi Kumar Sahu, MSc[26,32]; Elapavalooru V.S.S.K. Babu, PhD[27,32]; Rajeev Sukumaran, PhD[28,32]; Aiswarya R Nair, MD[28] Prajeesh Kooloth-Valappil, MSc[28,32]; Anoop Puthiyamadam, MSc[28,32]; Adarsh Velayudhanpillai, MSc[28]; Kalpana Chodankar, MBBS[29,32]; Samir Damare,PhD[29,32]; Yennapu Madhavi, PhD[30,32]; Ved Varun Aggarwal, PhD[31,32]; Sumit Dahiya, MSc[31,32]; Anurag Agrawal, PhD[1,32]; Debasis Dash, PhD[1,32]*; Shantanu Sengupta, PhD[1,32]*

#Co-first authors with equal contribution

**Corresponding authors*

Affiliation

1. *CSIR-Institute of Genomics and Integrative Biology, New Delhi-110025, India*
2. *CSIR-Advanced Materials and Processes Research Institute, Bhopal- 462026, Madhya Pradesh, India*
3. *CSIR-Central Building Research Institute, Roorkee- 247667, Uttarakhand, India*
4. *CSIR-Centre for Cellular Molecular Biology, Hyderabad-500007, Telangana, India*
5. *CSIR-Central Drug Research Institute, Lucknow-226031, Uttar Pradesh, India*
6. *CSIR-Central Electrochemical Research Institute, Karaikudi- 630003, Tamil Nadu, India*
7. *CSIR-Central Electronics Engineering Research Institute, Pilani- 323031, Rajasthan, India*
8. *CSIR-Central Food Technological Research Institute, Mysore-570005, Karnataka, India*
9. *CSIR-Central Institute of Medicinal Aromatic Plants, Lucknow- 226015, Uttar Pradesh, India*
10. *CSIR-Central Institute of Mining and Fuel Research, Dhanbad-826001, Jharkhand, India*
11. *CSIR-Central Leather Research Institute, Chennai-600020, Tamil Nadu, India*
12. *CSIR-Central Scientific Instruments Organization, Chandigarh-160030, India*
13. *CSIR-Central Salt Marine Chemicals Research Institute, Bhavnagar-364002, Gujarat, India*
14. *CSIR Fourth Paradigm Institute, Bengaluru-560037, India*
15. *CSIR- Headquarters, Rafi Marg, New Delhi-110011, India*
16. *CSIR-Institute of Himalayan Bioresource Technology, Palampur-176061*, *Himachal Pradesh, India*
17. *CSIR-Indian Institute of Chemical Biology, Kolkata-700032, West Bengal, India*
18. *CSIR-Indian Institute of Chemical Technology, Hyderabad-500007, Telangana, India*
19. *CSIR-Indian Institute of Integrative Medicine, Jammu-180001, Jammu and Kashmir, India*
20. *CSIR-Indian Institute of Petroleum, Dehradun-248005, Uttarakhand, India*
21. *CSIR-Institute of Minerals and Materials Technology, Bhubaneswar-751013, Odisha, India*
22. *CSIR-Institute of Microbial Technology, Chandigarh-160036, India*
23. *CSIR- National Aerospace Laboratories, Bengaluru-560017, Karnataka, India*
24. *CSIR-National Chemical Laboratory, Pune-411008, Maharashtra, India*
25. *CSIR-National Environmental Engineering Research Institute, Nagpur-440020, Maharashtra, India*
26. *CSIR-North - East Institute of Science and Technology, Jorhat-785006, Assam, India*
27. *CSIR-National Geophysical Research Institute, Hyderabad- 500007, Telangana, India*
28. *CSIR-National Institute for Interdisciplinary Science and Technology, Thiruvananthapuram- 695019, Kerala, India*
29. *CSIR-National Institute of Oceanography, Goa-403004, India*
30. *CSIR-National Institute of Science, Technology and Development Studies, New Delhi-110012, India*
31. *CSIR-National Physical Laboratory, New Delhi-110012, India*
32. *Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201002, Uttar Pradesh, India*

Contact Information of Corresponding Authors:

1. Dr. Shantanu Sengupta
   Email: shantanus@igib.res.in
   Contact No:+919891670220
   Address: CSIR-IGIB, Mathura Road,110025, Delhi, India

2. Dr. Debasis Dash

Email: ddash@igib.in
Contact No:+919810355975
Address: CSIR-IGIB, Mathura Road,110025, Delhi, India

Abstract

Background-India has been amongst the most affected nations during the SARS-CoV2 pandemic, with sparse data on country-wide spread of asymptomatic infections and antibody persistence. This longitudinal cohort study was aimed to evaluate SARS-CoV2 sero-positivity rate as a marker of infection and evaluate temporal persistence of antibodies with neutralization capability and to infer possible risk factors for infection.

Methods- Council of Scientific and Industrial Research, India (CSIR) with its more than 40 laboratories and centers in urban and semi-urban settings spread across the country piloted the pan country surveillance. 10427 adult individuals working in CSIR laboratories and their family members based on voluntary participation were assessed for antibody presence and stability was analyzed over 6 months utilizing qualitative Elecsys SARS CoV2 specific antibody kit and GENScript cPass SARS-CoV2 Neutralization Antibody Detection Kit. Along with demographic information, possible risk factors were evaluated through self to be filled online forms with data acquired on blood group type, occupation type, addiction and habits including smoking and alcohol, diet preferences, medical history and transport type utilized. Symptom history and information on possible contact and compliance with COVID 19 universal precautions was also obtained.

Findings–1058 individuals (10·14%) had antibodies against SARS-CoV2. A follow-up on 346 sero-positive individuals after three months revealed stable to higher antibody levels against SARS-CoV2 but declining plasma activity for neutralizing SARS-CoV2 receptor binding domain and ACE2 interaction. A repeat sampling of 35 individuals, at six months, revealed declining antibody levels while the neutralizing activity remained stable compared to three months. Majority of sero-positive individuals (75%) did not recall even one of nine symptoms since March 2020. Fever was the most common symptom with one-fourth reporting loss of taste or smell. Significantly associated risks for sero-positivity (Odds Ratio, 95% CI, p value) were observed with usage of public transport (1·79, 1·43 – 2·24, 2·81561E-06), occupational responsibilities such as security, housekeeping personnel etc. (2·23, 1·92 – 2·59, 6·43969E-26), non-smokers (1·52, 1·16 – 1·99, 0·02) and non-vegetarianism (1·67, 1·41 – 1·99, 3·03821E-08). An iterative regression analysis was confirmatory and led to only modest changes to estimates. Predilections for sero-positivity was noted with specific ABO blood groups -O was associated with a lower risk.

