

https://dspace.org/

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission dynamics and social contact patterns

Kleynhans, Jacoba Wilhelmina (Jackie)

2024
https://hdl.handle.net/10539/37577

Downloaded from a DSpace repository. https://dspace.org/

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission dynamics and social contact patterns

Jacoba Wilhelmina (Jackie) Kleynhans

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 27 March 2023

DECLARATION

I, Jackie Kleynhans, declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



(Signature of candidate)

27th day of March 2023 in Sandringham, Johannesburg

DEDICATION

To every frontline worker who put others' lives before their own, who had to face devastation every day, and still go back the next.

"Public Health is the science and the art of preventing disease, prolonging life, and promoting physical health and efficiency through organized community efforts for the sanitation of the environment, the control of community infections, the education of the individual in principles of personal hygiene, the organization of medical and nursing services for the early diagnosis and preventive treatment of disease, and the development of the social machinery which will ensure to every individual in the community a standard of living adequate for the maintenance of health."

Winslow CE. 1920. The untilled fields of public health. Science. 51(1306):23-33.

PREFACE

Not many epidemiologists would have the opportunity to work on a pandemic so early in their careers, never mind the opportunity to make it the topic of their PhD thesis. My involvement in the initial response to SARS-CoV-2 and subsequent research studies was a (hopefully) once-in-a-lifetime opportunity for which I will always be grateful, although it placed a toll on my mental and emotional wellbeing.

The 5th of March 2020, when the first COVID-19 case was identified in South Africa, will be embedded in my memory forever. I was driving to work, after three days of working from home on one of my PhD objectives. At the time my PhD was still focussing on influenza and contact patterns, and I had been intending to visit our collaborators in Italy in February. My trip was first postponed and then cancelled due to the devastating spread of SARS-CoV-2 in Italy. So I had three days of virtual calls to try and make up for this. On my way to work that Thursday morning, I received the call that the first case had been identified in South Africa. We had been preparing for this for weeks, with multiple meetings, planning sessions and press releases. But still, that first case threw me off balance; I quickly realized what all the gaps were that we had not considered yet. After a very hushed meeting, I was asked to contact the index case patient and do the subsequent contact tracing. The next day I did a house call with one of our clinical staff members to collect a sample and clinical information on one of the contacts. What my training did not prepare me for, is the fact that I was not just counting numbers, but engaging with people, people worried about what will happen to them, their families… I received calls at 5 am from a husband worried that his wife had been exposed. School headmasters worried if an event should continue. It is not as simple as collecting a swab, filling in a case investigation form and just moving on. The next few weeks were a blur of late nights, early mornings and trying to stay on top of case numbers as our data systems had time to catch up.

No workshop can ever truly prepare you for a pandemic. There are so many subtleties, and soft skills, that you need to acquire. Understanding people's reactions when they refuse to be in the same meeting room as you after hearing you were exposed to a confirmed case, allowing yourself and others to break down and vent, to make mistakes, learn from them and then move on and do better. I cannot entirely put in words how much I have learned in the last three years, the experience I have gained, and the professional and personal growth I experienced. But I will forever be grateful that I have had the opportunity to build my career on this experience, and for the amazing mentors that surrounded me during this time, allowing me to grow.

ACKNOWLEDGEMENTS

I would like to start by thanking Dr Melinda Suchard who gave me my very first opportunity at the National Institute for Communicable Diseases (NICD) after it had been my dream to one day work here. Melinda was also the one who encouraged me to pursue a career in epidemiology, setting up a meeting between me and the person who would later become my PhD supervisor, Prof Cheryl Cohen.

I would also like to thank the entire Centre for Respiratory Diseases and Meningitis (CRDM) team and all study collaborators, field teams and participants who made the collection and processing of the data possible. Thank you to all colleagues both within CRDM and the broader NICD for all the corridor conversations and motivation to keep going. Thank you to Dr Sibongile Walaza who has been not just a great manager but a mentor, allowing me so many opportunities to develop my skills. Thank you to Dr Monica Birkhead for taking the time to review this narrative. Thank you for the time I spent in the electron microscopy unit, it is still the most memorable and enjoyable of all my internship placements. Thank you to the examiners who offered their time to critically review this thesis, and for their valuable contributions to further strengthen this body of work.

This work would not have been possible without my amazing supervisors, Professor Cheryl Cohen and Associate Professor Stefano Tempia. 'Supervisors' is not the correct word, but rather mentors. Cheryl and Stefano, thank you for all the discussions, and for all your insights from which I could learn. You have never once hesitated to share your knowledge with me, always sending opportunities my way to further advance my development. Thank you for your kindness and support, not just during my PhD, but my entire public health journey so far.

Thank you to all my friends and family for your support and interest through the years. Thank you to my brother, Carel, for challenging me intellectually from a very young age, and for our long talks about the world, politics, life, pandemics... Thank you to my sister, Anca, for being my greatest cheerleader and dearest friend, for all our chats, and for getting me out of a rut even when I did not realize I was in one.

Dankie Ma en Pa vir al julle ondersteuning deur die jare en vir al die geleenthede wat julle vir my geskep het. Sonder julle sou ek nooit hier kon gewees het nie.

To my amazing husband, Pieter, thank you for being my best friend and soulmate. Thank you for keeping me sane, for listening when I vent, and for taking such good care of me, especially during those first few weeks of the pandemic. Thank you for your never-ending support and patience.

Soli Deo Gloria!

PRESENTATIONS ARISING FROM THIS RESEARCH PROJECT

1. PHASA 2021 (Virtual, 15 – 17 February 2021, oral presentation): “A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018”
2. WHO Solidarity II (Virtual, 12 March 2021, oral presentation): “A household cohort study of SARS-CoV-2 seroprevalence in a rural and an urban community of South Africa, July 2020 – February 2021”
3. WHO AFRO seminar on COVID-19 sero-epidemiological studies (Virtual, 28 June 2021, oral presentation): “Longitudinal SARS-CoV-2 seroprevalence in a rural and urban community household cohort in South Africa, during the first and second waves July 2020-March 2021”
4. Monthly REDCap Africa Call (Virtual, 24 November 2021, oral presentation): “Household transmission of SARS-CoV-2 from HIV-infected and HIV-uninfected adult index cases in South Africa, 2020-2021 Preliminary Results”
5. Faculty of Health Sciences Research Day (Johannesburg, 15 September 2022, poster): “Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study”
6. WHO AFRO seminar on COVID-19 sero-epidemiological studies (Virtual, 31 October 2022, oral presentation): “Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021”

PUBLICATIONS ARISING FROM THIS RESEARCH PROJECT

1. **Kleynhans** J, Tempia S, McMorrow ML, et al. A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018. *BMC Public Health* 2021; **21**(1): 1055.
2. **Kleynhans** J, Tempia S, Wolter N, et al. SARS-CoV-2 Seroprevalence in a Rural and Urban Household Cohort during First and Second Waves of Infections, South Africa, July 2020–March 2021. *Emerging Infectious Disease journal* 2021; **27**(12): 3020.
3. **Kleynhans** J, Tempia S, Wolter N, et al. SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa. *Emerging Infectious Disease journal* 2022; **28**(5): 1055.
4. **Kleynhans** J, Walaza S, Martinson NA, et al. Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study. *Clinical Infectious Diseases* 2022.
5. **Kleynhans J**, Dall'Amico L, Gauvin L, et al. Association of close-range contact patterns with SARS-CoV-2 household transmission, South Africa, 2020-2021. medRxiv 2022. 12.22.22283843. Preprint.

Candidate's role in all publications:

Conceptualization of studies, database management, data curation, analyses, draft and review of final manuscripts. For a full description of the student's role in each project, please refer to appendix A.

OTHER RELATED PUBLICATIONS

1. Cohen C, **Kleynhans** J, von Gottberg A, et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020-21. *The Lancet Infectious Diseases* 2022; **22**(6): 821-34.
2. Sun K, Tempia S, **Kleynhans** J, et al. SARS-CoV-2 transmission, persistence of immunity, and estimates of Omicron's impact in South African population cohorts. *Science Translational Medicine* 2022; **14**(659): eabo7081.
3. McCarthy KM, Tempia S, Kufa T, **Kleynhans J**, et al. The importation and establishment of community transmission of SARS-CoV-2 during the first eight weeks of the South African COVID-19 epidemic. *EClinicalMedicine* 2021; **39**: 101072.
4. Sun K, Tempia S, **Kleynhans J**, et al. Rapidly shifting immunologic landscape and severity of SARS-CoV-2 in the Omicron era in South Africa. *Nature Communications* 2023; **14**(1): 246.
5. Dall'Amico L, **Kleynhans J**, Gauvin L, et al. Estimating household contact matrices structure from easily collectable metadata. rXiv 2022: arXiv.2210.07034. Preprint.

ABSTRACT

Background

Understanding the community burden and transmission dynamics of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can assist to make informed decisions for prevention policies.

Methods

From August through October 2018, before the SARS-CoV-2 pandemic, we performed a cross-sectional contact survey nested in a prospective household cohort in an urban (Jouberton, North West Province) and a rural community (Agincourt, Mpumalanga Province) in South Africa to measure contact rates in 535 study participants. Participants were interviewed to collect details on all contact events (within and outside of the household).

During the SARS-CoV-2 pandemic we enrolled 1211 individuals from 232 randomly selected households in the same urban and rural community, and followed the cohort prospectively for 16 months (July 2020 through November 2021), collecting blood every two months to test for SARS-CoV-2 antibodies. Using these longitudinal SARS-CoV-2 seroprevalence estimates and comparing these with reported laboratory-confirmed cases, hospitalizations and deaths, we investigated the community burden and severity of SARS-CoV-2.

We also performed a case-ascertained household transmission study of symptomatic SARS-CoV-2 index cases living with HIV (LWH) and not LWH (NLWH) in two urban communities (Jouberton, North West Province and Soweto, Gauteng Province) from October 2020 through September 2021. We enrolled 131 SARS-CoV-2 index cases at primary healthcare clinics. The index cases and their 457 household contacts were followed up for six weeks with thrice weekly visits to collect nasal swabs for SARS-CoV-2 testing on reverse transcription real-time polymerase chain reaction (rRT-PCR), irrespective of symptoms. We assessed household cumulative infection risk (HCIR), duration of virus detection and the interval between index and contact symptom onset (serial interval). By collecting high-resolution household contact patterns in these households using wearable sensors, we assessed the association between contact patterns and SARS-CoV-2 household transmission.

Results

During the contact survey, we observed an overall contact rate of 14 (95% confidence interval (CI), 13-15) contacts per day, with higher contact rates in children aged 14-18 years (22, 95%CI 8-35) compared to children <7 years (15, 95%CI 12-17). We found higher contact rates in the rural site (21, 95%CI 14-28) compared to the urban site (12, 95%CI 11-13).

When comparing the household cohort seroprevalence estimates to district SARS-CoV-2 laboratory-confirmed infections, we saw that only 5% of SARS-CoV-2 infections were reported to surveillance. Three percent of infections resulted in hospitalization and 0.7% in death. People LWH were not more likely to be seropositive for SARS-CoV-2 (odds ratio [OR] 1.0, 95%CI 0.7–1.5), although the sample size for people LWH was small (159/1131 LWH).

During the case-ascertained household transmission study for SARS-CoV-2, we estimated a HCIR of 59% (220/373) in susceptible household members, with similar rates in households with an index LWH and NLWH (60% LWH vs 58% NLWH). We observed a higher risk of transmission from index cases aged 35–59 years (adjusted OR [aOR] 3.4, 95%CI 1.5–7.8) and ≥60 years (aOR 3.1, 95% CI 1.0–10.1) compared with those aged 18–34 years, and index cases with a high SARS-CoV-2 viral load (using cycle threshold values (C_t) <25 as a proxy, aOR 5.3, 95%CI 1.6–17.6). HCIR was also higher in contacts aged 13–17 years (aOR 7.1, 95%CI 1.5–33.9) and 18–34 years (aOR 4.4, 95% CI 1.0–18.4) compared with <5 years. Through the deployment of wearable sensors, we were able to measure high-resolution within-household contact patterns in the same households. We did not find an association between duration (aOR 1.0 95%CI 1.0-1.0) and frequency (aOR 1.0 95%CI 1.0-1.0) of close-proximity contact with SARS-CoV-2 index cases and household members and transmission.

Conclusion

We found high contact rates in school-going children, and higher contact rates in the rural community compared to the urban community. These contact rates add to the limited literature on measured contact patterns in South Africa. The burden of SARS-CoV-2 is underestimated in national surveillance, highlighting the importance of serological surveys to determine the true burden. Under-ascertainment of cases can hinder containment efforts through isolation and contact tracing. Based on seroprevalence estimates in our study, people LWH did not have higher SARS-CoV-2 community attack rates.

In the household transmission study, we observed a high HCIR in households with symptomatic index cases, and that index cases LWH did not infect more household members compared to people NLWH. We found a correlation between age and SARS-CoV-2 transmission and acquisition, as well as between age and contact rates.

Although we did not observe an association between household contact patterns and SARS-CoV-2 transmission, we generated SARS-CoV-2 transmission parameters and community and household contact data that can be used to parametrize infectious disease models for both SARS-CoV-2 and other pathogens to assist with forecasting and intervention assessments. The availability of robust data is important in the face of a pandemic where intervention strategies have to be adapted continuously.

TABLE OF CONTENTS

Declaration ... ii

Dedication ... iii

Preface ... v

Acknowledgements ... vi

Presentations arising from this research project ... vii

Publications arising from this research project ... viii

Other related publications ... ix

Abstract ... x

Table of contents ... xii

List of abbreviations ... xv

List of figures ... xvi

List of tables ... xvi

1. Introduction ... 1

2. Background ... 2

2.1. Epidemiology of SARS-CoV-2 ... 2

2.1.1. Emergence ... 2

2.1.2. Containment and control strategies ... 3

2.1.3. SARS-CoV-2 in South Africa ... 4

2.2. SARS-CoV-2 burden and transmission dynamics ... 6

2.2.1. Burden ... 6

2.2.2. Routes of transmission ... 8

2.2.3. Transmission dynamics ... 9

2.2.4. SARS-CoV-2 and HIV co-infection ... 10

2.3. Social contact patterns ... 11

2.3.1. Measuring contact patterns ... 11

2.3.2. The role of social contact patterns in SARS-CoV-2 transmission ... 13

2.4. Problem statement and justification for research ... 13

2.5. Aim and objectives 14
2.6. Conceptual framework 14
3. Methods 16
3.1. The PHIRST study: Measuring contact patterns using contact diaries 16
3.2. The PHIRST-C study: Measuring burden through seroprevalence 17
3.3. SA-S-HTS: Investigating SARS-CoV-2 transmission dynamics 18
3.4. Ethics 21
4. Results and Discussion 21
4.1. Combining data sources to estimate the burden of SARS-CoV-2 21
4.2. The household as an important setting for SARS-CoV-2 transmission 25
4.3. The influence of study design on the assessment of household transmission 26
4.4. Leveraging contact patterns to assess and predict SARS-CoV-2 transmission and to inform intervention strategies 29
4.5. The importance of not neglecting one epidemic when facing another: SARS-CoV-2 and HIV co-infection 33
4.6. Generating data for action 34
5. Limitations 36
6. Conclusion 37
7. References 38
7.1. Contact details for personal communications listed 51
8. Appendices 52
8.1. Appendix A: Role of the student 52
8.2. Appendix B: Approval of title change 55
8.3. Appendix C: Ethical clearance certificates 56
8.3.1. Main thesis protocol clearance 56
8.3.2. PHIRST clearance 57
8.3.3. Inclusion of SARS-CoV-2 in PHIRST (PHIRST-C) 59
8.3.4. PHIRST-C re-certification 60

8.3.5. SA-S-HTS clearance 61

8.4. Appendix D: Copies of manuscripts comprising PhD 62

8.4.1. A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018 62

8.4.2. SARS-CoV-2 Seroprevalence in a Rural and Urban Household Cohort during First and Second Waves of Infections, South Africa, July 2020–March 2021 79

8.4.3. SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa 101

8.4.4. Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study 118

8.4.5. Association of close-range contact patterns with SARS-CoV-2 household transmission, South Africa, 2020-2021 157

LIST OF ABBREVIATIONS

2019-nCoV	New coronavirus
ART	Anti-retroviral treatment
CIR	Case-to-infection ratio
COVID-19	Coronavirus disease 2019
DATCOV	Electronic hospital surveillance system for COVID-19
FIR	Fatality-to-infection ratio
GPS	Global positioning system
HCIR	Household cumulative infection risk
HIR	Hospitalization-to-infection ratio
HIV	Human immunodeficiency virus
IgA	Immunoglobulin A
k	Dispersion parameter
LWH	Living with HIV
MERS-CoV	Middle East respiratory syndrome virus
NICD	National Institute for Communicable Diseases
NMC	Notifiable medical condition
NMCSS	Notifiable medical conditions surveillance system
NPI	Non-pharmaceutical interventions
PHEIC	Public Health Emergency of International Concern
PHIRST	The Prospective Household observational cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa
PHIRST-C	The Prospective Household study of SARS-CoV-2, Influenza, and Respiratory Syncytial virus community burden, Transmission dynamics, and viral Interaction in South Africa
PLWH	People living with HIV
PNLWH	People not living with HIV
R_0	Basic reproductive number
R_e	Effective reproductive number
RNA	Ribonucleic acid
rRT-PCR	Reverse transcription real-time polymerase chain reaction
RSV	Respiratory syncytial virus
SARS-CoV	Severe acute respiratory syndrome coronavirus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SA-S-HTS	The South African SARS-CoV-2 household transmission study
SIR	Secondary infection risk
TB	*Mycobacterium tuberculosis*
VOC	Variant of concern
WHO	World Health Organization

LIST OF FIGURES

Figure 1. Laboratory-confirmed SARS-CoV-2 cases, hospitalizations and in-hospital deaths, South Africa, 1 March 2020 - 30 September 2022. 5
Figure 2. The conceptual framework for community burden and transmission dynamics of SARS-CoV-2. 15
Figure 3. Timing of data collection relative to laboratory-confirmed SARS-CoV-2 cases (in thousands) reported to Notifiable Medical Conditions Surveillance System (NMCSS), South Africa, 1 March 2020 - 30 September 2022. 20
Figure 4. Timing of blood collection, district SARS-CoV-2 weekly incidence and seroprevalence in all ages during blood draw in a) a rural community and b) an urban community, March 2020 – November 2021, South Africa. 22
Figure 5. SARS-CoV-2 seroprevalence at each blood collection, by age group, in a) a rural community and b) an urban community, South Africa, March 2020–November 2021. 23
Figure 6. SARS-CoV-2 a) case-to-infection (CIR), hospitalization-to-infection (HIR) and b) in-hospital fatality-to-infection (FIR IH) and excess death fatality-to-infection (FIR ED) ratios in a rural and an urban community during the first (Wuhan), second (Beta) and third (Delta) waves of infections, March 2020 - November 2021, South Africa. 25

LIST OF TABLES

Table 1. Methods commonly utilized to measure social contact patterns 11
Table 2. Studies that contributed data to this thesis. 19

1. Introduction

This Degree of Doctor of Philosophy (PhD) thesis was originally titled “Influenza transmission and social contact patterns”, and was aimed at investigating the role of contact patterns in the transmission of the influenza virus. As such, the first objective, measuring community contact rates and time use in two communities, was completed in 2018-2019. In 2020, public health priorities shifted from influenza to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and we decided to do the same for this thesis (appendix B). The remaining three objectives were then focused on the burden and transmission of SARS-CoV-2. The first objective, although completed before the SARS-CoV-2 pandemic, is still included as part of the body of work as contact patterns are important in the context of both SARS-CoV-2 transmission and the planning of intervention strategies.

The body of work submitted for the PhD degree consists of five manuscripts (four published in peer-reviewed journals and one full draft in peer-review, appendix D) and this integrating narrative chapter. The integrating narrative background section provides a literature review on SARS-CoV-2 emergence, burden and transmission, as well as contact patterns and their role in the transmission of SARS-CoV-2, as reported in the literature; and the problem statement and objectives for this study. This is followed by a summary of methods used for each of the four objectives, and a results and discussion section that integrates the results from the four objectives and five resulting manuscripts, discussing important considerations for SARS-CoV-2 burden and transmission, as well as future pandemics. The discussion is structured as six themes:

1. Combining data sources to estimate the burden of SARS-CoV-2
2. The household as an important setting for SARS-CoV-2 transmission
3. The influence of study design on the assessment of household transmission
4. Leveraging contact patterns to assess and predict SARS-CoV-2 transmission and to inform intervention strategies
5. The importance of not neglecting one epidemic when facing another: SARS-CoV-2 and HIV co-infection
6. Generating data for action

For the review of this body of work, the background in this integrating native can be read first, followed by the five manuscripts (see appendix D), and then returning to this chapter to review the methods summary, results and discussion.

2. Background

2.1. Epidemiology of SARS-CoV-2

2.1.1. Emergence

Coronaviruses are enveloped, positive-sense, single-stranded RNA viruses in the *Coronaviridae* family, a diverse family of viruses that can infect mammals, fish and avian species [1]. Human coronaviruses, for example, HCoV-229E and HCoV-OC43 cause mild respiratory tract infections, but only three coronaviruses have been highly pathogenic and caused outbreaks, all zoonotic in origin and emerging in the last 20 years [1]. Severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in November 2002 in Foshan City, Guangdong province, in China, and subsequently spread to 29 countries and regions, causing 8,096 cases and 774 deaths before the outbreak was declared over in July 2003 [2]. Sporadic cases were subsequently seen, but none after January 2004 [2]. Chinese horseshoe bats were found to be the natural reservoirs of SARS-CoV, with civet cats or masked palm civets likely responsible for the initial spill over to humans in wet animal markets [2]. Middle East respiratory syndrome virus (MERS-CoV) on the other hand was first detected in Saudi Arabia in June 2012, where direct spill-over from dromedary camels to humans occurred [3]. Unlike SARS-CoV, most infections arise from animal-to-human transmission, with a few subsequent outbreaks caused by human-to-human transmission mainly in hospital settings [3]. Although sporadic cases still occur, no large-scale outbreaks have been reported in recent years. Although most cases have been reported from Saudi Arabia, 2,260 cases and 803 deaths have been reported from 27 countries across the globe [3].

In December 2019 a cluster of pneumonia disease cases was reported from Wuhan, Hubei province, China [4]. Cases were linked to a seafood market [4], although not clear if the market was where the first human infection occurred, or only where the first cases were detected [5]. Although initially called new coronavirus (2019-nCoV), it was later classified as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent for coronavirus disease 2019 (COVID-19) [4]. Bats were identified as the likely reservoir for SARS-CoV-2, with the intermediate host still unknown [5]. Reverse transcription real-time polymerase chain reaction (rRT-PCR) assays were rapidly developed for diagnostics, and although a lockdown restricting movement in to and out of Wuhan was implemented on 23 January 2020, the virus had already spread to other parts of China, a situation exacerbated by the fact that Wuhan was a transport hub and there was a surge in travel due to the Chinese New Year celebrations [5]. Although contact tracing, isolation and social distancing was implemented in China, which successfully interrupted the chains of transmission and kept case numbers low [5], the virus had already spread globally where containment efforts were less successful. On 13 January 2020, the first case of COVID-19 was confirmed outside of China, in

Thailand, and by 30 January 2020 when the virus had already spread to 18 countries and 83 cases had been reported, the World Health Organization (WHO) declared the outbreak a Public Health Emergency of International Concern (PHEIC) [6]. COVID-19 was officially classified as a pandemic in March 2020, by which time it had already spread to 114 countries, with more than 118,000 cases and 4,291 deaths reported [7].

Like other RNA viruses, SARS-CoV-2 is prone to introduce errors during the replication of its genetic material, despite its proofreading ability [8]. Furthermore, when individuals are infected with more than one variant of the virus, an exchange of genetic material (recombination) can take place [8]. If these mutations provide a competitive advantage, the new mutant may become the dominant circulating form of the virus [8]. Strains with mutations that increase transmission, disease severity or immune-evasion are called variants of concern (VOC) and five have been identified so far: Alpha, Beta, Delta, Gamma, and the most recently emerged, Omicron [8]. The first documented samples of Beta were from South Africa in May 2020,[9] indicating that the variant likely emerged here. The Omicron variant was also first described from South Africa and Botswana in November 2021,[10] but the country of origin is still unclear due to Omicron sequences being detected in several other countries around the same time.[9] Each VOC that has emerged so far, has different transmissibility and severity characteristics, for example, Delta being more infectious and severe than Alpha and Beta, and Omicron being less severe than Delta, but evading immunity to cause more re-infections [8].

2.1.2. Containment and control strategies

The initial approach to control this newly emerged virus was to contain infections through intense testing and contact tracing, and large-scale lockdowns restricting population movement. Most countries were not successful in containing the virus immediately after introduction, and in the absence of vaccines, targeted therapies and an entirely naïve population, the only possible approach to control SARS-CoV-2 was non-pharmaceutical interventions (NPI). The aim was to flatten the curve: to reduce the number of infections at a current time point and rather spread infections out over time to not overburden the healthcare system, and to allow time for treatment and vaccine technologies to be developed [11]. Large-scale lockdowns came at a considerable economic cost and were only feasible for short periods at a time [11,12]. Other NPIs utilized included mandatory face masks while in public spaces, social and physical distancing, restriction in alcohol sales and increased hand hygiene [13]. In a very fast-paced changing landscape of the pandemic, it was important to have data available to quickly assess best-suited interventions and their success in reducing infections. NPIs were shown to be effective to reduce viral spread but did not eliminate transmission [11]. Individuals with public-facing occupations, especially healthcare workers, were still at greater risk of SARS-CoV-2 infection even while NPIs were in use [11]. Outbreaks still occurred in crowded living spaces and communal-living facilities, for example, prisons and nursing homes [11]. Adherence to NPIs would also have

differed based on the setting, limiting the effectiveness [14]. The wearing of masks and face coverings for example was shown to reduce the spread of SARS-CoV-2 but depended on the quality of the mask/covering, and the correct fitting and use [14].

The first vaccines for SARS-CoV-2 became available only a year after SARS-CoV-2 had emerged. Although vaccines were highly effective against severe disease and death, they did not prevent infection, especially as new variants of SARS-CoV-2 emerged, which meant that vaccination was no 'silver bullet' [11]. Further challenges include waning immunity, vaccine hesitancy and universal access, especially so in Africa [15]. The waning of immunity and emergence of new SARS-CoV-2 variants necessitates ongoing booster vaccination, and the development of new vaccines including antigens from circulating variants. Variant-modified vaccines incorporating the Beta, Delta, Omicron BA.1 or Omicron BA.4/5 spike proteins (including combined formulations) have been developed, but may not be available in all settings, and ongoing development would still be needed as new variants may emerge [16]. The requirement of regular booster doses places an additional burden on healthcare systems, further reducing access, and could also lead to communities losing confidence in vaccination. A universal coronavirus vaccine would resolve the issue of the continued evolution of SARS-CoV-2 but comes with both financial and technical challenges [17].

2.1.3. SARS-CoV-2 in South Africa

Surveillance for SARS-CoV-2 in South Africa started in early February 2020 when screening for fever was implemented for returning travellers at all ports of entry [18]. The first case of SARS-CoV-2 in South Africa was announced on 5 March 2020, three weeks after Egypt reported the first SARS-CoV-2 case in Africa [18]. After the detection of an additional 60 cases, a national state of disaster was announced and stayed in effect until 4 April 2022 [19]. This national state of disaster allowed for the framework to guide restrictions and NPIs in five different levels, level five being the most restrictive and only allowing the operation of and travel for essential services, and level one being the least restrictive [19]. Level five (a national lockdown) was implemented from 27 March - 30 April 2020 [19].

South Africa quickly established several surveillance platforms to track SARS-CoV-2 cases, hospitalizations and deaths. As part of the national state of disaster, COVID-19 has been declared a notifiable medical condition (NMC), with all laboratory diagnoses being reported to the notifiable medical conditions surveillance system (NMCSS) of the National Institute for Communicable Diseases (NICD). Further surveillance included community screening and testing, contact tracing and testing of symptomatic contacts, a comprehensive national electronic hospital surveillance system for COVID-19 (DATCOV) [20] and the notification of COVID-19 deaths from provincial authorities to the National Department of Health [18]. The NMCSS, DATCOV and COVID-19 deaths surveillance systems are not linked and operate independently. Linking data between these sources is theoretically possible but

would require manual linking through personal identifiers which may not be captured in a standard manner between the systems. Although the NMCSS includes all PCR-confirmed SARS-CoV-2 infections captured on laboratory information systems, the reporting of rapid antigen tests on these systems may not be consistent, and any infections not confirmed with a diagnostic test would also not be included. It will be important to understand the scale of underreporting in this surveillance system to estimate the burden of and population immunity to SARS-CoV-2. SARS-CoV-2 testing was also incorporated into the existing national syndromic surveillance for pneumonia (public sector hospitalizations), the out-patient influenza-like illness syndromic surveillance (public sector) and the viral watch (private general practitioners) surveillance systems [21]. Genomic [9] and wastewater surveillance were also established [22].

The surveillance systems allowed the careful tracking of South Africa's five waves of infection, each dominated by a different VOC: starting with the ancestral (also referred to as Wuhan) strain in wave 1 and followed by Beta-, Delta-, Omicron BA.1 and BA.2 and Omicron BA.4 and BA.5 in the next four waves [9] (Figure 1). By the end of September 2022, South Africa had reported 4.02 million laboratory-confirmed cases, with a cumulative incidence risk of 66,800 cases per 1 million population [23], 540,000 hospitalizations [20] and over 102,000 deaths [24].

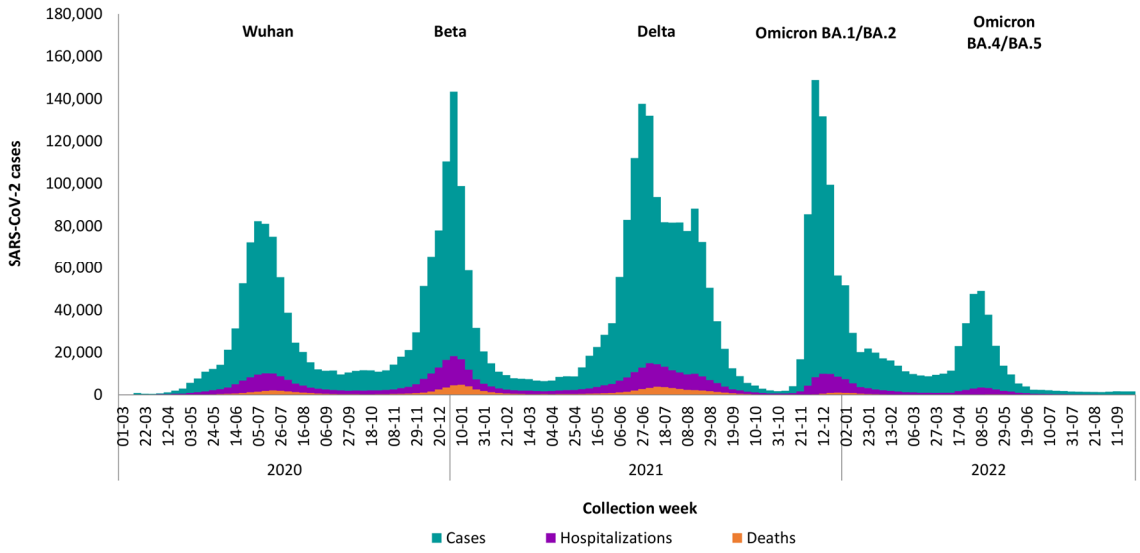


Figure 1. Laboratory-confirmed SARS-CoV-2 cases, hospitalizations and in-hospital deaths, South Africa, 1 March 2020 - 30 September 2022. Cases as reported to the Notifiable Medical Conditions Surveillance System (NMCSS) [23] and hospitalizations and deaths reported to the national electronic hospital surveillance system for COVID-19 (DATCOV) [20]. The variant which dominated the wave is indicated in the text above the curve [9].

2.2. SARS-CoV-2 burden and transmission dynamics

2.2.1. Burden

The burden of a disease is the combined impact of the disease on a population, which encompasses the prevalence, risk factors and outcomes of the disease, and the resulting financial implications [25]. The Global Burden of Diseases study summarizes the burden of disease from premature mortality and disability as disability-adjusted life years (DALYs), which includes years of life lost (YLL) and years lost due to disability (YLD) [26]. To calculate DALYs, one requires the number of cases and deaths, as well as the duration and severity of illness. The number of SARS-CoV-2 infections in a population can be estimated using the prevalence of SARS-CoV-2 antibodies (seroprevalence) in the population. These seroprevalence estimates can be used to calculate the total number of infections, and in turn, can be compared to the number of SARS-CoV-2 cases detected through surveillance (case-to-infection ratio, CIR) to estimate the number of infections reported to surveillance. To estimate severity, the number of infections can be compared to the number of SARS-CoV-2-related hospitalizations (hospitalization-to-infection ratio, HIR) to estimate the proportion of infections resulting in hospitalization and to the number of SARS-CoV-2-related deaths (fatality-to-infection ratio, FIR) to estimate the proportion of infections resulting in deaths.

By the end of September 2022, 614 million cases of SARS-CoV-2 infection had been recorded globally, with more than 6.5 million deaths [24]. The majority of cases had been reported from Europe and the Americas, with only 9.3 million cases and 174,500 deaths recorded from Africa [24]. During the initial phases of the pandemic, a major concern was that African countries' already fragile healthcare systems may be overwhelmed once widespread community transmission occurred, a phenomenon that occurred in Italy during the first wave of infections [27], and during the second, Delta dominated, wave in Delhi, India [28]. This, however, was not the case in Africa, and questions were raised if this was due to limited transmission, host genetics, a younger population, climate or under-ascertainment of cases.

While reported case numbers may point to Africa being spared a high case burden, seroprevalence studies rather show instead that SARS-CoV-2 was indeed circulating, and by comparing the number of infections based on seroprevalence to reported case numbers, it was apparent that infections rather went undiagnosed due to barriers to access to care, and limited testing capacities [29]. According to a World Bank Group Poverty and Equity survey in 11 sub-Saharan African countries during the early stages of the pandemic, half of the respondents who were unable to access healthcare indicated it was due to financial constraints. Furthermore, testing strategies focussed on symptomatic individuals, whilst a high proportion of infections are asymptomatic [29,30]. Although African healthcare systems did not collapse entirely, they were placed under major stress in many

settings [15,31], with resulting disruptions in services such as childhood vaccination, surveillance, and HIV and *Mycobacterium tuberculosis* (TB) testing and treatment programmes [32,33,34]. This highlights the difficulty in measuring burden when relying on only reported case numbers, with seroprevalence estimates indicating higher infection rates compared to reported case numbers, and the importance of active surveillance programs and seroprevalence studies to determine the true burden of SARS-CoV-2 infections. It is important to measure population SARS-CoV-2 seroprevalence to understand the existing exposure to SARS-CoV-2 in the specific setting when planning the relaxation of NPIs and implementation of interventions like vaccination. Furthermore, doing seroprevalence studies in longitudinal cohorts allow for serial comparisons of antibody responses relative to reported laboratory-confirmed SARS-CoV-2 through successive waves of infection, and to monitor the durability of antibody titres. Such data were still lacking in South Africa at the initiation of this study.

There is heterogeneity in SARS-CoV-2 seroprevalence in Africa based on the timing of the surveys relative to waves of infection, and on the population and area under investigation, thus; making it challenging to compare infection rates across studies. African seroprevalence estimates from one meta-analysis ranged from 1% in Togo (May 2020) to 45% in Nigeria healthcare workers (April 2020), with an estimated overall seroprevalence of 16% (including surveys done from April 2020 to November 2020) [29]. A meta-analysis using only WHO Unity-aligned study protocols reported a 3% SARS-CoV-2 seroprevalence in Africa by June 2020 [35]. In areas outside of the African continent, SARS-CoV-2 seroprevalence ranged from 3% in the Western Pacific region [36], with the lowest (0%) and highest (8%) seroprevalence in Australia (July 2020) [37] and Korea (June 2020) [38] respectively, to 15% in the Eastern Mediterranean [5] where the lowest (1%) and highest (38%) seroprevalence was recorded from Saudi Arabia (May 2020) [39] and Pakistan (July 2020) [40], respectively.

A meta-analysis mainly including high-income countries showed that the severity of the disease is greatly dependent on age: severe illness was estimated to occur in only 0.1% of SARS-CoV-2 infections in children younger than 10 years and increased with age to 30% in adults aged 80 years and older [41]. FIRs range from 0.0005% in young children to 18% in the elderly [41].

Seroprevalence studies were also performed in South Africa, mainly nested in blood donor samples, a population-based, cross-sectional study performed in the Gauteng province and a cross-sectional household survey in the KwaZulu-Natal, North West and Western Cape provinces. By May 2021 (after the Beta [second] wave), the national anti-SARS-CoV-2 antibody prevalence in blood donors (individuals >16 years) was 47% [42]. The estimates varied across provinces, from 32% in the Northern Cape province to 52% in the KwaZulu-Natal province. In the two study sites from which data was collected for this thesis, seroprevalence was 49% in the North West province and 48% in the Mpumalanga province [42]. When comparing the estimated number of infections based on seroprevalence in blood donors to reported case numbers, the case-to-infection (CIR) ranged from

16% in the Western Cape province, to 6% in the KwaZulu-Natal province, and 5% in both of the provinces of sites used on in this thesis (the North West and Mpumalanga provinces) [42]. Seroprevalence was higher in individuals of the black race compared to the white race, with no differences observed based on sex [42]. FIR based on blood donor seroprevalence (up to March 2021, also after the Beta wave) compared to reported excess deaths, increased with age, ranging from 0.014% in children 10-19 to 6.8% in adults ≥80 years, and differed based on province, ranging from 0.62% in the Mpumalanga province to 0.72% in the Western Cape province. In the North West province it was 0.64% and in the Gauteng province 0.62%. [43]

Results from the population-based, cross-sectional study done in Gauteng, showed that by January 2021 (during the Beta wave), seroprevalence was 19%, similar across all age groups [44]. Seroprevalence varied across districts, highest in the City of Johannesburg district at 26% and lowest in the City of Tshwane district at 13% [44]. Factors associated with previous infection were female sex, working in healthcare or production, and having more than one co-morbidity [44]. The CIR was estimated at 11% and the FIR based on COVID-19 recorded deaths was 0.28%, with excess deaths estimated at 0.67% [44].

