Household transmission of SARS-CoV-2 in the United States: living density, viral load, and disproportionate impact on communities of color

Supplementary Material

SUPPLEMENTAL METHODS

Study Design

The CO-HOST Study evaluated SARS-CoV-2 transmission in the households of individuals who tested positive and quarantined at home. Here we describe the pre-planned primary analysis of the secondary attack rate and risk factors associated with SARS-CoV-2 transmission in the household setting in the southern United States. Study follow-up started in April 2020 and ended in November 2020.

Ethics, standards and informed consent

The study was approved by the Institutional Review Board at the University of North Carolina and is registered as an observational study at clinicaltrials.gov (NCT04445233). All participants (or their parents/guardians) gave written, informed consent. Minors over the ages of 7 provided assent.

Role of the Funding source

None

Study setting

Index cases were recruited from April to October 2020 after testing at the Respiratory Diagnostic Center (RDC) at the University of North Carolina School of Medicine in Chapel Hill, a drive through testing venue on the medical center campus. Three index cases and three household contacts (all from different households) also enrolled in a treatment study through the RDC in which they were randomized to receive either the oral drug EIDD-2801 (molnupiravir) or placebo (NCT04405570). Participants were visited between 3-4 times at their private homes using a mobile unit van and returned to the Respiratory Diagnostic Center for the final study visit.

Study objectives and outcomes

The primary objective was to evaluate the secondary household attack rate among household members of persons quarantined in their home after testing positive for SARS-CoV-2.

The primary study endpoint was SARS-CoV-2 infection in the household contacts as determined by real-time PCR of nasopharyngeal or nasal swabs for SARS-CoV-2 at any of the timepoints or evidence of seroconversion during the study based on anti-SARS-CoV-2 antibody testing.

A secondary objective was to assess individual and household risk factors associated with SARS-CoV-2 transmission in the household.

Symptom evaluation and severity classification

Within the baseline questionnaire, participants were asked to quantify the number of days symptoms were present for any symptom reported. Symptom diaries consisted of a checklist of 14 symptoms in both English

and Spanish that the participant completed if they answered “Yes” to “Are you having any symptoms?”. Index cases completed daily symptom diaries until no symptoms were reported for two consecutive days. Household contacts received daily symptom evaluations until study day 21 to evaluate for symptoms that could be consistent with new infection. If the participant failed to fill out a symptom diary for two consecutive days, they were contacted by study staff and symptom evaluation was performed over the phone. Questionnaires and symptom diaries for minors were sent to and completed by their designated guardians.

Participants were categorized as having asymptomatic, or having mild, moderate, or severe/hospitalized symptoms based on a modified CDC definition[19]. Symptoms were quantified cumulatively over the course of the study to capture both number of symptoms per day and duration of symptoms. For example, if an individual reported 1 symptom on day 1, 3 symptoms on day 2, and 1 symptom on day 3 they would be considered to have 5 cumulative symptoms. Participants were considered asymptomatic if they reported at most 1 symptom for the duration of their enrollment. Additionally, the symptom reported could not be anosmia, dyspnea, or new or worsening cough. Participants were considered to have mild illness if they reported between 2 and 10 cumulative symptoms or if they reported 1 symptom of anosmia, dyspnea, or new or worsening cough. Participants were considered to have moderate illness if they reported 11 or more cumulative symptoms but were not hospitalized. Patients were categorized as having severe illness if they were hospitalized or died during the course of the study.

Statistical analysis

Potential risk factors for secondary transmissions within the household, including characteristics of index cases, households, and household contacts were examined. Statistical significance was tested by either Fisher’s exact or chi-square test when appropriate for categorical variables and Mann-Whitney test for continuous variables. For index cases, we determined whether demographic characteristics, comorbidities, self-reported mask-wearing behavior, duration and severity of symptoms, viral load and antibody status at enrollment, and educational status were associated with any secondary cases in the household (Table 2). We evaluated whether household-level fisk factor including household size, living space, living density (composite of household size and living space), and home ownership were associated with one or more secondary cases in the household among those who were PCR-negative at baseline enrollment (Table 3).Finally, for household contacts who were PCR-negative at baseline enrollment, we determined whether demographic characteristics, comorbidities, relationship to the index case, self-reported mask-wearing behavior, shared activities with the index case, and educational status were associated with acquisition of infection during study follow-up (Table S4). The odds ratio (OR) and corresponding 95% CI for potential risk factors using logistic regression with a random intercept to account for within-household correlation (for analysis of household contacts) are presented.