Interpretation – In a first-of-its-kind study from India, we report the sero-positivity in a country-wide cohort and identify variable susceptible associations for contacting infection. Serology and Neutralizing Antibody response provides much-sought-for general insights on the immune response to the virus among Indians and will be an important resource for designing vaccination strategies.

Funding- Council of Scientific and Industrial Research, India (CSIR)

Introduction

The World Health Organization declared SARS-CoV2 infection as a pandemic on March 11 2020.[1] Within two weeks, India announced a lockdown strategy that severely influenced the growth of the pandemic which was initially very focal in the large cities, gathering pace and spreading to smaller cities and towns as the nation unlocked for societal and economic considerations.

Early literature pointed towards asymptomatic transmission of SARS-CoV2 and raised the need for extended testing.[2,3] While, RT-PCR was an undisputed choice for establishing a positive infection, sero-surveillance revealed that many more were probably getting infected without manifesting symptoms.[4,5] Initial estimates of asymptomatic infection rate from the West were around 40-45%.[5]

In India, the first case of Covid was reported on January 30, 2020.[6] Serological surveys have confirmed that spread beyond the Indian megacities was minimal in early May-June, with less than 1% sero-positivity outside the designated containment zones, suggesting that the lockdown had been effective in limiting the spread.[7] This was not without human and economic cost. By the end of June, migrant workforces caught in the cities during the lockdown were sent to their rural homes, which might have been a contributing factor in the subsequent rapid, multi-focal rise in cases at the end of August. Existing studies from India, have been either limited to specific focal geographies or localities.[8-10] A pan India study has not been reported during this time and thus the present study was designed to assess spread of infection until September 2020, with three-monthly reassessment of anti SARS-CoV2 antibody response along with neutralization antibody trend in those found to be sero-positive. Council of Scientific and Industrial Research (CSIR) in its more than forty constituent laboratories and centers spread all over the country, representing a wide range of ethnicities, geo-social habitats and occupational exposures, initiated a longitudinal cohort study (Phenome-India Cohort) to estimate the burden of COVID-19 and to assess antibody stability. Here, we report results from phase 1 of this study; provide sentinel data about the pattern of infection spread and different characteristics of SARS-CoV2 infection across India.

We find that large cities have high sero-positivity rates, consistent with local surveys, but the majority of India continues to have low to moderate sero-positivity. Most sero-positive individuals had occupational exposure risk and significant symptoms were recalled by only 25%, with 75% recalling no symptoms at all.

Methods

Study Design, Sampling and Data Collection

The longitudinal cohort study was approved by Institutional ethics committee of CSIR-IGIB. 10427 adult individuals working in CSIR laboratories and their family members enrolled for the study based on voluntary participation. Informed consent was obtained from all the participants and the samples were collected maintaining all

recommended precautions. Blood samples (6 ml) were collected in EDTA vials from each participant. Plasma was separated by centrifugation at 1800 g for 15 minutes. Separated plasma were stored at -80°C until they were tested for antibodies. Most of the labs/centers transported the samples from their respective centers and labs to CSIR-IGIB, New Delhi in dry ice for sero-testing. All the participants were requested to fill an online questionnaire, which included information on date of birth, gender, blood group, type of occupation, history of Diabetes, Hypertension, Cardiovascular Disease, Liver and Kidney Disease, diet preferences, mode of travel, contact history, and hospital visits. These forms were then downloaded in MS-Excel data format and merged with registration forms filled at the time of sample collection based on unique ID's.

Procedure

An Electro-chemiluminescence Immunoassay (ECLIA)- Elecsys Anti-SARS-CoV2 kit (Roche Diagnostics) to qualitatively detect antibodies against SARS-CoV2 was used; and the assay was performed using a Cobas e411 (Roche Diagnostics) as per the manufacturer's protocol. This approved assay measures antibodies to SARS-CoV2 nucleocapsid antigen and is considered a method of choice when a single test is to be deployed with high sensitivity and specificity.[11] A COI $\geq$ 1 was considered sero-positive. To check if the antibodies had neutralization capability, we tested samples with COI $\geq$1 (as determined by ECLIA mentioned above), for neutralizing antibody (NAB) response directed against the spike protein using GENScript cPass™ SARS-CoV2 Neutralization Antibody Detection Kit (Genscript, USA), according to the manufacturer's protocol. This is a blocking ELISA used for qualitative detection of total neutralizing antibodies against SARS-CoV2 virus in plasma. A value of 20% or above was considered to have neutralizing ability.

Data and Statistical Analysis

Registration form data was utilized to assess the sero-positivity city wise. All the variables were assessed for responses type and blank fields i.e. responses which were not provided by the participants of the survey. Based on multiple response types for each variable, categories were made to assign the response to either of the categories. For visualization; ggpubr (v0.4.0), ggrepel (v0.8.2), ggplot2 (v3.3.2) packages were used. No data imputation was carried out. Chi-square test was performed to evaluate variables which had significant association with outcome of being tested positive ($p<0{\cdot}05$) along with Odds Ratio (OR) with 95% Confidence Interval (CI). An adjusted p value was obtained through Bonferroni Correction method for multiple comparison testing. Following the chi-square test an iterative logistic regression was carried out on balanced dataset. Variance Inflation Factor (VIF) was separately evaluated to assess multi-collinearity. Statistical analysis and model development was carried out with visualization in R programing environment version 3.6.1 and MS-Excel 2016.; faraway (v1.0.7) package was utilized for estimation of VIF.