In the cross-sectional household survey, also after the Beta wave of infection, overall seroprevalence was 45.2%, lowest in the North West province (39.8%) and highest in the KwaZulu-Natal province (50.2%) [45]. Factors associated with being seropositive included age, sex, race, being overweight or obese and low socioeconomic status [45]. In the North West province, the CIR, HIR, in-hospital FIR and excess death FIR were 4.4%, 2.5%, 0.3% and 0.3%, respectively [45].

2.2.2. Routes of transmission

Although zoonotic (animal to human), faecal-oral and fomite transmission of SARS-CoV-2 is possible and has been described in some instances, the main route of transmission for SARS-CoV-2 is respiratory [46,47]. When considering respiratory transmission, two main concepts are important: droplet-mediated transmission where the virus is present in drops of respiratory secretions that are expelled when talking, coughing or sneezing, and which travel only about 1 meter and fall to the floor due to their weight; or through aerosols where the virus can linger in the air for an extended duration of time [48]. For SARS-CoV-2, most evidence points to droplet-mediated transmission, where close proximity is important for transmission to take place, but aerosol transmission is also possible [46,48]. Ventilation seems to be key, and poor ventilation had resulted in high transmission events in households and public areas [46]. An important consideration, and one that added to the challenge of containing the virus, is that individuals can transmit SARS-CoV-2 irrespective of symptom status, and the risk of transmission is highest one day before symptom onset [46].

2.2.3. Transmission dynamics

The basic reproductive number (R_0) for an infectious disease is the average number of secondary cases that result from an infected individual in a completely susceptible population [46], and the effective reproductive number (R_e) considers infections once the population consists of both susceptible and non-susceptible individuals. R_e estimates for SARS-CoV-2 vary greatly, based on the current variant circulating, the NPIs in place and the existing population immunity. Considering the most recently emerged variants, the average R_e for Omicron was estimated at 3.4 (range 0.9 – 9.4), which was 2.5 times higher than the R_e for Delta [49].

A second important factor used to describe the transmissibility of a disease is the dispersion parameter (*k*), which describes the variation in the infectiousness of individuals. As with SARS-CoV and MERS-CoV, there is overdispersion in the transmissibility of SARS-CoV-2, which means there is great variation in secondary attack rates at the individual level, with some cases leading to no secondary transmission, while others may lead to super spreading events [46]. Due to this overdispersion, a small proportion of SARS-CoV-2 infections (10-20%) lead to the greatest proportion of secondary cases (80%) [46,50]. The occurrence of a super-spreading event depends on several factors, for example, the infectiousness of the index case (i.e. the number of infectious virions expelled), the contact patterns of the case (i.e. having a large number of contacts or more contacts with susceptible individuals), the timing of the contacts relative to infectiousness, as well as the environment (i.e. poor ventilation or crowding) [50,51].

The household has proven to be an important location for transmission [46,52] since secondary infection risk (SIR, the number of secondary cases arising from an index case) could be more than three times higher in household and family contacts compared to other contacts such as those that occurred in the workplace [53]. Household transmission can still drive epidemics as NPI may reduce transmission in the community, while household transmission can still occur and spill over to other households [54]. The household therefore constitutes a suitable environment in which to measure household cumulative infection risk (HCIR, the number of cases arising after the virus is introduced in the household, irrespective if these result from tertiary or subsequent chains of transmission), episode durations and serial intervals. However, a limitation of household transmission studies is that super spreading events, which may be driving most transmission in the community, cannot be considered in these small contact networks [46]. Within the household, it is possible to identify risk factors for transmission (characteristics of the index case increasing the transmission to a household contact) and for acquisition (characteristics of the household contact making them more susceptible to be infected) of SARS-CoV-2.

HCIR estimates differ based on timing, SARS-CoV-2 variant, previous SARS-CoV-2 infection and vaccination status of the index case and household contacts. The initial HCIR estimate from a meta-analysis was 19% [53] but was found to be higher with newly emerging variants: 36% for Alpha, 23% Beta, 30% for Delta and 43% for Omicron [55]. Estimates included in the meta-analysis varied greatly due to the different timing of the studies included, but most importantly the differences in methodology between these studies. HCIR estimates had not been available yet for South Africa at the initiation of our case-ascertained household transmission study, and these estimates are important to inform interventions to reduce the transmission of the virus. With South Africa's high burden of the human immunodeficiency virus (HIV) [56], doing household transmission studies can also shed light on the role of HIV-infection in SARS-CoV-2 transmission.

2.2.4. SARS-CoV-2 and HIV co-infection

South Africa has the fourth highest prevalence of people living with HIV (PLWH) in sub-Saharan Africa, with an estimated 19% of the adult population aged 15–49 years living with HIV [56]. Due to the impact of HIV on the immune system, PLWH may be at higher risk for SARS-CoV-2 infection and severe disease [57]. Few studies report the influence of HIV infection on SARS-CoV-2 transmission, duration of shedding or clinical severity, with most data available from high-income countries and little evidence from sub-Saharan Africa where most PLWH reside [56]. Prior research suggests that SARS-CoV-2 prevalence and COVID-19 diagnosis rates are similar between PLWH and people not living with HIV (PNLWH) [58,59]. This may point to no additional risk for SARS-CoV-2 acquisition in PLWH, but without a direct measurement (in a randomized control or observational study) of the acquisition of infection after exposure (irrespective of symptoms) it would not be possible to accurately determine this. Irrespective of the risk of transmission or acquisition, PLWH are at greater risk for hospitalization [57,59,60,61] and death [59,61,62,63,64] when infected with SARS-CoV-2, with the risk increasing with a decline in CD4+ T cells [57,59,64]. A meta-analysis found PLWH with COVID-19 had a 49% higher risk to be hospitalized and a 76% increased risk of death [61]. Data from South Africa showed that PLWH with COVID-19 had a 34% higher in-hospital mortality than individuals not living with HIV (LWH) [64]. Furthermore, the risk of in-hospital mortality was double in PLWH that were immunosuppressed compared to PLWH that were not immunosuppressed [64]. Limited data are available on the role of HIV in the transmission of SARS-CoV-2, which is also difficult to assess if not all contacts of cases are tested for SARS-CoV-2, irrespective of symptoms. Household transmission studies would be useful to assess both transmission from index cases LWH, and acquisition by contacts LWH. Although unknown at the inception of this study, we now know that PLWH with COVID-19 who are not virally suppressed for HIV, or immune compromised, shed SARS-CoV-2 for longer periods, which could also lead to increased secondary transmission, and evolution of the virus leading to new variants [30,57,65].

2.3. Social contact patterns

2.3.1. Measuring contact patterns

Methods utilized to investigate social mixing patterns include inferring contact patterns from secondary data like mobile phone usage or locations, the direct measurements of contacts using self-administered contact diaries (paper- or web-based), and interviews to complete contact diaries and proximity sensors, with each method having advantages and disadvantages (Table 1) [66]. Contact diaries make use of either online or paper-based surveys that request participants to report any contact events for a specified period, while proximity sensors measure any interactions between individuals wearing these sensors, eliminating the issue of recall or reporting bias [66].

Table 1. Methods commonly utilized to measure social contact patterns

Method	Description	Advantages	Disadvantages
Mobile phone usage [67]	Using mobile network tower locations that mobile phones are connected to, to infer interactions based on location	• No resources are needed for field deployment • Provides data in absence of contact survey data	• Very low-resolution data • Limited to individuals using mobile phones (i.e. adults and higher income groups) • Access to data needs special permissions • No metadata (e.g. ages) are available • High-resolution location data (home/work) not available
Global positioning system (GPS, mobility data) [68]	Using either a GPS-logging device or smart phone GPS to infer interactions based on location	• No resources are needed for field deployment • Higher resolution than phone usage • Provides data in absence of contact survey data	• Low-resolution data • Limited to individuals using mobile phones (i.e. adults and higher income groups) • GPS loggers reliant on individuals remembering to charge and use • The signal resolution of GPS can be low
Contact diaries (self-administered) [69,70,71,72]	Individuals complete a paper-based or electronic contact diary	• Large sample size • The study period can be longer	• Cannot be used in illiterate communities • Electronic diaries cannot be used in populations with no access to devices • Recall bias • Over-estimates contact duration • Under-estimates contact frequency
Contact diaries (interviews) [68,69,70,72]	Individuals are interviewed to complete a contact diary of contacts from the previous day	• Can be used in illiterate communities • Possibly less recall bias than self-administered diaries	• Recall bias • Field staff are required thus increasing the cost • Typically, only collect information for one day • Smaller sample size than self-administered • Over-estimates contact duration • Under-estimates contact frequency
Proximity sensors [70,72,73]	Wearable devices send and receive radio frequency waves to directly measure close proximity (<1.5m) interactions	• Direct measurement of close interactions that could lead to transmission • Removes recall bias	• Only measures contacts between individuals wearing sensors • Only measures face-to-face contacts • Deployments are logistically challenging • Depends on compliance • Hardware costs • Typically smaller sample size than contact diaries

Proximity sensors have been employed to measure contact events in workplaces [74], conferences [70,74,75,76], workshops [77], schools [69,78,79,80,81,82,83], universities [84], museums [75], hospitals [85,86,87,88,89] and households [90,91]. High-resolution contact patterns have become a topic of interest when considering infectious disease transmission, have been used in combination with virological data to investigate influenza transmission in a hospital [89], and more indirectly to model transmission of infectious diseases using collected contact data and known infectious disease parameters [76,85]. Specifically, in the context of SARS-CoV-2 transmission, proximity sensors have been used to measure contact patterns on a cruise ship, enabling the simulation of SARS-CoV-2 outbreaks to measure the effectiveness of NPIs [92].

Low-resolution methodologies like mobile phone and GPS tracking can be used to track the movement of individuals (how often and far do individuals travel?) and simulate how an infection could disseminate through a population, country, or even between countries, but does not provide details on how a pathogen might spread on an individual level [68]. Using contact diaries, more resolution can be obtained about individual-level contact patterns (who is contacted, how often, where and for how long?), although this method is subject to biases in the measurements [72]. With proximity sensors, the actual interaction between two individuals is measured and provides the exact time of the interaction and how long it occurred, although only logged for close-range, face-to-face interactions of two individuals both wearing the sensors. With the availability of these very high-resolution data, there are several ways (parameters) to summarize the data. The contact rate (total number of contact events, i.e. frequency) is a common measure used in both contact diaries and proximity studies [82]. The number of people contacted and duration of contact are also parameters that can be used to describe social contact patterns [82] [87]. A contact matrix is a summary measure to describe the average (or median) time in contact or the number of contact events between groups, often done using age categories [87] [83]. Published contact matrices provide a summary of contact patterns in a population that can be used for infectious disease transmission modelling.

Although social contact studies using proximity sensors have not been done in South Africa, contact diaries have been utilized to investigate contact patterns in Western Cape communities and one community in the KwaZulu-Natal Province. One study in the Western Cape province [93] and the KwaZulu-Natal province [94] study conducted interviews to assess contacts, and the other from the Western Cape province utilized a self-administered diary and an interview to assess contact events in all age groups [95]. It is unknown if the contact patterns observed in these communities are similar to those in other South African communities, and these contact studies were not performed in the context of virus transmission. It is important to have robust estimates of contact rates in multiple communities in South Africa to assist in the mathematical modelling of not only SARS-CoV-2

transmission, but also the transmission of other pathogens, and the impact of possible interventions on transmission.

2.3.2. The role of social contact patterns in SARS-CoV-2 transmission

Close proximity interactions play an important role in the transmission dynamics of SARS-CoV-2 [46]. This necessitated the implementation of both physical distancing (keeping a physical distance of >2 meters from others) and social distancing (limiting the interaction opportunities with others) in the NPI arsenal used globally to combat the spread of SARS-CoV-2. An Italian study showed the reduction of contacts with the implementation of restrictions, which translated to a decrease in the SARS-CoV-2 reproductive number [96]. While mobility and contact survey data showed that the implementation of NPIs led to a reduction in community contacts, for example, contacts at school, workplaces and social settings [97,98], and in turn opportunity for infection, it increased household contacts [54]. Although households are a focal point for SARS-CoV-2 transmission [52,97], especially during peaks of NPI restrictions when movement outside of the household is limited [99] and can in turn lead to spill over to other households [54], high-resolution household contact patterns have not yet been investigated in the context of SARS-CoV-2 transmission.

2.4. Problem statement and justification for research

Social contact patterns are an important consideration when investigating the transmission of SARS-CoV-2. In South Africa, a robust estimation of contact patterns within different communities and households is lacking, and the linking of these contact patterns to SARS-CoV-2 transmission events is poorly understood globally. To work towards filling this gap, we measured both household and community contact patterns to assess the association of contact patterns with SARS-CoV-2 transmission, and also to add to the limited literature on contact patterns in South Africa.

To estimate the burden of SARS-CoV-2, the number of infections (i.e. community attack rates) and the severity of these infections, need to be established, in essence building the burden of disease pyramid. We estimated the number of SARS-CoV-2 infections in two communities using seroprevalence and related this to hospitalizations and deaths. After gaining a better understanding of the community attack rates and severity of SARS-CoV-2, we investigated the transmission dynamics of SARS-CoV-2, focusing on the household as this forms a suitable environment to closely track transmission and acquisition of the virus.

Understanding the community attack rates of SARS-CoV-2 and drivers of SARS-CoV-2 transmission in the household, especially contact patterns, can help inform NPIs for future SARS-CoV-2 resurgences and potentially possibly future emerging pathogens with pandemic potential. Furthermore, due to

the high burden of HIV in sub-Saharan Africa, it is important to understand if there are differences in SARS-CoV-2 transmission in PLWH compared to PNLWH.

2.5. Aim and objectives

The overarching aim of this thesis was to describe the SARS-CoV-2 community attack rates and household transmission dynamics in PLWH and PNLWH and to assess the association of household contact patterns with SARS-CoV-2 transmission. In the study performed before the SARS-CoV-2 pandemic (objective 1), we focussed on describing contact patterns, which could be used in future studies for modelling transmission dynamics. In the studies performed during the SARS-CoV-2 pandemic (objectives 2-4), we described the burden and transmission dynamics of SARS-CoV-2.

1. To quantify within-household and community age-specific contact rates in a rural and an urban community in South Africa, using a contact diary survey (before the SARS-CoV-2 pandemic)
2. To assess differences in SARS-CoV-2 community attack rates in PLWH and PNLWH and disease severity across waves of the pandemic using seroprevalence
3. To determine the HCIR of SARS-CoV-2 in household contacts of index cases LWH and NLWH and assess factors associated with transmission, acquisition, duration of SARS-CoV-2 positivity and serial interval
4. To assess the association between contact patterns and the transmission of SARS-CoV-2 within households

2.6. Conceptual framework

The basis of this study stems from the disease burden pyramid, where the main public health endpoints usually measured (death, hospitalizations and diagnoses) form only the tip of the iceberg, but transmission is driven by the bottom of the pyramid: infections (Figure 2). Especially for SARS-CoV-2, this is important as the symptomatic fraction is small [30], and reported case numbers are likely an underestimate of the true burden. Furthermore, asymptomatic individuals can transmit the virus [46], which makes both symptomatic and asymptomatic infections important drivers of transmission. For the second objective, we estimated the SARS-CoV-2 community attack rate using seroprevalence in two communities, stratifying by age and HIV infection. This allowed us to estimate the bottom of the pyramid, including all infections, irrespective of symptoms. From there we will assess the proportion of infections resulting in hospitalization and death.

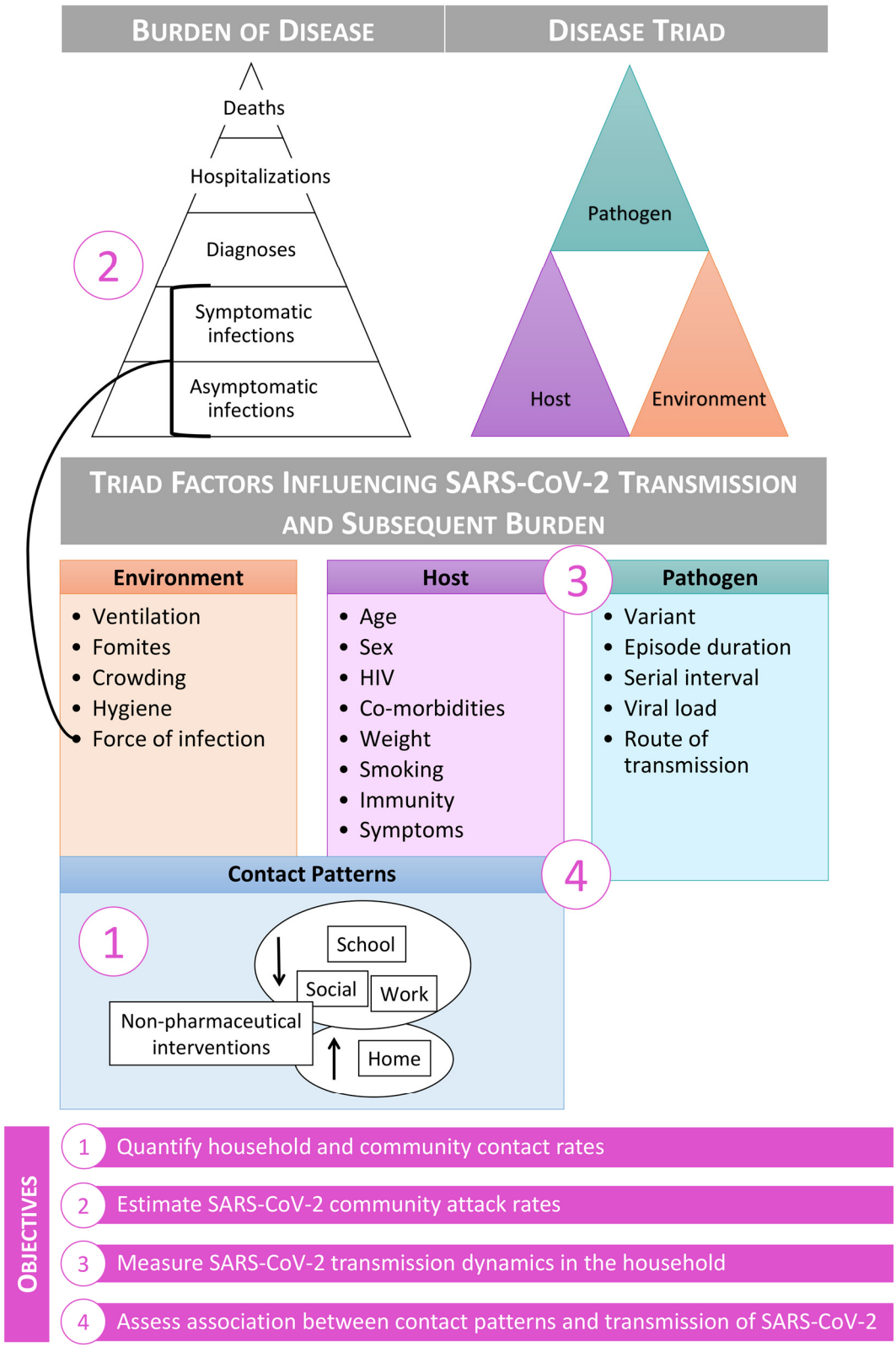


Figure 2. The conceptual framework for community burden and transmission dynamics of SARS-CoV-2. Factors influencing transmission are coloured based on the disease triad: environment (orange); host (purple); and pathogen (teal). Contact patterns (blue) span both environment and host, with non-pharmaceutical interventions reducing community contacts and increasing household contacts. Both symptomatic and asymptomatic infections can lead to transmission, translating to the force of infection in the community. Circled numbers (pink) refer to the objectives of this study listed at the bottom of the figure.

When considering the transmission dynamics of SARS-CoV-2, one has to consider the interplay of the human host, SARS-CoV-2 (pathogen) and the environment in the disease triad (Figure 2). Community transmission (bottom of burden pyramid) can be considered the force of infection for the introduction of SARS-CoV-2 in the household. We will consider the force of infection to form part of the environment category is we consider it from the perspective of an individual being exposed to infections from their environment. Also included in this conceptual framework but not investigated in this study is ventilation within indoor spaces visited, contact with fomites that could lead to infection, crowding within the spaces visited within the environment and hygiene and sanitation practices within the community. Following the introduction of SARS-CoV-2 in the household, one needs to consider host demographics, social habits and underlying conditions (host category), along with viral transmissibility (pathogen category), when considering the transmission from index cases and the acquisition by household contacts, which in turn can seed community transmission [46,54]. In the third objective of this study, we investigated the transmission dynamics of SARS-CoV-2 in the household. When considering the interaction of the host within the environment, contact patterns play an important role in transmission [46], specifically considering that NPIs may be reducing contacts in the community but potentially increasing contacts in the household [54]. As part of the first and last objectives in this study, we described both community and household social contact patterns in two communities before the pandemic (objective one), and the association of close-proximity events within the household with SARS-CoV-2 transmission (objective four).

3. Methods

Data used for this thesis were generated from three different studies, summarized in Table 2 and collection points indicated in Figure 3. Methodologies for each objective are summarized below and detailed in the respective manuscripts included as appendices.

3.1. The PHIRST study: Measuring contact patterns using contact diaries

The **P**rospective **H**ousehold observational cohort study of **I**nfluenza, **R**espiratory **S**yncytial virus (RSV) and other respiratory pathogens community burden and **T**ransmission dynamics in South Africa (The PHIRST Study) was a three-year study where 50 households at each of the urban (Jouberton, North West Province) and rural (Agincourt, Mpumalanga Province) sites were enrolled and followed up for 10 months from 2016 through 2018, with nasopharyngeal swabs collected twice a week (irrespective of symptoms) and tested on rRT-PCR for influenza and RSV. Since PHIRST was primarily aimed at investigating the household transmission of influenza, households were eligible if there were ≥3 or more household members, and if 80% of the household consented to be included in the study.

Households with <3 household members would not have been effective to include considering the minimum amount of data that could be added, and households where <80% of members were included would have resulted in an incomplete picture for household transmission. Methods for the PHIRST study are described in detail elsewhere [100]. For objective one of this thesis, to measure contact patterns (done before the SARS-CoV-2 pandemic), we nested a cross-sectional contact survey in the 2018 cohort from August through October 2018. We interviewed the 535 consenting study participants about their time use and contacts encountered on one randomly assigned day of the week. During the interview, we collected the age and sex of the person in contact with, if the contact occurred in an enclosed space or not, whether there was physical contact, if the person contacted was a member of the participant's household, total time spent with the person during the day, where the contact took place and how often the participant came into contact with the person. We described self-reported contact rates as the median number of people contacted per day, assessed differences in contact rates based on participant characteristics using quantile regression and assessed differences in contact rates based on contact characteristics within age groups using a Poisson model [101].

3.2. The PHIRST-C study: Measuring burden through seroprevalence

In light of the SARS-CoV-2 pandemic, the PHIRST cohort was approached and re-enrolled in 2020 in the **P**rospective **H**ousehold study of SARS-**C**oV-2, **I**nfluenza, and **R**espiratory **S**yncytial virus community burden, **T**ransmission dynamics, and viral Interaction in South Africa (PHIRST-C), also described in detail elsewhere [30]. We enrolled 100 households at each study site and collected two monthly blood samples from July 2020 through November 2021. For objective two of this thesis, we determined previous SARS-CoV-2 infection by using the Roche Elecsys anti-SARS-CoV-2 assay (Roche Diagnostics) to detect antibodies against the SARS-CoV-2 nucleocapsid protein. We considered a cut-off index (COI) >1.0 as an indication of prior infection (i.e., seropositivity). We calculated seroprevalence (by age and HIV status) by dividing the number of seropositive individuals by the number of individuals tested and adjusted for assay sensitivity and specificity using a Bayesian model as previously validated for SARS-CoV-2 and described [102]. To obtain credible intervals, Bayesian inference with 10,000 posterior draws was performed [102]. Using data from national SARS-CoV-2 surveillance, we divided the provincial surveillance laboratory-confirmed cases, hospitalizations and deaths by the age-standardized seroprevalence estimates to calculate CIR, HIR, and in-hospital and excess death FIR. The percentage under-reporting to national surveillance was calculated as 100 – CIR. We compared the seroprevalence and ratios between sites and across the first three waves of infection [103,104].

3.3. SA-S-HTS: Investigating SARS-CoV-2 transmission dynamics

For objectives three and four of this thesis, we performed a case-ascertained prospective observational household transmission study (The South African SARS-CoV-2 household transmission study, SA-S-HTS). We screened for SARS-CoV-2 cases aged ≥18 years with symptom onset ≤5 days before screening at primary public health clinics in Jouberton (North West Province) and Soweto (Gauteng Province). We approached households of individuals who tested positive for SARS-CoV-2 with known HIV status, COVID-19 symptom onset <7 days before household enrolment and no household members reporting symptoms in the 14 days before index screening. We enrolled 131 index cases and 457 household contacts and followed up the households for 6 weeks with thrice weekly visits to collect nasal swabs for SARS-CoV-2 testing (irrespective of symptoms). For objective three of this thesis, we calculated HCIR, episode duration and serial interval (stratified by households where the index was LWH and NLWH) as described below:

Calculations used to determine household transmission dynamics

$$HCIR = \frac{infected\ household\ contacts^{*}}{number\ of\ suceptable\ household\ contacts^{**}}$$

$$Episode\ duration = date\ of\ last\ positive\ swab - date\ on\ symptom\ onset$$

$$Serial\ interval = date\ of\ symptom\ onset\ in\ contact - date\ of symptom\ onset\ in\ index$$

HCIR: Household cumulative infection risk. *SARS-CoV-2 infection based on rRT-PCR on a nasal swab. **Susceptibility based on seronegative baseline SARS-CoV-2 serology. If seropositive for SARS-CoV-2 antibodies with no positive swab, not included in the denominator, if seropositive with a positive swab, included in numerator and denominator.

We opted to calculate HCIR instead of SIR, as there could have been more than one chain of transmission within the household, and the methodology used in this study cannot account for this and would lead to an over-estimation of the SIR. Although previous studies have opted to use only clinical data (symptom onset and end dates) for the calculation of episode duration, the duration of SARS-CoV-2 shedding could continue after symptoms have resolved [105]. One advantage of our study design is that we had the date of the last positive swab, and could therefore calculate the episode duration up to the cessation of shedding. However, for the serial interval, we opted to use only clinical data. Although we had the date of first positivity in the secondary cases, the date of first positivity for the index was not known. We therefore only calculated serial intervals for symptomatic secondary cases, using the date of symptom onset in the index to the date of symptom onset in the secondary case.

Table 2. Studies that contributed data to this thesis. Data included in the thesis indicated in bold face text

	PHIRST	PHIRST-C	SA-S-HTS
Design	Prospective household cohort study	Prospective household cohort study	Prospective household transmission study
Primary objective of original study	Estimate the community burden of influenza and RSV	Estimate the community burden of SARS-CoV-2	Determine duration of shedding, temporal variations in SARS-CoV-2 viral load, HCIR, serial interval and antibody responses by HIV-infection
PhD contribution	Objective 1 (Quantify within-household and community age-specific contact rates)	Objective 2 (Assess differences in SARS-CoV-2 community attack rates by HIV-infection and disease severity across waves of the pandemic)	Objective 3 (Determine HCIR by index HIV- infection, assess factors associated with transmission) and 4 (Assess association between contact patterns and transmission of SARS-CoV-2 within households)
Selection	Random selection of households	Random (PHIRST + additional households)	Case-ascertained
Sites	Jouberton (urban, NW) and Agincourt (rural, MP)	Jouberton (urban, NW) and Agincourt (rural, MP)	Jouberton (NW) & Soweto (GP)
Period	2016-**2018***	**2020-2021**	**2020-2021**
Sample size	50 households/site/year	100 households/site	60 households/site
Age distribution	At least 50% of households with at least one child <5 years of age	No age criterion for additional households	Index case ≥18 years, no age criterion for household contacts
Follow-up**	10 months (synchronous)	13 months (synchronous)	6 weeks (asynchronous)
Contact survey	**Once off, 2018**	None	None
Contact sensors	2 weeks, 3 deployments	None	**2 weeks, 1 deployment**
Viral pathogens tested for	Influenza, RSV	**SARS-CoV-2**, influenza, RSV	**SARS-CoV-2**
Nasal swab frequency	2x week	2x week	**3x week**
Serology testing frequency	Pre- & post- RSV & influenza seasons	**2 monthly**	**Baseline and 6 weeks**
Ethics***	150808	150808	M2008114
Candidate role	Protocol amendment, project management, training, data management & analysis	Protocol review, project management, training, data management & analysis	Protocol review, project management, training, data management & analysis

PHIRST: The Prospective Household observational cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa.
PHIRST-C: The Prospective Household study of SARS-CoV-2, Influenza, and Respiratory Syncytial virus community burden, Transmission dynamics, and viral Interaction in South Africa
SA-S-HTS: The South African SARS-CoV-2 household transmission study.
HCIR: Household cumulative infection risk
GP: Gauteng province, MP: Mpumalanga province, NW: North West Province.
*Only data from 2018 was used from the PHIRST study.
**Synchronous: Households enrolled and followed up at the same time; Asynchronous: Households enrolled and followed up during different periods, i.e. staggered follow-up for a fixed period after the clinical presentation of the index case.
***University of Witwatersrand Human Research Ethics Committee reference numbers for original study protocol, overall study protocol reference number: M191120.

To assess factors associated with HCIR, we used logistic regression accounting for within-site and household clustering using a mixed-effects hierarchical regression model. To assess factors associated with a time-to-event analysis (serial interval and episode duration), we used a multilevel mixed-effects survival model with Weibull accelerated failure time analysis.

During the first two weeks of follow-up, we deployed wearable proximity sensors to measure close-range proximity events (face-to-face interactions) between household members. For objective four of this thesis, we calculated the duration, frequency and average duration of close-range proximity events of each household member with the SARS-CoV-2 index case. We assessed the association of contact parameters with SARS-CoV-2 transmission using mixed effects logistic regression accounting for index and household member characteristics. We also constructed contact matrices by combining the median duration and frequency of close-range proximity events for all participants between each age group, respectively.

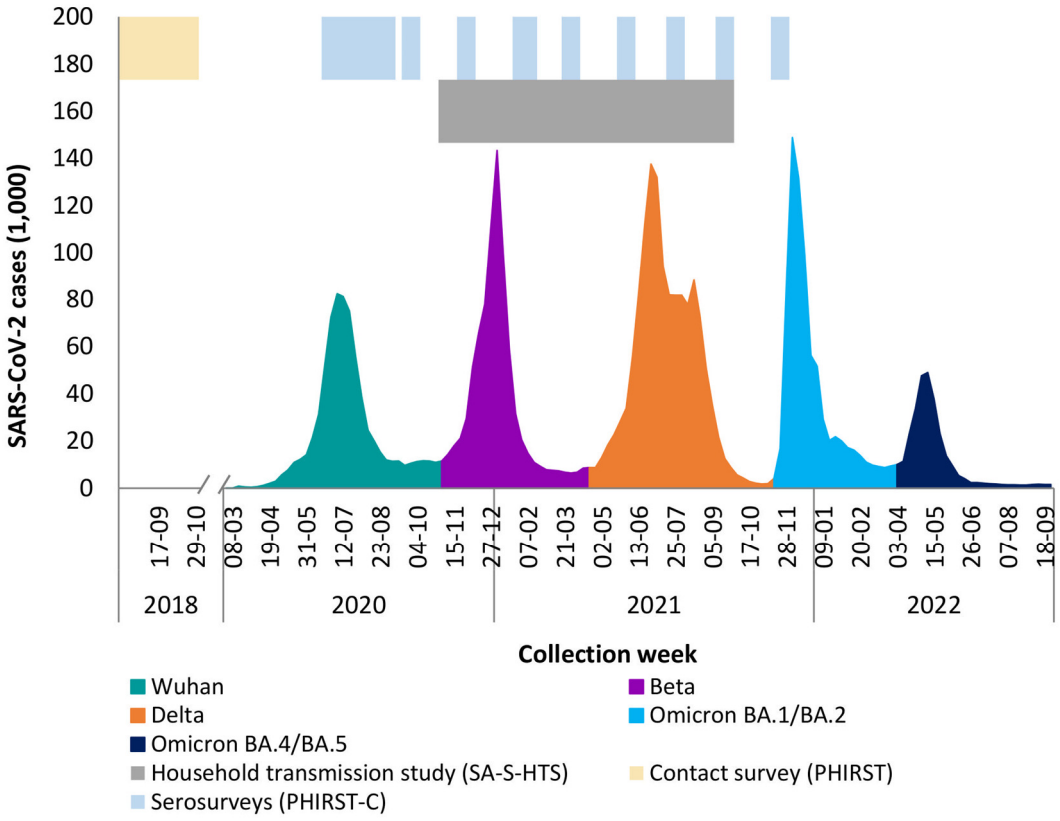


Figure 3. Timing of data collection relative to laboratory-confirmed SARS-CoV-2 cases (in thousands) reported to Notifiable Medical Conditions Surveillance System (NMCSS), South Africa, 1 March 2020 - 30 September 2022. Data collection periods indicated as pastel-coloured areas for the contact survey used for objective one (light yellow), serosurveys used for objective two (light blue), and household transmission study used for objectives three and four (light grey). SARS-CoV-2 case numbers[23] shaded by variant which dominated the wave[9].

Furthermore, we assessed factors associated with HCIR using logistic regression, and factors associated with prolonged episode duration and serial interval using Weibull survival models [106]. For

objective four of this thesis, we described age-based household contact patterns and assessed the association of the median number of contacts, the median duration of contacts, and the average duration of contact events with the index case and all SARS-CoV-2-infected household members, with the transmission of SARS-CoV-2. We used the Wilcoxon rank-sum test and logistic regression, controlling for individual characteristics, to assess the association of contact parameters with SARS-CoV-2 transmission (appendix D5).

3.4. Ethics

The overall study protocol for this thesis was approved by the University of the Witwatersrand Human Research Ethics Committee (reference M191120). Written informed consent was obtained from all study participants or their parent or legal guardian for children younger than 18 years, and additional assent from children aged 7-17 years. Individual study protocols were also approved by the University of the Witwatersrand Human Research Ethics Committee for the PHIRST and PHIRST-C studies (150808) and the SA-S-HTS study (M2008114).

4. Results and Discussion

In this section, we highlight key results from the five manuscripts and discuss six overarching themes in the context of SARS-CoV-2 burden and transmission. We start with how the availability of multiple data sources allows for the estimation of disease burden, then move on to the household transmission of SARS-CoV-2 and how study design needs to be considered when interpreting results from these studies. We also discuss how contact patterns influence SARS-CoV-2 transmission, the role of HIV in SARS-CoV-2 transmission, and lastly the importance of readily available data to make evidence-based decisions during a pandemic.

4.1. Combining data sources to estimate the burden of SARS-CoV-2

During the SARS-CoV-2 pandemic, seroprevalence estimates proved to be crucial to monitor the exposure of the population to the virus and planning for public health action. By collecting longitudinal serology samples in two communities over the first three waves of infection (Figure 4), before SARS-CoV-2 vaccination was widely available, we were able to monitor the seroprevalence of SARS-CoV-2 antibodies after each of the Wuhan, Beta and Delta waves of infection, and observe the different infection levels between two communities in South Africa (Figure 5) [103,104].

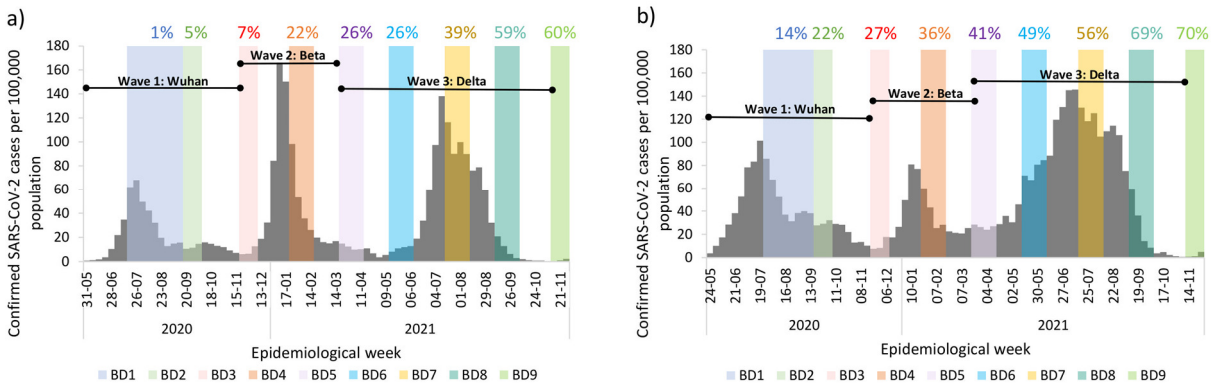


Figure 4. Timing of blood collection, district SARS-CoV-2 weekly incidence and seroprevalence in all ages during blood draw in a) a rural community and b) an urban community, March 2020 – November 2021, South Africa. Lines indicate the period used for analysis, and coloured text above each blood draw (BD) period shows SARS-CoV-2 seroprevalence in all ages during the draw. District incidence based on laboratory-confirmed SARS-CoV-2 infections (reverse transcription polymerase chain reaction) reported to National Notifiable Medical Conditions Surveillance System [104]. The epidemiological week is expressed as day-month of the week's starting date.