Supplementary Material

Laboratory analyses

All samples collected during the study were placed into a cooler on ice immediately after collection and transported within two hours to a Biosafety Laboratory (BSL) Level 2+ laboratory. Study participants who were hospitalized or left the household for other reasons were still followed until Day 28 to record outcomes, but sample collection was suspended.

qRT-PCR SARS-CoV-2 viral quantification

NP and NMT swab samples were tested using a Centers for Disease Control (CDC) RT-qPCR protocol authorized by the Food and Drug Administration (FDA) for emergency use that consists of three unique assays: two targeting regions of the virus' nucleocapsid gene (N1, N2) and one targeting human RNase P gene (RP) (Catalog # 2019-nCoVEUA-01, Integrated DNA Technologies) [1]. Details of assay implementation and calculation of the limit of detection are described elsewhere [2]. Briefly, samples were designated positive if all three PCRs were positive (N1 and N2 for virus, RP for adequate sampling). The viral load of each sample, in copies/uL, was extrapolated from standard curves generated for each viral assay (N1 and N2) using serial dilutions of the nCoVPC plasmid control (2 to 100,000 viral RNA copies/uL). The average copies/uL between the N1 and N2 assays was used as the final quantitative viral load. Probit analysis yielded a limit of detection (LOD) for the N1 and N2 assays of 9 and 13 copies/uL, respectively. Thus, the average LOD between the two assays, 11 copies/uL, was used as the cutoff for sample positivity. Based on the sample collection and RNA extraction volumes as well as volume of template RNA used in the RT-qPCR (5uL), the reported viral load represents the number of viral RNA copies per 5 uL of VTM or Shield sample.

Serology:

Rapid Test

The BioMedomics COVID-19 IgM/IgG Rapid Test is a point-of-care lateral flow immunoassay (LFIA)[3] that has been validated as a research tool [4]. Approximately 20 microliters of finger prick blood was obtained via a capillary sampler and dispensed on the sample port of the device. Two to three drops of buffer/developer solution were applied and results were read after 10 minutes by trained study staff. Positive, weak positive, and negative bands for IgM and IgG were recorded and a photograph was stored. A second reader reviewed the photographs blinded to the field results and consensus was reached on discrepant readings.

Immunoassay to detect antibodies against the receptor binding domain (RBD) of the spike protein

Plasma samples were heat inactivated at 56°C for 30 minutes, then total Ig binding to the receptor binding domain (RBD) of the SARS-CoV-2 spike protein was measured using a previously described enzyme-linked immunosorbent (ELISA) assay [5,6]. Briefly, biotinylated recombinant antigen produced in mammalian cells consisting of SARS-2 Spike RBD is captured on a 96-well ELISA plate coated with streptavidin. The serum sample at 1:40 dilution is incubated with the RBD-captured wells, and bound antigen detected using HRP conjugated anti-goat total (IgG, IgM and IgA) antibody on a microplate reader. This in-house ELISA was previously

evaluated on a large panel of well characterized samples and shown to have high sensitivity and specificity for detecting SARS-CoV-2 infection [5,6].

D614G genotyping

A real-time PCR assay targeting a 107 bp region encompassing the D614G mutation in the SARS-CoV-2 spike protein receptor binding domain associated with increased viral load [7] was designed to evaluate the prevalence of 614G mutants. 5ul of RNA was reverse transcribed using the Invitrogen SuperScript III First-Strand Synthesis System for RT-PCR kit (Thermofisher Scientific). 2.5ul cDNA was then placed in 22.5uL of qPCR master mix with Roche FastStart Universal Probe Master (ROX) along with primers and probes listed in **Table S1**. Positive control plasmids for mutant (MT) and wild-type (WT) sequences were synthesized by Genewiz (inserts listed in **Table S1**) and used to set the appropriate Ct threshold for positivity in each run. Samples were considered WT if detected only by WT probe; MT if detected only by MT probe or if detected by both MT and WT probes with MT Ct >3 cycles lower than WT Ct; or mixed (containing both WT and MT virus) if detected by both with Ct difference of <3 cycles.