Role of the funding source

The sponsor of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Variable extent of sero-positivity and asymptomatic infections in the Indian subcontinent

Humoral response against SARS-CoV2 is documented to be against both the nucleocapsid (NC) and the spike (S) proteins of the virus.[12,13] It has been shown that for SARS-CoV2 sero-surveillance studies anti-NC antibody response is more suitable due to its higher prevalence post-infection[13-15]. A humoral response against the SARS-CoV2 S is better suitable for assessing presence of neutralizing antibody response, as the receptor binding domain (RBD) of S protein interacts with angiotensin converting enzyme 2 (ACE2), its cognate receptor on host cell surface.[12,13,16-19]

Thus to achieve a cross-sectional assessment on the sero-prevalence, we measured anti-SARS-CoV2 NC antibody in a total of 10427 individuals from CSIR labs and centres spread over 17 States and two Union Territories with varied population densities (Figure 1 and eTable 1). We found that the overall sero-prevalence was 10·14%. The sero-prevalence within the cohort correlated with the population density, confirming the effect of population density on infection spread dynamics (eFigure 1).

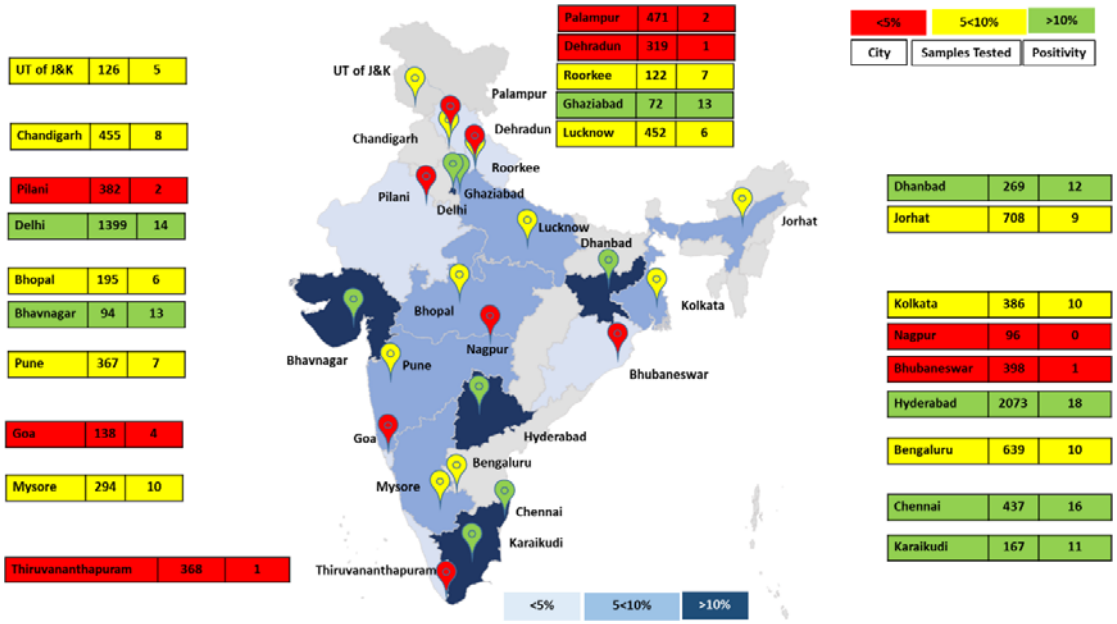

Figure 1: City wise total samples collected and sero-positivity. (For states with multiple cities, city sero-positivity was averaged except for the state of Maharshtra as Nagpur had sero-positivity of zero. India map may not be to scale and is for repsentation purposes only. Sero-positivity is rounded off)

We found 95% of the sero-positive individuals also had significant neutralizing activity.

To determine whether sero-positivity could be a surrogate of future transmission of the virus, we plotted the rate of new cases before and after the date of sero-sampling in a district (data obtained from covid19india.org) versus seropositivity. Higher sero-positivity predicted declining new cases, while the largest increases were seen exclusively in areas of low sero-positivity (Figure 2 and eFigure 2)

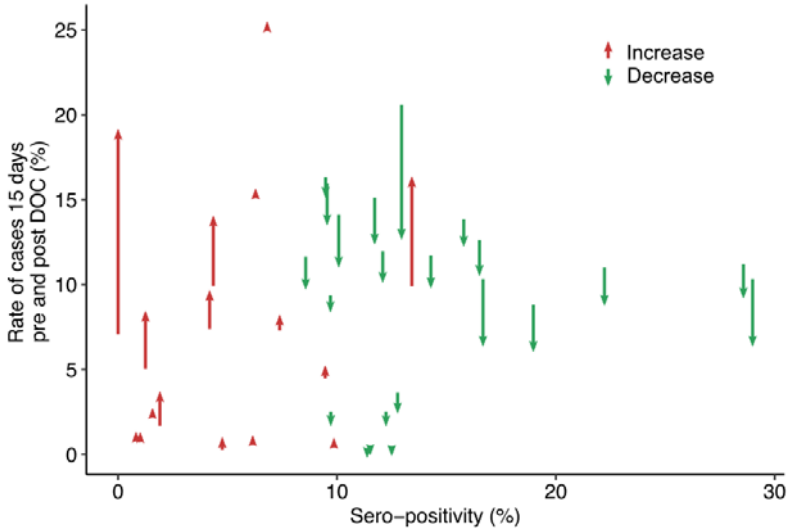

*Figure 2: Rate of Cases 15 days prior and 15 days after DOC are shown as an increase or decrease marked with arrow direction and colour for all labs/centres. Labs/centers with low sero-positivity are mostly showing an increase while it's the opposite for labs/centers with high initial sero-positivity. (Details in eTable 1). *For a few cities, district wise data for number of samples tested was not avaibale and hence state data has been utilzed as a surrogate).*

Parameters associated with susceptibility to infection

As noted earlier, the CSIR-cohort represented a wide range of ethnicity subclasses, geo-social habitats, varied patterns of human interactions and occupational exposures, thus enabling us to assess the relative contributions of those parameters towards susceptibility to infection. Apart from gender and age, distribution of the other variables recorded in CSIR-cohort (prevalence of smoking, diet, physiological parameters like ABO blood group type) were not much different from the national averages.[20,21]