From 553 individuals in the rural community and 499 individuals from the urban community, we saw seroprevalence at the first blood draw (collections up to 17 September 2020), to be only 1% (95% credible interval [95% CrI] 0%–2%) in the rural community and 15% (95% CrI 12%–18%) in the urban community (Figure 5). By looking at seroprevalence estimates from samples collected from 23 November through 12 December 2020 (blood draw 3), we were able to see how the rural community had lower infection levels in wave one (7%, 95% CrI 5%–9%) compared to the urban community (27%, 95% CrI 23%–31%), likely due to the isolated nature of the rural community. Furthermore, we were able to monitor how this changed after the Beta-dominated (second) wave when seroprevalence more than tripled in the rural community (26%, 95% CrI 22%–29%), likely due to the influx of people during the December summer holidays, and the emergence of the Beta variant [103]. After the Delta-dominated (third) wave, from samples collected from 22 March through 11 April 2021 (blood draw 5), we observed more similar estimates between the two communities, 60% (95% CrI 56%–64%) in the rural community and 70% (95% CrI 66%–74%) in the urban community (Figure 5) [104].

We could also describe the changing seroprevalence between age groups through the waves, and how this differed between the two communities (Figure 5). At the end of the Wuhan-dominated (first) wave, seroprevalence was highest in individuals aged 19-34 years in the rural community (12%, 95% CrI 4%-23%), and in those aged 35-59 years in the urban community [103]. Although children initially had lower attack rates (based on seroprevalence) in the Wuhan- and Beta-dominated waves, by the end of the Delta-dominated (third) wave, the highest seroprevalence in both the rural and urban community was in those aged 13-18 years, at 80% (95% CrI 70%-88%) and 83% (95% CrI 73%-90%), respectively [104].

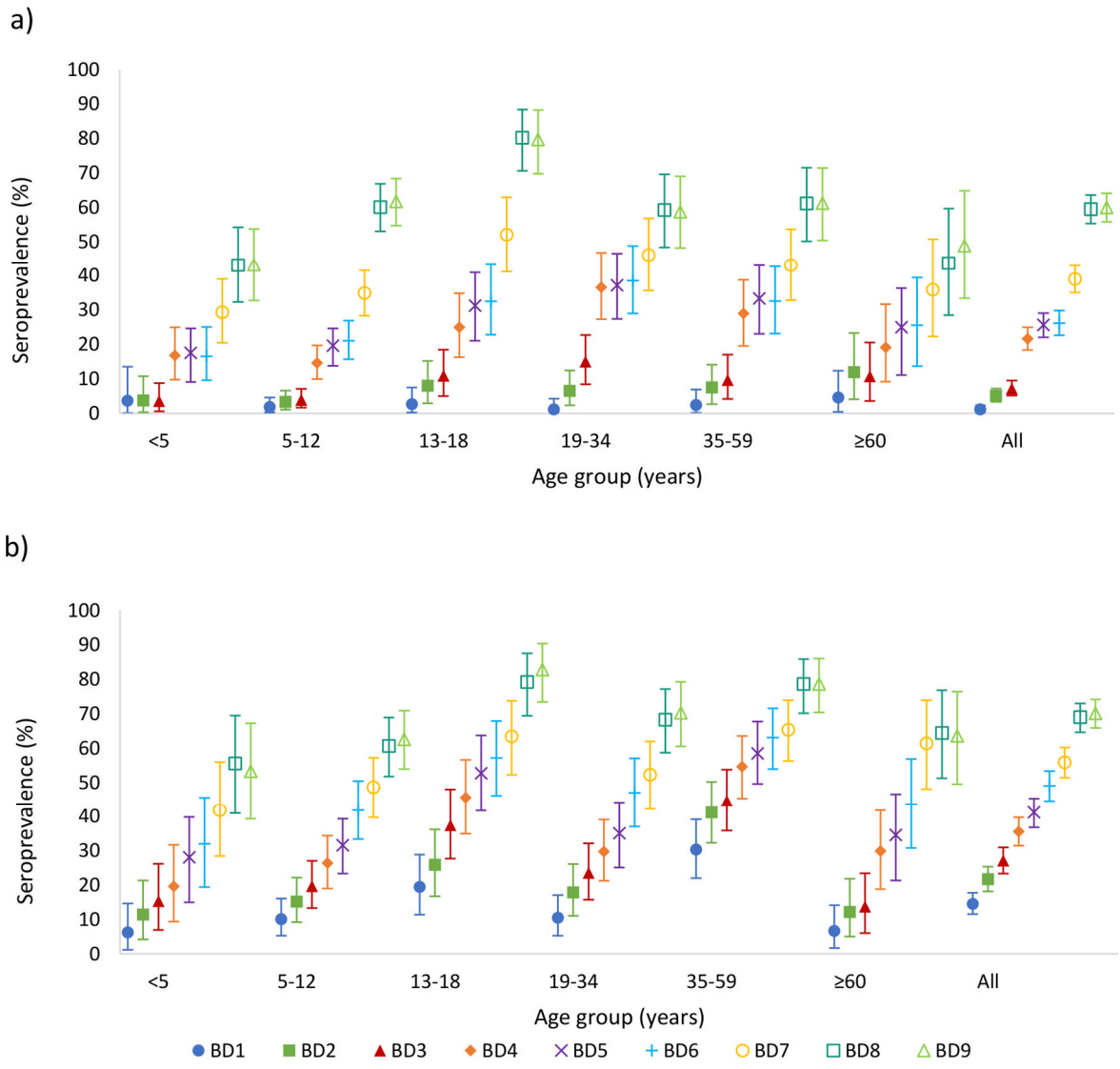


Figure 5. SARS-CoV-2 seroprevalence at each blood collection, by age group, in a) a rural community and b) an urban community, South Africa, March 2020–November 2021. Baseline blood draw (BD1) collected 20 July – 17 September 2020; second draw (BD2) 21 September – 10 October 2020; third draw (BD3) 23 November – 12 December 2020; fourth draw (BD4) 25 January – 20 February 2021; fifth draw (BD5) 22 March – 11 April 2021; sixth draw (BD6) 20 May – 9 June 2021; seventh draw (BD7) 19 July – 5 August 2021; eighth draw (BD8) 13–25 September 2021; ninth draw (BD9) 15–27 November 2021. Error bars represent 95% credible intervals. Seroprevalence estimates adjusted for the sensitivity and specificity of the assay.[104]

Although the longitudinal serosurveys in the well-defined and characterized PHIRST-C cohort have the advantage of being able to monitor infection trends over time, and to link this to rRT-PCR confirmed infections, they come at great cost and with logistical challenges, which results in a smaller sample size, and possible lack of generalizability both to other communities and to comparability with other published studies. The diversity between just these two communities does highlight the importance of serosurveys in different populations within different countries, especially in the early

stages of a pandemic where exposure may still be low and containment efforts are still being considered. Age-based seroprevalence estimates after each wave also allow for the identification of age groups where exposure may have been low, and additional interventions (for example vaccination) may be needed to reduce severe disease in naïve populations.

As seroprevalence estimates increase due to wide-spread transmission, longitudinal surveys would become even more important (compared to cross-sectional surveys) in distinguishing recent from past infections. By comparing antibody levels between our first and fifth blood draws we saw that 93% of participants infected during the first wave remained seropositive after more than six months, although we did not account for possible repeat infections within this period [103]. The additional advantage of a cohort with longitudinal seroprevalence studies is that antibody kinetics can be investigated and analytic approaches or thresholds put in place to distinguish between previous and recent infections. This has been done for the PHIRST-C cohort, where boosting of antibody levels was used to infer re-infections during the Omicron wave in South Africa, and we found that the attack rate for Omicron was around 60%, with more than half of infections being re-infections [107].

In additional to estimating the number of infections in a population, seroprevalence data can also be leveraged to investigate the severity of a disease. We did this by comparing the seroprevalence estimates from the PHIRST-C cohort to provincial surveillance data, to calculate CIR, HIR and FIRs. By comparing hospitalization and fatality ratios to the number of infections rather than laboratory-confirmed cases, more robust estimates can be obtained to assess the severity of disease, and also to compare this across settings, countries and even other pathogens. We observed CIRs ranging from 3% to 5%, HIRs ranging from 0.4% to 3% and FIRs ranging from 0.1% to 0.7% across the first three waves of infections and between the two communities (Figure 6) [103,104]. FIRs fall within the same range as reported for Gauteng Province, South Africa (0.3% to 0.7%) [44], and all three estimates are in line with global estimates of 7%, 1% and 0.4%, and sub-Saharan estimates of 0.7%, 0.6% and 0.2% for CIF, HIR and FIRs [108]. South African CIR estimates are higher than for the rest of Africa, likely due to higher testing capacity. Being able to calculate these ratios is only possible if both seroprevalence and surveillance data are available, and highlights the importance of the routinely measured public health endpoints and surveillance platforms. And beyond that, the availability of these data, and collaborations between groups to be able to evaluate these crucial epidemiological parameters, enable monitoring of potential changes in the severity of disease both at the beginning and through a pandemic. For example, the seroprevalence estimates after the third SARS-CoV-2 were useful in interpreting the lower severity of the fourth, Omicron BA.1/2-dominated wave in South Africa [109,110].

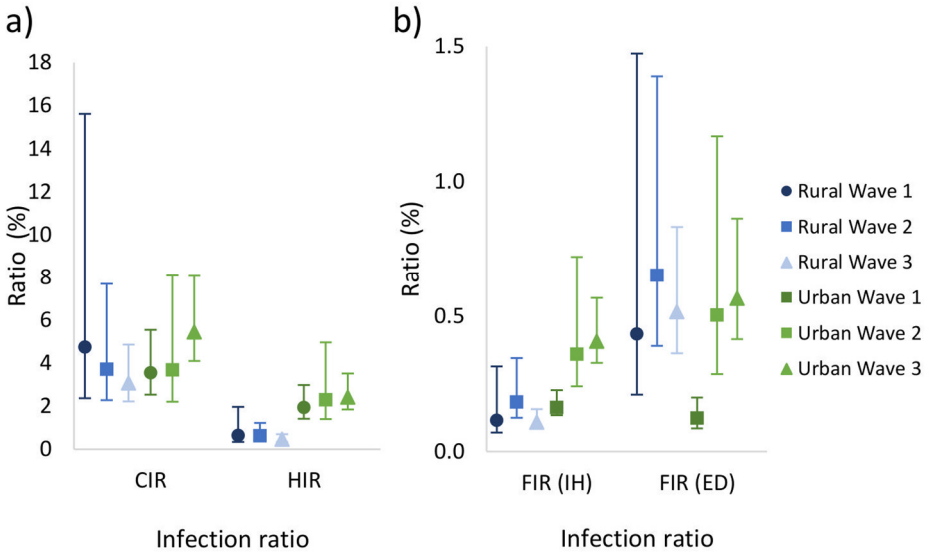


Figure 6. SARS-CoV-2 a) case-to-infection (CIR), hospitalization-to-infection (HIR) and b) in-hospital fatality-to-infection (FIR IH) and excess death fatality-to-infection (FIR ED) ratios in a rural and an urban community during the first (Wuhan), second (Beta) and third (Delta) waves of infections, March 2020 - November 2021, South Africa.

4.2. The household as an important setting for SARS-CoV-2 transmission

Transmission of SARS-CoV-2 occurs both within the community, for example at workplaces, schools, social gatherings, and within the household. Although seroprevalence can be used to estimate the community attack rate of SARS-CoV-2, it is difficult to study transmission dynamics on a community level, as the study population would have to be very large. The household is therefore a convenient unit to study transmission dynamics as it forms a micro-unit or contact network. Furthermore, while NPIs are in place that reduce the number of contacts in the community, household contacts increase [54] making the household an even more important location for the transmission of SARS-CoV-2. From our case-ascertained household transmission study (SA-S-HTS), including 131 index cases and 457 household contacts in two communities in South Africa, with a median household size of 5, we observed a household cumulative infection risk of 59% [106], which is amongst the higher global estimates reported so far [55]. This was likely due to the inclusion of only symptomatic index cases and the fact that more than half of included households (68%) experienced crowding (more than two people sleeping in one room) [53].

As with the influenza virus [111], age plays an important role in the transmission dynamics of SARS-CoV-2 in the household. Index cases that were aged 35-59 years (aOR 3.4, 95% CI 1.5-7.8) and ≥60 years (aOR 3.1, 95% CI 1.0-10.1), compared to index cases aged 18-34 years, resulted in higher household transmission. These individuals also had longer episode durations: 18.6 days (aOR 0.4,

95%CI 0.2-0.6) in index cases aged 35-59 years and 27.3 days (aOR 0.2, 95%CI 0.1-0.5) in index cases aged ≥60 years, compared to those aged 18-34 years who had a mean episode duration of 16.5 days. Contacts aged 13-17 years (aOR 7.1, 95% CI 1.5-33.9) and 18-34 years (aOR 4.4, 95% CI 1.0-18.4) compared to <5 years, had a higher risk of acquisition in the household. Lower acquisition in children <5 years could be due to less interaction with infected household members, but this was not investigated in our study. Contacts 35-59 years (aHR 0.3, 95% CI 0.1-0.9) and ≥60 years 11 (aHR 0.2, 95% CI 0.0-0.8), had longer serial intervals compared to those aged 18-34 years. The longer serial intervals could be related to the longer episode duration in these age groups. The higher HCIR when the index cases were older than 35 years, and contacts aged 13-34 years, can translate to spill-over of infections to places of employment and acquisition in schools, further driving community transmission [106].

There was also an increased risk of transmission (HCIR) when index case minimum C_t values were lower than 25 (as a proxy for high viral load, aOR 7.5, 95% CI 2.2-26.0) compared to C_t>35 [106]. Although C_t values cannot directly be used to infer the shedding of live virus and therefor potential for transmission, infectiousness is likely higher when lower C_t values are observed. Results from the PHIRST-C household cohort study showed that lower C_t values were also associated with symptomatic illness [30]. We also observed longer episode durations when minimum C_t values were lower (aHR 0.6, 95% CI 0.2-0.9) [106].

Although our study, and other literature that became available recently, assist in better understanding the transmission dynamics of SARS-CoV-2, there are still gaps that exist. Entering a period where most of the population has already been exposed to the virus, or vaccinated, as evident from the 97% seroprevalence estimates in South African blood donors in March 2022 [112], we need to better understand the role of the immunological response to infection, the timing of antibody waning and how this relates to the risk of infection and severe disease. Studies so far have shown that although newly emerged variants may be able to evade both natural- and vaccine-induced immunity, individuals are still protected from severe disease and death [113,114]. Re-infections were also found to result in lower household transmission compared to primary infections [30]. However, the longevity of this protection is still to be determined.

4.3. The influence of study design on the assessment of household transmission

Due to the forced proximity of individuals within a household and the high number of interactions, the household is a convenient environment to study the transmission dynamics of SARS-CoV-2. One still needs to consider, however, how study design may influence measured outcomes, and keep this in mind when interpreting results. A review by Tsang *et al*. describes how study design for influenza

transmission studies may influence results, and highlights possible gaps and biases [115]. As we performed two household transmission studies for SARS-CoV-2 with different study designs; one a cohort study with a random selection of households (PHIRST-C) and the other a case-ascertained household transmission study (SA-S-HTS) - we were able to compare findings between the two studies. In the PHIRST-C study, households were selected at random from two communities and followed up for 13 months, testing participants for SARS-CoV-2 twice a week using rRT-PCR on nasal swabs, irrespective of symptoms. With this sampling frame, we could identify index cases (the first person in the household testing positive for SARS-CoV-2) in both symptomatic and asymptomatic individuals. The point of infection and transmission could also be determined at high resolution, albeit within the 3-4-day gap between sample collection. In the case-ascertained SA-S-HTS study on the other hand, we screened for SARS-CoV-2 infection at public health clinics in two different communities and enrolled index cases who were the first person within their household to present with COVID-19-compatible symptoms, with symptom onset ≤5 days prior. The household was then enrolled within the next two days and followed up for six weeks, with thrice weekly SARS-CoV-2 testing, irrespective of symptoms. This results in the inclusion of only households with symptomatic index cases, and due to the delay between symptom onset and presenting for care (up to five days) and the further delay to household enrolment (up to seven days from initial symptom onset), transmission could have already occurred at the point of household enrolment and contact testing, so that transmission is rather calculated retrospectively than the prospective measurement of PHIRST-C. These study design differences could have contributed to differences in measured outcomes.

The generation interval (time between infection of the index to infection of contact) of 6.5 days in PHIRST-C [30] compared well with the 6-day serial interval (symptom onset in the index to symptom onset in case) in our case-ascertained SA-S-HTS [106]. We observed a longer episode duration in SA-S-HTS compared to PHIRST-C: 19 days [106] compared to 12 days [30]. This could be due to SA-S-HTS only including symptomatic index cases, and the presence of symptoms was found to be associated with longer episode durations in PHIRST-C [30]. We saw similar factors associated with transmission and acquisition, although PHIRST-C had the additional advantage of showing high transmission risk from children aged <5 years [30], whereas the case-ascertained SA-S-HTS only included adult index cases. The most notable difference was the HCIR estimates, with 58% in SA-S-HTS [106] and 23% in PHIRST-C [30]. Even when considering the HCIR in only the community that overlapped between the two studies (Jouberton, North West province), the HCIR was still higher in SA-S-HTS (55%) [106] compared to PHIRST-C (29%) [30]. We hypothesize this was likely driven by the fact that index cases in SA-S-HTS were not just symptomatic, but symptomatic enough to seek care. Although the presence of symptoms was not associated with higher HCIR estimates in PHIRST-C [30], the severity of symptoms

was not considered. A meta-analysis on SARS-CoV-2 household transmission did however find increased transmission from symptomatic index cases [53]. In both studies, a lower minimum C_t value during the index episode was associated with increased HCIR [30,106]. It was reassuring that this association was still found in SA-S-HTS since the lowest possible C_t value in the episode could have been before seeking care, or between screening and household enrolment, and would have been missed.

Although HCIR estimates from SA-S-HTS are more comparable with other published studies [55], PHIRST-C may be more representative of community-wide transmission dynamics, considering that the largest proportion of infections are asymptomatic [30] and this will echo in most households. However, a PHIRST-C study design is expensive and logistically challenging. While a case-ascertained study design like SA-S-HTS lends itself to relatively easy case identification and shorter follow-up periods, a study design such as PHIRST-C where households are randomly selected and followed up for long periods enriches for a higher proportion of participants that are unemployed compared to the community unemployment rates (35% vs 73% in rural and 30% vs 64% in urban community) [100]. In SA-S-HTS we specifically had difficulty in keeping individuals in follow-up when their quarantine period was over and they had to return to work. So where the PHIRST-C study design may bias more towards unemployed individuals, the SA-S-HTS study design may bias more towards individuals with access to care, and who may be more likely to need a medical note or SARS-CoV-2 test results for work. The true reflection within the community may be a mid-point between these two, but would likely not be generalizable to other communities. But do we expect transmission dynamics to be different based on employment status? Although community attack rates may be higher in those who are employed (due to higher possible contact rates), estimates such as secondary attack rate within the household, serial interval and duration of infection likely do not differ based on employment status, as employment status was not associated with any of these outcomes for SA-S-HTS.

A major limitation of the SA-S-HTS study design is the possible misclassification of the index case. The classification of the index case in a case-ascertained study relies solely on who in the household presented with symptoms first. If the true index case was asymptomatic, the presumptive index case would be incorrectly classified. The PHIRST-C study design overcomes this by being able to identify the (almost) exact point where the index became infected (or at least the point where the virus became detectable in the nose). Both study designs also have a further complication: discerning the secondary transmission chain from tertiary and quaternary transmission chains and further introductions of the virus into the household from the community. For both the PHIRST-C and the SA-S-HTS we relied on SARS-CoV-2 variant classifications to exclude households of mixed clusters, where there may have been secondary introductions. The problem however, is that at the peak of an

infection wave, while a dominant variant is circulating, there is a higher probability for the same variant to be introduced within the household, rather than a different variant, and therefore only relying on variant classification may not be sufficient. One may want to consider that multiple introductions of SARS-CoV-2 into the household is a rare event, but this may not be the case while the force of infection from the community is high (periods of high community circulation), and further investigation of this occurrence would be needed. Both post-secondary transmission chains and same-variant subsequent introductions into households were not considered in either PHIRST-C or SA-S-HTS, but could theoretically be done with full genome sequencing data and the modelling of transmission chains based on shedding viral kinetics. This has been done for the influenza virus [116] and could be explored further in the future.

Another important consideration is the time and effort to setup a randomly selected prospective household cohort. The implementation of the original PHIRST study took many months, with site collaborators having a presence in these communities for years before, allowing better community acceptance of the study. The extraordinary feat of getting PHIRST-C up-and-running within the first four months of SARS-CoV-2 being detected in South Africa was thanks to the collaborations and systems already in place due to the original PHIRST study that ran from 2016 through 2018. This highlights the importance of collaborations, having access to previous cohorts and having protocols and experience in place to initiate new studies were results are urgently needed. Even so, resources would not always be available to do the perfect study, and therefore case-ascertained studies are still useful. It is important, however, to understand the limitations of the study design, and to use context from other settings and studies when interpreting the results, being careful not to overstate findings. For example, in our case-ascertained SA-S-HTS, HCIR was very high, but this was not the situation for all households in the community where the virus was introduced, but only those with symptomatic medically-attended index cases.

4.4. Leveraging contact patterns to assess and predict SARS-CoV-2 transmission and to inform intervention strategies

We did not see an association between household contact patterns and SARS-CoV-2 transmission (appendix D5). As described in the manuscript, there may be several reasons for this, including that airborne transmission plays a larger role in household transmission, that contact rates are already so high in the household that transmission would have occurred irrespective, that infection status is rather determined by host factors, or due to study limitations. Firstly, we will further explore these study limitations with possible solutions to improve the study design for future investigations.

Although the case-ascertained study design for SA-S-HTS ensured households were followed up for a shorter period during which SARS-CoV-2 was circulating within the household, which reduced cost

and simplified logistics as it would be difficult to deploy sensors for the long periods needed to capture transmission events in a cohort study, it meant that there was a delay between the onset of illness, and more specifically the point of the index case becoming infectious, to the time that the household was enrolled. In our study, the median number of days from index symptom onset to household enrolment was 3 days (range 1-7 days), but the risk of infection is highest just before symptom onset [54], which meant that most of the infections likely already occurred by the time that we enrolled household contacts and started measuring contact events. There may not be a way to circumvent this limitation in this study design for the collection of high-resolution contact patterns, but a more simplified approach would be to interview household members at enrolment to collect retrospective contact data, as done previously, mostly during contact tracing activities [53]. This does however have other limitations as well, including recall bias (Table 1).

If specifically wanting to consider high-resolution contacts as measured with proximity sensors, a random selection of households with prospective follow-up and frequent testing, such as the PHIRST-C study design may be more useful. The deployment of the sensors would have to be at the peak of an infection wave where the probability of the introduction of the virus into the household is higher, to reduce the time required for sensors to be deployed. Deploying sensors for extended periods can be logistically challenging as battery changes are needed (M. Makhasi 2022, *personal communication 22 June*), compliance may reduce with extended deployments, and the higher risk of sensors being damaged or lost increases (and with it all collected data), increases. On the latter point, headway is being made to allow remote/wireless data transfer from the sensors to devices stationed in the household that could mitigate this issue (C. Cattuto 2022, *personal communication 11 January*). One option is to deploy the sensors at a single point during the follow-up period, irrespective if the virus is circulating in the household at that time, and to use these contact data collected to infer contact patterns when the virus is introduced in the household. The problem, however, is that contact patterns may differ once infection status is known, based on seasonality (winter vs summer) or during holiday periods.

Another consideration is to deploy sensors and measure transmission in other settings, such as workplaces, schools and hospitals. Successful deployments of proximity sensors have been done in all of these environments [69,74,89], and with the addition of frequent rRT-PCT testing, can be used to investigate SARS-CoV-2 transmission dynamics in these settings. These settings may also be a more suitable environment in which to assess the association of contact events with the transmission, if the reason for not detecting an association between contact patterns and SARS-CoV-2 in the household was due to the high threshold of contacts in the household (as less frequent contacts are expected in these settings than in households).

Future studies making use of proximity sensors in the South African setting would also need to consider some challenges that arose during our study. Before deployment, community engagement was required to ensure the electronic devices (which have a blinking light) would not be considered a threat to the local population which could lead to conflict for study participants. The deployment of these sensors is resource intensive and requires a dedicated project manager who can work on a Linux or Mac operating system and use command line coding for the formatting and timestamping of sensors, as well as the download of data which could be a timeous process depending on the number of sensors and duration of the deployment. The labelling of sensors and linking of sensors to participant identifiers should be done with care and the utmost attention to detail. A dedicated field team is needed to explain the study to participants in detail, obtain consent for participation, setup the sensor by putting in the battery, check functionality and package the sensors before it is issued to the participants. This cannot be done ahead of time as this will lead to the logging of nonsense contacts. The packaging and method of wearing should be customized to the requirements of the population. We made use of custom-designed polyvinyl chloride (PVC) pouches which are resistant to perspiration and, to a limited degree, rain. Sensors were also placed in small plastic bags before placing them in the PVC pouch to further protect the sensor from water. Although this worked to an extent, some sensors (and data) were still lost due to water damage. Pouches were stapled shut to prevent tampering by participants. Sensors can be worn as a badge pinned to the shirt on the chest, or on a lanyard around the neck, depending on the preference of the participant. Compliance could be a challenge and field teams need to visit participants at least every 2-3 days to confirm the sensors are still being worn and to replace any non-functional sensors or batteries. The longer these sensors are in the field, the higher probability for them to get lost or damaged, which would lead to complete data loss.

Something important that was not considered in our study, was ventilation within the areas of contact. This is especially relevant if aerosol transmission plays a greater role in the transmission of SARS-CoV-2 in the household compared to droplet-mediated transmission, since aerosol transmission would not need close-proximity contact. Although aerosol transmission of SARS-CoV-2 was initially only considered to be important in aerosol-generating procedures [46], more recent studies have shown that airborne transmission is likely a major pathway for SARS-CoV-2 transmission [48,117]. A modelling study simulating the risk of SARS-CoV-2 airborne transmission within a home showed that the risk of infection to other household members was high even if the infected individual was isolated in a separate room [118]. Considering the likely importance of aerosol transmission for SARS-CoV-2, ventilation is an important variable to measure and has been done for TB studies [119]. By better understanding the role of ventilation in transmission dynamics, better suggestions can be made to implement transmission mitigation strategies. Outside of the household,

this could also assist in safely keeping workplaces, schools, restaurants and bars open during future outbreaks of respiratory pathogens.

We were able to describe, albeit at a low resolution, contact rates and time use in both a rural and an urban community in South Africa [101]. In addition to this, we also measured high-resolution within-household contact patterns (appendix D5). Contact rates are crucial for the parameterization of transmission models, and may be very different between communities and different countries [120]. It is therefore important to obtain estimates of local contact rates when doing modelling. In South Africa, very limited contact data is available, and currently available estimates are from eight Western Cape province communities [93,95] and one community in the KwaZulu-Natal province [94]. Our study added two more communities, from two additional provinces for which contact data are available.

As seen in previous contact surveys within South Africa [95] and from around the world [120], the highest contact rates were observed amongst school-going children, observed in both our contact survey including household and community contacts [101] and the high-resolution within-household contacts (appendix D5). However, the absolute number of contacts measured in our contact survey was different from other contact rates from South Africa, and other countries. This highlights the importance of having local contact rates available for transmission modelling. We have the additional advantage of having high-resolution testing data for influenza, RSV, streptococcus and diphtheria for this cohort, which will in future studies be used with the contact data to build transmission models. Another consideration, however, is that contact patterns may not just be driven by the community characteristics, but also by timing, specifically considering work or school days, holiday periods and seasons. This highlights a limitation of our study where only one day was investigated for each participant. A consideration for future contact surveys in our setting is to do longitudinal surveys, instead of a cross-sectional study, still measuring a day at a time, but repeating this multiple times during the year, ensuring the capture of different seasons, days of the weeks and holiday periods.

We described that the elderly, who are at risk for more severe disease [108], have lower contact rates than other age groups, and that most time spent in contact was at the home, but school-going children have higher contact rates and many community contacts [101]. One of the approaches to mitigate SARS-CoV-2 transmission was the closure of schools which resulted in major disruptions of the education system, with a major toll on child well-being and economic impact on caregivers [121]. During the initial waves of infection, infection rates in children were low, with higher attack rates in adults of working age, with more infections in children in subsequent waves as adults were already exposed [122]. Considering the social and economic cost of school closures, this is not a reasonable control measure for outbreak situations, especially since the severity of illness is low amongst these age groups [108]. Control measures should rather focus on reducing transmission while keeping

schools open through the use of NPIs, for example ensuring adequate ventilation in classrooms [123]. One control measure that has also been considered for the influenza virus [124], is the possibility of targeting those age-groups driving transmission i.e. children, for vaccination, and not those at risk for severe disease, i.e. the elderly or those with underlying conditions. With many countries, and especially South Africa now showing high seroprevalence estimates for SARS-CoV-2, the virus will soon be considered endemic, and it will be important to establish which individuals are now the main drivers of transmission, and consider these individuals for intervention strategies like vaccination. The problem, however, is that current SARS-CoV-2 vaccines protect against severe disease, but show lower effectiveness against infection, especially when considering newly emerged immune escape variants and the waning of protection [125,126,127,128], making this a non-feasible option for SARS-CoV-2 until vaccines are available that are more effective against infection. Mucosal immunoglobulin A (IgA) antibodies in the respiratory tract may be crucial to this but are not elicited from the injectable vaccines currently in use for SARS-CoV-2. Intranasally administered vaccines like those that have been developed for influenza, could overcome this problem [129], and several candidates for SARS-CoV-2 are currently in development [130].

4.5. The importance of not neglecting one epidemic when facing another: SARS-CoV-2 and HIV co-infection

Initially, there was great concern about the effect that SARS-CoV-2 would have once introduced in sub-Saharan Africa, especially considering the high burden of HIV [56] and the possibility of increased susceptibility and severity of illness in PLWH [57]. It was therefore important to assess the effect of HIV in SARS-CoV-2 transmission dynamics. Although we did not observe a difference in seroprevalence between PLWH and PNLWH (aOR 1, 95% CI 0.7-1.5) after adjusting for age and sex [103], or an increased risk in transmission or acquisition within the household, we did find that index cases LWH had longer episode durations at high viral load (aHR 0.4 95% CI 0.1-0.9), and contacts LWH had longer serial intervals (aHR 0.2, 95% CI 0.0-1.4) [106]. This could point to PLWH shedding the virus for longer, which could in turn also lead to transmission outside of the household, with contacts being susceptible to infection even when index cases may have been shedding the virus at a lower viral load. Along with this, previous studies have also shown that PLWH that are infected with SARS-CoV-2 are at higher risk for severe disease and death [59]. Longer episode durations in PLWH can also lead to the evolution of new SARS-CoV-2 variants [131].

This underlines how crucial it is to maintain HIV testing and treatment programmes in the face of other challenges that may place a major strain on the healthcare system. Although anti-retroviral treatment (ART) dispensing did not seem to decline overall during the initial stages of the pandemic in South Africa [132], when considering data on the provincial level, some provinces did observe

reductions.[33] HIV testing and initiation into care also declined [33,132]. The SARS-CoV-2 pandemic also had effects on TB programmes, with a drop in both testing and case detection rates during each wave of infection, and modelling suggests that an increase in TB burden may follow the pandemic [34]. This underlines how healthcare programmes should not function in silos, even in the face of healthcare emergencies, and that situations like the pandemic should rather be seen as an opportunity to provide holistic care than shifting resources to one problem whilst creating a vacuum somewhere else.

The similar HCIR in households with an index case LWH and not immunocompromised and households where the index case was NLWH, may point to a similar risk of transmission from people NLWH and people LWH that are on treatment and virally suppressed. However, we observed a signal for higher household transmission if the index case was LWH and immunocompromised, although this was not significant (aOR 2.5, 95% CI 0.4-15.3) [106]. Based on our initial sample size calculations, we planned to enrol 174 contacts exposed to a SARS-CoV-2 index case LWH, but we only managed to enrol 103, even after extending the study period in the Soweto site. Sample size calculations were based on the HIV prevalence in individuals with influenza-like illness, and the situation proved to be very different in SARS-CoV-2, where the HIV prevalence was not as high as initially anticipated. Adding to this, only a small proportion (14%, 4/28) of PLWH enrolled were immunosuppressed. To effectively measure the difference in transmission between people NLWH and people LWH who are immunocompromised, one would have to increase the sample size of the latter group. Enriching the sample population with PLWH may be difficult. One option is to do screening at clinics providing HIV testing and treatment, although this may again enrich the sample for asymptomatic cases, which will then be a different population than PNLWH. If wanting to specifically enrich for PLWH that are immunosuppressed, this may be even more challenging as these are the individuals that are probably more reluctant to seek care. Once they do, they may be more likely to be hospitalized, which would make the household unsuitable for enrolment in a household transmission study. In a South African study looking at the shedding of SARS-CoV-2 in hospitalized patients, 40% were PLWH, and 37% of these individuals were immunocompromised [65].

4.6. Generating data for action

Early in 2020 when the emergence of SARS-CoV-2 was classified as a PHEIC, there was a vacuum of knowledge on the transmission dynamics of SARS-CoV-2. Initial estimates were based on the behaviour of SARS-CoV and MERS-CoV, which were soon found to be different from SARS-CoV-2 [133,134]. Thus a race began to generate and distribute results as quickly as possible to make informed decisions on containment, transmission mitigation, and distribution of resources. Although on one hand this made it challenging to do research as the landscape was constantly shifting, and adaption

was needed as new information became available or variants emerged, on the other hand, it created an environment for collaboration, transparency and sharing of data. Protocols were developed and shared internationally for quick adaptation to local settings, allowing standardized collection of data [35,135]. Funds were made available for SARS-CoV-2 studies and genomic surveillance was prioritized as variants emerged, something that was not done at this scale for other pathogens before 2020. No longer was it possible to perform a study and release results only a year or two later. A shift occurred for real-time data analysis and timely sharing of results through preprint servers, with all SARS-CoV-2 related publications made available as open access, and sharing of data directly with incident management teams and policy makers. This was no different for the data generated in our study. Seroprevalence estimates were shared as soon as available with the South African modelling consortium [136] and along with results from transmission studies, to the South African SARS-CoV-2 incident management team and peers from around the world at WHO meetings. These data, along with data generated from other studies, were used to model SARS-CoV-2 transmission, identify target groups for vaccination, and plan for the gradual relaxation of interventions in South Africa. Thanks to political interest, funding and collaboration efforts, there was likely more progress made in two years for SARS-CoV-2 research than in a decade for many other pathogens.

Seroprevalence estimates, transmission parameters and contact rates are crucial for mathematical modelling, which allows further assessment of transmission dynamics where studies are not feasible, forecasting of disease burden to allow for planning and allocation of resources, and assessment of the possible effect of interventions to reduce the transmission of SARS-CoV-2 [107,137,138]. Initial strategies to control the spread of SARS-CoV-2 relied on extreme interventions to reduce the contact between individuals and resultant transmission, leading to the lockdowns of complete populations in some areas [137]. Although initially appropriate to reduce the surge of infections and allow the healthcare system to prepare, the downstream economic and social devastation caused highlighted how these interventions should only be implemented in extreme circumstances and for short periods [12]. As data become available policies must be adapted in a timely fashion based on the evidence on hand. Based on CIR from South Africa and elsewhere [108], as well as estimates on the small symptomatic fraction of infections [30], and the fact that most transmission occurs just before symptom onset [54], contact tracing and testing of symptomatic contacts was not an effective containment measure in our setting, and although stopped in South Africa in February 2022 [139], this measure could have been abolished earlier to redirect resources elsewhere to assist in the overburdened health system or other neglected health programs.

It is important to remember that pandemics are dynamic, and quick adaption is needed based on evidence generated. For example, at the start of and outbreak, it may make sense to immediately quarantine all household members of contacts of a case, in an attempt to get ahead of the chain of

transmission. However, as widespread community transmission occurs, specifically in the case where the virus has high rates of asymptomatic infection, and low hospitalization and death ratios in otherwise healthy individuals, the priority should shift to protect those at risk for severe disease from infection. With subsequent waves of infection and the uncoupling of infections from hospitalizations and deaths [109,110], all NPI can be relaxed, something that has been done in South Africa and globally, albeit slower than possibly needed.

5. Limitations

Each study included in this body of work had inherent constraints. The contact survey used for objective 1 had a smaller sample size than other published contact studies, and might not be representative of the community. An offset, however, is that households were randomly selected, and the cohort was very well described, with longitudinal virological data on influenza and RSV, which will be combined in future studies for modelling transmission dynamics. We only measured contacts for one day, and acknowledge that contact patterns will be different between different days and periods. We opted for this approach to reduce bias from reporting fatigue. In future studies in similar longitudinal cohorts, contact studies can be nested throughout the study period for different days for each participant.

We had a relatively small sample size for our seroprevalence study compared to other published seroprevalence studies. However, being a prospective cohort study, we were able to collect samples from individuals over 16 months and track changes in seroprevalence over time. Households were also randomly selected and we were able to include persons of all ages. This is a benefit over doing seroprevalence studies from residual blood donor samples where one is faced with selection bias and restricted to individuals 16 years and older, although seroprevalence studies from blood donor studies are costly. A further limitation is that we did not calculate CIR, HIR and FIR by age, but rather calculated age-adjusted estimates.

Lastly, as discussed in section 4.3, results from the case-ascertained household transmission study would not reflect the transmission dynamics for SARS-CoV-2 in all households, but specifically in households were index cases were symptomatic. It also becomes difficult in household transmission studies to disentangle transmission chains, and all cases in household contacts may not have arisen from the presumptive index case. The use of binomial models such as logistic regression (as in this study) to calculate SIR has been criticized for not being able to distinguish between multiple chains of transmission [140], hence the use of HCIR and not SIR in the current study. Future analyses of these data can consider making use of longitudinal chain binomial models or pairwise survival analysis.