SARS-CoV-2 sequencing

The majority (n=28) of cDNA libraries were generated using ARTIC Network amplicons [8] to generate cDNA followed by library construction with a QIAGEN® (Hilden, Germany) QIAseq FX kit, as described in [9]. Paired-end libraries were sequenced on an Illumina MiSeq at the UNC High-Throughput Sequencing Facility. Nine additional libraries (including six that did not sequence well originally) were constructed using a Swift Biosciences (Ann Arbor, MI) Normalase® Amplicon Panel (SNAP) and sequenced on an Illumina iSeq 100.

Following demultiplexing, libraries underwent adapter and quality trimming according to default parameters for paired-end reads in Trim Galore! [10]. Trimmed fastq files were converted to unaligned BAM format, trimmed of primer sequences, aligned to the Wuhan reference sequence, and assembled into fasta format using the Broad Institute viral NGS pipelines [11] implemented in Docker Desktop.

The resulting fasta files were aligned via MAFFT v7.450 [12][13] implemented in Geneious Prime® 2021. Alignments were uploaded to Nextclade [14] to assign clades and evaluate final sequence quality. Relatedness between viral sequences was also assessed via phylogenomic analysis in MrBayes v3.2.6 [15] implemented in Geneious Prime® 2021 using default parameters and setting the Wuhan reference sequence as the outgroup. Samples from the same household were considered to be related if they were assigned to the same larger clade by Nextclade as well as the same clade in MrBayes. All sequences included in this analysis are available on GISAID under the accession numbers EPI_ISL_3088340 to EPI_ISL_3088373 and EPI_ISL_3247163 to EPI_ISL_3247165.

Household transmission of SARS-CoV-2 in the United States: living density, viral load, and disproportionate impact on communities of color

Supplementary Material

Data entry, handling, storage and security

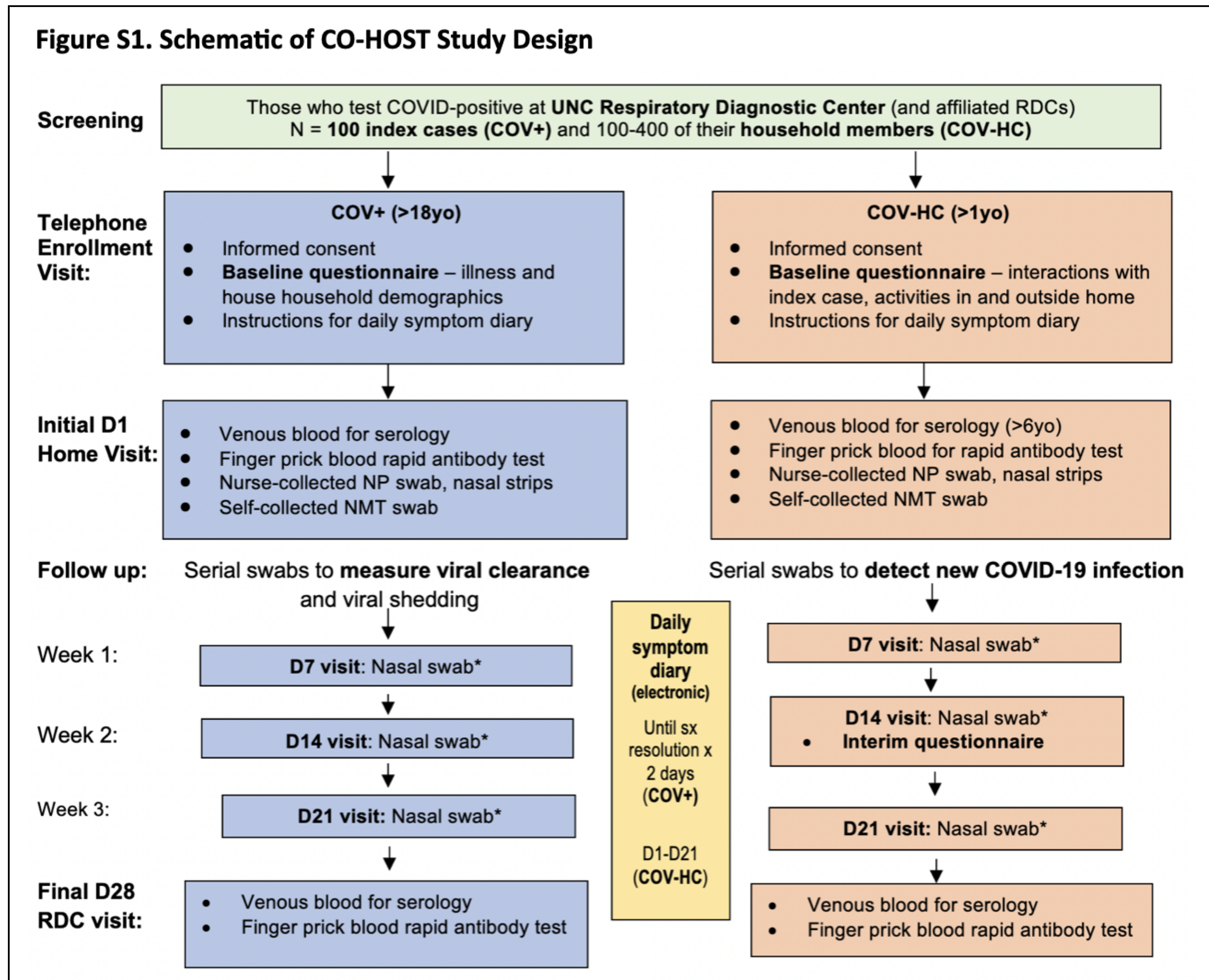
After giving written consent, the participants were given a study identification number, which was used in all future datasets for participant anonymity. Collected data were entered in real-time using electronic Case Report Forms (eCRF) developed on a REDCap (Research Electronic Data Capture) database. Any data collected on paper format was entered by a study staff member and then checked by the study coordinator. Daily symptom diaries were entered directly into the REDCap database by the participants and were checked by study staff for completion and inconsistencies. Laboratory related data were extracted directly from laboratory equipment and uploaded to the database. The study was conducted in compliance with Good Clinical Practice.