Parameter (Category)	Males			Females			Total		
	No.	OR(95% CI)	P value (Adjusted)	No.	OR(95% CI)	P value (Adjusted)	No.*	OR(95% CI)	P value (Adjusted)
Gender (M/F)	6705	-	-	2573	-	-	9278	1.3 (1.11-1.53)	0.01210436
Occupation (OS/S)	6457	2.05 (1.73-2.42)	1.24E-16	2487	2.75 (1.95-3.89)	2.47E-08	8953	2.23 (1.92-2.59)	6.44E-26
Smoking (No/Yes)	6379	1.62 (1.23-2.14)	0.00582434	2400	-	-	8788	1.52 (1.15-1.99)	0.02573872
Diet type (NV/V)	6345	1.78 (1.45-2.19)	3.21E-07	2382	1.33 (0.97-1.82)	0.510086527	8736	1.67 (1.41-1.99)	3.04E-08

Mode of Transport (Public/Private)	6375	1.91 (1.44-2.55)	7.31E-05	2398	1.83 (1.26-2.69)	0.010667048	8787	1.79 (1.43-2.24)	2.82E-06
Alcohol (Yes/No)	6379	0.81(0.68-0.96)	0.50699433	2400	1.14 (0.67-1.92)	1	8788	0.90 (0.76-1.05)	1
Diabetes (Yes/No)	6382	1.0 (0.75-1.32)	1	2400	1.71 (1.02-2.88)	0.332124018	8791	1.13 (0.88-1.45)	1
HT (Yes/No)	6381	0.87 (0.66-1.15)	1	2400	1.40 (0.86-2.25)	1	8790	0.99 (0.78-1.26)	1
CVD (Yes/No)	6382	1.09 (0.54-2.20)	1	2400	-	-	8791	1.0 (0.51-1.93)	1

*Table 1 Demographics of data available for different variables (M- Males, F-Females, OS- Outsourced Staff, S-Staff, NV-Non-Vegetarian, V-Vegetarian, HT-Hypertension, CVD-Cardiovascular Disease, OR-Odds Ratio, CI-Confidence interval. *Male and female numbers not adding to total as gender field was not available or filled by these participants)*

Gender distribution in our dataset was not the same (72% males and 28% females) as country wide distribution (51·5% males and 48·5% females).[22] Gender data was available for 9278 individuals where we observed a differential sero-prevalence between genders, 10·43% among males, vs 8·20% among females. While, the OR was greater than 1 (OR 1·30 95% CI 1·11 -1·53, p=0·012), the association was not significant on iteration model with balanced dataset.

Out of 861 sero-positive individuals who also provided data on symptomatology, 647 individuals (75·3%) did not recall any of the nine symptoms asked for. Most commonly reported symptoms were those of a mild disease similar to seasonal influenza with fever (~50 %) as the most common reported symptom. Loss of smell or taste was less frequently reported (~25%). Shortness of breath was reported by only 2·6% of the symptomatic people (eFigure 3).

Data for 7496 individuals was available for their Blood Group type. Blood Group (BG) distribution of individuals in our study was similar to national reference based on a recent systematic review (eTable 2).[20] Sero-prevalence was highest for blood group type AB (10·19%) followed by group B (9·94%), group O (7·09%) and the lowest for group A (6·97%). Blood O was observed to be protective, with an odds ratio of 0·76 (95 % CI 0·64 -0·91, p=0·018) vs Non O blood group types, while B appeared to be associated with higher sero-positivity 1·36 (95 % CI 1·15 -1·61, p=0·001). Rh factor was not found to have significant association with sero-positivity (p=0·35).

Interestingly, non-smokers in the cohort recorded a higher sero-prevalence of 10·11%, as opposed to smokers (6·88%). This association was found to be statistically significant in our cohort with an odds ratio calculated for the entire dataset of 1·52 (95 % CI 1·15 – 1·99, p=0·02) and for males at 1·62 (95 % CI: 1·23-2·14, p=0·005). Non-vegetarians had a sero-prevalence of 11%, while sero-positivity among vegetarians was 6·86%. This association was found to be statistically significant in our cohort with an odds ratio calculated for the entire dataset of 1·67 (95 % CI 1·41 – 1·99, p=3·03821E-08) and for males at 1·69 (95 % CI: 1·39-2·06, p= 3·20862E-07). From the responses obtained from the online forms, occupation was clubbed into two broad Categories-Staff and Outsourced Staff. The staff group (S) consisted of faculty, students, office staff and family members, while the outsourced staff (OS) comprised of individuals involved in security services, housekeeping,

horticulture etc. Data of 8953 individuals was available for their occupation type. Staff had a positivity of 7·98 percent, while positivity for Outsourced Staff was 16·22%. Outsourced Staff (OS) were thus at higher risk of infection, as expected from their occupation type, with an odds ratio of 1·23 (95 % CI 1·92 – 2·59, p=6·43969E-26) for the entire dataset and also when seen separately for male (2·05; 95 % CI:1·73- 2·42, p=1·2408E-16) and female subgroups (2·75; 95 % CI:1·95 – 3·89, p=2·46898E-08) were analyzed separately. Sero-positivity was 15·59 percent amongst public transport users, but 9·33% amongst private transport users with an OR of 1·79 (95 % CI 1·43 – 2·24, p=2·81561E-06) for combined, 1·91 (95 % CI: 1·44- 2·55, p=7·31089E-05) for males and 1·83 (95 % CI:1·26–2·67, p=0·010667048) for females. The results for odds ratio and iteratively run regression model for entire dataset and separately for male and female gender are shown in Figure 3.