6. Conclusion

In conclusion, we found that 95% of SARS-CoV-2 infections during the first three waves of infection in two South African communities, were not reported to national surveillance and that when considering the number of infections, HIRs ranged from 0.4% to 3%, and FIR from 0.1% to 0.7%. The burden of SARS-CoV-2 is different between the urban and rural communities, and between age groups through the successive waves of SARS-CoV-2 infection, but we did not see a difference between SARS-CoV-2 seroprevalence between PLWH and people NLWH. We observed a high HCIR in households with symptomatic adult SARS-CoV-2 index cases, with higher transmission from older index cases, and higher acquisition by individuals aged 13-34 years [106]. PLWH had longer episode durations, and longer serial intervals, although this did not translate to a higher HCIR in households where the index case was LWH [106]. We did not observe an association between household social contact patterns and SARS-CoV-2 transmission but were able to generate high-resolution household contact data (appendix D5), as well as community contacts in two communities in South Africa [101]. Age played an important role in social contact patterns, we observed higher contact rates in children and the rural community. These data can be used for future infectious disease transmission modelling for SARS-CoV-2 and other pathogens [101].

Continuous monitoring of population SARS-CoV-2 seroprevalence will be important to monitor immunity to the virus and to signal when vaccination campaigns would be needed to boost population immunity. For this however, a standardized correlate of protection of SARS-CoV-2 should be defined, for which research is already underway [141]. Furthermore, it will be essential to monitor changes in the severity of SARS-CoV-2 infection as the virus becomes endemic, but also as new variants arise. The transmission dynamics of SARS-CoV-2 should also be monitored as it becomes endemic.

Although hindsight is always easier than foresight, and it is simpler to make recommendations after worse-case scenarios did not occur, the lessons learned in the SARS-CoV-2 pandemic must be remembered and applied for future emerging pathogens and pandemics – something that is predicted to occur more often. A crucial aspect still lacking is a full account of the impact that NPIs had on the economy, education system and other sectors, and for studies to be done to determine the long-term effects of these interventions. This will assist in making informed decisions in future pandemics, weighing the full cost and benefit of these measures. We recommend continuing to build on the collaborations initiated during the pandemic, and the expansion of techniques employed for swift data generation and publication to be used in other research areas. Performing robust studies to generate data for policies is crucial going forward, along with the quick adaption of policies once solid evidence becomes available.

7. References

1. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: Implications for SARS-CoV-2. *Nature Reviews Microbiology*. 19(3):155-70. DOI: https://doi.org/10.1038/s41579-020-00468-6.
2. Hui DSC, Zumla A. 2019. Severe acute respiratory syndrome: Historical, epidemiologic, and clinical features. *Infectious Disease Clinics of North America*. 33(4):869-89. DOI: https://doi.org/10.1016/j.idc.2019.07.001.
3. Bleibtreu A, Bertine M, Bertin C, Houhou-Fidouh N, Visseaux B. 2020. Focus on middle east respiratory syndrome coronavirus (MERS-CoV). *Médecine et Maladies Infectieuses*. 50(3):243-51. DOI: https://doi.org/10.1016/j.medmal.2019.10.004.
4. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 579(7798):270-3. DOI: https://doi.org/10.1038/s41586-020-2012-7.
5. World Health Organization [Internet]. 2020. Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). Cited 10 August 2022. Available from: https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19).
6. World Health Organization [Internet]. 2020. Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-nCoV). Cited 10 August 2022. Available from: https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).
7. World Health Organization [Internet]. 2020. WHO director-general's opening remarks at the media briefing on COVID-19 - 11 March 2020. Cited 10 August 2022. Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.
8. Solomon M, Liang C. 2022. Human coronaviruses: The emergence of SARS-CoV-2 and management of COVID-19. *Virus Research*. 319:198882. DOI: https://doi.org/10.1016/j.virusres.2022.198882.
9. National Institute for Communicable Diseases [Internet]. 2022. Tracking SARS-CoV-2 variants. Cited 21 July 2022. Available from: https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/sars-cov-2-genomic-surveillance-update/.

10. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. 2022. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*. 603(7902):679-86. DOI: https://doi.org/10.1038/s41586-022-04411-y.
11. Koelle K, Martin MA, Antia R, Lopman B, Dean NE. 2022. The changing epidemiology of SARS-CoV-2. *Science*. 375(6585):1116-21. DOI: https://doi.org/10.1126/science.abm4915.
12. Yanovskiy M, Socol Y. 2022. Are lockdowns effective in managing pandemics? *International Journal of Environmental Research and Public Health*. 19(15):9295. DOI: https://doi.org/10.3390/ijerph19159295.
13. South African Government [Internet]. 2020. COVID-19 risk adjusted strategy. Available from: https://sacoronavirus.co.za/covid-19-risk-adjusted-strategy/.
14. Leech G, Rogers-Smith C, Monrad JT, Sandbrink JB, Snodin B, Zinkov R, et al. 2022. Mask wearing in community settings reduces SARS-CoV-2 transmission. *Proceedings of the National Academy of Sciences*. 119(23):e2119266119. DOI: https://doi.org/10.1073/pnas.2119266119.
15. Hemmeda L, Shabani MM, Kolawole BO, Muhammad SA, Fatima K, Siddiqui A, et al. 2022. Third wave of COVID-19 pandemic in Africa: Challenges and recommendations. *Annals of Medicine and Surgery*. 80:104314. DOI: https://doi.org/10.1016/j.amsu.2022.104314.
16. Khoury DS, Docken SS, Subbarao K, Kent SJ, Davenport MP, Cromer D. 2023. Predicting the efficacy of variant-modified COVID-19 vaccine boosters. *Nature Medicine*. Published online ahead of print. DOI: https://doi.org/10.1038/s41591-023-02228-4.
17. Cohen J. 2022. Making broader coronavirus vaccines is a struggle. *Science*. 377(6606):566-7. DOI: https://doi.org/10.1126/science.ade2363.
18. McCarthy KM, Tempia S, Kufa T, Kleynhans J, Wolter N, Jassat W, et al. 2021. The importation and establishment of community transmission of SARS-CoV-2 during the first eight weeks of the South African COVID-19 epidemic. *EClinicalMedicine*. 39:101072. DOI: https://doi.org/10.1016/j.eclinm.2021.101072.
19. South African Government [Internet]. 2022. Newsroom. Cited 10 August 2022. Available from: https://www.gov.za/newsroom?title_field_value=corona&field_gcis_speech_category_tid=All&field_gcis_speech_government_lvl_tid=All&field_gcis_speech_date_value_1%5Bmin%5D%5Bdate%5D=&field_gcis_speech_date_value_1%5Bmax%5D%5Bdate%5D=.
20. National Institute for Communicable Diseases [Internet]. 2022. Weekly hospital surveillance (DATCOV) update. Cited 10 August 2022. Available from: https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-hospital-surveillance-datcov-update/.
21. National Institute for Communicable Diseases [Internet]. 2022. Weekly respiratory pathogens report. Cited 10 August 2022. Available from: https://www.nicd.ac.za/diseases-a-z-

index/disease-index-covid-19/surveillance-reports/weekly-respiratory-pathogens-surveillance-report-week/.

22. National Institute for Communicable Diseases [Internet]. 2022. Wastewater-based epidemiology for SARS-CoV-2 surveillance in South Africa. Cited 10 August 2022. Available from: https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-reports/wastewater-based-epidemiology-for-sars-cov-2-in-south-africa/.
23. National Institute for Communicable Diseases [Internet]. 2022. COVID-19 weekly epidemiology brief. Cited 21 July 2022. Available from: https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-epidemiological-brief/.
24. World Health Organization [Internet]. 2022. WHO coronavirus (COVID-19) dashboard. Cited 10 August 2022. Available from: https://covid19.who.int/.
25. Murray CJL. 2022. The global burden of disease study at 30 years. *Nature Medicine*. 28(10):2019-26. DOI: https://doi.org/10.1038/s41591-022-01990-1.
26. Murray CJL, Lopez AD. 1997. Global mortality, disability, and the contribution of risk factors: Global burden of disease study. *The Lancet*. 349(9063):1436-42. DOI: https://doi.org/10.1016/S0140-6736(96)07495-8.
27. Bosa I, Castelli A, Castelli M, Ciani O, Compagni A, Galizzi MM, et al. 2022. Response to COVID-19: Was Italy (un)prepared? *Health Economics, Policy and Law*. 17(1):1-13. DOI: https://doi.org/10.1017/S1744133121000141.
28. Dhar MS, Marwal R, VS R, Ponnusamy K, Jolly B, Bhoyar RC, et al. 2021. Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. *Science*. 374(6570):995-9. DOI: https://doi.org/10.1126/science.abj9932.
29. Hajissa K, Islam MA, Hassan SA, Zaidah AR, Ismail N, Mohamed Z. 2022. Seroprevalence of SARS-CoV-2 antibodies in Africa: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*. 19(12):7257. DOI: https://doi.org/10.3390/ijerph19127257.
30. Cohen C, Kleynhans J, von Gottberg A, McMorrow ML, Wolter N, Bhiman JN, et al. 2022. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: Results of the PHIRST-c cohort study, South Africa, 2020-21. *The Lancet Infectious Diseases*. 22(6):821-34. DOI: https://doi.org/10.1016/S1473-3099(22)00069-X.
31. Jassat W, Mudara C, Ozougwu L, Tempia S, Blumberg L, Davies MA, et al. 2021. Difference in mortality among individuals admitted to hospital with COVID-19 during the first and second waves in South Africa: A cohort study. *The Lancet Global Health*. 9(9):e1216-e25. DOI: https://doi.org/10.1016/S2214-109X(21)00289-8.
32. Cardoso Pinto AM, Ranasinghe L, Dodd PJ, Budhathoki SS, Seddon JA, Whittaker E. 2022. Disruptions to routine childhood vaccinations in low- and middle-income countries during the

COVID-19 pandemic: A systematic review. *Frontiers in Pediatrics*. 10:979769. DOI: https://doi.org/10.3389/fped.2022.979769.

33. Roy K, Himelfarb A, Karrah K, Porterfield L, Paremoer L, Serag H, et al. 2022. The social, behavioral, and ethical modalities of COVID-19 on HIV care in South Africa: A systematic review. *International Journal of Environmental Research and Public Health*. 19(15):9766. DOI: https://doi.org/10.3390/ijerph19159766.
34. Dheda K, Perumal T, Moultrie H, Perumal R, Esmail A, Scott AJ, et al. 2022. The intersecting pandemics of tuberculosis and COVID-19: Population-level and patient-level impact, clinical presentation, and corrective interventions. *The Lancet Respiratory Medicine*. 10(6):603-22. DOI: https://doi.org/10.1016/S2213-2600(22)00092-3.
35. Lewis HC, Ware H, Whelan M, Subissi L, Li Z, Ma X, et al. 2022. SARS-CoV-2 infection in Africa: A systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *BMJ Global Health*. 7(8):e008793. DOI: https://doi.org/10.1136/bmjgh-2022-008793.
36. Azami M, Moradi Y, Moradkhani A, Aghaei A. 2022. SARS-CoV-2 seroprevalence around the world: An updated systematic review and meta-analysis. *European Journal of Medical Research*. 27(1):81. DOI: https://doi.org/10.1186/s40001-022-00710-2.
37. Coatsworth N, Myles PS, Mann GJ, Cockburn IA, Forbes AB, Gardiner EE, et al. 2021. Prevalence of asymptomatic SARS-CoV-2 infection in elective surgical patients in Australia: A prospective surveillance study. *ANZ Journal of Surgery*. 91(1-2):27-32. DOI: https://doi.org/10.1111/ans.16564.
38. Song SK, Lee DH, Nam JH, Kim KT, Do JS, Kang DW, et al. 2020. IgG seroprevalence of COVID-19 among individuals without a history of the coronavirus disease infection in daegu, korea. *Journal of Korean Medical Science*. 35(29). DOI: https://doi.org/10.3346/jkms.2020.35.e269.
39. Banjar A, Al-Tawfiq JA, Alruwaily A, Alserehi H, Al-Qunaibet A, Alaswad R, et al. 2021. Seroprevalence of antibodies to SARS-CoV-2 among blood donors in the early months of the pandemic in Saudi Arabia. *International Journal of Infectious Diseases*. 104:452-7. DOI: https://doi.org/10.1016/j.ijid.2021.01.028.
40. Younas A, Waheed S, Khawaja S, Imam M, Borhany M, Shamsi T. 2020. Seroprevalence of SARS-CoV-2 antibodies among healthy blood donors in karachi, pakistan. *Transfusion and Apheresis Science*. 59(6). DOI: https://doi.org/10.1016/j.transci.2020.102923.
41. Herrera-Esposito D, de los Campos G. 2022. Age-specific rate of severe and critical SARS-CoV-2 infections estimated with multi-country seroprevalence studies. *BMC Infectious Diseases*. 22(1):311. DOI: https://doi.org/10.1186/s12879-022-07262-0.

42. Vermeulen M, Mhlanga L, Sykes W, Coleman C, Pietersen N, Cable R, et al. 2021. Prevalence of anti-SARS-CoV-2 antibodies among blood donors in South Africa during the period January-May 2021. *Research Square*. DOI: https://doi.org/10.21203/rs.3.rs-690372/v2.
43. Mhlanga L, Vermeulen M, Grebe E, Welte A. 2021. SARS-CoV-2 infection fatality rate estimates for South Africa. *Research Square*. DOI: https://doi.org/10.21203/rs.3.rs-707813/v2.
44. Mutevedzi PC, Kawonga M, Kwatra G, Moultrie A, Baillie V, Mabena N, et al. 2022. Estimated SARS-CoV-2 infection rate and fatality risk in Gauteng province, South Africa: A population-based seroepidemiological survey. *International Journal of Epidemiology*. 51(2):404-17. DOI: https://doi.org/10.1093/ije/dyab217.
45. Wolter N, Tempia S, von Gottberg A, Bhiman JN, Walaza S, Kleynhans J, et al. 2022. Seroprevalence of severe acute respiratory syndrome coronavirus 2 after the second wave in South Africa in human immunodeficiency virus–infected and uninfected persons: A cross-sectional household survey. *Clinical Infectious Diseases*. 75(1):e57-e68. DOI: https://doi.org/10.1093/cid/ciac198.
46. Meyerowitz EA, Richterman A, Gandhi RT, Sax PE. 2020. Transmission of SARS-CoV-2: A review of viral, host, and environmental factors. *Annals of Internal Medicine*. 174(1):69-79. DOI: https://doi.org/10.7326/M20-5008.
47. Geng Y, Wang Y. 2022. Stability and transmissibility of SARS-CoV-2 in the environment. *Journal of Medical Virology*. Published online ahead of print. DOI: https://doi.org/10.1002/jmv.28103.
48. Wang CC, Prather KA, Sznitman J, Jimenez JL, Lakdawala SS, Tufekci Z, et al. 2021. Airborne transmission of respiratory viruses. *Science*. 373(6558):eabd9149. DOI: https://doi.org/10.1126/science.abd9149.
49. Wang S, Zhang F, Wang Z, Du Z, Gao C. 2022. Reproduction numbers of SARS-CoV-2 Omicron subvariants. *Journal of Travel Medicine*. Published online ahead of print. DOI: https://doi.org/10.1093/jtm/taac108.
50. Chen PZ, Koopmans M, Fisman DN, Gu FX. 2021. Understanding why superspreading drives the COVID-19 pandemic but not the H1N1 pandemic. *The Lancet Infectious Diseases*. 21(9):1203-4. DOI: https://doi.org/10.1016/S1473-3099(21)00406-0.
51. Kuylen EJ, Torneri A, Willem L, Libin PJK, Abrams S, Coletti P, et al. 2022. Different forms of superspreading lead to different outcomes: Heterogeneity in infectiousness and contact behavior relevant for the case of SARS-CoV-2. *PLoS Computational Biology*. 18(8):e1009980. DOI: https://doi.org/10.1371/journal.pcbi.1009980.
52. Hsu CY, Wang JT, Huang KC, Fan ACH, Yeh YP, Chen SLS. 2021. Household transmission but without the community-acquired outbreak of COVID-19 in Taiwan. *Journal of the Formosan Medical Association*. 120:S38-S45. DOI: https://doi.org/10.1016/j.jfma.2021.04.021.

53. Madewell ZJ, Yang Y, Longini IM, Jr., Halloran ME, Dean NE. 2020. Household transmission of SARS-CoV-2: A systematic review and meta-analysis. *JAMA Network Open*. 3(12):e2031756-e. DOI: https://doi.org/10.1001/jamanetworkopen.2020.31756.
54. Nande A, Adlam B, Sheen J, Levy MZ, Hill AL. 2021. Dynamics of COVID-19 under social distancing measures are driven by transmission network structure. *PLoS Computational Biology*. 17(2):e1008684. DOI: https://doi.org/10.1371/journal.pcbi.1008684.
55. Madewell ZJ, Yang Y, Longini IM, Jr, Halloran ME, Dean NE. 2022. Household secondary attack rates of SARS-CoV-2 by variant and vaccination status: An updated systematic review and meta-analysis. *JAMA Network Open*. 5(4):e229317-e. DOI: https://doi.org/10.1001/jamanetworkopen.2022.9317.
56. UNAIDS [Internet]. 2020. Unaids data 2020. Cited 10 August 2022. Available from: https://www.unaids.org/sites/default/files/media_asset/2020_aids-data-book_en.pdf.
57. Ambrosioni J, Blanco JL, Reyes-Urueña JM, Davies MA, Sued O, Marcos MA, et al. 2021. Overview of SARS-CoV-2 infection in adults living with HIV. *The Lancet HIV*. 8(5):e294-e305. DOI: https://doi.org/10.1016/S2352-3018(21)00070-9.
58. Friedman EE, Devlin SA, McNulty MC, Ridgway JP. 2021. SARS-CoV-2 percent positivity and risk factors among people with HIV at an urban academic medical center. *PloS One*. 16(7):e0254994. DOI: https://doi.org/10.1371/journal.pone.0254994.
59. Tesoriero JM, Swain C-AE, Pierce JL, Zamboni L, Wu M, Holtgrave DR, et al. 2021. COVID-19 outcomes among persons living with or without diagnosed HIV infection in New York State. *JAMA Network Open*. 4(2):e2037069-e. DOI: https://doi.org/10.1001/jamanetworkopen.2020.37069.
60. Spinelli MA, Lynch KL, Yun C, Glidden DV, Peluso MJ, Henrich TJ, et al. 2021. SARS-CoV-2 seroprevalence, and IgG concentration and pseudovirus neutralising antibody titres after infection, compared by HIV status: A matched case-control observational study. *The Lancet HIV*. 8(6):S2352-3018(21)00072-2. DOI: https://doi.org/10.1016/S2352-3018(21)00072-2.
61. Danwang C, Noubiap JJ, Robert A, Yombi JC. 2022. Outcomes of patients with HIV and COVID-19 co-infection: A systematic review and meta-analysis. *AIDS Research and Therapy*. 19(1):3. DOI: https://doi.org/10.1186/s12981-021-00427-y.
62. Ssentongo P, Heilbrunn ES, Ssentongo AE, Advani S, Chinchilli VM, Nunez JJ, et al. 2021. Epidemiology and outcomes of COVID-19 in HIV-infected individuals: A systematic review and meta-analysis. *Scientific Reports*. 11(1):6283-95. DOI: https://doi.org/10.1038/s41598-021-85359-3.
63. Dong Y, Li Z, Ding S, Liu S, Tang Z, Jia L, et al. 2021. HIV infection and risk of COVID-19 mortality: A meta-analysis. *Medicine*. 100(26):e26573. DOI: https://doi.org/10.1097/MD.0000000000026573.

64. Jassat W, Cohen C, Tempia S, Masha M, Goldstein S, Kufa T, et al. 2021. Risk factors for COVID-19-related in-hospital mortality in a high HIV and tuberculosis prevalence setting in South Africa: A cohort study. *The Lancet HIV*. 8(9):e554-e67. DOI: https://doi.org/10.1016/S2352-3018(21)00151-X.
65. Meiring S, Tempia S, Bhiman JN, Buys A, Kleynhans J, Makhasi M, et al. 2022. Prolonged shedding of SARS-CoV-2 at high viral loads amongst hospitalised immunocompromised persons living with HIV, South Africa. *Clinical Infectious Diseases*. 75(1):e144-e56. DOI: https://doi.org/10.1093/cid/ciac077.
66. Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DAT. 2012. Close encounters of the infectious kind: Methods to measure social mixing behaviour. *Epidemiology and Infection*. 140(12):2117-30. DOI: https://doi.org/10.1017/S0950268812000842.
67. Tizzoni M, Bajardi P, Decuyper A, Kon Kam King G, Schneider CM, Blondel V, et al. 2014. On the use of human mobility proxies for modeling epidemics. *PLoS Computational Biology*. 10(7):e1003716. DOI: https://doi.org/10.1371/journal.pcbi.1003716.
68. Paz-Soldan VA, Reiner RC, Jr., Morrison AC, Stoddard ST, Kitron U, Scott TW, et al. 2014. Strengths and weaknesses of global positioning system (GPS) data-loggers and semi-structured interviews for capturing fine-scale human mobility: Findings from Iquitos, Peru. *PLoS Neglected Tropical Diseases*. 8(6):e2888. DOI: https://doi.org/10.1371/journal.pntd.0002888.
69. Smieszek T, Barclay VC, Seeni I, Rainey JJ, Gao H, Uzicanin A, et al. 2014. How should social mixing be measured: Comparing web-based survey and sensor-based methods. *BMC Infectious Diseases*. 14:136. DOI: https://doi.org/10.1186/1471-2334-14-136.
70. Smieszek T, Castell S, Barrat A, Cattuto C, White PJ, Krause G. 2016. Contact diaries versus wearable proximity sensors in measuring contact patterns at a conference: Method comparison and participants' attitudes. *BMC Infectious Diseases*. 16:341. DOI: https://doi.org/10.1186/s12879-016-1676-y.
71. Hoang TV, Coletti P, Kifle YW, Kerckhove KV, Vercruysse S, Willem L, et al. 2021. Close contact infection dynamics over time: Insights from a second large-scale social contact survey in Flanders, Belgium, in 2010-2011. *BMC Infectious Diseases*. 21(1):274. DOI: https://doi.org/10.1186/s12879-021-05949-4.
72. Hoang TV, Coletti P, Melegaro A, Wallinga J, Grijalva C, Edmunds J, et al. 2019. A systematic review of social contact surveys to inform transmission models of close contact infections. *Infectious Diseases*. 30(5):723-36. DOI: https://doi.org/10.1097/EDE.0000000000001047.
73. Barrat A, Cattuto C, Tozzi AE, Vanhems P, Voirin N. 2014. Measuring contact patterns with wearable sensors: Methods, data characteristics and applications to data-driven simulations of infectious diseases. *Clinical Microbiology and Infection*. 20(1):10-6. DOI: https://doi.org/10.1111/1469-0691.12472.

74. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J-F, Vespignani A. 2010. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS One*. 5(7):e11596-e. DOI: https://doi.org/10.1371/journal.pone.0011596.
75. Isella L, Stehlé J, Barrat A, Cattuto C, Pinton J-F, Van den Broeck W. 2011. What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of Theoretical Biology*. 271(1):166-80. DOI: http://dx.doi.org/10.1016/j.jtbi.2010.11.033.
76. Stehle J, Voirin N, Barrat A, Cattuto C, Colizza V, Isella L, et al. 2011. Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine*. 9:87. DOI: https://doi.org/10.1186/1741-7015-9-87.
77. Barrat A, Cattuto C, Colizza V, Pinton J-F, Broeck WVd, Vespignani A. 2008. High resolution dynamical mapping of social interactions with active RFID. *arXiv*. preprint arXiv:0811.4170. DOI: https://doi.org/10.1371/journal.pone.0011596.
78. Leecaster M, Toth DJA, Pettey WBP, Rainey JJ, Gao H, Uzicanin A, et al. 2016. Estimates of social contact in a middle school based on self-report and wireless sensor data. *PloS One*. 11(4):e0153690. DOI: https://doi.org/10.1371/journal.pone.0153690.
79. Fournet J, Barrat A. 2014. Contact patterns among high school students. *PloS One*. 9(9):e107878. DOI: https://doi.org/10.1371/journal.pone.0107878.
80. Guclu H, Read J, Vukotich CJ, Jr., Galloway DD, Gao H, Rainey JJ, et al. 2016. Social contact networks and mixing among students in K-12 schools in Pittsburgh, PA. *PloS One*. 11(3):e0151139. DOI: https://doi.org/10.1371/journal.pone.0151139.
81. Mastrandrea R, Fournet J, Barrat A. 2015. Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PloS One*. 10(9):e0136497. DOI: https://doi.org/10.1371/journal.pone.0136497.
82. Salathe M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. 2010. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences of the United States of America*. 107(51):22020-5. DOI: https://doi.org/10.1073/pnas.1009094108.
83. Stehle J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton JF, et al. 2011. High-resolution measurements of face-to-face contact patterns in a primary school. *PloS One*. 6(8):e23176. DOI: https://doi.org/10.1371/journal.pone.0023176.
84. Smieszek T, Burri EU, Scherzinger R, Scholz RW. 2011. Collecting close-contact social mixing data with contact diaries: Reporting errors and biases. *Epidemiology and Infection*. 140(4):744-52. DOI: https://doi.org/10.1017/S0950268811001130.
85. Hornbeck T, Naylor D, Segre AM, Thomas G, Herman T, Polgreen PM. 2012. Using sensor networks to study the effect of peripatetic healthcare workers on the spread of hospital-

associated infections. *Journal of Infectious Diseases*. 206(10):1549-57. DOI: https://doi.org/10.1093/infdis/jis542.

86. Isella L, Romano M, Barrat A, Cattuto C, Colizza V, Van den Broeck W, et al. 2011. Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. *PloS One*. 6(2):e17144. DOI: https://doi.org/10.1371/journal.pone.0017144.
87. Machens A, Gesualdo F, Rizzo C, Tozzi AE, Barrat A, Cattuto C. 2013. An infectious disease model on empirical networks of human contact: Bridging the gap between dynamic network data and contact matrices. *BMC Infectious Diseases*. 13:185. DOI: https://doi.org/10.1186/1471-2334-13-185.
88. Vanhems P, Barrat A, Cattuto C, Pinton JF, Khanafer N, Regis C, et al. 2013. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PloS One*. 8(9):e73970. DOI: https://doi.org/10.1371/journal.pone.0073970.
89. Voirin N, Payet C, Barrat A, Cattuto C, Khanafer N, Regis C, et al. 2015. Combining high-resolution contact data with virological data to investigate influenza transmission in a tertiary care hospital. *Infection Control and Hospital Epidemiology*. 36(3):254-60. DOI: https://doi.org/10.1017/ice.2014.53.
90. Kiti MC, Tizzoni M, Kinyanjui TM, Koech DC, Munywoki PK, Meriac M, et al. 2016. Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors. *EPJ Data Sci*. 5:21. DOI: https://doi.org/10.1140/epjds/s13688-016-0084-2.
91. Ozella L, Gesualdo F, Tizzoni M, Rizzo C, Pandolfi E, Campagna I, et al. 2018. Close encounters between infants and household members measured through wearable proximity sensors. *PloS One*. 13(6):e0198733. DOI: https://doi.org/10.1371/journal.pone.0198733.
92. Pung R, Firth JA, Spurgin LG, Chang A, Kong J, Wong J, et al. 2022. Using high-resolution contact networks to evaluate SARS-CoV-2 transmission and control in large-scale multi-day events. *Nature Communications*. 13(1):1956. DOI: https://doi.org/10.1038/s41467-022-29522-y.
93. Dodd PJ, Looker C, Plumb ID, Bond V, Schaap A, Shanaube K, et al. 2016. Age- and sex-specific social contact patterns and incidence of *mycobacterium tuberculosis* infection. *American Journal of Epidemiology*. 183(2):156-66. DOI: https://doi.org/10.1093/aje/kwv160.
94. McCreesh N, Dlamini V, Edwards A, Olivier S, Dayi N, Dikgale K, et al. 2021. Impact of the COVID-19 epidemic and related social distancing regulations on social contact and SARS-CoV-2 transmission potential in rural South Africa: Analysis of repeated cross-sectional surveys. *BMC Infectious Diseases*. 21:928-. DOI: https://doi.org/10.1186/s12879-021-06604-8.
95. Johnstone-Robertson SP, Mark D, Morrow C, Middelkoop K, Chiswell M, Aquino LDH, et al. 2011. Social mixing patterns within a South African township community: Implications for respiratory disease transmission and control. *American Journal of Epidemiology*. 174(11):1246-55. DOI: https://doi.org/10.1093/aje/kwr251.

96. Trentini F, Manna A, Balbo N, Marziano V, Guzzetta G, O'Dell S, et al. 2022. Investigating the relationship between interventions, contact patterns, and SARS-CoV-2 transmissibility. *Epidemics*. 40:100601. DOI: https://doi.org/10.1016/j.epidem.2022.100601.
97. Aleta A, Martin-Corral D, Bakker MA, Pastore YPA, Ajelli M, Litvinova M, et al. 2022. Quantifying the importance and location of SARS-CoV-2 transmission events in large metropolitan areas. *Proceedings of the National Academy of Sciences of the United States of America*. 119(26):e2112182119. DOI: https://doi.org/10.1073/pnas.2112182119.
98. Liu CY, Berlin J, Kiti MC, Del Fava E, Grow A, Zagheni E, et al. 2021. Rapid review of social contact patterns during the COVID-19 pandemic. *Epidemiology*. 32(6):781-91. DOI: https://doi.org/10.1097/ede.0000000000001412.
99. Aleta A, Ferraz de Arruda G, Moreno Y. 2020. Data-driven contact structures: From homogeneous mixing to multilayer networks. *PLoS Computational Biology*. 16(7):e1008035. DOI: https://doi.org/10.1371/journal.pcbi.1008035.
100. Cohen C, McMorrow ML, Martinson NA, Kahn K, Treurnicht FK, Moyes J, et al. 2021. Cohort profile: A prospective household cohort study of influenza, respiratory syncytial virus and other respiratory pathogens community burden and transmission dynamics in South Africa, 2016–2018. *Influenza and Other Respiratory Viruses*. 15(6):789-803. DOI: https://doi.org/10.1111/irv.12881.
101. Kleynhans J, Tempia S, McMorrow ML, von Gottberg A, Martinson NA, Kahn K, et al. 2021. A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018. *BMC Public Health*. 21:1055. DOI: https://doi.org/10.1186/s12889-021-11136-6.
102. Larremore DB, Fosdick BK, Bubar KM, Zhang S, Kissler SM, Metcalf CJE, et al. 2021. Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys. *eLife*. 10:e64206. DOI: https://doi.org/10.7554/eLife.64206.
103. Kleynhans J, Tempia S, Wolter N, von Gottberg A, Bhiman J, Buys A, et al. 2021. SARS-CoV-2 seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. *Emerging Infectious Disease journal*. 27(12):3020. DOI: https://doi.org/10.3201/eid2712.211465.
104. Kleynhans J, Tempia S, Wolter N, von Gottberg A, Bhiman J, Buys A, et al. 2022. SARS-CoV-2 seroprevalence after third wave of infections, South Africa. *Emerging Infectious Disease journal*. 28(5):1055. DOI: https://doi.org/10.3201/eid2805.220278.
105. Corsini Campioli C, Cano Cevallos E, Assi M, Patel R, Binnicker MJ, O'Horo JC. 2020. Clinical predictors and timing of cessation of viral RNA shedding in patients with COVID-19. *Journal of Clinical Virology*. 130:104577. DOI: https://doi.org/10.1016/j.jcv.2020.104577.

106. Kleynhans J, Walaza S, Martinson NA, Neti M, von Gottberg A, Bhiman JN, et al. 2022. Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: A case-ascertained, prospective observational household transmission study. *Clinical Infectious Diseases*. Published online ahead of print. DOI: https://doi.org/10.1093/cid/ciac640.
107. Sun K, Tempia S, Kleynhans J, Gottberg Av, McMorrow ML, Wolter N, et al. 2022. Rapidly shifting immunologic landscape and severity of SARS-CoV-2 in the Omicron era in South Africa. *medRxiv : the preprint server for health sciences*.2022.08.19.22278993. DOI: https://doi.org/10.1101/2022.08.19.22278993.
108. Barber RM, Sorensen RJD, Pigott DM, Bisignano C, Carter A, Amlag JO, et al. 2022. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: A statistical analysis. *The Lancet*. 399(10344):2351-80. DOI: https://doi.org/10.1016/S0140-6736(22)00484-6.
109. Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. 2022. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: A data linkage study. *The Lancet*. 339(10323):437-46. DOI: https://doi.org/10.1016/S0140-6736(22)00017-4.
110. Madhi SA, Kwatra G, Myers JE, Jassat W, Dhar N, Mukendi CK, et al. 2022. Population immunity and COVID-19 severity with Omicron variant in South Africa. *New England Journal of Medicine*. 386(14):1314-26. DOI: https://doi.org/10.1056/NEJMoa2119658.
111. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Hellferscee O, et al. 2021. Asymptomatic transmission and high community burden of seasonal influenza in an urban and a rural community in South Africa, 2017-18 (PHIRST): A population cohort study. *The Lancet Global Health*. 9(6):e863-e74. DOI: https://doi.org/10.1016/S2214-109X(21)00141-8.
112. Bingham J, Cable R, Coleman C, Glatt TN, Grebe E, Mhlanga L, et al. 2022. Estimates of prevalence of anti-SARS-CoV-2 antibodies among blood donors in South Africa in March 2022. *Research Square*. DOI: https://doi.org/10.21203/rs.3.rs-1687679/v1.
113. Soleimanian S, Alyasin S, Sepahi N, Ghahramani Z, Kanannejad Z, Yaghobi R, et al. 2022. An update on protective effectiveness of immune responses after recovery from COVID-19. *Frontiers in Immunology*. 13:884879. DOI: https://doi.org/10.3389/fimmu.2022.884879.
114. Torresi J, Edeling MA, Nolan T, Godfrey DI. 2022. A complementary union of SARS-CoV-2 natural and vaccine induced immune responses. *Frontiers in Immunology*. 13:914167. DOI: https://doi.org/10.3389/fimmu.2022.914167.
115. Tsang TK, Lau LLH, Cauchemez S, Cowling BJ. 2016. Household transmission of influenza virus. *Trends in Microbiology*. 24(2):123-33. DOI: https://doi.org/10.1016/j.tim.2015.10.012.

116. Tsang TK, Cowling BJ, Fang VJ, Chan K-H, Ip DKM, Leung GM, et al. 2015. Influenza A virus shedding and infectivity in households. *The Journal of infectious diseases*. 212(9):1420-8. DOI: https://doi.org/10.1093/infdis/jiv225.
117. Duval D, Palmer JC, Tudge I, Pearce-Smith N, O'Connell E, Bennett A, et al. 2022. Long distance airborne transmission of SARS-CoV-2: Rapid systematic review. *BMJ*. 377:e068743. DOI: https://doi.org/10.1136/bmj-2021-068743.
118. Picard C-F, Cony Renaud Salis L, Abadie M. 2022. Home quarantine: A numerical evaluation of SARS-CoV-2 spread in a single-family house. *Indoor Air*. 32(5):e13035. DOI: https://doi.org/10.1111/ina.13035.
119. Wood R, Morrow C, Ginsberg S, Piccoli E, Kalil D, Sassi A, et al. 2014. Quantification of shared air: A social and environmental determinant of airborne disease transmission. *PloS One*. 9(9):e106622. DOI: https://doi.org/10.1371/journal.pone.0106622.
120. Prem K, Cook AR, Jit M. 2017. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology*. 13(9):e1005697. DOI: https://doi.org/10.1371/journal.pcbi.1005697.
121. Lange S, Altrock C-M, Gossmann E, Fegert JM, Jud A. 2022. COVID-19: What price do children pay? An analysis of economic and social policy factors. *International Journal of Environmental Research and Public Health*. 19(13):7604. DOI: https://doi.org/10.3390/ijerph19137604.
122. Nathanielsz J, Toh ZQ, Do LAH, Mulholland K, Licciardi PV. 2022. SARS-CoV-2 infection in children and implications for vaccination. *Pediatric Research*. Published online ahead of print. DOI: https://doi.org/10.1038/s41390-022-02254-x.
123. Ferrari S, Blázquez T, Cardelli R, Puglisi G, Suárez R, Mazzarella L. 2022. Ventilation strategies to reduce airborne transmission of viruses in classrooms: A systematic review of scientific literature. *Building and Environment*. 222:109366. DOI: https://doi.org/10.1016/j.buildenv.2022.109366.
124. Sandmann FG, van Leeuwen E, Bernard-Stoecklin S, Casado I, Castilla J, Domegan L, et al. 2022. Health and economic impact of seasonal influenza mass vaccination strategies in European settings: A mathematical modelling and cost-effectiveness analysis. *Vaccine*. 40(9):1306-15. DOI: https://doi.org/10.1016/j.vaccine.2022.01.015.
125. Tan CY, Chiew CJ, Pang D, Lee VJ, Ong B, Lye DC, et al. 2022. Vaccine effectiveness against Delta, Omicron BA.1 and BA.2 in a highly vaccinated Asian setting: A test-negative design study. *Clinical Microbiology and Infection*. Published online ahead of print. DOI: https://doi.org/10.1016/j.cmi.2022.08.002.
126. Lau JJ, Cheng SMS, Leung K, Lee CK, Hachim A, Tsang LCH, et al. 2023. Real-world COVID-19 vaccine effectiveness against the Omicron BA.2 variant in a SARS-CoV-2 infection-naive population. *Nature Medicine*. 29(2):348-57. DOI: https://doi.org/10.1038/s41591-023-02219-5.