Data availability

Data is available on request for any interested researchers to allow replication of results provided all ethical requirements are met.

SUPPLEMENTAL TABLES/FIGURES:

Figure S1. Schematic of CO-HOST Study Design



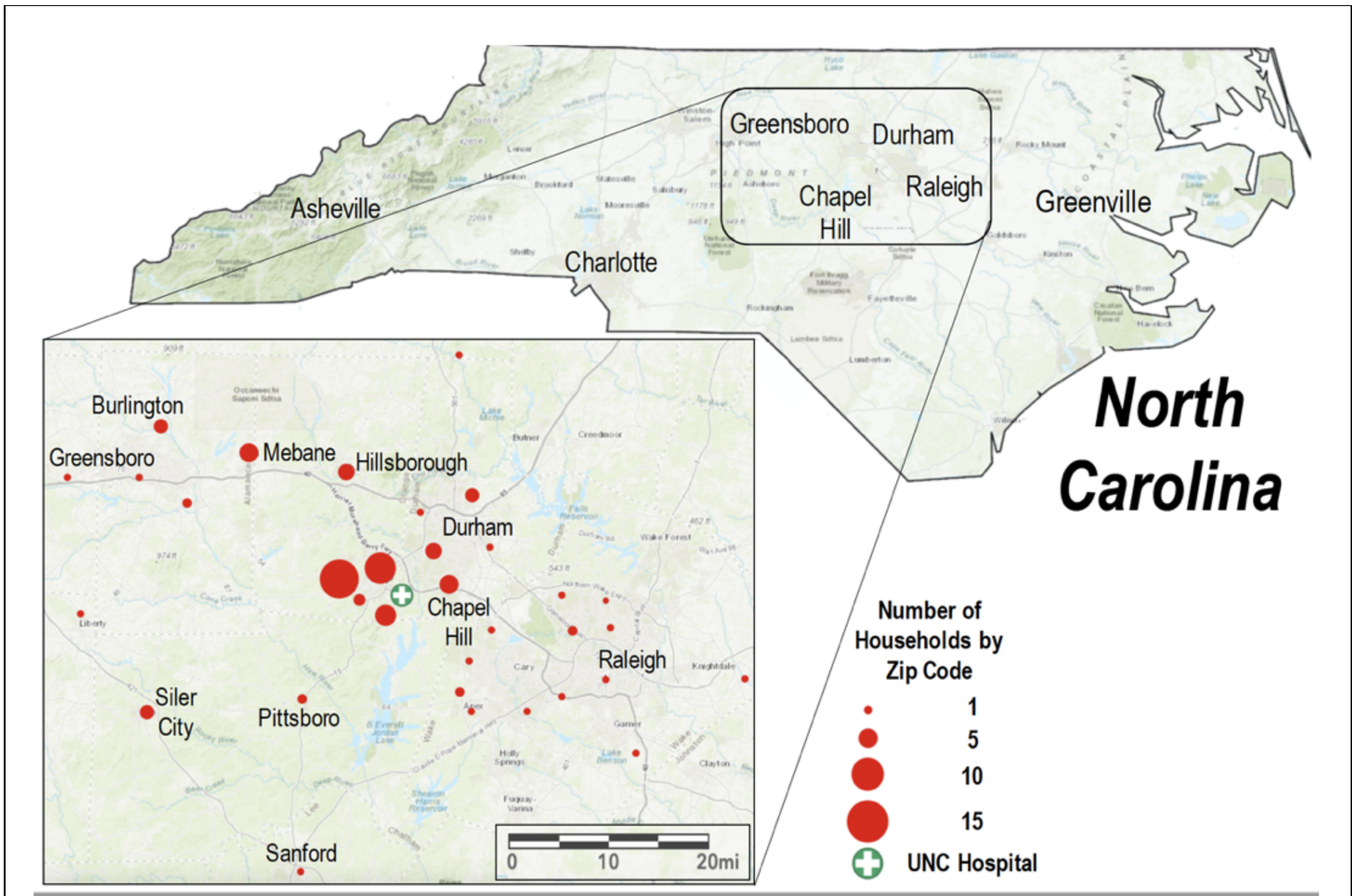


Figure S2. Geography of 100 households enrolled and followed in the UNC CO-HOST study. Figure prepared using ArcGIS (Esri, Redlands, California).

Supplementary Material

Table S1. Comorbidities of study participants

INDIVIDUALS	Index cases (n)	Index cases (%)	Household contacts (n)	Household contacts (%)
Underlying Conditions for Adults >18y*	***83***	%	***116***	%
Cancer	3	**3.6**	0	**0.0**
Chronic lung disease	1	**1.2**	2	**1.7**
Asthma	8	**9.6**	17	**15**
Daily smoker	2	**2.4**	11	**9.4**
Diabetes (n = 82 index)	4	**4.8**	10	**8.6**
Hypertension	12	**14**	23	**20**
Heart disease (n = 82 index)	1	**1.2**	2	**1.7**
HIV	0	**0.0**	0	**0.0**
Chronic kidney disease	1	1.2	0	0
Chronic liver disease	0	**0.0**	0	**0.0**
Recent (within past 2 weeks) or current pregnancy	3	**3.6**	5	**4.3**
BMI >30 (n = 78 index, 103 HC)	28	**36**	38	**37**