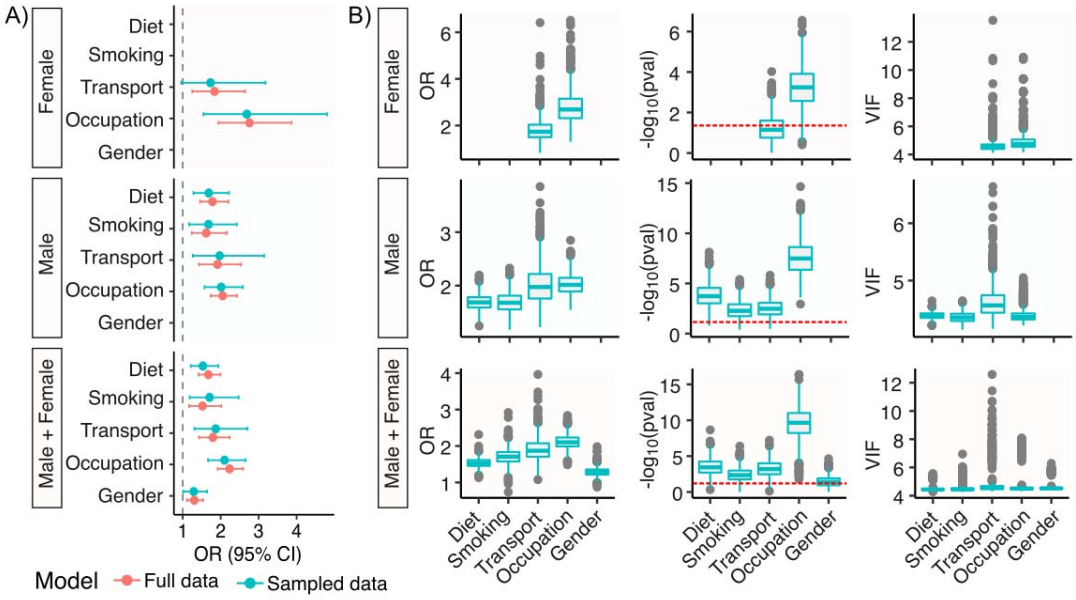

Figure 3 A) Odds Ratio of full dataset and sampled data set obtained from model. B) Odds, p value and VIF for sampled dataset with iterations on regression model. (For Diet: Non-Vegetarian against Vegetarian, For Smoking: Non-Smoking against Smoking, For Transportation: Public against Private, For Occupation: Outsourced Staff against Staff, For Gender: Male against Female). 1300, 1100 and 1000 iterations were run for Female, Male +Female and Males respectively.

Presence or absence of Diabetes, Hypertension and Cardiovascular Disease were not found to be significantly associated with sero-outcomes (p>0·05, Table 1).

Kinetics of humoral response to SARS-CoV2

The sero-survey was initiated in June 2020 at CSIR-IGIB and was conducted in other Institutes between July and September. Of the 346 individuals whose samples were available at three months, anti-nucleocapsid antibody levels remained mostly steady to higher, with only five (1·4%) becoming sero-negative (Figure 4A). In contrast, Neutralizing antibody values decreased 197 (56·9 percent) individuals after three months, with levels dropping below threshold (<20 percent) in 11 individuals (Figure 4B). However, in the 35 individuals tested at six months; the anti-

nucleocapsid antibody levels declined while neutralizing antibody levels were mostly unaltered at six months compared to the values at three months (Figure 4C and D). This indicates that although there is an overall decline in both nucleocapsid antibody and neutralizing antibody levels at six months, in most cases they still are above the threshold value and could potentially confer protection from infection even at six months.

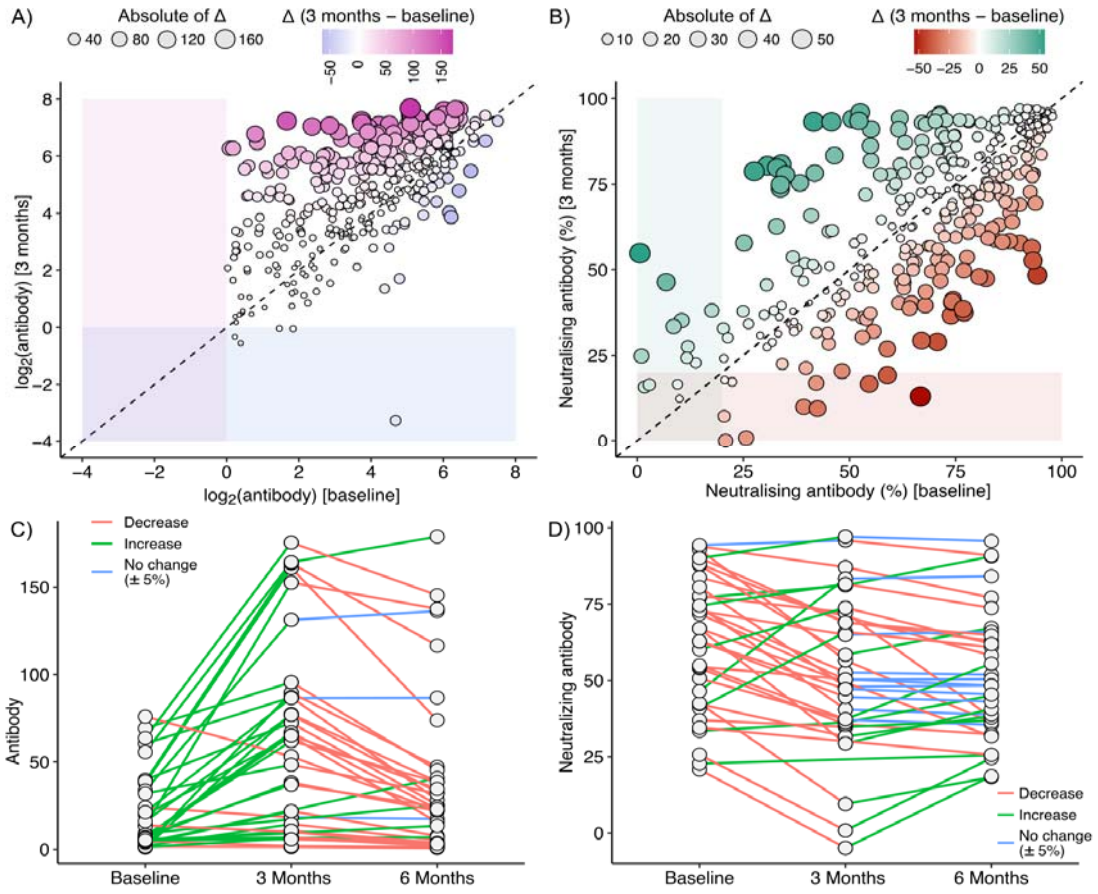

Figure 4: Antibody levels (A) and Neutralizing Antibody percent (B) level at baseline (x-axis) and after 3 months (y-axis). Antibody levels (C) and Neutralizing Antibody (D) percent (y-axis) at baseline, 3 months and after 6 months (x-axis) for 35 individuals depicting the trend.