127. Link-Gelles R, Ciesla AA, Roper LE, Scobie HM, Ali AR, Miller JD, et al. 2023. Early estimates of bivalent mrna booster dose vaccine effectiveness in preventing symptomatic SARS-CoV-2 infection attributable to Omicron ba.5- and xbb/xbb.1.5-related sublineages among immunocompetent adults - increasing community access to testing program, United States, December 2022-January 2023. *MMWR. Morbidity and mortality weekly report*. 72(5):119-24. DOI: https://doi.org/10.15585/mmwr.mm7205e1.
128. Tartof SY, Xie F, Yadav R, Wernli KJ, Martin ET, Belongia EA, et al. 2023. Prior SARS-CoV-2 infection and COVID-19 vaccine effectiveness against outpatient illness during widespread circulation of SARS-CoV-2 Omicron variant, US flu ve network. *medRxiv : the preprint server for health sciences*.2023.01.10.23284397. DOI: https://doi.org/10.1101/2023.01.10.23284397.
129. Russell MW, Mestecky J. 2022. Mucosal immunity: The missing link in comprehending SARS-CoV-2 infection and transmission. *Frontiers in Immunology*. 13:957107. DOI: https://doi.org/10.3389/fimmu.2022.957107.
130. World Health Organization [Internet]. 2022. COVID-19 vaccine tracker and landscape. Cited 8 October 2022. Available from: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines.
131. Van Egeren D, Novokhodko A, Stoddard M, Tran U, Zetter B, Rogers MS, et al. 2021. Controlling long-term SARS-CoV-2 infections can slow viral evolution and reduce the risk of treatment failure. *Scientific Reports*. 11(1):22630. DOI: https://doi.org/10.1038/s41598-021-02148-8.
132. Dorward J, Khubone T, Gate K, Ngobese H, Sookrajh Y, Mkhize S, et al. 2021. The impact of the COVID-19 lockdown on HIV care in 65 South African primary care clinics: An interrupted time series analysis. *The Lancet HIV*. 8(3):e158-e65. DOI: https://doi.org/10.1016/S2352-3018(20)30359-3.
133. Baselga M, Güemes A, Alba JJ, Schuhmacher AJ. 2022. SARS-CoV-2 droplet and airborne transmission heterogeneity. *Journal of Clinical Medicine*. 11(9):2607. DOI: https://doi.org/10.3390/jcm11092607.
134. Koh WC, Naing L, Chaw L, Rosledzana MA, Alikhan MF, Jamaludin SA, et al. 2020. What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate and associated risk factors. *PloS One*. 15(10):e0240205. DOI: https://doi.org/10.1371/journal.pone.0240205.
135. World Health Organization [Internet]. 2022. Unity studies: Early investigation protocols. Cited 10 August 2022. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/early-investigations.
136. South African COVID-19 Modelling Consortium [Internet]. 2022. SACMC epicemic explorer. Available from: https://www.sacmcepidemicexplorer.co.za/.

137. Aleta A, Martín-Corral D, Pastore y Piontti A, Ajelli M, Litvinova M, Chinazzi M, et al. 2020. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*. 4(9):964-71. DOI: https://doi.org/10.1038/s41562-020-0931-9.
138. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. 2020. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *The Lancet Infectious Diseases*. 20(5):553-8. DOI: https://doi.org/10.1016/S1473-3099(20)30144-4.
139. National Department of Health Republic of South Africa [Internet]. 2022. Changes to COVID-19 quarantine, isolation and contact tracing. Cited 2 September 2022. Available from: https://www.nicd.ac.za/wp-content/uploads/2022/02/Circular-on-Isolation-Quarantine-and-Contact-tracing.pdf.
140. Sharker Y, Kenah E. 2021. Estimating and interpreting secondary attack risk: Binomial considered biased. *PLoS Computational Biology*. 17(1):e1008601. DOI: https://doi.org/10.1371/journal.pcbi.1008601.
141. Misra A, Theel ES. 2022. Immunity to SARS-CoV-2: What do we know and should we be testing for it? *Journal of Clinical Microbiology*. 60(6):e00482-21. DOI: https://doi.org/10.1128/jcm.00482-21.

7.1. Contact details for personal communications listed

1. Mvuyo Makhasi. Database manager. Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases. Email: mvuyom@nicd.ac.za
2. Ciro Cattuto. Scientific Director. ISI Foundation. Email: ciro.cattuto@isi.it

8. Appendices

8.1. Appendix A: Role of the student

Objective 1 (Quantify within-household and community age-specific contact rates in a rural and an urban community in South Africa using a contact diary survey before the SARS-CoV-2 pandemic)

- Conceptualization of the study (in collaboration with colleagues)
- Protocol development
    - Procedures for the contact diary aspect of the PHIRST study
    - Protocol amendment (PHIRST) to the ethics committee
    - Developed data collection tools (contact and time use survey)
    - Developed standard operating procedures
- Managed logistics for fieldwork
- Training of field workers to conduct interviews
- Site visits to monitor fieldwork
- Data management
    - Development of all databases and data quality tools
    - Weekly checks and liaising with sites for query resolution
- Data analysis
    - Identification of best-suited analytic approach (in collaboration with a statistician)
    - Development of all analytic codes (Stata)
    - Execution of analytic codes and interpretation of results
    - Visualization of results
- Manuscript preparation
    - Draft full manuscript
    - Liaising with co-authors for review and addressing all co-author comments
    - Submission to journal
    - Addressing the reviewer's comments and resubmitting to the journal
    - Review proofs

Objective 2 (Assess differences in SARS-CoV-2 community attack rates in PLWH and PNLWH and disease severity across waves of the pandemic using seroprevalence)

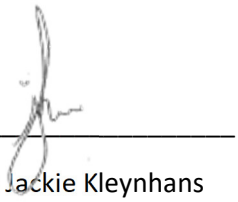
- Conceptualization of the study (in collaboration with colleagues)
- Protocol review for adaption from PHIRST to PHIRST-C
    - Study procedures
    - Data collection tools
    - Developed standard operating procedures
- Assisted with the coordination of fieldwork for PHIRST-C
    - Liaising with sites to ensure required consumables are available
    - Planning of survey dates
    - Arranging transport of specimens
- Data management for PHIRST-C
    - Development of all databases and data quality tools
    - Weekly checks and liaising with sites for query resolution
    - Uploading of results to databases

- Training for PHIRST-C
    - Training of field staff on study procedures and database use
- Data analysis
    - Requests for data from DATCOV (hospitalization and deaths) and SAMRC (excess deaths)
    - Development of all analytic codes (Stata) except for adjusted seroprevalence estimation (based on previously published analytic approach)
    - Development of an Excel spreadsheet for the calculation of infection ratios
    - Execution of analytic codes and interpretation of results
    - Visualization of results
- Manuscript preparation
    - Draft full manuscript
    - Liaising with co-authors for review and addressing all co-author comments
    - Submission to journal
    - Addressing the reviewer's comments and resubmitting to the journal
    - Review proofs

Objectives 3 & 4 (Determine the household cumulative infection risk (HCIR) of SARS-CoV-2 in household contacts of index cases LWH and NLWH; and assess factors associated with transmission, acquisition, duration of SARS-CoV-2 positivity and serial interval; and assess the association between contact patterns and the transmission of SARS-CoV-2 within households)

- Conceptualization of the study (in collaboration with colleagues)
- Protocol development
    - Review of procedures
    - Review collection tools
    - Developed standard operating procedures
- Data management
    - Development of all databases and data quality tools
    - Running checks and liaising with sites for query resolution
    - Uploading of results to databases
- Training of field staff on study procedures and database use
    - Training of proximity coordinator on deployment and data download of contact sensors
- Data analysis
    - Identification of best-suited analytic approach (in collaboration with a statistician)
    - Adaption of analysis codes (Stata) from a previous study for objective 3
    - Development of all analytic codes (R) for objective 4
    - Liaising with ISI foundation for the decryption and cleaning of contact sensor data
    - Execution of analytic codes and interpretation of results
    - Visualization of results
- Manuscript preparation
    - Draft full manuscript
    - Liaising with co-authors for review and addressing all co-author comments
    - Submission to journal
    - Addressing reviewer's comments and resubmitting to journal (objective 3 only)
    - Review proofs (objective 3 only)

All co-authors have been informed that the papers are to be used in this PhD thesis and agree on their use in the thesis.



Candidate: Jackie Kleynhans



Supervisor Cheryl Cohen



Supervisor Stefano Tempia

8.2. Appendix B: Approval of title change

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Private Bag 3 Wits, 2050
Fax: 027117172119
Tel: 02711 7172076

Reference: Mrs Sandra Benn
E-mail: sandra.benn@wits.ac.za

12 March 2021
Person No: 1307593
TAA

Mrs JW Kleynhans
P.O. Box 29364
Sandringham
2131
South Africa

Dear Mrs Jacoba Kleynhans

Doctor of Philosophy: Change of title of research

I am pleased to inform you that the following change in the title of your Thesis for the degree of **Doctor of Philosophy** has been approved:

From:	**Influenza transmission and social contact patterns**
To:	**Severe acute respiratory syndrome corona virus 2 (SARS-CoV-2} transmission dynamics and social contact patterns.**

Yours sincerely

Mrs Sandra Benn
Faculty Registrar
Faculty of Health Sciences

8.3. Appendix C: Ethical clearance certificates

8.3.1. Main thesis protocol clearance

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R49 Ms J Kleynhans

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M191120

NAME: (Principal Investigator)	Ms J Kleynhans
DEPARTMENT:	School of Public Health Medical School University
PROJECT TITLE:	*Severe acute respiratory syndrome corona virus 2 (SARS-Cov-2) transmission dynamics and social contact patterns* Change of study title noted on 2022/09/16
DATE CONSIDERED:	2019/11/29
DECISION:	Approved unconditionally
CONDITIONS:	
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Professor C Cohen and Dr S Tempia
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2019/11/29

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **November** and therefore reports and re-certification will be due in the month of **November** each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator

2022/09/16
Date

8.3.2. PHIRST clearance

University
of the Witwatersrand,
Johannesburg

Human Research Ethics Committee: (Medical)
FWA Registered No IRB 00001223

SECRETARIAT: Suite 189, Private Bag x2600, Houghton 2041, South Africa Tel: +27-11-274 9200 Fax +27-11-274 9281

30 September 2015

FAXED & COURIERED

Dr J Moyes

National Institute for Communicable Diseases
1 Modderfontein Road
Sandringham
Johannesburg
2131

Fax: 011 882 9979

Dear Dr Moyes,

PROTOCOL: PHIRST - A PROSPECTIVE HOUSEHOLD OBSERVATIONAL COHORT STUDY OF INFLUENZA, RESPIRATORY SYNCYTIAL VIRUS AND OTHER RESPIRATORY PATHOGENS COMMUNITY BURDEN AND TRANSMISSION DYNAMICS IN SOUTH AFRICA (THE PHIRST STUDY)

ETHICS REFERENCE NO: 150808

RE : FINAL ETHICS APPROVAL

This is to certify that the above-mentioned trial was reviewed by the University of the Witwatersrand, Human Research Ethics Committee (HREC), and the Protocol Review Committee (PRC) on: 28 August 2015.

The University of the Witwatersrand, Human Research Ethics Committee Approval Granted for the above mentioned study is valid for five years. Where required by Sponsor to have approval on a more frequent basis it remains the responsibility of the Sponsor and Investigator to apply for continuing review and approval, or for the duration of the Trial.

1. THIS APPROVAL IS SUBJECT TO THE FOLLOWING PROVISOS:

* A copy of the MCC Approval and/or MCC Notification letter must be submitted to the Ethics Regulatory Office Secretariat before the study commences / or where an Amendment may be implemented (IF MCC APPROVAL / NOTIFICATION IS APPLICABLE). It remains the responsibility of the Principal Investigator and/or Sponsor to ensure that the relevant approvals are in place.

* The study is conducted according to the protocol submitted to the University of the Witwatersrand, Human Research Ethics Committee. Any amendments to the protocol must first be submitted to the Human Research Ethics Committee for approval.

* During the study, the University of the Witwatersrand, Human Research Ethics Committee is informed immediately of :
  - Any Unexpected Serious Adverse Events or Unexpected Adverse Drug Reactions, which, in the Investigator and/or the Sponsor's opinion are suspected to be related to the study drug. (Refer to POL-IEC-001 and SOP-IEC-005, Item 3.4).
  - Any data received during the trial which, may cast doubt on the validity of the continuation of the study .

* The University of the Witwatersrand, Human Research Ethics Committee is notified of any decision to discontinue the study and the reason stated.

* The investigators authorised by this approval participate in this study. Additional investigators shall be submitted to the University of the Witwatersrand, Human Research Ethics Committee for approval prior to their participation in the study.

* In the event of an authorised Investigator ceasing to participate in the study, the University of the Witwatersrand Human Research Ethics Committee must be informed and the reason for such cessation given.

2. PRINCIPLES OF INFORMED CONSENT:

* The University of the Witwatersrand, Human Research Ethics Committee requires that in all studies, the Principles of Informed Consent are adhered to. This applies to volunteers as well as patients.

3. PROGRESS REPORTS:

* The University of the Witwatersrand, Human Research Ethics Committee requests that the MCC Progress Reports be submitted twice a year either in March and September or six monthly from start of study to the HREC Secretariat Office - 011 274 9281 and a report of the final results, at the conclusion of the study. (IF APPLICABLE)

4. REIMBURSEMENT TO PATIENTS FOR TRANSPORT:

* The Human Research Ethics Committee: (Medical) is in agreement that reimbursement per visit is according to the Medicines Control Council of SA and that reimbursement should be appropriate according to the situation.

5. TRANSPORT AND STORAGE OF BLOOD AND TISSUE SAMPLES IN SOUTH AFRICA:

* If blood specimens are to be stored for future analysis and is planned that such analysis will be done outside Wits, then the blood must be stored at a facility in South Africa agreed with the relevant IRB, with release of sub-samples only once projects have been approved by the local Research Ethics Committee applicable to where the analysis will be done as well as by the Wits Human Research Ethics Committee: (Medical).

6. GENETIC TESTING

* The Human Research Ethics Committee: Medical; will not approve open-ended genetic testing as this does not fit the Human Research Ethics Committee criteria.

7. GOOD CLINICAL PRACTICE

The South African Department of Health, Medicines Control Council requires Good Clinical Practice (GCP) Training for all Investigators in Clinical Trials, and that GCP training be renewed every three (3) years.

As yet, there are no National Guidelines for the content of GCP courses. Until these are available the Wits Human Research Ethics Committee (Medical) will note courses completed by Investigators without approval of the content of the individual courses.

8. THE SUPPORTING APPROVAL DOCUMENTS ARE ATTACHED:

8.1 Ethics Approval Form signed by the Chairperson of the HREC - Kindly return the copy of the Approval Form signed by the Principal Investigator /(s) per fax: 011 274 9281 for our records (this is applicable with the initial Approval).

8.2 Protocol Review Committee Approval Signature page signed by the Acting Chairperson of the PRC.

8.3 List of members present at the HREC meeting held as per INDEPENDENT ETHICS COMMITTEE APPROVAL FORM

9. WE AWAIT YOUR RESPONSES AS REQUESTED: Ensure to have these documents forwarded at the earliest for the HREC records.

* MCC Approval letter and/or letter of Notification before the above study may commence / or where an Amendment may be implemented (IF MCC APPROVAL / NOTIFICATION IS APPLICABLE). It remains the responsibility of the Principal Investigator and/or Sponsor to ensure that the relevant approvals are in place.
* Copy of Independent Ethics Declaration Approval Form signed by the Principal Investigator. (this is applicable with the initial Approval).
* Kindly forward the above to the undersigned at fax: 011 274 9281 at your earliest convenience.

Please ensure that all Investigators have valid GCP training before participating in the above-mentioned study.

Noted: The study will not be conducted in any health care facilities and does not require PRC clearance. The site will apply for provincial ethics clearance on receipt of Wits HREC clearance.

The above has been noted for the Ethics Committee information and records.

KINDLY FORWARD TO THE RELEVANT INVESTIGATORS / CRA / SPONSOR / STUDY CO-ORDINATORS - WHERE APPLICABLE

Regards,
DR MERRYLL VORSTER
For and on behalf of the Human Research Ethics Committee: (Medical)

8.3.3. Inclusion of SARS-CoV-2 in PHIRST (PHIRST-C)

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG | HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

23rd May, 2020

Prof. C. Cohen/Dr J. Moyes
National Institute of Communicable Diseases
Sandringham
Sent by E mail to: cherylc@nicd.ac.za; jossmoyes@gmail.com

Dear Prof Cohen and Dr Moyes

Protocol Ref No's-
M150808:
Protocol Title: ***Prospective Household study of Influenza and Respiratory Syncitial virus community burden, Transmission dynamics and viral interaction in South Africa (the PHIRST Study)***
M160667:
Protocol Title: ***Essential communicable diseases surveillance and outbreak investigation activities of the NICD***

I refer to your HREC application dated 7th May, 2020, in which an amendment is sought for your study entitled "Prospective Household study of Influenza and Respiratory Syncitial virus community burden, Transmission dynamics and viral interaction in South Africa", Ethics clearance M150808. This amendment will include additions in the context of the SARS-CoV-2 pandemic.

I confirm that the HREC, Medical finds the study to be covered within the scope of the blanket clearance granted to the NICD under clearance M160667, this being valid until 17th July, 2021. I would request with the present submission that you amend the Information leaflet 1: consent for adults (page 6) with regards to ensuring confidentiality as indicated (see attached); and also provide an assent form for children to sign or provide a mark where capable. Please provide copies of the amended documents to the HREC, Medical. Please also be aware that M150808 will expire in August, 2020 and will require a short progress report and motivation for renewal.

Yours sincerely

Clement Penny (PhD)
Chair HREC (Medical)

8.3.4. PHIRST-C re-certification

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Human Research Ethics Committee: (Medical)
FWA Registered No IRB 00001223

SECRETARIAT: Suite 189, Private Bag x2600, Houghton 2041, South Africa Tel: +27-11-274 9200 Fax: +27-11-274 9281

28 August 2020

EMAILED

Prof C Cohen,
Centre Head: Epidemiology
Centre for Respiratory Diseases & Meningitis
National Institute for Communicable Diseases
a division of the National Health Laboratory
1 Modderfontein Rd Sandringham
2193

Email: cherylc@nicd.ac.za

Dear Prof Cohen,

PROTOCOL: PHIRST - A PROSPECTIVE HOUSEHOLD STUDY OF SARS-COV-2, INFLUENZA, AND RESPIRATORY SYNCYTIAL VIRUS COMMUNITY BURDEN, TRANSMISSION DYNAMICS AND VIRAL INTERACTION IN SOUTH AFRICA (THE PHIRST-C STUDY)

ETHICS REFERENCE NO: 150808

RE : RE-CERTIFICATION OF CLINICAL PROTOCOL

We acknowledge receipt of your letter dated 27 August 2020 with the following documentation pertaining to the above-captioned trial.

This is to certify that the above-mentioned trial was initially reviewed by the University of the Witwatersrand, Human Research Ethics Committee (HREC): on 28 August 2015 and Approved on 30 September 2015.

The re-certification application for the above-mentioned trial has been reviewed at the HREC meeting dated: 28 August 2020 and is hereby re-certified until the next HREC meeting dated 27 August 2021.

The University of the Witwatersrand, HREC Initial Approval Granted for the above mentioned study is valid for five years. Where required by Sponsor to have approval on a more frequent basis it remains the responsibility of the Sponsor and Investigator to apply for continuing review and approval, or for the duration of the Trial.

PLEASE NOTE: - ANNUAL REVIEW
· Re-certification Applications require formal review and renewal by the Full HREC at the monthly meetings
· Two copies of Re-certification Applications Forms required for circulation to HREC to be tabled as Agenda item
· As per the OHRP Guidelines:
- The date of annual Re-certification will be 1(one) year after the date of the convened meeting where the study was initially reviewed. Re-certification must occur within 1(one) year of the date of the meeting
· The Re-certification dates will not reflect against the protocol amendment approval dates.

Expiry: ☑ **Expiry Date:** 27 Aug 2021

The above has been noted for the Ethics Committee information and records.

KINDLY FORWARD TO THE RELEVANT INVESTIGATORS / CRA / SPONSOR / STUDY CO-ORDINATORS - WHERE APPLICABLE

Regards,

PROF CLEMENT PENNY

For and on behalf of the Human Research Ethics Committee: (Medical)

8.3.5. SA-S-HTS clearance

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

2020/12/24

Professor C Cohen, et al
School of Public Health and
National Institute for Communicable Diseases
Sandringham

Sent by e-mail to: Sibongilew@nicd.ac.za

Dear Dr Walaza

Re: **Protocol Ref No: M2008114**
Protocol Title: *Household transmission of SARS-CoV-2 from HIV-infected and HIV-uninfected adult index cases in South Africa*
Principal Investigator: Professor C Cohen, et al

Thank you for your letter of 2020/12/07.

I confirm that the amendments listed therein, which are too numerous to repeat here, have been noted and approved.

Thank you for keeping us informed.

Yours Sincerely

.....
Mr I Burns
For the Human Research Ethics Committee (Medical)

.....
Dr CB Penny, Chairperson, Human Research Ethics Committee (Medical)

Works2000/Iain0007/Acknowledge.docx

Kleynhans et al. BMC Public Health (2021) 21:1055
https://doi.org/10.1186/s12889-021-11136-6

BMC Public Health

RESEARCH Open Access

A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018

Jackie Kleynhans[1,2*], Stefano Tempia[2,3,4,5], Meredith L. McMorrow[3,4,6], Anne von Gottberg[1,7], Neil A. Martinson[8,9,10], Kathleen Kahn[11], Jocelyn Moyes[1,2], Thulisa Mkhencele[1], Limakatso Lebina[8], F. Xavier Gómez-Olivé[11], Floidy Wafawanaka[11], Azwifarwi Mathunjwa[1], Cheryl Cohen[1,2] and the PHIRST group[1,7,8,11]

Abstract

Background: Describing contact patterns is crucial to understanding infectious disease transmission dynamics and guiding targeted transmission mitigation interventions. Data on contact patterns in Africa, especially South Africa, are limited. We measured and compared contact patterns in a rural and urban community, South Africa. We assessed participant and contact characteristics associated with differences in contact rates.

Methods: We conducted a cross-sectional study nested in a prospective household cohort study. We interviewed participants to collect information on persons in contact with for one day. We described self-reported contact rates as median number people contacted per day, assessed differences in contact rates based on participant characteristics using quantile regression, and used a Poisson model to assess differences in contact rates based on contact characteristics within age groups. We also calculated cumulative person hours in contact within age groups at different locations.

Results: We conducted 535 interviews (269 rural, 266 urban), with 17,252 contacts reported. The overall contact rate was 14 (interquartile range (IQR) 9–33) contacts per day. Those ≤18 years had higher contact rates at the rural site (coefficient 17, 95% confidence interval (95%CI) 10–23) compared to the urban site, for those aged 14–18 years (13, 95%CI 3–23) compared to < 7 years. No differences were observed for adults. There was a strong age-based mixing, with age groups interacting more with similar age groups, but also interaction of participants of all ages with adults. Children aged 14–18 years had the highest cumulative person hours in contact (116.3 rural and 76.4 urban).

Conclusions: Age played an important role in the number and duration of contact events, with children at the rural site having almost double the contact rate compared to the urban site. These contact rates can be utilized in mathematical models to assess transmission dynamics of infectious diseases in similar communities.

Keywords: Contact diaries, Urban, Rural, Infectious disease modelling

* Correspondence: jackiel@nicd.ac.za
[1]Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa
[2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
Full list of author information is available at the end of the article

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Contact patterns within communities impact disease transmission and may inform development of interventions to reduce transmission [1]. To evaluate the impact of both pharmaceutical and non-pharmaceutical interventions on disease transmission, mathematical models are utilized when intervention studies may be unethical or too expensive to perform [2]. Mathematical modelling results are only as reliable as the data used for parametrization and the validity of the assumptions made. Disease transmission models usually require accurate age-related contact patterns [3].

Although sub-Saharan Africa has the highest burden of infectious diseases globally, [4] data on contact patterns within the region are limited, making modelling infectious disease transmission challenging [5]. Contact studies have been carried out in Kenya, [6, 7] Uganda, [8, 9] Senegal, [10] Zimbabwe [11] and South Africa [12, 13]. The South African studies were performed in Cape Town communities, and it is unknown how contact patterns reported in Cape Town compare to other South African communities.

The Prospective Household cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa (PHIRST study) was a three-year, randomly selected, household-level community cohort study conducted at two sites, from 2016 through 2018 [14]. The primary objectives of the study were to estimate the community burden of and assess the transmission dynamics of influenza, respiratory syncytial virus (RSV), *Bordetella pertussis* and *Streptococcus pneumoniae* colonization. To better understand the role of contact patterns on the transmission of these organisms, we nested a contact diary study in the final year of the PHIRST study.

We aimed to measure and compare contact patterns in a rural and an urban community in South Africa using interviewer completed contact diaries, and to assess participant and contact characteristics that influence differences in contact rates.

Methods

Study population

From August to October 2018, participants in a prospective household cohort study in two sites were invited to participate in the contact survey. One hundred and seventeen households, 56 households with 282 participants at the rural site, and 61 households with 285 participants at the urban site, were included. In short, PHIRST was a three-year household-level community cohort study conducted in a rural community in Bushbuckridge Municipality (Mpumalanga Province) and an urban area in the Matlosana Municipality (North West Province) from 2016 through 2018. A random selection of approximately 50 households were enrolled each year from 2016 through 2018. The rural site formed part of a health and socio-demographic surveillance site (HDSS) at the Medical Research Council (MRC)/University of Witwatersrand Rural Public Health and Health Transitions Research Unit, Agincourt. At the rural site, a list of households with 3 or more members were requested from the HDSS and at the urban site, a list of 450 global positioning system (GPS) coordinates within the Jouberton geographical area was generated. Households from these lists were approached sequentially and were eligible for enrolment if it consisted of 3 or more household members (sharing at least 4 meals a week) and > 80% of household members consented to participant in the study. As part of the main study, households were visited twice weekly for 8–10 months for collection of upper respiratory specimens, symptom and healthcare consultation information from all participating household members. Sample size calculations for the cohort study were based on the primary objective that focussed on transmission dynamics of influenza and RSV, and was based on 10% risk of infection, with 95% confidence intervals and 5% precision [14].

Data collection

We interviewed participants of the PHIRST study once to ascertain information on all the persons they were in close contact with for one day. For each participant's contact diary, we collected age (or age group if age in years was unknown: < 7, 7–13, 14–18, 19–64, or ≥ 65 years); sex; if contact occurred in enclosed space or not; and whether there was physical contact for each person contacted on the day. We also ascertained if the contact was a member of the participant's household; total time spent with the contact during the day (< 5 min, 5–14 min, 15–59 min, 1–4 h, > 4 h); where the contact took place (home, school, work, transport or other) and how often the participant came into contact with the person (daily/almost daily, once or twice a week, once or twice a month, less than once a month, or never met before).

We assigned each household a random day of the week to measure contact events. Fieldworkers visited each household the day prior to the measurement day to explain study procedures to all household members. Fieldworkers informed participants of what information would be collected, provided a blank notebook and pen and encouraged participants to make notes that might be of use during the interview. The day following the measurement day, the field worker visited the household again and conducted an interview with each participant to complete the contact survey (Additional file 4). A time use survey was completed at the same time to assist with recall of contacts. If the entire household or a

participant within the household was not available for interview on the selected day, a more convenient day was chosen, based on the age distribution of individuals allocated to the days of the week, giving preference to ensure that there were children aged < 5 years and elderly individuals ≥65 years allocated to each day of the week. The investigation spanned from 14 August – 14 October and 2 August – 26 September at the rural and urban site, respectively. There were 2 public holidays within this period, which were not investigated for any participants. The school holiday was from 29 September – 8 October, in which period no investigations were done, except for 9 surveys from the rural site performed on the first day of the holiday which fell on a Saturday. We collected both contact events within the household and the community using the contact diaries and captured data on a REDCap database [15, 16].

Contact classification and contact rates

We investigated contact rate as the primary outcome, which we defined as the median number of people contacted per day per group (based on characteristics of the study participant such as age, sex, etc.). A contact was defined as a two-way conversations of at least three words at a distance not requiring voices being raised (< 2 m between individuals), with or without physical contact, or where physical contact took place without conversation. We further classified contacts as physical and non-physical, where physical contacts included kissing, touching, or hand shaking.

We analyzed the data using Stata 14.2 (StataCorp, College Station, Texas, USA). We accounted for within-household clustering using the Taylor-linearised variance estimation ("svy" Stata function) when calculating contact rates. Ninety-five percent confidence intervals for median contact rates were obtained using the 'epctile' function. Stratified contact rates are only presented where data were available for four or more participants within the stratum.

Participant characteristics associated with contact rate

To assess differences in contact rates based on participant characteristics, or day investigated, we used quantile regression, with 500 bootstrap replications and accounting for clustering at site and household level [17]. Quantile regression allows univariate and multivariable comparison of percentiles (including medians) between groups or categories [18]. This analytical approach was selected a priori due to the right skewed nature of contact data. Contact events are not normally distributed, but rather follow a power-low distribution where several individuals have few contacts and a few have large numbers of contacts [19], making the use of linear regression unsuitable for the analysis of contact events. The quantile regression model outputs a coefficient that quantifies the increase/decrease in median number of contact events for each category of the evaluated predictor when compared to the reference category. Predictors were first assessed on univariate analysis, and those with a p-value of < 0.2 were included in multivariable analyses. We did separate analysis for children (aged 0–18 years) and adults (≥19 years). We assessed site, age, relationship to head of household, sex, if day investigated was a weekday (Monday-Friday) or weekend (Saturday/Sunday), day of the week investigated and HIV status. For adults we also assessed education, employment, alcohol consumption, and cigarette smoking. Backwards elimination was done by excluding categories with the highest p-values until all remaining predictors had a p-value of < 0.05.

Age-related mixing patterns and contact characteristics associated with differences in contact rates

To investigate age-related mixing patterns, we calculated the total number of contacts each age group had with a specified age group (< 7, 7–13, 14–18, 19–64 and ≥ 65 years) and then calculated the proportion of contacts with that age group out of total contacts for the day.

To assess differences in contact rates based on the characteristics of the contact (contact type, contact with a household member, duration of contact, frequency of contact and location of contact), we utilized a Poisson model for each age group described above. Lack of over-dispersion was assessed using a negative binomial model, hence a Poisson model was retained for the analysis.

Cumulative person hours in contact

As a secondary outcome, we estimated the total time spent in contact with persons. We assigned a minute value to each category of total time spent with the contact, based on the mid-point for the first 4 categories, and the lower point for the last. We assigned: < 5 min = 2.5, 5–14 min = 7.5, 15–59 min = 30, 1–4 h = 120, > 4 h = 240. For each age category and location, the contact took place we calculated cumulative person hours in contact = (summed total minutes for each contact / participant in age category)/60. Since contact could have been with multiple people simultaneously, the time spent in contact could be more than the 24-h reporting period. Where more than one location was selected for the contact, the time spent counted towards both locations.

Results

Study population

We conducted 535 interviews from August through October 2018, representing 95% of the 2018 PHIRST cohort (N = 565), of which 269/280 (96%) were conducted in the rural, and 266/285 (93%) in the urban site

(Additional file 1). The average household size was five individuals at each site (range 3–11 at the rural site; 4–10 at the urban site). There were more females ($n = 337$, 63%) included than males. Most adults were unemployed (127/225, 56% Additional file 2). The most common relationship to the head of the household was being their child (165/423, 39%), followed by being a grandchild (91/423, 22%).

Contact rates

Overall, 17,252 contacts were reported by 535 participants, of which 59% (10,304) were reported from the rural site. Many individuals reported a low number of contacts and few individuals reported a high number of contacts. The overall contact rate was 14 (IQR 9–33, Fig. 1, Additional file 2), with the highest number of contacts from one participant reported as 153. There were 58/269 (21.6%) participants at the rural site and 20/266 (7.5%) at the urban site that reported more than 50 contacts for the day. We observed a higher contact rate at the rural site compared to the urban site (21, IQR 11–46 vs 12, IQR 7–18, $p = 0.001$). The highest contact rate was observed in children aged 14–18 years, with a contact rate of 22 contacts (IQR 12–91) per day (Fig. 1, Additional file 2). For both sites, higher contact rates were reported for Wednesdays and Thursdays (Additional file 2).

Participant characteristics associated with contact rate

Using quantile regression (median difference) site, age and day of the week investigated were found to be associated with differences in contact rates for children (aged ≤18 years, Table 1). There was a higher contact rate at the rural site compared to the urban site (coefficient 17, 95% CI 10 to 24). Compared to those < 7 years old, children 7–13 years and 14–18 years had a higher contact rate (9, 95% CI 4–14 and 13, 95% CI 4–22). Compared to Saturdays, there were more contacts on Wednesdays (18, 95% CI 4–32) and Thursdays (12, 95% CI 1–23). HIV status was not associated with number of contacts for children or adults. For adults there were no factors associated with number of contacts on multivariable analysis (Table 2).

Age-related mixing patterns and contact characteristics associated with differences in contact rates

There was a strong age-based assortment (mixing), with age groups interacting with similar age groups, but also a higher interaction of participants of all ages with other adults (Fig. 2).

Using Poisson regression, we found children aged < 7 years were more likely to have contact with children their own age (risk ratio (RR) 8.2, 95%CI 5.8–11.6) than those aged ≥65 years, to have contact with the same people daily (RR 16.1 95%CI 6.7–38.7) compared to once or twice a month, and to have longer contacts (> 4 h, RR 2.9 95%CI 1.9–4.4) compared to those lasting < 5 min (Table 3). Children aged 7–13 and 14–18 years had a similar contact profile, except for being more likely to have contact with household members (RR 1.3 95%CI 1.2–1.4 and RR 1.2 95%CI 1.1–1.4, respectively) than

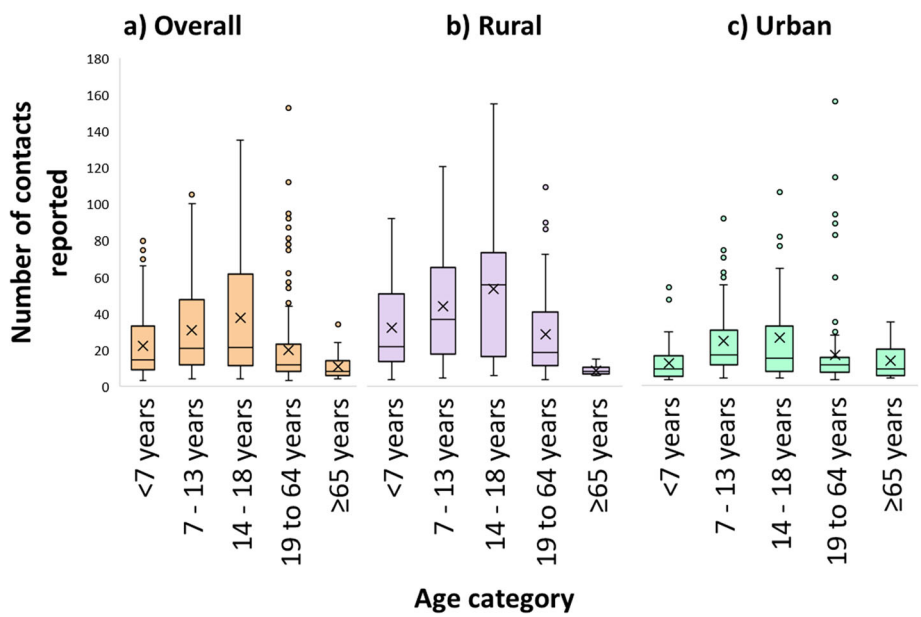


Fig. 1 Contact rate by age group, South Africa, 2018. Physical and conversational contacts (**a**) overall, (**b**) rural and (**c**) urban site. Line represents the median, cross represents the average, box represents the 25th and 75th percentile, whiskers represent 1st and 99th percentile, circles indicate outliers

Table 1 Participant characteristics associated with differences in median number of people contacted for participants ≤18 years, South Africa, 2018

	Univariate		Multivariable	
	Coefficient	**95% CI**	**Coefficient**	**95% CI**
Site				
Rural	15	3 to 27	17	10 to 24
Urban	Reference	–	Reference	–
Age				
< 7 years	Reference	–	Reference	–
7–13 years	6	1 to 11	9	4 to 14
14–18 years	7	– 7 to 21	13	4 to 22
Relationship to head of household				
Grandchild	1	– 11 to 13		
Child	2	– 10 to 14		
Other child*	Reference	–		
Gender				
Male	2	–4 to 8		
Female	Reference	–		
Type of day investigated				
Weekend day	Reference	–		
Weekday	12	3 to 21		
Day of week investigated				
Monday	5	–1 to 11	1	–7 to 9
Tuesday	7	–4 to 18	2	–7 to 11
Wednesday	35	7 to 63	18	4 to 32
Thursday	27	11 to 43	12	1 to 23
Friday	9	–6 to 24	Omitted**	–
Saturday	Reference	–	Reference	–
Sunday	3	0 to 6	1	–4 to 6
HIV Status				
Negative	Reference	–		
Positive	5	–28 to 38		

95% CI – 95% confidence interval
*Other child – Includes relationships of participants younger than 18 years to the head of household (n): brother (10), sister (11), wife (1), cousin (5), niece (6), nephew (9), other relative (10), not related (2)
**omitted because of collinearity

non-household members, and those aged 14–18 years being less likely to have contacts of 1–4 h duration (RR 0.6 95%CI 0.5–0.8) compared to contacts lasting < 5 min. Adults aged 19–64 years also had shorter contacts and were less likely to see people daily compared to once or twice a month (RR 0.3 95%CI 0.2–0.3). Adults ≥65 years were more likely to come into contact with someone they never met before (RR 5.8 95% CI 1.8–18.9) than someone they see once or twice a month, and less likely to have contacts at home (RR 0.3 95%CI 0.2–0.4) than in public transport (Table 3).