BMI 25-29.9 (n = 78 index, 103 HC)	23	**29**	31	**30**
Underlying Conditions for Adults >50y*	***22***	%	***38***	%
Asthma	1	**4.5**	2	**5.3**
Daily smoker	0	**0.0**	5	**13**
Diabetes	4	**18**	6	**16**
Hypertension	5	**23**	19	**50**
BMI >30 (n = 22 index, 35 HC)	10	**45**	13	**37**
BMI 25-29.9 (n = 22 index, 35 HC)	5	**23**	13	**37**
BMI >30 and one or more co-morbidity (adults >18y)	16	**19**	22	**19**
BMI >30 and one or more co-morbidity (adults >50y)	7	**32**	12	**32**

Table S2. Household demographics

HOUSEHOLDS	(n=91)	(%)
Household Size		
2 People	24	26
3 People	20	22
4 People	21	23
5 or more people	26	29
Home Ownership (n = 89)		
Single-family home/townhome occupied by owner	56	**63**
Single-family home/townhome occupied by renter	22	**25**
Apartment occupied by renter	10	**11**
Other	1	**1.1**
Rooms in the House*		
2 or fewer rooms	10	**11**
3-5 rooms	38	**42**
6 or more rooms	43	**47**
**including bedrooms, kitchen, and common rooms, but not bathrooms or garage*		
Living Space (n = 86)		
<500 sq feet (<46.5 sq m)	3	**3.5**
500-1000 sq feet (46-93 sq m)	17	**20**
1000-2000 sq feet (93-186 sq m)	26	**30**
>2000 sq feet (>186 sq m)	40	**47**