Discussion

Sero-surveillance studies involving longitudinal cohorts are helpful in determining infection dynamics, antibody response dynamics, progression towards herd immunity and likelihood of major outbreaks in a population. Such information is invaluable towards a well-informed vaccination program, especially in a setting where shortages are expected. Although India has been among one of the most affected nations in this SARS-CoV2 pandemic, such data is not yet available. The present study, which recruited individuals from 24 cities, provides an important and timely snapshot across multiple geographies. It confirms that by September 2020, there was wide but

uneven spread of the pandemic, with the epidemic subsequently growing more rapidly in regions with low sero-positivity in September. Overall sero-prevalence in the CSIR-cohort is different between general staff containing a high representation of people working from home and using private transport, and outsourced staff performing on-site duties like sanitation and security and using public transport. Higher sero-prevalence among outsourced staff and public transport users in our cohort, is more likely to be representative of general population of cities and towns that are part of the cohort. Millions of migrant workers trapped in Indian cities during the lockdown returned to villages in June in packed public transport. Outsourced workers, with highest seropositivity, reflect the high infection rate of this subgroup and it can be reasonably assumed that the pandemic had already reached rural India by September 2020.

Apart from the sero-positivity rate, our data also reveals important associations between demographic, physiological, lifestyle-related and occupational attributes with susceptibility to infection. The workforce in our cohort comprised of adult population and no major difference was observed in sero-prevalence amongst different age groups. Males were found to be more susceptible, in agreement with other published reports[23]. However, there were fewer females in our study and many of the occupational responsibilities with higher chances of exposure, like that of security personnel, were skewed towards males. On iteratively ran regression models we found gender was not a predictor for sero-positivity.

ABO blood group type has been shown to be associated with SARS-CoV2 infection, but the results are variable in different studies. Most studies found that O group is associated with lower risk of infection or severity and Blood Group A was reported to be high risk in some studies[24-27]. In a meta-analysis authored by Golinelli et al[28], they observed positive association with A blood group, while blood group O was to be associated with lesser positivity using a random effects model. Another study from India observed blood group O to be associated with less mortality while blood group B with higher mortality when they analyzed the national data available[29]. A complex molecular interaction is said to play a significant role, and the molecular pathways for the same need to be elucidated for the effects observed specially with protective effect of blood group O, which was also confirmed by our study.

In regard to diet preferences, while an association was observed overall and in males, it needs to be corroborated with further elaborative studies. It has been proposed that a fiber-rich diet may play an important role in COVID-19 through anti-inflammatory properties by modification of gut microbiota.[30] Vegetarian diet is known to have high fiber content and protective effect through microbial alteration, but an implied effect on immune-biology and lung biology is not yet elucidated.[30,31] A recent review highlighted the role of trace elements, nutraceuticals and probiotics in COVID-19.[32] These, through their immune-modulatory property exert an anti-viral effect. However, these observations should not advocate the usage or restriction of any diet type.

Our finding that smokers are less likely to be sero-positive is the first report from general population and part of growing evidence that despite COVID-19 being a respiratory disease, smoking may be protective. Two studies from France and similar reports from Italy, New York and China reported lower infection rate among

smokers.[33-37] While it was said the ACE 2 expression is higher and favorable to virus entry in smokers, increased mucous production through goblet cells may be acting as a first line of defense.[38,39] Effect of increased nicotine receptor expression was also questioned.[38,40] Hence, there is a need for focused mechanistic studies to understand the effect of smoking and nicotine on SARS-CoV2 infection. Smoking is known to be severely detrimental to health and associated with multiple diseases and this observation should not be taken to be an endorsement, especially given that the association is not proved to be causal.

Serial follow up of antibody response provides important insights and there are limited studies published as of date in regard to it.[41-43], specially from the country. A longitudinal design of the study enabled us to follow up sero-positive individuals every three months. While general antibodies to NC antigen are well maintained over three months, neutralizing antibodies begin to decline. However, after six months, there is a decline in antibodies to NC antigen while neutralizing antibody levels mostly remain stable after the initial decline. The persistence of antibodies is likely to confer protection against reinfection with the same strain at least till six months. We had 11 known cases of RT-PCR confirmed infection that were antibody negative. We speculate that SARS-CoV2 NC antigen, being similar to other corona viruses, incites a stronger and long-lasting antibody production due to prior reactive cells, while, the RBD domain of the spike protein is a necessary target for the neutralization assay and possibly, being more unique, incites a weaker response. A longitudinal cohort enables to monitor both antibody and neutralization antibody response in temporal domain and thus enable an effective strategy for vaccination planning. Herd immunity does not seem to be in sight and an effective vaccine may be necessary to control this pandemic.

Conclusion

CSIR, India, piloted an assessment of sero-positivity against SARS-CoV2 in over 40 labs and centers spread across the country. The aggregate seropositivity of 10·14% suggested that more than a hundred million Indians were infected by September 2020, primarily in larger cities but rapidly spreading all over the nation. Use of private transport, lower-exposure occupations, smoking, vegetarianism and A or O blood groups appeared to be protective, using seropositivity as a surrogate for infection. Antibody levels were mostly stable at three months, but observed to start declining by six months, although the levels were still above the detection threshold.