Cumulative person hours in contact

A total of 339.8 and 257.5 cumulative person hours in contact were reported for the rural and urban site, respectively. Children aged 14–18 years at both sites had the highest cumulative person hours in contact (116.3 h rural and 76.4 h urban, Fig. 3 and Additional file 3). For children between 7 and 18 years the most person hours in contact were at school (75.9–116.3 h). Children < 7 years at the rural site spent more time at school (47.1 h) compared to home (17.4 h), whereas those at the urban site had 12.4 person hours in contact at school and 20.0 person hours in contact at home. Adults (19–64 years) had most of their person hours in contact at home (rural 17 h, urban 17.8 h) or another location (rural 13.7 h and urban 11.5 h). The elderly (≥65 years) had the lowest number of person hours in contact of all the ages, most being at home (rural 13.4 h, urban 19.8 h). The elderly at the urban site had the highest person hours in contact during transport (7.4 h) than other ages.

Discussion

In this study including participants from randomly selected households, site, age and the day of the week were independently associated with differences in contact rates in children ≤18 years, estimated with contact diary survey data representing > 17,000 contacts. We observed higher contact rates among participants in the rural site, among children aged less than 14 years and on Wednesdays and Thursdays. Children reported seeing the same individuals daily, whereas adults and the elderly had more variation in how often they were in contact with the same individual.

The difference contact rates observed between the rural and urban site (21 and 12, respectively) in our study is in contrast to what was seen in Zimbabwe where individuals at the urban site had a greater number of contacts than the rural site (median of 11 vs. 10, respectively). Although the number of contacts at the urban site for both South Africa and Zimbabwe was similar, there was a lower number of contacts at the rural site in Zimbabwe compared to South Africa. This could also be influenced by cultural, household composition and geographical differences between the sites. We observed that these differences were driven by children. The contact rate recorded in a prior study of the Cape Town community in South Africa was more comparable to the rural site, with 20 in Cape Town compared to 21 contacts at the rural and 12 at the urban site in our study [13].

A contact study performed in rural Uganda found the mean number of contacts per person as 7, [8] and in Kenya 18 contacts were observed per day, the trend of having more contacts in a rural setting compared to a semi urban setting was also observed [6]. This highlights

Table 2 Participant characteristics associated with differences in median number of people contacted for participants > 18 years, South Africa, 2018

	Univariate	
	Coefficient	**95% CI**
Site		
Rural	2	0 to 4
Urban	Reference	–
Age		
19–64 years	2	0 to 4
≥65 years	Reference	–
Relationship to head of household		
Head of home	3	−1 to 7
Parent	Reference	–
Grandchild	3	−3 to 9
Child	3	−1 to 7
Other adult*	2	−2 to 6
Gender		
Male	0	−2 to 2
Female	Reference	–
Type of day investigated		
Weekend day	2	0 to 4
Weekday	Reference	–
Day of week investigated		
Monday	1	−3 to 5
Tuesday	1	−2 to 4
Wednesday	3	−1 to 7
Thursday	2	−2 to 6
Friday	Reference	–
Saturday	4	0 to 8
Sunday	3	−1 to 7
Education		
No schooling	2	0 to 4
Primary	1	−1 to 3
Some secondary	Reference	–
Completed secondary	2	0 to 4
Post-secondary	3	−1 to 7
Employment		
Employed	2	0 to 4
Unemployed	Reference	–
Drinking alcohol		
No	2	1 to 3
Yes	Reference	–
Smoking cigarettes		
No	2	1 to 3
Yes	Reference	–
HIV Status		
Negative	1	0 to 2
Positive	Reference	–

95% CI – 95% confidence interval

* Other adult – Includes relationships of participants 18 years and older to the head of household (n): brother (7), sister (7), wife (19), husband (4), cousin (1), niece (4), nephew (2), other relative (3), not related (4)

the variability that can be observed in contact rates between different settings, and most likely the underlying culture and social-demographic situation may dictate contact rates, but could also be attributed to differences in study methods. Surveys where participants are informed in advance of information that will be collected generally results in less recall bias, and self-completed contact diaries can allow a large sample size as less interviewers are needed, but have a lower response rate [20].

We observed a strong age-based mixing pattern as with many prior contact studies, [6, 11, 21] which was most pronounced for the 14 to 18 year old age group. The highest contact rate was observed in children, especially those of primary school age (7–13 years). This also holds true for studies done in Britain, [22] China, [21] Kenya [6] and Zimbabwe [11]. Certain age groups were also more or less likely to have contact with other age groups, have contact with individuals for certain durations and frequencies and have contacts at certain locations. This information can be useful when looking at targeted intervention strategies to reduce disease transmission in specific age groups, for example providing transport for the elderly where safe physical distancing can be practiced.

The cumulative person hours in contact was also highest in children and dominated by school contacts. Schools have previously been described as a focal point for contact and transmission events for respiratory viruses [23–25]. School closure has been shown to reduce influenza transmission – whether done so for holidays or in response to an outbreak [26–28]. However, the implementation of reactive school closures is challenging, and comes at a social and economic cost [28]. Adaption of schools to accommodate physical and social distancing is a promising alternative [29]. During the coronavirus disease 2019 (COVID-19) pandemic, school closures were also used as part of a multi-faceted approached to curb the spread of the Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), [30, 31] although it seems that

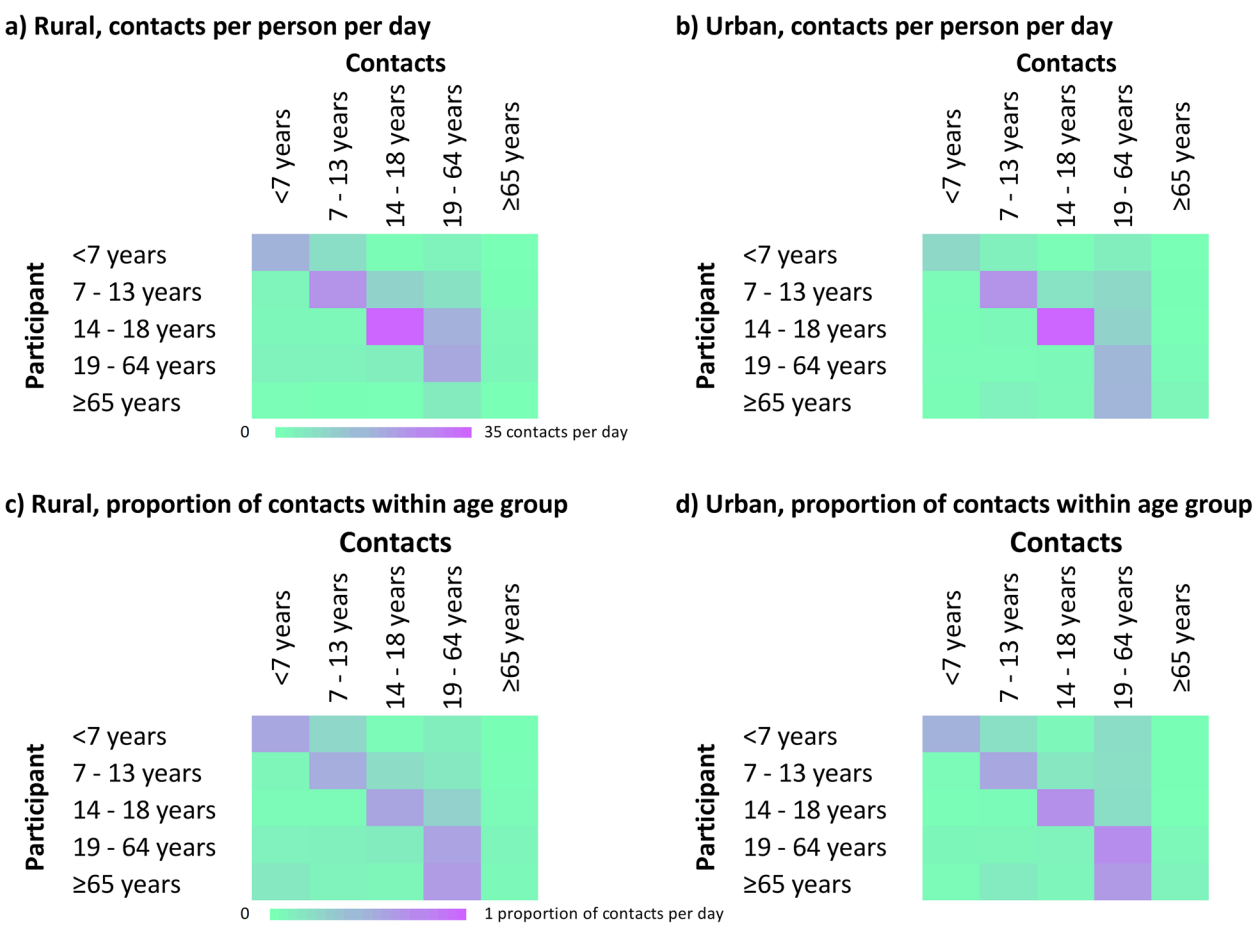


Fig. 2 Age-related contact matrix for total number of people contacted, South Africa, 2018. Physical, conversational and confined contacts by (**a**, **c**) rural and (**b**, **d**) urban site, showing (**a**, **b**) number of contacts and (**c**, **d**) proportion of contacts within age groups, with participant age on y axis and contact age on x axis

specifically for SARS-CoV-2, schools may not be a specific driver for transmission [32].

Sex, education status, employment, alcohol use and smoking cigarettes were not found to be independent factors associated with number of contacts in our study. Employment was found to influence the number of contacts in Britain, but was also dependent on type of employment [22]. A previous study within a South African community in Cape Town found differences in contact rates based on employment status and level of education, although different methods were utilized [13]. A greater number of contacts was observed for employed individuals, and those with secondary school education [13]. The high proportion of study participants which were unemployed, and low proportion with secondary education in our study may have contributed to not observing differences between these groups. This also limits the generalizability of our results to other communities.

Our study had limitations. Compared to other contact surveys preformed, we had a smaller sample size which could lead to results not being representative for the community. The advantage however is that households were randomly selected. The characterization of contact patterns in the PHIRST cohort will also be used in future research on influenza transmission patterns in these households. We only surveyed one day per participant and therefore cannot describe variability in day-to-day contacts by individual participants. We used this sampling frame to prevent fatigue and subsequent reduction in reported number of contacts as seen in other studies [33]. We were not able to measure differences in holiday and non-holiday periods. Relationships within the households were complex and capturing function within the household (breadwinner, child-caregiver, etc.) might have led to differences in contact rates, but the information was not collected for the PHIRST cohort. As a general limitation of the contact survey, we could have missed some contact

Table 3 Contact characteristics associated with differences between contact rates within age groups, South Africa, 2018

	Risk ratio (95% confidence interval)					
	All ages	**< 7 years**	**7–13 years**	**14–18 years**	**19–64 years**	**≥65 years**
Type of contact						
Physical	**1.2 (1.1 to 1.2)**	**1.5 (1.4 to 1.6)**	1.0 (0.9 to 1.0)	**0.8 (0.7 to 0.9)**	0.9 (0.9 to 1.0)	**0.3 (0.2 to 0.4)**
Non-physical	Reference					
Age of contact						
< 7 years	**6.1 (5.5 to 6.8)**	**8.2 (5.8 to 11.6)**	**0.6 (0.4 to 0.8)**	**0.5 (0.3 to 0.8)**	**0.3 (0.2 to 0.3)**	**0.3 (0.1 to 0.7)**
7–13 years	**8.5 (7.7 to 9.5)**	**2.9 (2.0 to 4.1)**	**3.5 (2.7 to 4.4)**	**0.4 (0.2 to 0.6)**	**0.2 (0.2 to 0.2)**	0.4 (0.2 to 1.0)
14–18 years	**6.7 (6.0 to 7.5)**	**0.6 (0.4 to 0.8)**	**1.8 (1.4 to 2.4)**	**6.3 (4.4 to 9.1)**	**0.2 (0.2 to 0.3)**	**0.2 (0.1 to 0.5)**
19–64 years	**11.9 (10.7 to 13.2)**	1.2 (0.8 to 1.7)	0.9 (0.7 to 1.1)	**1.9 (1.3 to 2.8)**	0.9 (0.8 to 1.0)	1.3 (0.6 to 2.9)
≥65 years	Reference					
Gender of contact						
Male	Reference					
Female	1.2 (1.1 to 1.2)	1.0 (0.9 to 1.1)	0.8 (0.8 to 0.9)	1.0 (0.9 to 1.1)	**1.2 (1.1 to 1.3)**	1.1 (0.8 to 1.5)
Not reported	0.1 (0.1 to 0.1)	1.4 (1.2 to 1.7)	0.4 (0.3 to 0.5)	1.2 (0.9 to 1.5)	1.3 (1.1 to 1.6)	NEO
Contact is household member						
No	**3.9 (3.7 to 4.0)**	0.9 (0.8 to 1.0)	**1.3 (1.2 to 1.4)**	**1.2 (1.1 to 1.4)**	**0.8 (0.8 to 0.9)**	**0.4 (0.3 to 0.5)**
Yes	Reference					
Total time spent with contact during the day						
< 5 m	Reference					
5-15 m	**2.5 (2.2 to 2.9)**	1.3 (0.8 to 2.1)	1.4 (1.0 to 2.1)	0.8 (0.6 to 1.1)	0.9 (0.8 to 1.1)	0.6 (0.2 to 1.8)
15 m-1 h	**8.7 (7.7 to 9.9)**	**2.0 (1.3 to 3.1)**	**1.6 (1.1 to 2.2)**	0.8 (0.6 to 1.1)	**0.8 (0.7 to 0.9)**	0.8 (0.3 to 2.2)
1-4hts	**16.8 (14.8 to 19.1)**	**2.9 (1.9 to 4.5)**	**2.5 (1.8 to 3.5)**	**0.6 (0.5 to 0.8)**	**0.5 (0.4 to 0.6)**	0.7 (0.3 to 1.7)
> 4 h	**22.1 (19.5 to 25.0)**	**2.9 (1.9 to 4.4)**	**2.3 (1.6 to 3.2)**	1.0 (0.7 to 1.3)	**0.4 (0.4 to 0.5)**	0.5 (0.2 to 1.3)
Frequency of contact						
Daily/almost daily	**34.6 (30.8 to 38.8)**	**16.1 (6.7 to 38.7)**	**8.9 (4.9 to 16.1)**	1.4 (1.0 to 2.0)	**0.3 (0.2 to 0.3)**	0.8 (0.3 to 2.7)
Once or twice a week	**5.7 (5.1 to 6.5)**	**5.5 (2.2 to 13.3)**	**4.4 (2.4 to 8.1)**	1.3 (0.9 to 1.9)	**0.7 (0.6 to 0.8)**	2.4 (0.8 to 7.9)
Once or twice a month	Reference					
< 1 a month	**0.7 (0.6 to 0.9)**	0.8 (0.2 to 3.5)	0.4 (0.1 to 1.3)	No observations	1.2 (1.0 to 1.4)	0.5 (0.0 to 4.4)
Never met before	**1.8 (1.6 to 2.1)**	**3.9 (1.5 to 10.0)**	**4.5 (2.4 to 8.4)**	**0.5 (0.3 to 0.8)**	**0.8 (0.7 to 0.9)**	**5.8 (1.8 to 18.9)**
Location of contact						
Home	**3.6 (3.3 to 3.9)**	1.2 (1.0 to 1.5)	**2.1 (1.6 to 2.8)**	1.1 (0.8 to 1.5)	**0.8 (0.7 to 0.9)**	**0.3 (0.2 to 0.4)**
School	**8.2 (7.7 to 8.8)**	**1.5 (1.2 to 1.8)**	**3.7 (2.8 to 4.9)**	**2.6 (2.0 to 3.5)**	0.1 (0.1 to 0.1)	No observations
Work	**0.8 (0.7 to 0.8)**	No observations	0.0 (0.0 to 0.1)	0.0 (0.0 to 0.1)	2.0 (1.7 to 2.3)	0.0 (0.0 to 0.1)
Transport	Reference					
Other	**5.1 (4.7 to 5.5)**	0.8 (0.6 to 1.0)	**2.2 (1.7 to 2.9)**	1.2 (0.9 to 1.7)	0.9 (0.8 to 1.1)	0.1 (0.1 to 0.1)

Cells indicate risk ratio with 95% confidence interval in parenthesis. Reference group chosen as category with least number of contacts when combining all ages, except where lowest number were in a category where some age groups have zero contacts, then second lowest used (frequency)

events, especially in younger children and shorter contacts due to recall bias. We believe by employing the time-use survey to guide participants and interviewers through the day, missed contact events would have been reduced.

Conclusions

We used contact diaries to describe the contact patterns in a rural and urban South African community. Age played an important role in the number and duration of contact events, with children having the highest contact rates; most of this contact time occurred at school. Children at the rural site included in our study had almost twice the contact rate than children from the urban site, and contact rates also differed based on day of the week. Our results, compared to previous studies, highlight how contact rates may differ based on the setting, and this

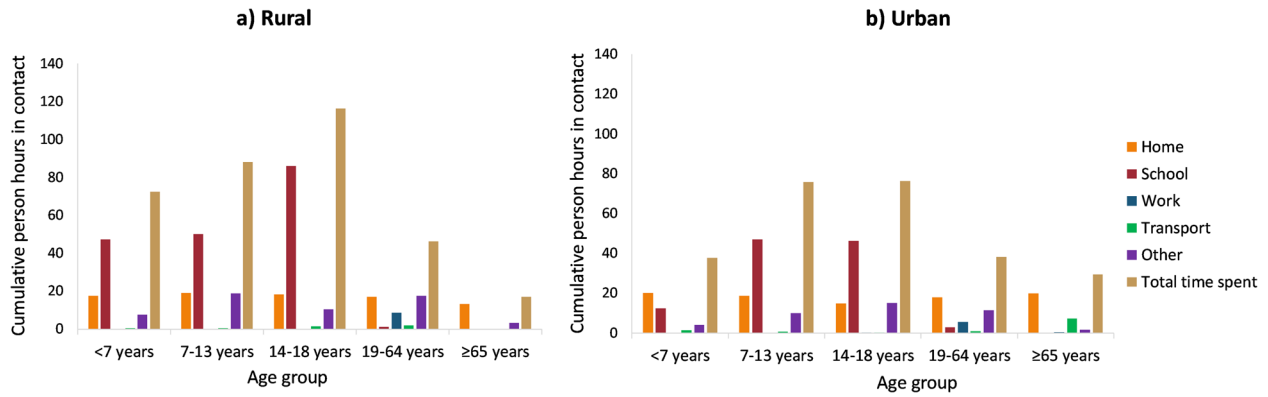


Fig. 3 Cumulative person hours in contact reported by age, location and site, South Africa, 2018

could influence results when building transmission models. These data will be useful when building infectious disease models in similar communities to assess transmission dynamics and possible intervention strategies. A specific approach to explore would be interventions to reduce contacts, and therefore disease transmission, within schools.

Abbreviations

95% CI: 95% confidence interval; COVID-19: Coronavirus disease 2019; IQR: Interquartile range; PHIRST study: Prospective Household cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa; RR: Risk ratio; RSV: Respiratory syncytial virus; SARS-CoV-2: Severe acute respiratory syndrome corona virus 2 (SARS-CoV-2)

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-021-11136-6.

Additional file 1. Study population for contact survey at rural and urban site, South Africa, 2018

Additional file 2. Description of study population and contact rates at rural and urban site, South Africa, 2018.

Additional file 3. Cumulative person hours in contact reported by age, location and site, South Africa, 2018.

Additional file 4. Contact survey and time use survey.

Additional file 5. Contact survey data. (CSV 5071 kb)

Acknowledgements

We would like to thank Matebejane Moroe, and Sizzy Ngobeni for the coordination of field activities and Jennifer Kabasia, Joseph Lechalaba, Munei Nthulana, Mpho Sekwele, Nakedi Maredi, Precious Lubisi, Patrick Mashiloane, Themby Nkovana for their devotion to the project and conducting the interviews.

The PHIRST Group (excluding named authors)

Amelia Buys[1], Angela Mathee[12], Brigitte Language[13], Lorens Maake[1], Florette Treurnicht[1], Katlego Mothlaoleng[8], Maimuna Carrim[1], Nicole Wolter[1,7], Orienka Hellferscee[1], Ryan G Wagner[11], Stuart Piketh[13].

[12]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa.

[13]Unit for Environmental Science and Management, School of Geo- and Spatial Science, North-West University, Potchefstroom, South Africa.

Disclaimer

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the funding agencies.

Authors' contributions

ST, MLM, AvG, NAM, KK, JM, TM, CC contributed to conception and design of protocol of the PHIRST study. JK, ST, CC conceptualised the study. CC, ST provided supervision. NAM, KK, LL, FXG, AM was responsible for site supervision and management. JM, AM did project administration. JK, TM managed databases and curated data. JK did formal analysis, ST, MLM, NAM, KK, LL, FXG, FW, CC provided project resources. JK drafted the original manuscript. The author(s) read and approved the final manuscript.

Funding

The study was funded by a cooperative agreement between the National Institute for Communicable Diseases of the National Health Laboratory Service and the United States Centers for Disease Control and Prevention (grant number 1U01IP001048).

Availability of data and materials

All data generated and analysed during this study are included in this published article and its supplementary information files (additional file 5).

Declarations

Ethics approval and consent to participate

The study was approved by the University of Witwatersrand Human Research Ethics Committee (Reference 150808 and M191120). All methods were carried out in accordance with relevant guidelines and regulations. All participants provided written informed consent before participating in the study, if subjects were < 18 years, written informed consent was obtained from a parent and/or legal guardian, with children aged 7–17 years also providing assent.

Consent for publication

Not applicable.

Competing interests

CC has received grant support from Sanofi Pasteur, Advanced Vaccine Initiative, U.S. Centers for Disease Control and Prevention, and payment of travel costs from Parexel. All other authors have no competing/conflict of interest. All other authors have no competing/conflict of interest.

Author details

[1]Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. [2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [3]Influenza Division, Centers for Disease Control and Prevention, Atlanta, GA, USA. [4]Influenza Program, Centers for Disease Control and Prevention,

Pretoria, South Africa. [5]MassGenics, Duluth, Georgia, USA. [6]United States Public Health Service, Rockville, MD, USA. [7]School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [8]Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. [9]Johns Hopkins University Center for Tuberculosis Research, Baltimore, MD, USA. [10]Department of Science and Technology/National Research Foundation Centre of Excellence for Biomedical Tuberculosis Research, University of the Witwatersrand, Johannesburg, South Africa. [11]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [12]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa. [13]Unit for Environmental Science and Management, School of Geo- and Spatial Science, North-West University, Potchefstroom, South Africa.

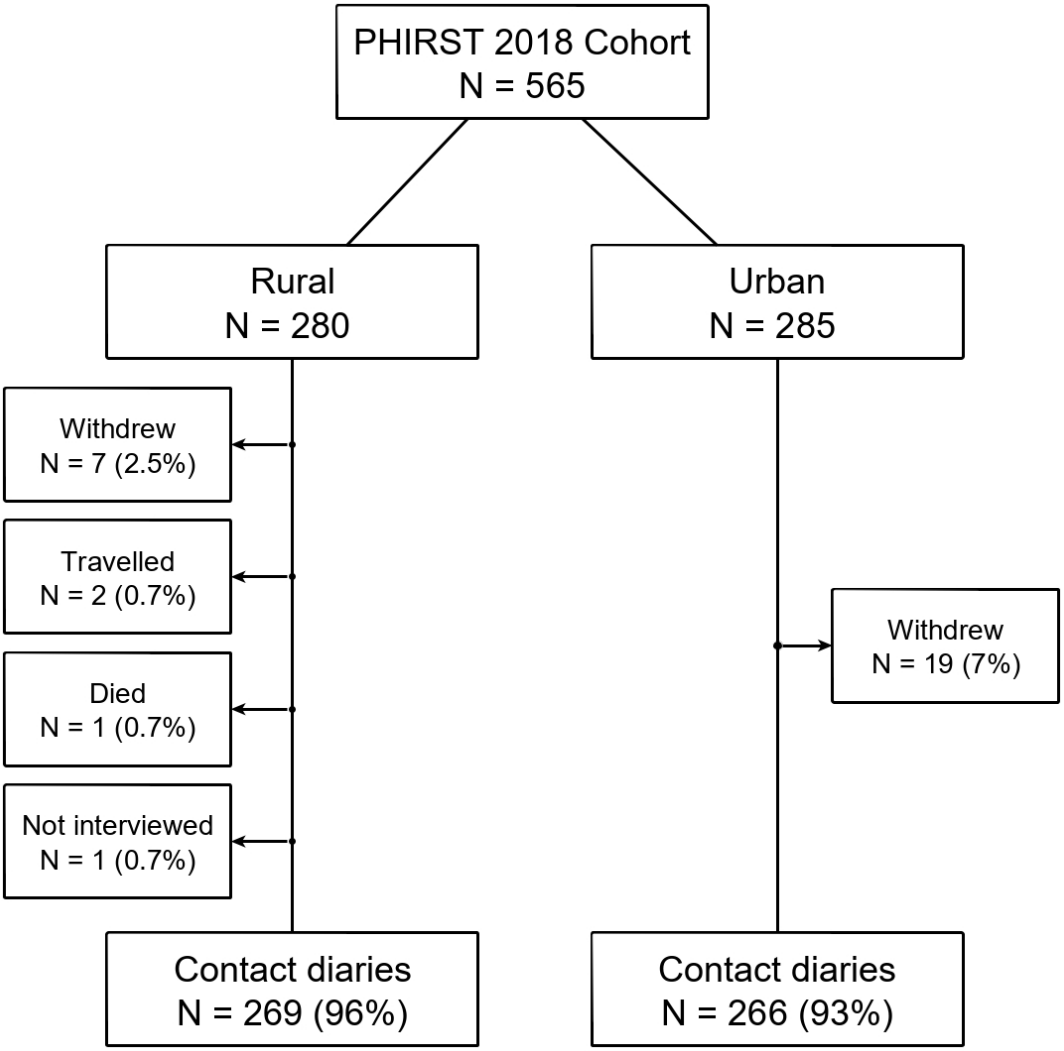
Received: 14 January 2021 Accepted: 24 May 2021
Published online: 03 June 2021

References

1. Yin Q, Shi T, Dong C, Yan Z. The impact of contact patterns on epidemic dynamics. PLoS One. 2017;12(3):e0173411. https://doi.org/10.1371/journal.pone.0173411.
2. Willem L, Verelst F, Bilcke J, Hens N, Beutels P. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). BMC Infect Dis. 2017;17:612.
3. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. Virulence. 2013;4(4):295–306. https://doi.org/10.4161/viru.24041.
4. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the global burden of disease study 2016. Lancet (London, England). 2017;390(10100):1151–210.
5. Arregui S, Aleta A, Sanz J, Moreno Y. Projecting social contact matrices to different demographic structures. PLoS Comput Biol. 2018;14(12):e1006638. https://doi.org/10.1371/journal.pcbi.1006638.
6. Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. PLoS One. 2014;9(8):e104786. https://doi.org/10.1371/journal.pone.0104786.
7. Kiti MC, Tizzoni M, Kinyanjui TM, Koech DC, Munywoki PK, Meriac M, et al. Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors. EPJ Data Sci. 2016;5(1):21. https://doi.org/10.1140/epjds/s13688-016-0084-2.
8. le Polain de Waroux O, Cohuet S, Ndazima D, Kucharski AJ, Juan-Giner A, Flasche S, et al. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. BMC Infect Dis. 2018;18(1):172. https://doi.org/10.1186/s12879-018-3073-1.
9. le Polain de Waroux O, Flasche S, Kucharski AJ, Langendorf C, Ndazima D, Mwanga-Amumpaire J, et al. Identifying human encounters that shape the transmission of Streptococcus pneumoniae and other acute respiratory infections. Epidemics. 2018;25:72–9. https://doi.org/10.1016/j.epidem.2018.05.008.
10. Potter GE, Wong J, Sugimoto J, Diallo A, Victor JC, Neuzil K, et al. Networks of face-to-face social contacts in Niakhar, Senegal. PLoS One. 2019;14(8):e0220443. https://doi.org/10.1371/journal.pone.0220443.
11. Melegaro A, Del Fava E, Poletti P, Merler S, Nyamukapa C, Williams J, et al. Social contact structures and time use patterns in the Manicaland Province of Zimbabwe. PLoS One. 2017;12(1):e0170459. https://doi.org/10.1371/journal.pone.0170459.
12. Dodd PJ, Looker C, Plumb ID, Bond V, Schaap A, Shanaube K, et al. Age- and sex-specific social contact patterns and incidence of mycobacterium tuberculosis infection. Am J Epidemiol. 2016;183(2):156–66. https://doi.org/10.1093/aje/kwv160.
13. Johnstone-Robertson SP, Mark D, Morrow C, Middelkoop K, Chiswell M, Aquino LDH, et al. Social mixing patterns within a south African township community: implications for respiratory disease transmission and control. Am J Epidemiol. 2011;174(11):1246–55. https://doi.org/10.1093/aje/kwr251.
14. Cohen C. Protocol: a prospective household observational cohort study of influenza, respiratory syncytial virus and other respiratory pathogens community burden and transmission dynamics in South Africa (the PHIRST study). 2018. https://www.nicd.ac.za/wp-content/uploads/2020/08/Community-Burden-RSV-and-Influenza-protocol-version-ammend-13-July-201._.pdf.
15. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208. https://doi.org/10.1016/j.jbi.2019.103208.
16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81. https://doi.org/10.1016/j.jbi.2008.08.010.
17. Parente PMDC, Santos Silva JMC. Quantile regression with clustered data. J Econ Method. 2016;5(1):1–15. https://doi.org/10.1515/jem-2014-0011.
18. Lê Cook B, Manning WG. Thinking beyond the mean: a practical guide for using quantile regression methods for health services research. Shanghai Arch Psychiatry. 2013;25(1):55–9. https://doi.org/10.3969/j.issn.1002-0829.2013.01.011.
19. Danon L, House TA, Read JM, Keeling MJ. Social encounter networks: collective properties and disease transmission. J R Soc Interface. 2012;9(76):2826–33. https://doi.org/10.1098/rsif.2012.0357.
20. Hoang TV, Coletti P, Melegaro A, Wallinga J, Grijalva C, Edmunds J, et al. A systematic review of social contact surveys to inform transmission models of close contact infections. Infect Dis. 2019;30(5):723–36.
21. Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and urban areas of southern China. Proc Biol Sci. 2014;281(1785):20140268.
22. Danon L, Read JM, House TA, Vernon MC, Keeling MJ. Social encounter networks: characterizing Great Britain. Proc R Soc Lond B Biol Sci. 2013;280(1765):20131037. https://doi.org/10.1098/rspb.2013.1037.
23. Chen SC, You ZS. Social contact patterns of school-age children in Taiwan: comparison of the term time and holiday periods. Epidemiol Infect. 2014;143(6):1139–47. https://doi.org/10.1017/S0950268814001915.
24. Eames KT, Tilston NL, Brooks-Pollock E, Edmunds WJ. Measured dynamic social contact patterns explain the spread of H1N1v influenza. PLoS Comput Biol. 2012;8(3):e1002425. https://doi.org/10.1371/journal.pcbi.1002425.
25. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. Am J Epidemiol. 2006;164(10):936–44. https://doi.org/10.1093/aje/kwj317.
26. Litvinova M, Liu QH, Kulikov ES, Ajelli M. Reactive school closure weakens the network of social interactions and reduces the spread of influenza. Proc Natl Acad Sci U S A. 2019;116(27):13174–81. https://doi.org/10.1073/pnas.1821298116.
27. Jackson C, Vynnycky E, Mangtani P. The relationship between school holidays and transmission of influenza in England and Wales. Am J Epidemiol. 2016;184(9):644–51. https://doi.org/10.1093/aje/kww083.
28. Cauchemez S, Ferguson N, Wachtel C, Tegnell A, Saour G, Duncan B, et al. Closure of schools during an influenza pandemic. Lancet Infect Dis. 2009;9(8):473–81. https://doi.org/10.1016/S1473-3099(09)70176-8.
29. Uscher-Pines L, Schwartz HL, Ahmed F, Zheteyeva Y, Meza E, Baker G, et al. School practices to promote social distancing in K-12 schools: review of influenza pandemic policies and practices. BMC Public Health. 2018;18(1):406. https://doi.org/10.1186/s12889-018-5302-3.
30. Auger KA, Shah SS, Richardson T, Hartley D, Hall M, Warniment A, et al. Association between statewide school closure and COVID-19 incidence and mortality in the US. JAMA. 2020;324(9):859–70. https://doi.org/10.1001/jama.2020.14348.
31. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. Lancet Child Adolesc Health. 2020;4(5):397–404. https://doi.org/10.1016/S2352-4642(20)30095-X.
32. Lewis D. Why schools probably aren't COVID hotspots. Nature. 2020;587(7832):17. https://doi.org/10.1038/d41586-020-02973-3.
33. Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DAT. Close encounters of the infectious kind: methods to measure social mixing behaviour. Epidemiol Infect. 2012;140(12):2117–30. https://doi.org/10.1017/S0950268812000842.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Additional file 1. Study population for contact survey at rural and urban site, South Africa, 2018

A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018

Jackie Kleynhans, Stefano Tempia, Meredith L. McMorrow, Anne von Gottberg, Neil A. Martinson, Kathleen Kahn, Jocelyn Moyes, Thulisa Mkhencele, Limakatso Lebina, F. Xavier Gómez-Olivé, Floidy Wafawanaka, Azwifarwi Mathunjwa, Cheryl Cohen for the PHIRST group

Table. Description of study population and contact rates at rural and urban site, South Africa, 2018.

95%CI – 95% confidence interval

NEO – Not enough observations, contact rate only calculated for strata with 4 or more participants.

NA – Not applicable, the number of participants who reported smoking at the rural site were only 3, and therefore contact rate could not be calculated. We therefore also did not calculate the overall contact rate.

Other adult – Includes relationships of participants 18 years and older to the head of household (n): brother (7), sister (7), wife (19), husband (4), cousin (1), niece (4), nephew (2), other relative (3), not related (4).

Other child – Includes relationships of participants younger than 18 years to the head of household (n): brother (10), sister (11), wife (1), cousin (5), niece (6), nephew (9), other relative (10), not related (2).

	Overall		Rural		Urban	
	Participants (%)	**Contact rate (95%CI)**	**Participants (%)**	**Contact rate (95%CI)**	**Participants (%)**	**Contact rate (95%CI)**
Overall	535 (100)	14 (13-15)	269 (100)	21 (14-28)	266 (100)	12 (11-13)
Age						
< 7 years	134 (25)	15 (12-17)	86 (32)	19 (11-27)	48 (18)	9 (7-11)
7-13 years	122 (23)	21 (15-27)	60 (22)	32 (19-45)	62 (23)	17 (12-21)
14-18 years	56 (10)	22 (8-35)	32 (12)	49 (27-70)	24 (9)	15 (5-24)
19-64 years	208 (39)	12 (11-13)	85 (32)	16 (9-23)	123 (46)	11 (10-12)
≥65 years	15 (3)	8 (4-12)	6 (2)	7 (3-11)	9 (3)	9 (0-18)
Household role						
Head of home	112 (21)	13 (10-15)	56 (21)	13 (10-15)	56 (21)	10 (8-12)
Parent	10 (2)	9 (6-11)	1 (0)	NEO	5 (2)	9 (6-12)
Grandchild (18 years or older)	12 (2)	12 (1-22)	3 (1)	12 (1-22)	9 (3)	12 (2-22)
Child (18 years or older)	40 (7)	12 (8-16)	18 (7)	12 (8-16)	22 (8)	12 (9-15)
Other adult	51 (10)	11 (10-12)	14 (5)	11 (10-12)	37 (14)	11 (9-13)
Grandchild (younger than 18 years)	91 (17)	17 (13-21)	48 (18)	17 (13-21)	43 (16)	12 (8-16)
Child (younger than 18 years)	165 (31)	18 (14-22)	94 (35)	18 (14-22)	71 (27)	15 (12-18)
Other child (younger than 18 years)	54 (10)	16 (2-30)	35 (13)	16 (2-30)	19 (7)	11 (8-14)
Sex						
Male	198 (37)	15 (13-17)	95 (35)	26 (14-38)	103 (39)	13 (12-14)
Female	337 (63)	14 (13-15)	174 (65)	21 (15-27)	163 (61)	11 (10-12)
Type of day investigated						
Weekday	364 (68)	13 (11-15)	170 (63)	16 (13-19)	194 (73)	10 (8-12)
Weekend day	171 (32)	16 (13-18)	99 (37)	33 (23-42)	72 (27)	13 (12-14)
Day of week investigated						
Monday	81 (15)	14 (12-16)	24 (9)	34 (22-46)	57 (21)	12 (10-14)
Tuesday	76 (14)	14 (12-16)	32 (12)	15 (5-24)	44 (17)	13 (10-16)
Wednesday	47 (9)	29 (9-49)	29 (11)	51 (29-73)	18 (7)	19 (9-29)
Thursday	103 (19)	21 (11-31)	58 (22)	42 (32-51)	45 (17)	13 (9-17)
Friday	57 (11)	13 (7-19)	27 (10)	20 (5-35)	30 (11)	12 (7-16)
Saturday	94 (18)	13 (10-16)	56 (21)	16 (14-18)	38 (14)	11 (7-15)
Sunday	77 (14)	12 (9-15)	43 (16)	16 (7-25)	34 (13)	9 (7-11)
Education level						
<18 years	310 (58)	17 (14-20)	177 (66)	28 (18-38)	133 (50)	13 (11-15)

None	18 (3)	12 (6-18)	14 (5)	12 (-1-25)	4 (2)	NEO
Primary	45 (8)	13 (6-20)	20 (7)	26 (14-38)	25 (9)	9 (5-13)
Some secondary	95 (18)	11 (10-12)	22 (8)	12 (9-14)	73 (27)	11 (9-13)
Secondary completed	64 (12)	13 (8-18)	36 (13)	22 (11-32)	28 (11)	12 (10-14)
Post-secondary	3 (1)	NEO	0 (0)	NEO	3 (1)	NEO
Employment						
Employed	98 (18)	17 (14-20)	32 (12)	28 (18-38)	66 (25)	13 (11-15)
Unemployed	127 (24)	13 (11-15)	60 (22)	18 (7-29)	67 (25)	12 (11-13)
<18 years	310 (58)	11 (9-13)	177 (66)	13 (6-20)	133 (50)	9 (7-11)
Alcohol use						
No	148 (28)	14 (11-17)	92 (34)	20 (11-29)	56 (21)	13 (9-16)
Yes	116 (22)	11 (10-12)	19 (7)	11 (-1-23)	97 (36)	11 (10-12)
<15 years	271 (51)	17 (14-20)	158 (59)	27 (17-37)	113 (42)	13 (11-15)
Smoke cigarettes						
No	221 (41)	NA	108 (40)	NEO	113 (42)	11 (9-13)
Yes	43 (8)	NA	3 (1)	NEO	40 (15)	12 (9-14)
<15 years	271 (51)	NA	158 (59)	NEO	113 (42)	13 (11-15)

A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018

Jackie Kleynhans, Stefano Tempia, Meredith L. McMorrow, Anne von Gottberg, Neil A. Martinson, Kathleen Kahn, Jocelyn Moyes, Thulisa Mkhencele, Limakatso Lebina, F. Xavier Gómez-Olivé, Floidy Wafawanaka, Azwifarwi Mathunjwa, Cheryl Cohen for the PHIRST group

Table. Cumulative person hours in contact reported by age, location and site, South Africa, 2018.