Supplementary Material

Table S3. Demographics of household contacts based on infection status: infected at baseline (PCR-positive) vs. secondary cases (PCR+ during study follow-up) vs. uninfected at study conclusion

HOUSEHOLD CONTACTS	Total (%)	Infected at baseline (%)	Secondary cases (%)	Uninfected (%)
	176	**73**	**33**	**70**
Male	87 (49%)	38 (52%)	11 (33%)	38 (54%)
Female	89 (51%)	35 (48%)	22 (67%)	32 (46%)
Race/Ethnicity				
White, non-Hispanic	96 (55%)	32 (44%)	18 (55%)	46 (66%)
Non-White	73 (42%)	38 (52%)	15 (45%)	20 (29%)
Black or African American	17 (9.7%)	8 (11%)	4 (12%)	5 (7.1%)
Hispanic/Latinx	54 (31%)	29 (40%)	10 (30%)	15 (21%)
Other, non-Hispanic	2 (1.1%)	1 (1.4%)	1 (3.0%)	0
Unknown	7 (4.0%)	3 (4.1%)	0	4 (5.7%)
Language				
Spanish speaking (yes)	28 (16%)	20 (27%)	2 (6.1%)	6 (8.6%)
Spanish speaking (no)	148 (84%)	53 (73%)	31 (94%)	64 (91%)
Age				
0-12y	38 (22%)	11 (15%)	8 (24%)	19 (27%)
13-17y	22 (13%)	8 (11%)	5 (15%)	9 (13%)
18-24y	23 (13%)	10 (14%)	4 (12%)	9 (13%)
25-49y	56 (32%)	27 (37%)	13 (39%)	16 (23%)
50-64y	27 (15%)	11 (15%)	2 (6.1%)	14 (20%)
>65y	10 (5.7%)	6 (8.2%)	1 (3.0%)	3 (4.3%)
Education (excluding <18y)				
Total Responses for Adults >18y	*113*	*53*	*18*	*42*
High school or lower	54 (48%)	32 (60%)	8 (44%)	14 (33%)
College degree	34 (30%)	11 (21%)	7 (39%)	16 (38%)
Graduate degree	25 (22%)	10 (19%)	3 (17%)	12 (29%)
Occupation (excluding <18y)				
Total Responses for Adults >18y	*116*	*54*	*20*	*42*
Education	6 (5.2%)	2 (3.7%)	3 (15%)	1 (2.4%)
Healthcare worker	9 (7.8%)	5 (9.3%)	1 (5.0%)	3 (7.1%)
Retail/hospitality/other frontline worker	22 (19%)	13 (24%)	3 (15%)	6 (14%)
Student	12 (10%)	4 (7.4%)	1 (5.0%)	7 (17%)
White collar worker	34 (29%)	14 (26%)	4 (20%)	16 (38%)
Other (trade and arts)	6 (5.2%)	2 (3.7%)	1 (5.0%)	3 (7.1%)
Not working outside the home	27 (23%)	14 (26%)	7 (35%)	6 (14%)
Co-Morbidities (excluding <18y)				
Diabetes	9 (7.8%)	4 (7.4%)	2 (10%)	3 (4.3%)
High blood pressure	24 (21%)	11 (20%)	10 (50%)	3 (4.3%)
Total BMI Responses for Adults >18y	*103*	*46*	*19*	*38*
BMI >30	38 (37%)	22 (47%)	9 (47%)	7 (18%)
BMI 25-29.9	31 (30%)	15 (32%)	3 (16%)	13 (34%)
BMI <25	34 (33%)	9 (19%)	7 (37%)	18 (47%)
BMI >30 and one or more co-morbidity	22 (19%)	13 (28%)	4 (21%)	5 (14%)

Table S4. Potential risk factors for SARS-CoV-2 infection among household contacts