Author Contribution

SSG, DD and AA conceptualized the idea and coordinated the research activity. SN, SSG, RajatU, AKB and SPradhan did all the sample analysis, storage and neutralization antibody experiments. VS, SSG, DD, NB, SN, RK, SPrakash, and RajatU contributed in Data Acquisition, Data Pre-Processing, Data Analysis. VS, RK, DD, SSG, and NB did statistical analysis and model development. RK, VS, SSG, DD and VKumar contributed in visualization and dashboard for analysis. RKhan, BSR, TKB, MA, GRC, AL, SKar, SRM, MNM, MS, SK, AS, BT, MV, GV, SRV, AjayA, DGupta, PMH, SPM, GM, PVR, APal, VKAgarwal, AKMaurya, RVKS, AKRaman, SRA, PK, SV, HKS, AKothari, AThakur, DSParihar, ASaifi, RJain, JKaur, AMishra,

AmarD, IG, ASingh, PSingh, RC, GC, PKK, RYadav, DSingh, SS, SSarkar, PB, SA, VSingh, SVerma, DSoni, SSeth, SVashisht, ST, FF, APS, ASharma, BS, YP, VHallan, VPatial, DamanS, NVT, PC, SKM , DGanguly, JSarkar, SR, BNKumar, AKKumar, SGG, PSJ, RChouhan, VLJ, NK, DG, GT, US, PS, SRC, RKumar, PG, AT, DSharma, RPRinge, MK, DS, MDharne, SGD, RJoshi, APP, SNM, AJK, VW, RY, AK, MB, AHC, SEK, KK, SP, JK, GNS, TP, PM, WR, PBharali, DO, RKSahu, EB, RKS, AN, PKV, AP, AVP, KC, SD, YM, VVA, SDahiya, GR contributed in lab/ centre co-ordination, enrolment of volunteers, sample collection and logistics. VS, SSG, DD and AA wrote the first draft. SSG, VS, DD, DGanguly, GRC, Manikandan and AA critically revised the manuscript. All the authors agreed and contributed to the final draft of the manuscript.

Declaration of Interest

We declare no conflict of interest.

Data Sharing Statement:

De-identified categorized data of the participants can be provided on request basis after approval of ethics committee of the institute/organisation to which the requesting person belongs, stating the potential need and usage and signing of an agreement with the corresponding authors on behalf of CSIR-IGIB. The request could be made to the corresponding authors after 90 days of publication.

Acknowledgment

SSG would like to acknowledge CSIR grant MLP 2007 for this work. SN, AKB and RajatU would like to thank CSIR for their fellowships. SPrakash would like to acknowledge CSIR grant MLP-2002(CSIR-IGIB) for the fellowship. NB would like to address CSIR grant GAP-0192(CSIR-IGIB) for this work. We would also like to thank Directors of all the CSIR Institutes for facilitating the study. The authors also thank; Pushpesh Ranjan, Jitendra, Neeraj Kumar, Abhijeet and Rajkumar from AMPRI. V.Santosh Kumar from CCMB. Dr. Chandra Prakash Pandey from CDRI. Vipul Sharma, Akansh Agarwal, Hansraj Choudhary, Vijay Chatterjee, Narendra Meena, Ved Prakash, Alok Mishra, Navin Singhal, Ankit Shukla and Sudeep Rathore from CEERI. Avilash S Rani C and Naveen Shashidhar Kumbar from CFTRI. Swachchha Majumdar from CGCRI. Dr Dnyaneshwar Umrao Bawankule, Dr Debabrata Chanda, Pankaj Shukla, Sanjay Singh, Dr Dayanandan Mani, Ravi Kumar, Pankaj Yadav, Parmanand Kumar from CIMAP. D C Sharma, Dr Neelam J Gupta, A K Jain and Sudhansu Bhagat from CRRI. Pankaj Pandey, Rajesh, Dr. Mohammed Faruq and Ajay Pratap Singh from IGIB. Yogita Singh and Karvan Kaushal from CSIO. Jaykumar Patel, Shrikant Khandare and Dr Kannan Srinivasan from CSMCRI. Dr. Prakash L, Ganapathi G Bhatt, Shashikala U, Shashidhar K N from NAL. Dr. Vidyadhar Mudkavi, Ravichandran C and Sunil Babu M G from 4PI.
Dr. Robin Singh, Mahesh S, Mohit Kumar Swarnakar, Dr. Pankaj Kulurkar from IHBT. Saikat Chakrabarti and Sandip Paul from IICB. Siva Ranjith and B Vijay Kumar from IICT. Sajad Ahmed from IIIM. Rene Christina, Neha Bansal and Ayan Banerjee from IIP. S.K. Goyal from NEERI-Delhi. Dr. P.R Meganathan and Dr. Shaik Basha from NEERI-Hyderabad.
Antara Sharma from NEIST. Dr. Shuchismita Benzwal and Chaitanya Dinesh from NGRI. Shana S Nair from NIIST. N. Anandavalli and P. Vasudevan from SERC.

Vibha Malhotra Swaney from TKDL. Rashmi Arya and Prafulla Malwadkar from URDIP. Ajeet Singh and Dr. R.K.Sinha from HRDC.

References

1. . https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.
2. Nishiura H, Kobayashi T, Miyama T, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis* 2020; **94**: 154-5.
3. Bai Y, Yao L, Wei T, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA* 2020; **323**(14): 1406-7.
4. Brown TS, Walensky RP. Serosurveillance and the COVID-19 Epidemic in the US: Undetected, Uncertain, and Out of Control. *JAMA* 2020; **324**(8): 749-51.
5. Oran DP, Topol EJ. Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative Review. *Ann Intern Med* 2020; **173**(5): 362-7.
6. Andrews M, Areekal B, Rajesh K, et al. First confirmed case of COVID-19 infection in India: A case report. *Indian Journal of Medical Research* 2020; **151**(5): 490-2.
7. Murhekar M, Bhatnagar T, Selvaraju S, et al. Prevalence of SARS-CoV-2 infection in India: Findings from the national serosurvey, May-June 2020. *Indian Journal of Medical Research* 2020; **152**(1): 48-60.
8. Malani A, Shah D, Kang G, et al. Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. *The Lancet Global Health*.
9. Ray A, Singh K, Chattopadhyay S, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in hospitalized patients at a tertiary referral center in North India. *medRxiv* 2020: 2020.08.22.20179937.
10. Satpati P, Sarangi SS, Gantait K, et al. Sero-surveillance (IgG) of SARS-CoV-2 among Asymptomatic General population of Paschim Medinipur District, West Bengal, India(Conducted during last week of July and 1st week of August 2020) - A Joint Venture of VRDL Lab (ICMR), Midnapore Medical College & Hospital & Department of Health and Family Welfare,Govt. of West Bengal, Paschim Medinipur. *medRxiv* 2020: 2020.09.12.20193219.
11. Krüttgen A, Cornelissen CG, Dreher M, Hornef MW, Imöhl M, Kleines M. Determination of SARS-CoV-2 antibodies with assays from Diasorin, Roche and IDvet. *Journal of virological methods* 2021; **287**: 113978.
12. Siracusano G, Pastori C, Lopalco L. Humoral Immune Responses in COVID-19 Patients: A Window on the State of the Art. *Front Immunol* 2020; **11**: 1049-.
13. Ren L, Zhang L, Chang D, et al. The kinetics of humoral response and its relationship with the disease severity in COVID-19. *Communications Biology* 2020; **3**(1): 780.