*Total number of contacts per group / number of participants in group

	Cumulative person hours in contact / average number of contacts* (hours per contact)									
	Rural					Urban				
	<7 years	**7 - 13 years**	**14 - 18 years**	**19 - 64 years**	**≥65 years**	**<7 years**	**7 - 13 years**	**14 - 18 years**	**19 - 64 years**	**≥65 years**
Home	17.4 / 6.0 (2.9)	19.0 / 6.8 (2.8)	18.1 / 7.2 (2.5)	17.0 / 6.6 (2.6)	13.4 / 5.3 (2.5)	20.0 / 6.0 (3.3)	18.5 / 6.1 (3.0)	14.8 / 5.8 (2.5)	17.8 / 5.8 (3.0)	19.8 / 5.8 (3.4)
School	47.1 / 18.1 (2.6)	50.0 / 28.1 (1.8)	86.2 / 41.7 (2.1)	1.2 / 2.2 (0.6)		12.4 / 7.7 (1.6)	47.0 / 24.7 (1.9)	46.2 / 30.5 (1.5)	3.0 / 1.6 (1.9)	
Work				8.7 / 3.2 (2.7)				0.3 / 18.5 (0.0)	5.7 / 0.0 (698.7)	
Transport	0.3 / 0.7 (0.5)	0.3 / 0.4 (0.9)	1.5 / 1.4 (1.1)	1.9 / 4.3 (0.4)		1.4 / 1.0 (1.5)	0.7 / 1.9 (0.4)	0.1 / 0.4 (0.2)	1.1 / 1.7 (0.6)	7.4 / 5.7 (1.3)
Other	7.7 / 6.0 (1.3)	18.7 / 11.6 (1.6)	10.5 / 9.8 (1.1)	17.3 / 17.6 (1.0)	3.4 / 2.3 (1.4)	4.0 / 2.0 (2.0)	9.9 / 5.9 (1.7)	15.1 / 6.8 (2.2)	11.5 / 7.9 (1.5)	1.6 / 11.7 (0.1)
All	72.6 / 30.8 (2.4)	88.1 / 46.9 (1.9)	116.3 / 60.2 (1.9)	46.1 / 33.8 (1.4)	16.8 / 7.7 (2.2)	37.7 / 16.6 (2.3)	75.9 / 38.6 (2.0)	76.4 / 62.0 (1.2)	38.2 / 17.0 (2.2)	29.2 / 23.1 (1.3)

A cross-sectional study measuring contact patterns using diaries in an urban and a rural community in South Africa, 2018

Jackie Kleynhans, Stefano Tempia, Meredith L. McMorrow, Anne von Gottberg, Neil A. Martinson, Kathleen Kahn, Jocelyn Moyes, Thulisa Mkhencele, Limakatso Lebina, F. Xavier Gómez-Olivé, Floidy Wafawanaka, Azwifarwi Mathunjwa, Cheryl Cohen for the PHIRST group

Appendix

Contact Diary 2
Time use survey 3

Initials of interviewer: _____

PHIRST Study: Contact Diary **Study ID: _____** **DOB: _____** **Date of investigated day: _____** **Day of week: _____**

Date of interview: _____ **Who was interviewed?** ☐ **Participant** ☐ **Caregiver** **If caregiver, provide Study ID: _____** **Page _____ of _____**

Age (or range*) in years	Gender		In enclosed space?		Did you talk?		Did you touch his/ her skin?		Member of household?		Total time spent with person during whole day					Where did you have contact? (tick all that apply)					How often do you have contact with this person in general?					If group of people of same age, sex and contact type, add number of people
	M	F	Yes	No	Yes	No	Yes	No	Yes	No	<5m	5-15m	15m–1hr	1-4hrs	>4 hrs	Home	School	Work	Trans-port	Other	Daily/al-most daily	1/2 times a week	1/2 times a month	<1 a month	Never met before	
10 - ( )	☐	☑	☐	☑	☐	☑	☑	☐	☑	☐	☐	☐	☐	☐	☑	☑	☐	☐	☐	☐	☑	☐	☐	☐	☐	
7 - (13)	☑	☐	☑	☐	☐	☑	☐	☑	☐	☑	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑	☐	☐	☐	☐	30
14 - (18)	☑	☐	☑	☐	☑	☐	☐	☑	☐	☑	☐	☑	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☑	
- ( )	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Initials of interviewer: _____

PHIRST Study: Contact Diary **Study ID: _____** **DOB: _____** **Date of investigated day: _____** **Day of week: _____**

Date of interview: _____ **Who was interviewed?** ☐ **Participant** ☐ **Caregiver** **If caregiver, provide Study ID: _____**

	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
Location of activity																								
Home	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Someone else's home	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Church	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
School	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Place of work	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Shebeen/pub/bar	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Shop/Market	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Vehicle/Taxi/Bus	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Other public area	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Personal																								
Sleep/rest	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Bathing/dressing	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Eating	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Running errands	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Household																								
Cooking	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Shopping	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Housekeeping	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Tend to children	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Work/Education																								
Work	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
School	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Looking for work	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Free time																								
Reading	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Socializing	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Sport	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Television	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Mobile phone	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Smoke break	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Tea/coffee break	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Playing	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Other activities																								
Travelling	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Practicing religion	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

Form checked by: _____ **Date Checked: _____**

SARS-CoV-2 Seroprevalence in a Rural and Urban Household Cohort during First and Second Waves of Infections, South Africa, July 2020–March 2021

Jackie Kleynhans, Stefano Tempia, Nicole Wolter, Anne von Gottberg, Jinal N. Bhiman, Amelia Buys, Jocelyn Moyes, Meredith L. McMorrow, Kathleen Kahn, F. Xavier Gómez-Olivé, Stephen Tollman, Neil A. Martinson, Floidy Wafawanaka, Limakatso Lebina, Jacques du Toit, Waasila Jassat, Mzimasi Neti, Marieke Brauer, Cheryl Cohen, for the PHIRST-C Group[1]

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections may be underestimated because of limited access to testing. We measured SARS-CoV-2 seroprevalence in South Africa every 2 months during July 2020–March 2021 in randomly selected household cohorts in 2 communities. We compared seroprevalence to reported laboratory-confirmed infections, hospitalizations, and deaths to calculate infection–case, infection–hospitalization, and infection–fatality ratios in 2 waves of infection. Post–second wave seroprevalence ranged from 18% in the rural community children <5 years of age, to 59% in urban community adults 35–59 years of age. The second wave saw a shift in age distribution of case-patients in the urban community (from persons 35–59 years of age to persons at the extremes of age), higher attack rates in the rural community, and a higher infection–fatality ratio in the urban community. Approximately 95% of SARS-CoV-2 infections were not reported to national surveillance.

The first laboratory-confirmed case of coronavirus disease (COVID-19) in South Africa was reported on March 5, 2020, and the country has since experienced 2 waves of COVID-19, the first peaking in July 2020 and the second in January 2021 (*1*). Across Africa, the second wave was more severe than the first (*2*), and specifically in South Africa, higher weekly incidence, hospitalizations, and deaths were reported for the second wave, compared with the first (*3*–*5*). The second wave in South Africa was coupled with the emergence of a new variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), B.1.351, also known as 501Y.V2 or Beta (*6*).

South Africa reported >1.6 million laboratory-confirmed cases by mid-May 2021 (*3*), but many cases go undiagnosed because of mild or absent symptoms or the lack of (or reluctance to access) care or testing. Data on the proportion of persons with serologic evidence of prior SARS-CoV-2 infection are critical to assess infection rates, calculate infection–hospitalization ratios (IHRs) and infection–fatality ratios (IFRs), compare infection prevalence between waves of infection and to guide public health responses (*7*). SARS-CoV-2 seroprevalence is higher in close contacts of case-patients and at-risk healthcare workers and lower in persons <20 years of age or ≥65 years

Author affiliations: National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa (J. Kleynhans, S. Tempia, N. Wolter, A. von Gottberg, J.N. Bhiman, A. Buys, J. Moyes, W. Jassat, M. Neti, C. Cohen); University of the Witwatersrand, Johannesburg (J. Kleynhans, S. Tempia, N. Wolter, A. von Gottberg, J.N. Bhiman, J. Moyes, C. Cohen); Centers for Disease Control and Prevention, Atlanta, Georgia, USA (S. Tempia, M.L. McMorrow); Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa (M.L. McMorrow); MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), University of the Witwatersrand, Johannesburg (K. Kahn, F.X. Gómez-Olivé, S. Tollman, F. Wafawanaka, J. du Toit); Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg (N.A. Martinson, L. Lebina); Johns Hopkins University Center for TB Research, Baltimore, Maryland, USA (N.A. Martinson); Ampath Pathology, Pretoria (M. Brauer)

DOI: https://doi.org/10.3201/eid2712.211465

[1]Additional members of the PHIRST-C group who contributed to this manuscript are listed at the end of this article.

of age, with no differences based on sex (*8*). Whether HIV infection increases the risk for SARS-CoV-2 infection is still unclear, and results from studies thus far have varied (*9*,*10*).

We describe the seroprevalence of SARS-CoV-2 in 2 household cohorts in a rural and an urban community at 5 timepoints from July 2020 to March 2021, during 2 epidemic waves. We compare disease prevalence between the first and second wave by comparing the seroprevalence by wave to reported laboratory-confirmed infections, hospitalizations, and deaths within the respective districts.

Methods

Study Population

We conducted a prospective study on a randomly selected household cohort in a rural community (Agincourt, Ehlanzeni District, Mpumalanga Province) and an urban community (Jouberton, Dr. Kenneth Kaunda District, North West Province) as part of the Prospective Household Study of SARS-CoV-2, Influenza, and Respiratory Syncytial Virus Community Burden, Transmission Dynamics, and Viral Interaction (PHIRST-C) study in South Africa. Methods for the cohort study are detailed in the Appendix (https://wwwnc.cdc.gov/EID/article/27/12/21-1465-App1.pdf). Recruitment to this study began in July 2020, and follow-up will continue through August 2021. Households that previously participated in the PHIRST study during 2016–2018 (*11*,*12*) and additional randomly selected households were eligible. Households with ≥3 household members of any age were enrolled if ≥80% of members consented.

The study was approved by the University of the Witwatersrand Human Research Ethics Committee (reference no. 150808). The US Centers for Disease Control and Prevention relied on local clearance (Institutional Review Board approval no. 6840).

Seroprevalence

We collected baseline data and blood (blood draw [BD] 1) at enrollment (July 20–September 17, 2020) and every 2 months thereafter: BD2, September 21–October 10; BD3, November 23–December 12, 2020; BD4, January 25–February 20, 2021; and BD5, March 22–April 11, 2021). We confirmed HIV status from medical records (if a person was HIV-infected) and by using a rapid test for participants with unknown or self-reported negative status. We determined previous SARS-CoV-2 infection by using the Roche Elecsys anti-SARS-CoV-2 assay (Roche Diagnostics, https://www.roche.ch/en/standorte/rotkreuz.htm) to detect antibodies against the SARS-CoV-2 nucleocapsid protein. We performed the assay on the Cobas e601 instrument (Roche Diagnostics), and we considered a cutoff index (COI) ≥1.0 as an indication of prior infection (i.e., seropositivity). We performed data analysis in Stata 14 (StataCorp, https://www.stata.com) (*13*). We adjusted seroprevalence estimates for sensitivity and specificity, as previously described (*14*), on the basis of the manufacturers' reported 99.5% sensitivity and 99.8% specificity (*15*). We obtained seroprevalence 95% credible intervals (CrIs) by using Bayesian inference with 10,000 posterior draws (*14*). We used Pearson's χ^2 test to assess the statistical significance of differences in SARS-CoV-2 seropositivity between the 2 communities and across BDs, waves of infection, and HIV status.

Calculation of Infection–Case Ratio, Infection–Hospitalization Ratio, and Infection–Fatality Ratio by Wave of Infection

To assess the prevalence of SARS-CoV-2 and compare the severity of illness between the 2 waves, we performed an ecologic study comparing estimated number of infections on the basis of seroprevalence in our cohort study to reported number of cases, hospitalizations, and in-hospital and excess deaths in the same district for each wave. We calculated the age- and sex-adjusted total number of infections, laboratory-confirmed cases, hospitalizations, deaths, infection–case ratio (ICR) (i.e., number of infections compared with laboratory-confirmed cases), IHR, and in-hospital and excess deaths IFR (Appendix) for the first (March 1–November 21, 2020) and second (November 22, 2020–March 27, 2021) wave of infection (Figure 1).

Comparison of Cases between First and Second Wave of Infection

We compared characteristics of participants who showed seroconversion during the first and second wave of infections by using unconditional logistic regression. We compared participants who showed seroconversion in wave 1 (BD3) with those who showed seroconversion in wave 2 (BD5, excluding BD3 seroconversions). For this analysis, we only included participants with a BD3 and BD5 paired serum sample. For the multivariable model, we assessed all variables that were significant at $p<0.2$ on univariate analysis and dropped nonsignificant factors ($p\geq0.05$) with manual backward elimination. We also compared the site, age, sex, and HIV status of persons with a BD

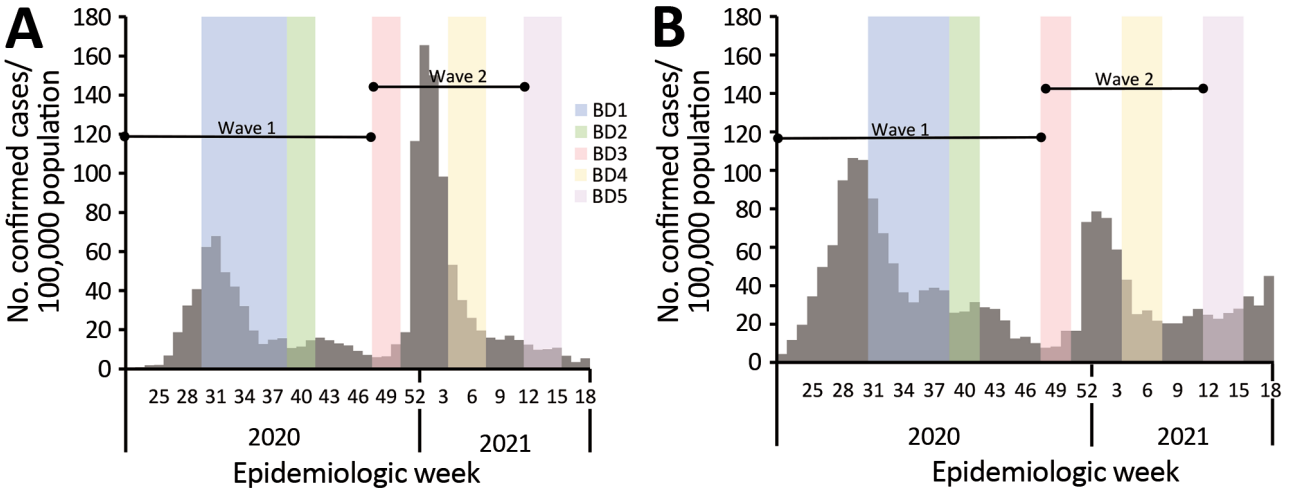


Figure 1. Timing of blood collection and weekly incidence of severe acute respiratory syndrome coronavirus 2 infection in the rural community district (A) and the urban community district (B), South Africa, March 2020–March 2021. BD, blood draw.

3+5 pair with those without a BD 3+5 pair by using logistic regression.

Persistence of SARS-CoV-2 Antibodies

For participants with 5 serum samples collected and who showed seroconversion during BD2 to BD5, we plotted COI values with the BD at which seroconversion took place as point 0. For participants who were seropositive at baseline, we plotted COI results from each BD. We calculated mean COI and the exact 95% CI at each point by using the Clopper–Pearson method. We assessed percentage of participants with COI ≥1 at each subsequent BD as number of participants with COI ≥1 divided by total number of participants who showed seroconversion during BD2 to BD5 with a serum sample at the timepoint.

Results

Study Population

In the rural community, we approached 185 households, 118 (64%) were enrolled, and 641/692 (92%) of household members consented, agreed to participate, or both. In the urban community, 352 households were approached, 114 (32%) enrolled, and 570/607 (93%) of household members consented, agreed to participate, or both. In both communities, the percentage of children, women or girls, and unemployed persons included in the cohort were higher than in district census data (Appendix Table 1). Median age was 13 (interquartile range 7–29) and 21 (interquartile range 10–43) years, and HIV prevalence was 14% (95% CI 11%–17%) in the rural community and 18% (95% CI 14%–21%) in the urban community.

Seroprevalence

Most (83% [n = 553]) participants who lived in the rural community and most (83% [n = 499]) who lived in the urban community had both BD3 and BD5 blood collected (Appendix). Seroprevalence, adjusted for assay sensitivity and specificity, in the rural community was lower at BD1 than in the urban community (1% [95% CrI 0%–2%] vs. 15% [95% CrI 12%–18%]; p<0.001), increasing after the first wave of infections (at BD3) to 7% (95% CrI 5%–9%) in the rural community and 27% (95% CrI 23%–31%) in the urban community (p<0.001) (Figure 2; Appendix). After the second wave (BD5), seroprevalence increased to 26% (95% CrI 22%–29%; p<0.001) in the rural community and to 41% (95% CrI 37%–45%; p<0.001) in the urban community (Appendix)..

At BD5, seroprevalence was highest in the 19–34 years age group (37% [95% CrI 28%–47%]) in the rural community and the 35–59 years age group (59% [95% CrI 49%–68%]) in the urban community (Figure 2; Appendix Table 3). The seroprevalence was lowest in children <5 years of age, 18% (95% CrI 10%–26%) in the rural community and 28% (95% CrI 17%–41%) in the urban community.

At BD5, SARS-CoV-2 seroprevalence was similar between HIV-infected and HIV-uninfected participants (Appendix Table 4). Persons who were HIV-positive were not more likely to be seropositive (adjusted odds ratio 1.0 [95% CI 0.7–1.5]).

Infection–Case Ratio, Infection–Hospitalization Ratio, and Infection–Fatality Ratio by District and Wave of Infection

During the first wave of infections (BD3) the age- and sex-adjusted seroprevalence at the rural site was

11.75% (95% CrI 3.42%–24.60%), resulting in an ICR of only 4.74% (95% CI 2.36%–15.62%). We observed a 0.64% (95% CI 0.34%–1.96%) IHR and an in-hospital IFR of 0.12% (95% CI 0.07%–0.31%) and an excess deaths IFR of 0.43% (95% CI 0.21%–1.47%) (Figure 3, 4).

The seroprevalence in the rural community was 22.43% (95% CrI 10.46%–37.67%) for the second wave. The ICR was 3.71% (95% CI 2.28%–7.68%), IHR was 0.61% (95% CI 0.40%–1.22%), in-hospital IFR was 0.18% (95% CI 0.12%–0.34%), and excess deaths IFR was 0.65% (95% CI 0.39%–1.39%) (Figure 3, 4).

In the urban community, the seroprevalence at BD3 was 29.58% (95% CrI 18.04%–43.20%). We found a 3.54% (95% CI 2.53%–5.55%) ICR and 1.93% (95% CI 1.41%–2.98%) IHR. The in-hospital IFR was 0.16% (95% CI 0.13%–0.23%) and excess deaths IFR was 0.12% (95% CI 0.09%–0.20%) (Figure 3, 4). During the second wave, the seroprevalence in the urban community was 15.19% (95% CrI 6.49%–26.96%), resulting in an ICR estimate of 3.67% (95% CI 2.21%–8.07%), an IHR of 2.29% (95% CI 1.39%–4.96%), an in-hospital IFR of 0.36% (95% CI 0.24%–0.72%), and an excess deaths IHR of 0.50% (95% CI 0.29%–1.17%) (Figure 3, 4). These estimates standardized to World Health Organization world population estimates are shown in Appendix Figure 2.

Comparison of Case-Patients between First and Second Wave of Infection

Compared with the urban community, persons in the rural community who showed seroconversion were 4.7 (95% CI 2.9–7.6) times more likely to show seroconversion during the second wave. Compared with persons 35–59 years of age, persons 5–12 years of age were 2.1 (95% CI 1.1–4.2) times more likely to show seroconversion in the second wave and persons ≥60 years of age were 2.8 (95% CI 1.1–7.0) times more likely to show seroconversion in the second wave (Table). When we stratified the analysis by site, this association was only detected in the urban community (Appendix Table 6). Persons who did not have a BD 3+5 pair were more likely to be <5 or 19–34 years of age (Appendix Table 7).

Persistence of SARS-CoV-2 Antibodies

Of the 72 participants who were seropositive at BD1 and with BDs 1–5 samples collected, 99% (71/72) still had a COI ≥1 by BD5. The mean COI at baseline for seropositive participants was 64, which increased to 125 at BD2 and dropped to 59 at BD5 (Figure 5, panel A). The participant who no longer had detectable SARS-CoV-2 antibodies at BD5 had a starting COI of 9. Of the 210 participants with BD 1–5 samples, 99% (167/169), 99% (70/71%), and 93% (41/44) still had a COI ≥1 in the first, second, and third BD after initial seroconversion, respectively (Figure 5, panel B). The participants who seroreverted had starting COIs ranging from 2 to 6, and none showed seroconversion again after reversion during the study period. The mean COI at the point of seroconversion was 48, which increased to 86 at the first BD after seroconversion and reduced to 61 at the third BD after seroconversion.

Discussion

We assessed SARS-CoV-2 seroprevalence in 1,211 persons living in 2 diverse communities in South Africa and show that laboratory-confirmed cases reported from study districts greatly underestimate the actual prevalence of SARS-CoV-2 infections. At baseline, seroprevalence was 1% and 15%, increasing to 7% and 27%, respectively, after the first wave, by March 2021. After the second epidemic wave, seroprevalence was

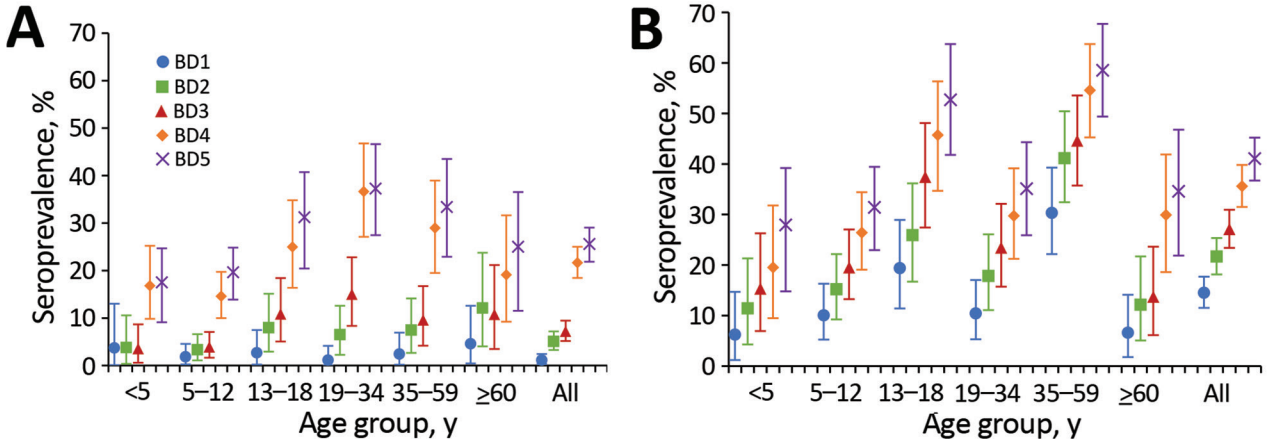


Figure 2. Seroprevalence of severe acute respiratory syndrome coronavirus 2 at each blood collection, by age group, in a rural community (A) and an urban community (B), South Africa, March 2020–March 2021.

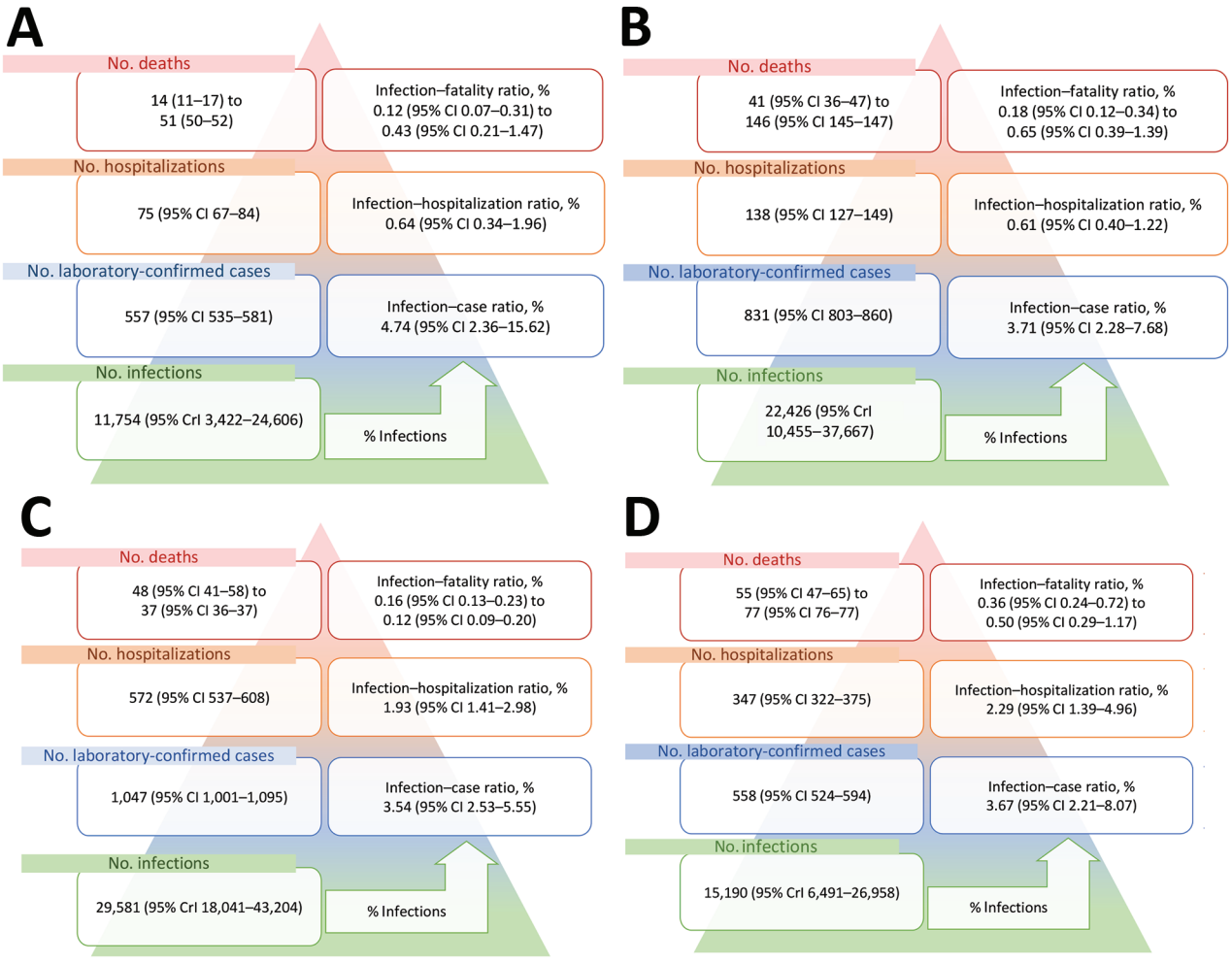


Figure 3. Age- and sex-standardized number of severe acute respiratory syndrome coronavirus 2 infections, laboratory-confirmed diagnoses, hospitalizations, and deaths per 100,000 population in a rural community during infection wave 1 (A) and wave 2 (B) and an urban community during infection wave 1 (C) and wave 2 (D), South Africa, March 2020–March 2021. CrI, credible interval.

26% in the rural community and 41% in the urban community. The highest seroprevalence was 59% in adults 35–59 years of age in the urban community, and the lowest was 18% in rural community children <5 years of age. During the second wave, compared with the first wave, the rural site was more affected, and infections in the second wave more likely affected children 5–12 and adults ≥60 years of age in the urban community. In the urban community, IFR was higher in the second wave (0.36%–0.50%) compared with the first (0.12%–0.16%), and numbers of infections were lower, suggesting possible increased severity associated with the emergence of novel variant B.1.351. Most persons who showed seroconversion maintained detectable SARS-CoV-2 antibodies in subsequent serum samples.

Low seropositivity was observed at the rural site at baseline, and seroprevalence remained low after the first wave of infections, reaching 7%, which was considerably lower than the 27% at the urban site at the same time. This observation could be related to the relatively isolated location and lower population density in the rural community compared with more densely populated urban community. Seroprevalence in the rural site increased to 26% after the second wave of infections within the district. This increase could have been attributable to possible increased transmissibility of the B.1.135 lineage that was circulating in the second wave (*16*), as well as additional transmission networks in the community during the December holiday period, when largescale urban-to-rural migration takes place as persons return home for year-end holidays. The urban site had fewer seroconversions in the second wave compared with the first, which may be attributable to existing immunity among persons in the community after the first wave.

As seen in previous studies (*8*), adults had the highest seroprevalence levels, although a relatively high seroprevalence of 18% and 28% persisted in children <5 years of age at the rural and urban community, respectively.

When comparing the characteristics of persons infected in the second wave to those infected in the first, persons infected in the second wave were more likely to be from the rural site and to be <13 or >60 years of age, compared with persons 35–59 years of age. The shift in age groups affected was only detected in the urban community, possibly because of the large number of adults infected during wave 1, whereas the number of infections in wave 1 in the rural community was lower.

A study conducted among blood donors in South Africa during the second wave found a seroprevalence of 32%–63% in 5 provinces of South Africa that have both rural and urban communities (W. Sykes et al., unpub. data, https://doi.org/10.21203/rs.3.rs-233375/v1). In our study, we observed a seroprevalence in adults ranging from 25% to 37% in rural households, and from 35% to 59% in urban households, suggesting that seroprevalence is heterogeneous between communities. In Kenya, the seroprevalence in blood donors during the country's first wave of infections was 4% and was also higher in urban communities (*17*). In a population-level household serosurvey conducted in Zambia during their first wave of infections, 11% of persons had evidence of SARS-CoV-2 infection (*18*).

Based on our estimates, only 4%–6% of cases were laboratory-confirmed, suggesting that substantially higher prevalence of infection was ascertained through serologic testing and that the differences may have been greater in the urban community than the rural community; however, more extensive studies are needed to assess whether this observation is consistent in other areas. Compared with the urban community, the rural community had less than half the rate of hospitalization (0.6% vs. 2.0%). These observations may be attributable to differences in referral and testing policies, health-seeking behavior, and access to care, as well as differences in circulating lineages within these districts.

A study comparing the severity of the first and second waves of infections in South Africa in hospitalized patients found a higher mortality rate in the second wave, compared with the first (*19*). At the urban site, the IFR was higher in the second wave (0.36%–0.50%) compared with the first (0.12%–0.16%), although no differences were observed in IHR between the 2 waves. The lower overall number of infections in the second wave in this site means that our finding of increased mortality is unlikely to be related to pressure on health services. The increased severity of the second wave may be related to increased severity of the B.1.135 variant, but further studies are needed to confirm this relation. The excess death IFR during the first wave in the urban site was smaller than the in-hospital IFR. This difference may be attributable to uncertainty on the process for excess death estimation, or that the 85% contribution of COVID-19 to excess deaths was an underestimation within the province. However, the in-hospital IFR followed the same trend of increase between wave 1 and 2 (0.16% to 0.36%). Although no significant increase was observed between the IFR in the rural community between wave

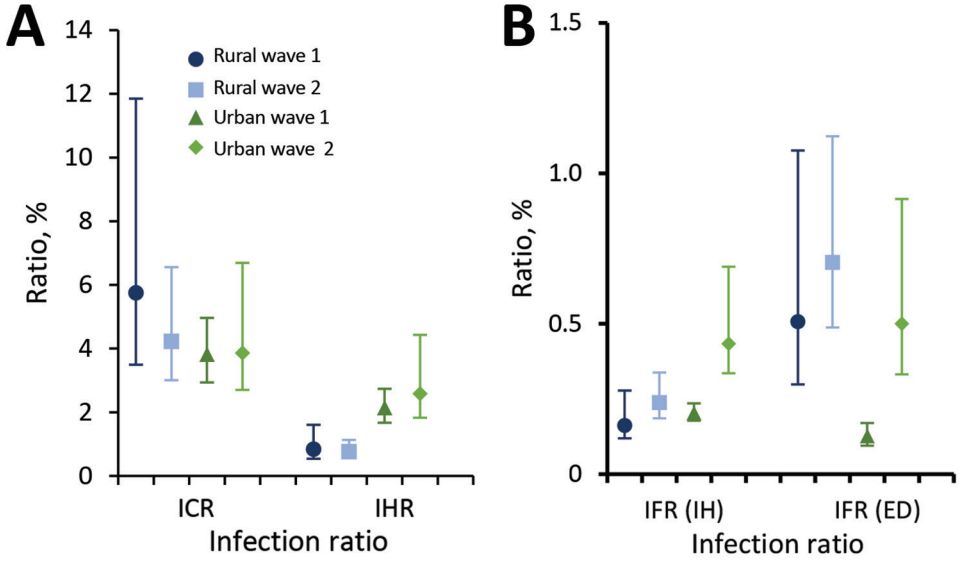


Figure 4. Severe acute respiratory syndrome coronavirus 2 infection–case and infection–hospitalization ratios (A) and in-hospital and excess deaths infection-fatality ratios (B) in a rural and urban community during the first and second wave of infections, South Africa, March 2020–March 2021. Vertical lines represent 95% CIs. Wave 1: March 1–November 21, 2020. Wave 2: November 2020 22–March 27, 2021. ED, excess deaths; ICR, infection–case ratio; IFR, infection–fatality ratio; IH, in-hospital; IHR, infection–hospitalization ratio.

Table. Comparison of participants with detectable SARS-CoV-2 antibodies after the first wave (blood draw 3) and second wave (blood draw 5), South Africa, July 2020–April 2021*

Characteristic	Infected in wave 1, no. (%)	Infected in wave 2, no. (%)	Univariate OR (95% CI)	Multivariable aOR (95% CI)
Site				
Rural	40/140 (29)	100/140 (71)	**4.9 (3.1–7.8)**	**4.7 (2.9–7.6)**
Urban	139/210 (66)	71/210 (34)	Referent	Referent
Sex				
M	65/123 (53)	58/123 (47)	Referent	
F	114/227 (50)	113/227 (50)	1.1 (0.7–1.7)	
Age group, y				
<5	9/25 (36)	16/25 (64)	**3.2 (1.3–8.0)**	2.7 (1.0–7.2)
5–12	30/74 (41)	44/74 (59)	**2.6 (1.4–4.9)**	**2.1 (1.1–4.2)**
13–18	36/64 (56)	28/64 (44)	1.4 (0.7–2.7)	1.3 (0.6–2.6)
19–34	34/67 (51)	33/67 (49)	1.7 (0.9–3.3)	1.3 (0.7–2.6)
35–59	59/92 (64)	33/92 (36)	Referent	Referent
≥60	11/28 (39)	17/28 (61)	**2.8 (1.2–6.6)**	**2.8 (1.1–7.0)**
HIV status				
Negative	139/271 (51)	132/271 (49)	1.0 (0.6–1.7)	
Positive	35/68 (51)	33/68 (49)	Referent	
CD4 count, cells/μL				
≥200	28/54 (52)	26/54 (48)	1.9 (0.2–21.7)	
<200	2/3 (67)	1/3 (33)	Referent	
Viral load, copies/mL				
<1,000	28/51 (55)	23/51 (45)	Referent	
≥1,000	2/7 (29)	5/7 (71)	3.0 (0.5–17.2)	
Other underlying illness‡				
No	161/316 (51)	155/316 (49)	1.1 (0.5–2.2)	
Yes	18/34 (53)	16/34 (47)	Referent	
Body mass index category				
Underweight	10/22 (45)	12/22 (55)	1.5 (0.6–3.9)	
Normal weight	65/141 (46)	76/141 (54)	1.5 (0.9–2.5)	
Overweight	46/84 (55)	38/84 (45)	1.1 (0.6–1.9)	
Obese	58/103 (56)	45/103 (44)	Referent	
Currently smoking‡				
No	109/190 (57)	81/190 (43)	Referent	
Yes	20/36 (56)	16/36 (44)	1.1 (0.5–2.2)	
Alcohol use‡				
No	88/167 (53)	79/167 (47)	2.0 (1.1–3.8)	
Yes	41/59 (69)	18/59 (31)	Referent	
Employment status§				
Unemployed	70/128 (55)	58/128 (45)	1.9 (0.5–6.4)	
Student	9/13 (69)	4/13 (31)	Referent	
Employed	25/41 (61)	16/41 (39)	1.4 (0.4–5.5)	

*Includes all participants with blood draw 3 and 5 serum pairs and who showed seroconversion at either draw. Bold indicates a statistically significant difference. aOR, adjusted odds ratio; OR, odds ratio; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
†Self-reported history of asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, organ transplant, immunosuppressive therapy, organ transplantation, cancer, liver disease, renal disease, or diabetes.
‡Among persons ≥15 years of age.
§Among persons ≥18 years of age.