HOUSEHOLD CONTACTS	All Household Contacts (n, %)	Infected (n, %)	Uninfected (n, %)	p-value
	103 (100%)	33 (32%)	70 (68%)	-
Relationship to index case				
Partner	23 (22%)	10 (30%)	13 (19%)	NS
Child	39 (38%)	10 (30%)	29 (41%)	NS
Parent	19 (18%)	3 (9%)	16 (23%)	0.10
Caregiver	32 (31%)	10 (30%)	22 (31%)	NS
Age				
<18y	41 (40%)	13 (39%)	28 (40%)	NS
18-50y	48 (47%)	18 (55%)	30 (43%)	
>50y	14 (14%)	2 (6%)	12 (17%)	
Sex				
Female	54 (52%)	22 (67%)	32 (46%)	0.05
Male	49 (48%)	11 (33%)	38 (54%)	
Shared activities prior to enrollment (missing n = 3)				
Sharing bedroom	32 (32%)	14 (42%)	18 (27%)	NS
Sharing bathroom	66 (66%)	28 (85%)	38 (57%)	0.003
Sharing meals	72 (72%)	25 (76%)	47 (70%)	NS
Sharing car rides	58 (58%)	21 (64%)	37 (55%)	NS
Mask wearing prior to enrollment (missing n = 1)				
Mask wearing at home	24 (26%)	6 (20%)	18 (30%)	NS
Race/Ethnicity				
White, non-Hispanic	64 (62%)	18 (55%)	46 (66%)	NS*
Black or African American	9 (9%)	4 (12%)	5 (7%)	
Other, non-Hispanic	4 (4%)	1 (3%)	3 (4%)	
Hispanic/Latinx	26 (25%)	10 (30%)	16 (23%)	
Co-morbidities for adults >18y (missing n = 5 for obesity)				
Diabetes	5 (8%)	1 (5%)	4 (10%)	NS
Obesity, BMI >30	16 (28%)	9 (47%)	7 (18%)	0.06
Education for adults >18y (missing n = 2)				
High school or lower	22 (37%)	8 (44%)	14 (33%)	NS
College degree	23 (38%)	7 (39%)	16 (38%)	
Graduate degree	15 (25%)	3 (17%)	12 (29%)	

p-values are adjusted for household clustering and only reported if ≤ 0.10
*compares White, non-Hispanic vs all other categories

Supplementary Material

REFERENCES

1. Center for Disease Control and Prevention. CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. Revis Biol Celular **2020**; 3:30.
2. Muller MS, Bhattarai Chhetri S, Basham C, et al. Practical strategies for SARS-CoV-2 RT-PCR testing in resource-constrained settings. 2021; Available at: http://medrxiv.org/lookup/doi/10.1101/2021.02.18.21251999.
3. Li Z, Yi Y, Luo X, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. J Med Virol **2020**; 92:1518–1524.
4. Naranbhai V, Chang CC, Beltran WFG, et al. High Seroprevalence of Anti-SARS-CoV-2 Antibodies in Chelsea, Massachusetts. J Infect Dis **2020**; 222:1955–1959.
5. Premkumar L, Segovia-Chumbez B, Jadi R, et al. The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. Sci Immunol **2020**; 5. Available at: http://dx.doi.org/10.1126/sciimmunol.abc8413.
6. Markmann AJ, Giallourou N, Bhowmik DR, et al. Sex disparities and neutralizing antibody durability to SARS-CoV-2 infection in convalescent individuals. medRxiv **2021**; Available at: http://dx.doi.org/10.1101/2021.02.01.21250493.
7. Korber B, Fischer WM, Gnanakaran S, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell **2020**; 182:812–827.e19.
8. Quick J. nCoV-2019 sequencing protocol v3 (LoCost). **2020**; Available at: https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye.pdf. Accessed 13 May 2021.
9. Sekizuka T, Itokawa K, Hashino M, et al. A Genome Epidemiological Study of SARS-CoV-2 Introduction into Japan. mSphere **2020**; 5. Available at: https://msphere.asm.org/content/5/6/e00786-20.abstract. Accessed 13 May 2021.
10. Babraham Bioinformatics - Trim Galore! Available at: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 13 May 2021.
11. viral-ngs: genomic analysis pipelines for viral sequencing — viral-ngs v2.1.19.0-384-gc1d835ff documentation. Available at: https://viral-pipelines.readthedocs.io/en/latest/index.html. Accessed 13 May 2021.
12. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol **2013**; 30:772–780.
13. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res **2002**; 30:3059–3066.
14. Nextclade. Available at: https://clades.nextstrain.org. Accessed 13 May 2021.
15. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001; 17:754–755. Available at: http://dx.doi.org/10.1093/bioinformatics/17.8.754.