14. Dhama K, Sharun K, Tiwari R, et al. COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. *Hum Vaccin Immunother* 2020; **16**(6): 1232-8.
15. Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clin Infect Dis* 2020; **71**(15): 778-85.
16. Wu F, Wang A, Liu M, et al. Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications. *medRxiv* 2020: 2020.03.30.20047365.
17. Wong SK, Li W, Moore MJ, Choe H, Farzan M. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *The Journal of biological chemistry* 2004; **279**(5): 3197-201.
18. Carrillo J, Izquierdo-Useros N, Ávila-Nieto C, Pradenas E, Clotet B, Blanco J. Humoral immune responses and neutralizing antibodies against SARS-CoV-2; implications in pathogenesis and protective immunity. *Biochemical and Biophysical Research Communications* 2020.
19. Ahmed SF, Quadeer AA, McKay MR. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* 2020; **12**(3).
20. Patidar GK, Dhiman Y. Distribution of ABO and Rh (D) Blood groups in India: A systematic review. *ISBT Science Series*; **n/a**(n/a).
21. Mohan P, Lando HA, Panneer S. Assessment of Tobacco Consumption and Control in India. *Indian Journal of Clinical Medicine* 2018; **9**: 1179916118759289.
22. Chandramouli C. https://www.censusindia.gov.in/2011-prov-results/prov_results_paper1_india.html.
23. Ortolan A, Lorenzin M, Felicetti M, Doria A, Ramonda R. Does gender influence clinical expression and disease outcomes in COVID-19? A systematic review and meta-analysis. *International Journal of Infectious Diseases* 2020; **99**: 496-504.
24. Zhao J, Yang Y, Huang H, et al. Relationship between the ABO Blood Group and the COVID-19 Susceptibility. *medRxiv* 2020: 2020.03.11.20031096.
25. Wu Y, Feng Z, Li P, Yu Q. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. *Clinica Chimica Acta* 2020; **509**: 220-3.
26. Göker H, Aladağ Karakulak E, Demiroğlu H, et al. The effects of blood group types on the risk of COVID-19 infection and its clinical outcome. *Turk J Med Sci* 2020; **50**(4): 679-83.
27. Severe Covid GG, Ellinghaus D, Degenhardt F, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med* 2020; **383**(16): 1522-34.
28. Golinelli D, Boetto E, Maietti E, Fantini MP. The association between ABO blood group and SARS-CoV-2 infection: A meta-analysis. *PLOS ONE* 2020; **15**(9): e0239508.
29. Padhi S, Suvankar S, Dash D, et al. ABO blood group system is associated with COVID-19 mortality: An epidemiological investigation in the Indian population. *Transfus Clin Biol* 2020; **27**(4): 253-8.
30. Conte L, Toraldo DM. Targeting the gut–lung microbiota axis by means of a high-fibre diet and probiotics may have anti-inflammatory effects in COVID-19 infection. *Therapeutic Advances in Respiratory Disease* 2020; **14**: 1753466620937170.
31. Trompette A, Gollwitzer ES, Yadava K, et al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature medicine* 2014; **20**(2): 159-66.

32. Jayawardena R, Sooriyaarachchi P, Chourdakis M, Jeewandara C, Ranasinghe P. Enhancing immunity in viral infections, with special emphasis on COVID-19: A review. *Diabetes & metabolic syndrome* 2020; **14**(4): 367-82.
33. Fontanet A, Tondeur L, Madec Y, et al. Cluster of COVID-19 in northern France: A retrospective closed cohort study. *medRxiv* 2020: 2020.04.18.20071134.
34. Makoto Miyara FT, Valérie POURCHER, Capucine Morelot-Panzini, et al. Low rate of daily active tobacco smoking in patients with symptomatic COVID-19. Qeios.; 2020.
35. Rossato M, Russo L, Mazzocut S, Di Vincenzo A, Fioretto P, Vettor R. Current smoking is not associated with COVID-19. *European Respiratory Journal* 2020; **55**(6): 2001290.
36. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020; **369**: m1966-m.
37. Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: could nicotine be a therapeutic option? *Intern Emerg Med* 2020; **15**(5): 845-52.
38. Polverino F. Cigarette Smoking and COVID-19: A Complex Interaction. *Am J Respir Crit Care Med* 2020; **202**(3): 471-2.
39. Cai G, Bossé Y, Xiao F, Kheradmand F, Amos CI. Tobacco Smoking Increases the Lung Gene Expression of ACE2, the Receptor of SARS-CoV-2. *Am J Respir Crit Care Med* 2020; **201**(12): 1557-9.
40. Oakes JM, Fuchs RM, Gardner JD, Lazartigues E, Yue X. Nicotine and the renin-angiotensin system. *Am J Physiol Regul Integr Comp Physiol* 2018; **315**(5): R895-R906.
41. Figueiredo-Campos P, Blankenhaus B, Mota C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in COVID-19 patients and healthy volunteers up to 6 months post disease onset. *European Journal of Immunology*; **n/a**(n/a).
42. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nature Microbiology* 2020; **5**(12): 1598-607.
43. Deshpande G, Sapkal G, Tilekar B, et al. Neutralizing antibody responses to SARS-CoV-2 in COVID-19 patients. *Indian Journal of Medical Research* 2020; **152**(1): 82-7.