1 and 2, the excess (maximum) IFR estimate (0.43%) was already high in wave 1, similar to the IFR for wave 2 in the urban community (0.50%). Considering the lower IHR in the rural community for both waves compared with the urban community, this observation may point toward lack of access to care or delays in seeking care. In-hospital SARS-CoV-2 mortality rates have previously also been shown to be higher in Mpumalanga Province where the rural community is located (*19*).

Our first wave in-hospital IFR estimates (0.12% rural, 0.16% urban) were similar to the age-adjusted 0.15% reported from India for the first wave of laboratory-confirmed deaths from SARS-CoV-2 infection (*20*). In addition, our first wave excess death IFR was higher in the rural (0.431%) and lower in the urban (0.12%) community compared with the age-adjusted 0.28% IFR excess deaths reported from Brazil during their first wave of infections (*21*).

Although previous studies have shown that antibodies against the SARS-CoV-2 nucleocapsid wanes more quickly than those against the spike protein (*22*,*23*), 93% of persons who showed seroconversion at BD2 still had detectable nucleocapsid antibodies 6 months later. Direct antigen-sandwich format assays, such as the Roche anti-N assay used in our study,

have been found to reliably detect antibodies in longitudinal samples (*24*). Of the few that seroreverted in this timeframe, the starting COI was low.

We did not observe a difference in SARS-CoV-2 seroprevalence in HIV-infected and HIV-uninfected persons at either site. Although our sample size was too small to detect small differences in a stratified analysis, we also did not observe a signal when using logistic regression. Because HIV causes immune suppression, a concern exists that HIV-infected persons may be more susceptible to SARS-CoV-2 infection (*9*,*10*). Although HIV infection may not increase susceptibility to infection, it has been demonstrated to be a risk factor for having onset of severe COVID-19 and death after infection (*9*,*25*).

Our study is limited by a small sample size, reducing the power for accurate seroprevalence estimates in small age strata, and inclusion of only 2 geographic sites, and therefore may not be representative of other districts and provinces in South Africa. Because we used seroprevalence to estimate infections by wave, we could have missed reinfections in the second wave. Based on data from the same cohort, these reinfections occurred in only a small portion (3%) of the cohort (C. Cohen et al., unpub. data, https://doi.org/10.1101/2021.07.20.21260855) and would have had a negligible influence on the infection ratios. ICR, IHR, and IFR formed part of an ecologic analysis, which is inherently prone to biases. Excess deaths in the first wave may be underestimated because the reporting period only started in June. Transmission dynamics within our cohort may not be similar to the transmission dynamics within the district. Seventeen percent of persons did not have a BD 3+5 pair, and bias could have been introduced if the seroprevalence were different for those without a BD 3+5 blood pair. Ongoing follow-up of this cohort will track future infections and monitor antibody waning, and compare these data to laboratory-confirmed infections and symptoms from twice-weekly follow-up.

A strength of our study is the collection of samples from prospectively followed-up persons from randomly selected households within the study communities and inclusion of persons of all ages. As a longitudinal study, our study provides the advantage

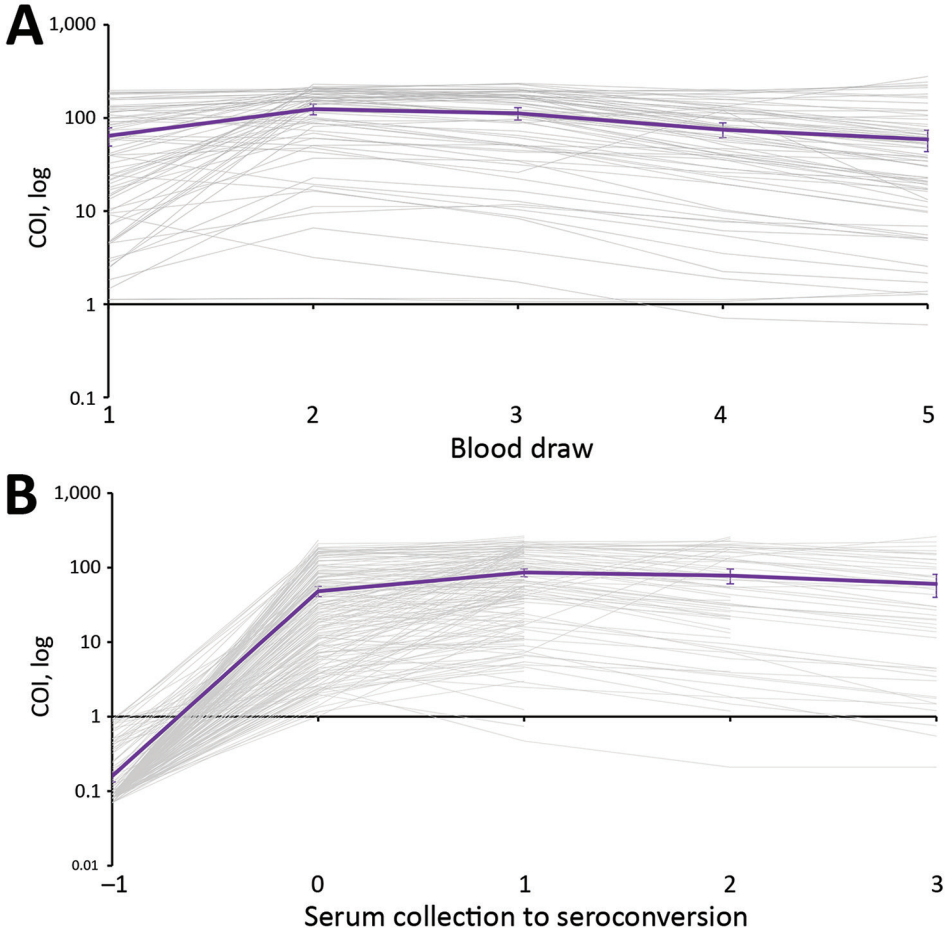


Figure 5. Cutoff index (COI) on Roche Elecsys (Roche Diagnostics, https://www.roche.ch/en/standorte/rotkreuz.htm) anti–severe acute respiratory syndrome coronavirus 2 assay for persons with blood draws 1–5 samples who were seropositive at baseline (A) or showed seroconversion during blood draws 2–5, South Africa, July 2020–April 2021. Purple line indicates mean COI with 95% CIs. COI values in panel B are aligned to first draw before seroconversion, COI, cutoff index.

of serial comparisons of antibody responses in relation to reported laboratory-confirmed SARS-CoV-2 infections within the community through 2 successive SARS-CoV-2 waves.

We estimate that ≈95% of SARS-CoV-2 infections in these 2 communities were not laboratory-confirmed and reported to the national surveillance system, which has major implications for contact tracing and isolation and other measures to contain infection. We observed heterogeneity between seroprevalence estimates based on pandemic wave, community, and age group, indicating the need for ongoing studies that include diverse settings.

Additional members of the PHIRST-C group who contributed to this manuscript were Kgaugelo Patricia Kgasago, Linda de Gouveia, Maimuna Carrim, Mignon du Plessis, Retshidisitswe Kotane, and Tumelo Moloantoa.

Acknowledgments

We thank all persons participating in the study and the field teams for their hard work and dedication to the study, especially Tumelo Moloantoa, Kgaugelo Kgasago, Nompumelelo Yende, and Sizzy Ngobeni for coordination of site activities; Charlotte Yaldwyn and Shingirai Chipungu for processing of collected bloods; the Centre for Respiratory Diseases and Meningitis Laboratory Team; Doreen Janse van Rensburg for managing logistics surrounding testing of collected bloods and result extraction; Nevashan Govender, Genevie Ntshoe, Andronica Moipone Shonhiwa, Darren Muganhiri, Itumeleng Matiea, Eva Mathatha, Fhatuwani Gavhi, Teresa Mashudu Lamola, Matimba Makhubele, Mmaborwa Matjokotja, Simbulele Mdleleni, and Masingita Makhubela from the national SARS-CoV-2 National Institute for Communicable Diseases Surveillance Team; Lincoln Darwin, Fazil Mckenna, Trevor Graham Bell, Ndivhuwo Munava, Muzammil Raza Bano, and Jimmy Khosa from National Institute for Communicable Diseases Information Technology Department; Rob Dorrington from the Centre for Actuarial Research for providing wave-specific excess death estimates; Ben Cowling, Kanta Subbarao, Juliet Pulliam, Melissa Rolfes, and Mosa Moshabela from the PHIRST-C Scientific Committee; and Andrew Whitelaw, June Fabian, Jennifer Verani, Lindiwe Qwabe, Banele Faku, and Saheen Methar from the PHIRST-C Safety Committee.

This work was supported by the National Institute for Communicable Diseases of the National Health Laboratory Service and the US Centers for Disease Control and Prevention (cooperative agreement no. 6U01IP001048-04-02 awarded to C.C.). C.C. reports receiving grant funds from US Centers for Disease Control and Prevention, Wellcome Trust, and South African Medical Research Council.

The investigators welcome enquiries about possible collaborations and requests for access to the dataset. Data will be shared after approval of a proposal and with a signed data access agreement. Investigators interested in more details about this study, or in accessing these resources, should contact the corresponding author.

About the Author

Ms. Kleynhans is an epidemiologist in the Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases. Her primary research interests include the epidemiology of respiratory diseases like influenza and COVID-19, vaccine impact studies, and modelling of infectious disease transmission dynamics.

References

1. National Institute for Communicable Diseases. COVID-19 testing summary. 2021 12 Feb 2021 [cited 2021 Feb 16]. https://www.nicd.ac.za/wp-content/uploads/2021/03/COVID-19-Testing-Summary-Week-8-2021.pdf
2. Salyer SJ, Maeda J, Sembuche S, Kebede Y, Tshangela A, Moussif M, et al. The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. Lancet. 2021;397:1265–75. https://doi.org/10.1016/S0140-6736(21)00632-2
3. National Institute for Communicable Diseases. COVID-19 weekly epidemiology brief, week 19. 2021 May 22 [cited 2021 May 22]. https://www.nicd.ac.za/wp-content/uploads/2021/05/COVID-19-Weekly-Epidemiology-Brief-week-19-2021.pdf
4. National Institute for Communicable Diseases. COVID-19 sentinel hospital surveillance update, week 18. 2021 May 14 [cited 2021 May 22]. https://www.nicd.ac.za/wp-content/uploads/2021/05/NICD-COVID-19-Weekly-Sentinel-Hospital-Surveillnace-update-Week-18-2021.pdf
5. South African Medical Research Council. Report on weekly deaths in South Africa, week 18. 2021 May 10 [cited 2021 May 22]. https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa
6. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med. 2021;27:622–5. https://doi.org/10.1038/s41591-021-01285-x
7. Murhekar MV, Clapham H. COVID-19 serosurveys for public health decision making. Lancet Glob Health. 2021;9:e559–60. https://doi.org/10.1016/S2214-109X(21)00057-7
8. Chen X, Chen Z, Azman AS, Deng X, Sun R, Zhao Z, et al. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. Lancet Glob Health. 2021;9:e598–609. https://doi.org/10.1016/S2214-109X(21)00026-7
9. Spinelli MA, Lynch KL, Yun C, Glidden DV, Peluso MJ, Henrich TJ, et al. SARS-CoV-2 seroprevalence, and IgG concentration and pseudovirus neutralising antibody titres after infection, compared by HIV status: a matched case-control observational study. Lancet HIV. 2021;8:e334–41. https://doi.org/10.1016/S2352-3018(21)00072-2

10. Ssentongo P, Heilbrunn ES, Ssentongo AE, Advani S, Chinchilli VM, Nunez JJ, et al. Epidemiology and outcomes of COVID-19 in HIV-infected individuals: a systematic review and meta-analysis. Sci Rep. 2021;11:6283. https://doi.org/10.1038/s41598-021-85359-3
11. Cohen C, McMorrow ML, Martinson NA, Kahn K, Treurnicht FK, Moyes J, et al. Cohort profile: a prospective household cohort study of influenza, respiratory syncytial virus and other respiratory pathogens community burden and transmission dynamics in South Africa, 2016–2018. Influenza Other Respir Viruses. 2021 Jul 23 [Epub ahead of print].
12. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Hellferscee O, et al.; PHIRST group. Asymptomatic transmission and high community burden of seasonal influenza in an urban and a rural community in South Africa, 2017-18 (PHIRST): a population cohort study. Lancet Glob Health. 2021;9:e863–74. https://doi.org/10.1016/S2214-109X(21)00141-8
13. South African Government. COVID-19: vaccine strategy. 2021[cited 2021 Feb 16]. https://sacoronavirus.co.za/2021/01/03/covid-19-vaccine-strategy
14. 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. eLife. 2021;10:e64206. https://doi.org/10.7554/eLife.64206
15. Roche. Elecsys anti-SARS-CoV-2. 2020 [cited 2021 Mar 21]. https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html
16. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature. 2021;592:438–43. https://doi.org/10.1038/s41586-021-03402-9
17. Uyoga S, Adetifa IMO, Karanja HK, Nyagwange J, Tuju J, Wanjiku P, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors. Science. 2021;371:79–82. https://doi.org/10.1126/science.abe1916
18. Mulenga LB, Hines JZ, Fwoloshi S, Chirwa L, Siwingwa M, Yingst S, et al. Prevalence of SARS-CoV-2 in six districts in Zambia in July, 2020: a cross-sectional cluster sample survey. Lancet Glob Health. 2021;9:e773–81. https://doi.org/10.1016/S2214-109X(21)00053-X
19. Jassat W, Mudara C, Ozougwu L, Tempia S, Blumberg L, Davies M-A, et al.; DATCOV author group. Difference in mortality among individuals admitted to hospital with COVID-19 during the first and second waves in South Africa: a cohort study. Lancet Glob Health. 2021;9:e1216–25. https://doi.org/10.1016/S2214-109X(21)00289-8
20. Murhekar MV, Bhatnagar T, Selvaraju S, Rade K, Saravanakumar V, Vivian Thangaraj JW, et al. Prevalence of SARS-CoV-2 infection in India: findings from the national serosurvey, May–June 2020. Indian J Med Res. 2020;152:48–60. https://doi.org/10.4103/ijmr.IJMR_3290_20
21. Silva AAMD, Lima-Neto LG, Azevedo CMPES, Costa LMMD, Bragança MLBM, Barros Filho AKD, et al. Population-based seroprevalence of SARS-CoV-2 and the herd immunity threshold in Maranhão. Rev Saude Publica. 2020;54:131. https://doi.org/10.11606/s1518-8787.2020054003278
22. Lumley SF, Wei J, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, et al.; Oxford University Hospitals Staff Testing Group. The duration, dynamics, and determinants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibody responses in individual healthcare workers. Clin Infect Dis. 2021;73:e699–709. https://doi.org/10.1093/cid/ciab004
23. Fenwick C, Croxatto A, Coste AT, Pojer F, André C, Pellaton C, et al. Changes in SARS-CoV-2 spike versus nucleoprotein antibody responses impact the estimates of infections in population-based seroprevalence studies. J Virol. 2021;95:e01828–20. https://doi.org/10.1128/JVI.01828-20
24. Di Germanio C, Simmons G, Kelly K, Martinelli R, Darst O, Azimpouran M, et al. SARS-CoV-2 antibody persistence in COVID-19 convalescent plasma donors: dependency on assay format and applicability to serosurveillance. Transfusion. 2021 Jun 15 [Epub ahead or print]. https://doi.org/10.1111/trf.16555
25. Western Cape Department of Health and National Institute for Communicable Diseases SA. Risk factors for coronavirus disease 2019 (COVID-19) death in a population cohort study from the Western Cape Province, South Africa. Clin Infect Dis. 2020 Aug 29 [Epub ahead of print].

Address for correspondence: Jackie Kleynhans, Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, 1 Modderfontein Rd, Sandringham, 2192, Johannesburg, South Africa; email: jackiel@nicd.ac.za

Article DOI: https://doi.org/10.3201/eid2712.211465

SARS-CoV-2 Seroprevalence in a Rural and Urban Household Cohort during First and Second Waves of Infections, South Africa, July 2020–March 2021

Appendix

Supplementary Methods

Study Population

PHIRST-C (A Prospective Household study of SARS-CoV-2, Influenza, and Respiratory Syncytial virus community burden, Transmission dynamics and viral interaction in South Africa) is a continuation from PHIRST (A Prospective Household observational cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa) (*1*). PHIRST-C is conducted in two communities with established influenza-like illness and pneumonia surveillance sites (Appendix Figure 1).

South Africa consists of nine provinces, which are divided into 52 districts, which form the second level of administration. Districts are further divided into municipalities. The rural site is located in the Bushbuckridge Municipality, Ehlanzeni District, Mpumalanga Province and forms part of a health and sociodemographic surveillance site (HDSS) at the Medical Research Council (MRC)/University of Witwatersrand Rural Public Health and Health Transitions Research Unit, Agincourt. During PHIRST, two of the 29 villages within the HDSS were selected according to convenience (proximity and burden of other studies within the site) each year between 2016–2018. A random selection of ≈50 households with a household size of 3 or more were approached for enrolment. The urban site is located in the Jouberton Township, Matlosana Municipality, Dr Kenneth Kaunda District, North West Province. A list of 450 global positioning system (GPS) coordinates were generated using Google Earth. Study staff approached the nearest house within 30 m of the coordinate point for enrolment. Approximately 50 households were enrolled each year during 2016–2018. At both sites, households were eligible if they consisted of 3 or more household members

(sharing at least four meals a week). All household members were approached for inclusion in the study, and households were eligible for inclusion if >80% of members consented to be enrolled.

For PHIRST-C, all households who participated in PHIRST from the urban site, and households from the 2017 and 2018 cohort in the rural site were approached for enrolment. To supplement the sample size, additional households at each site were approached using the same methods as for PHIRST. Villages in the rural site were restricted to those four used in 2017 and 2018 of PHIRST. Informed consent was obtained from participating adults or a parent/guardian for children <18 years of age. In addition to parent/guardian consent, assent was also obtained from children aged 7–17 years. When performing rapid HIV tests, both pre- and post-test counselling was provided, including referral to care for those who tested positive.

The Bushbuckridge Municipality within the Ehlanzeni District is considered as predominantly rural, with main industries being agriculture and tourism (*2*). The Ehlanzeni District has a population of 1,828,738, with sixty percent of the district's population being >18 years of age (Appendix Table 1). The City of Matlosana Municipality is considered to be 88.2% urban, and is located in the Dr Kenneth Kaunda District, which has a population of 797,715 (*3*). Sixty-five percent of the district population is aged >18 years (Appendix Table 1). At both sites, the percentage of children, females, and unemployed individuals were higher in the cohort than in the district (Appendix Table 1).

Calculation of Infection-Case Ratio (ICR), Infection-Hospitalization Ratio (IHR), and Infection-Fatality Ratio (IFR) by Wave of Infection

We calculated the age-adjusted total number of infections, laboratory-confirmed cases, hospitalizations, deaths, ICR (number of infections compared to laboratory-confirmed cases), IHR and in-hospital and excess death IFR for each wave of infection as described in below equations. We defined the first wave as March 1, 2020 (week 11) to November 21, 2020 (week 47), coinciding with the first case of SARS-CoV-2 reported in South Africa, and ending the week before BD3 started. We defined the second wave as November 22, 2020 (week 48) to March 27, 2021 (week 11), starting directly after the defined wave 1 period, and ending the week before BD5 started. Due to differences between the sex ratios of our cohort and the district population, the infection estimates were also adjusted for sex. Data sources used in these calculations are described in the next section. Age standardized estimates for the selected endpoints for each wave were obtained as follows:

$$Infections = \frac{\sum_i (s_i \times SAp_i)}{\sum_i (SAp_i)} \times 100{,}000$$

Where s_i is the seroprevalence in the cohort in the respective community and wave for age and sex group i and $p_{i(SA)}$ is the South African population for age and sex group i. Calculated for wave 1 as seroprevalence at blood draw 3, and for wave 2 as seroprevalence at blood draw 5, excluding those who seroconverted at blood draw 3. Estimates only included participants with a blood draw 3 and 5 pair, and adjusted for sensitivity and specificity of test (*4*).

$$Laboratory - confirmed\ cases = \frac{\sum_i \left((c_i \div p_{i(d)}) \times p_{i(SA)}\right)}{\sum_i \left(p_{i(SA)}\right)} \times 100{,}000$$

Where c_i is the number of laboratory-confirmed cases (RT-PCR and antigen-based tests) from the respective district reported to the NMCSS (wave 1: 3 March – 21 November 2020, wave 2: 22 November 2020 – 27 March 2021) in age group i, $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age group i.

$$Hospitalisations = \frac{\sum_i \left((h_i \div p_{i(d)}) \times p_{i(SA)}\right)}{\sum_i \left(p_{i(SA)}\right)} \times 100{,}000$$

Where h_i is the number of hospitalizations from the respective district reported to COVID-19 Sentinel Hospital Surveillance (DATCOV, wave 1: 5 March – 21 November 2020, wave 2: 22 November 2020 – 27 March 2021) (*5*) in age group i, $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age group i.

$$In - hospital\ deaths = \frac{\sum_i \left((d_i \div p_{i(d)}) \times p_{i(SA)}\right)}{\sum_i \left(p_{i(SA)}\right)} \times 100{,}000$$

Where d_i is the number of in-hospital deaths from the respective districts reported to DATCOV (wave 1: 5 March–21 November 2020, wave 2: 22 November 2020–27 March 2021) (*5*) in age group i, $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age group i.

$$Excess\ deaths = \frac{(ExD\ x\ 0.85)\ \times\ p_{(SA)}}{p_{(SA)}} \times 100{,}000$$

Where *ExD* is the rate of provincial excess deaths adjusted to the South African population reported by South African Medical Research Council (rural wave 1: 28 June–21 November 2020, urban wave 1: 21 June–21 November 2020, both communities wave 2: 22

November 2021–26 March 2021) (*6*), and $p_{(SA)}$ is the total South African population. According to estimates only 85% of excess deaths are attributable to COVID-19 (*6*).

For infections, hospitalizations, in-hospital (minimum) and excess (maximum) deaths, 95% CIs were calculated using the Clopper-Pearson method with a Poisson distribution.

$$Infection - case\ ratio\ (ICR) = \frac{Laboratory - confirmed\ cases}{Infections}\ x\ 100$$

$$Infection - hospitalization\ ratio\ (IHR) = \frac{Hospitalisations}{Infections}\ x\ 100$$

$$Infection - fatality\ ratio\ (IFR) = \frac{Deaths}{Infections}\ x\ 100$$

Confidence intervals for infection ratios were calculated as ratios from the 95% confidence intervals of infection, hospitalization and death rates. For international comparisons, we repeated the calculations to standardize to the WHO standard world population (*7*), using Sprague multipliers (*8*) to expand age groups to 1-year bands, and aggregate to the age groups used in this study.

Data Sources

Population Denominators (*9*)

Population numbers for each district, by age, were obtained from the StatsSA 2020 mid-year population estimates for 2020.

Notifiable Medical Conditions Surveillance System (NMCSS) (*10*)

Reverse transcription PCR (RT-PCR) testing in South Africa to detect SARS-CoV-2 RNA started on 28 January 2020. Rapid SARS-CoV-2 antigen testing was implemented in November 2020. All laboratories in the private and public sector performing SARS-CoV-2 tests automatically feed testing data from their lab information systems to the NMCSS, from where daily and weekly reporting on national, provincial and district SARS-CoV-2 infections are performed. No serology tests are used for this reporting. Results from antigen tests may be underestimated in the NMCSS. The number of cases reported to the NMCSS in the Ehlanzeni District (rural community) and Dr Kenneth Kaunda District (urban community) was used as an estimate of reported, laboratory-confirmed SARS-CoV-2 infections. Wave 1: 3 March 1 November 2020, wave 2: 22 November 2020–27 March 2021.

COVID-19 National Hospital Surveillance (DATCOV) (*5*)

DATCOV is a national hospital surveillance system to which all hospitals where COVID-19 admissions occurred in the public and private sector in South Africa report to. The case definition includes any person admitted with a positive RT-PCR test result for SARS-CoV-2. In-hospital outcome was available for all patients. The total number of hospitalizations and in-hospital deaths in each district reported to DATCOV between 3 March –21 November 2020 (wave 1) and 22 November 2020–27 March 2021 (wave 2) was used to calculate in the IHR and minimum IFR, respectively.

South African Medical Research Council (SAMRC) Report on Weekly Deaths (*6*)

The Burden of Disease Unit at the SAMRC produces a weekly report on excess deaths in South Africa. Deaths in South Africa are registered on the Department of Home Affair's National Population Register and includes all citizens with a South African identification number. The number of excess deaths were defined as the number of all-cause deaths in that week minus the number of deaths expected for the week based on 2014–2019 trends. Excess death estimates are adjusted for incomplete reporting, and death estimates in provinces are age-standardized to the national population. Eighty-five percent of the reported excess deaths were used to calculate the maximum IHR. In both provinces, excess deaths were only available starting 21 June (urban) and 28 June 2020 (rural), and not from March as for infections and hospitalizations, therefore the period for wave 1 in the rural community was defined as 28 June –21 November 2020, and as 21 June–21 November 2020 in urban communities for wave 1. For both communities wave 2 was defined as 22 November 2021–26 March 2021. The wave-specific excess death rates were kindly provided by Professor Rob Dorrington.

Differences in SARS-CoV-2 Seroprevalence by HIV Status and Age

To assess differences in seropositivity based on HIV status, we considered seroprevalence at the fifth blood draw and compared the seroprevalence of individuals at each site and age group by HIV status using the Pearson's chi-squared test, and furthermore also compared the seroprevalence by HIV status while controlling for site, sex and age using logistic regression. We calculated power using the two-sample proportion, likelihood-ratio test for a 95% CI, to detect an odds ratio of at least 1.8, assuming a 24% SARS-CoV-2 seroprevalence and 14% HIV prevalence. In the 19–34 year old group (n = 247) we only had 33% power.

References

1. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Hellferscee O, et al.; PHIRST Group. Asymptomatic transmission and high community burden of seasonal influenza in an urban and a rural community in South Africa, 2017–18 (PHIRST): a population cohort study. Lancet Glob Health. 2021;9:e863–74. PubMed https://doi.org/10.1016/S2214-109X(21)00141-8

2. Bushbuckridge Local Municipality. Bushbuckridge Local Municipality final integrated development plan 2020/21. 2020 [cited 2021 Feb 16]. https://www.cogta.gov.za/cgta_2016/wp-content/uploads/2021/02/Bushbuckridge-Municipality.pdf

3. City of Marlosana. Matlosana spatial development framework. 2009 [cited 2021 Feb 16]. http://www.matlosana.gov.za/Documents/Policies/SPLUMA/SDF%20FRAMEWORK%202009.pdf

4. 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. eLife. 2021;10:e64206. PubMed https://doi.org/10.7554/eLife.64206

5. National Institute for Communicable Diseases. COVID-19 sentinel hospital surveillance update, week 18. 2021 May 14 [cited 2021 May 22]. https://www.nicd.ac.za/wp-content/uploads/2021/05/NICD-COVID-19-Weekly-Sentinel-Hospital-Surveillnace-update-Week-18-2021.pdf

6. South African Medical Research Council. Report on weekly deaths in South Africa, week 18. 2021 May 10 [cited 2021 May 22]. https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa

7. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. 2001 [cited 2021 May 5]. https://www.who.int/healthinfo/paper31.pdf

8. Shryock HS, Larmon EA, Siegel JS; US Department of Commerce, Bureau of the Census. The methods and materials of demography. Washington: US Government Printing Office; 1980.

9. Statistics South Africa. Statistical release P0302: mid-year population estimates 2020. 2020 [cited 2021 May 14]. http://www.statssa.gov.za/?page_id=1854&PPN=P0302&SCH=72634

10. National Institute for Communicable Diseases. COVID-19 weekly epidemiology brief, week 19. 2021 May 22 [cited 2021 May 22]. https://www.nicd.ac.za/wp-content/uploads/2021/05/COVID-19-Weekly-Epidemiology-Brief-week-19-2021.pdf

Appendix Table 1. Comparison of characteristics between district (Ehlanzeni and Dr Kenneth Kaunda Districts) (9) and PHIRST-C cohort population, 2020, South Africa.

Characteristic	Ehlanzeni District	Rural community cohort	p-value	Dr Kenneth Kaunda Districts	Urban community cohort	p-value
Age group, y						
<5	203727/1828739 (11%)	99/668 (15%)	<0.01	75574/797716 (9%)	56/598 (9%)	0.93
5–12	319208/1828739 (17%)	210/668 (31%)	<0.01	121145/797716 (15%)	132/598 (22%)	<0.01
13–18	207344/1828739 (11%)	89/668 (13%)	0.11	82582/797716 (10%)	82/598 (14%)	0.01
19–34	491914/1828739 (27%)	108/668 (16%)	<0.01	212152/797716 (27%)	120/598 (20%)	<0.01
35–59	462982/1828739 (25%)	89/668 (13%)	<0.01	237146/797716 (30%)	122/598 (20%)	<0.01
≥60	143564/1828739 (8%)	46/668 (7%)	0.35	69117/797716 (9%)	58/598 (10%)	0.37
Female sex	961384/1828739 (53%)	427/668 (64%)	<0.01	405734/797716 (51%)	336/598 (56%)	0.01
Unemployed	187428/530865 (35%)	198/270 (73%)	<0.01	76654/257457 (30%)	202/315 (64%)	<0.01

Appendix Table 2. Number of sera collected (n) from consenting PHIRST-C participants (N) and percentage sampled during blood draws (BD) in the rural and urban community by blood draw, site, age group, March 2020 – March 2021, South Africa*

	No. (%)						
Characteristic	BD1	BD2	BD3	BD4	BD5	BD1–5	BD3,5
Rural							
Age group, y							
<5	25/102 (25)	49/102 (48)	78/102 (76)	87/102 (85)	89/102 (87)	17/102 (17)	76/102 (75)
5–12	148/215 (69)	169/215 (79)	195/215 (91)	202/215 (94)	201/215 (93)	136/215 (63)	193/215 (90)
13–18	69/94 (73)	72/94 (77)	80/94 (85)	82/94 (87)	78/94 (83)	61/94 (65)	76/94 (81)
19–34	87/117 (74)	88/117 (75)	91/117 (78)	91/117 (78)	95/117 (81)	62/117 (53)	89/117 (76)
35–59	76/91 (84)	77/91 (85)	80/91 (88)	81/91 (89)	82/91 (90)	66/91 (73)	78/91 (86)
≥60	40/49 (82)	39/49 (80)	44/49 (90)	45/49 (92)	42/49 (86)	35/49 (71)	41/49 (84)
All ages	445/668 (67)	494/668 (74)	568/668 (85)	588/668 (88)	587/668 (88)	377/668 (56)	553/668 (83)
Urban							
Age group, y							
<5	45/57 (79)	50/57 (88)	50/57 (88)	44/57 (77)	48/57 (84)	33/57 (58)	45/57 (79)
5–12	116/135 (86)	116/135 (86)	125/135 (93)	123/135 (91)	122/135 (90)	103/135 (76)	120/135 (89)
13–18	75/84 (89)	75/84 (89)	81/84 (96)	77/84 (92)	78/84 (93)	69/84 (82)	78/84 (93)
19–34	102/130 (78)	98/130 (75)	100/130 (77)	99/130 (76)	98/130 (75)	77/130 (59)	92/130 (71)
35–59	110/130 (85)	110/130 (85)	117/130 (90)	112/130 (86)	111/130 (85)	98/130 (75)	111/130 (85)
≥60	57/62 (92)	55/62 (89)	56/62 (90)	55/62 (89)	53/62 (85)	51/62 (82)	53/62 (85)
All ages	505/598 (84)	504/598 (84)	529/598 (88)	510/598 (85)	510/598 (85)	431/598 (72)	499/598 (83)

*Baseline (BD1) 20 July–17 September 2020, second draw (BD2) 21 September–10 October 2020, third draw (BD3) 23 November–12 December 2020, fourth draw (BD4) 25 January–20 February 2021, fifth draw (BD5) 22 March–11 April 2021.

Appendix Table 3. Individuals seropositive for SARS-CoV-2 antibodies with adjusted seroprevalence and 95% creditable intervals by blood draw (BD), site and age, July 2020 – April 2021, South Africa*

Characteristic	BD1	BD2	BD3	BD4	BD5
Rural					
Age group, y					
<5	0/25 (4, 0–13)	1/49 (4, 0–13)	2/78 (4, 1–9)	14/87 (17, 10–25)	15/89 (18, 10–26)
5–12	2/148 (2, 0–5)	5/169 (3, 0–5)	7/195 (4, 2–7)	29/202 (15, 10–20)	39/201 (20, 14–25)
13–18	1/69 (3, 0–7)	5/72 (8, 0–7)	8/80 (11, 5–18)	20/82 (25, 16–35)	24/78 (31, 22–42)
19–34	0/87 (1, 0–4)	5/88 (7, 0–4)	13/91 (15, 8–23)	33/91 (37, 27–47)	35/95 (37, 28–47)
35–59	1/76 (2, 0–7)	5/77 (7, 0–7)	7/80 (10, 4–17)	23/81 (29, 20–39)	27/82 (33, 23–44)
≥60	1/40 (5, 0–13)	4/39 (12, 0–13)	4/44 (11, 3–21)	8/45 (19, 9–32)	10/42 (25, 13–38)
All ages	5/445 (1, 0–2)	25/494 (5, 0–2)	41/568 (7, 5–9)	127/588 (22, 18–25)	150/587 (26, 22–29)
Urban					
Age group, y					
<5	2/45 (6, 1–15)	5/50 (11, 4–21)	7/50 (15, 7–26)	8/44 (20, 9–32)	13/48 (28, 17–41)
5–12	11/116 (10, 5–16)	17/116 (15, 9–22)	24/125 (20, 13–27)	32/123 (26, 19–34)	38/122 (31, 23–40)
13–18	14/75 (19, 11–29)	19/75 (26, 17–36)	30/81 (37, 27–48)	35/77 (46, 35–56)	41/78 (53, 42–64)
19–34	10/102 (10, 5–17)	17/98 (18, 11–26)	23/100 (23, 16–32)	29/99 (30, 21–39)	34/98 (35, 26–44)
35–59	33/110 (30, 22–39)	45/110 (41, 32–50)	52/117 (45, 36–54)	61/112 (55, 45–64)	65/111 (59, 49–68)
≥60	3/57 (7, 2–14)	6/55 (12, 5–22)	7/56 (14, 6–24)	16/55 (30, 19–42)	18/53 (35, 22–47)
All ages	73/505 (15, 12–18)	109/504 (22, 18–25)	143/529 (27, 23–31)	181/510 (36, 32–40)	209/510 (41, 37–45)

*Baseline (BD1) 20 July–17 September 2020, second draw (BD2) 21 September–10 October 2020, third draw (BD3) 23 November –12 December 2020, fourth draw (BD4) 25 January–20 February 2021, fifth draw (BD5) 22 March–11 April 2021.

Appendix Table 4. SARS-CoV-2 seroprevalence from the fifth blood collection (22 March – 11 April 2021), by site, age group and HIV status, South Africa

	Rural			Urban		
Age group, y	HIV-uninfected	HIV-infected	p-value	HIV-uninfected	HIV-infected	p-value
<5	15/87 (17)	0/1 (0)	0.65	12/46 (26)	0/1 (0)	0.55
5–12	37/195 (19)	0/2 (0)	0.49	35/116 (30)	1/3 (33)	0.91
13–18	21/66 (32)	0/3 (0)	0.24	38/73 (52)	3/4 (75)	0.37
19–34	22/65 (34)	10/23 (43)	0.41	26/73 (36)	7/19 (37)	0.92
35–59	12/36 (33)	12/39 (31)	0.81	37/61 (61)	28/47 (60)	0.91
≥60	6/27 (22)	4/10 (40)	0.28	17/46 (37)	1/7 (14)	0.24
All ages	113/476 (24)	26/78 (33)	0.07	165/415 (40)	40/81 (49)	0.11

Appendix Table 5. Comparison of participants with detectable SARS-CoV-2 antibodies at the fifth blood collection (22 March–11 April 2021), South Africa*

Characteristic	Seropositive	Multivariable aOR (95% CI)
Site		
Rural	150/587 (26)	Referent
Urban	209/510 (41)	1.9 (1.5–2.5)
Sex		
Male	130/441 (29)	Referent
Female	229/656 (35)	1.2 (0.9–1.6)
Age group, y		
<5	28/137 (20)	1.2 (0.7–1.9)
5–12	77/323 (24)	Referent
13–18	65/156 (42)	2.6 (1.5–4.5)
19–34	69/193 (36)	2.0 (1.2–3.4)
35–59	92/193 (48)	3.2 (1.8–5.5)
≥60	28/95 (29)	1.5 (0.8–2.9)
HIV status		
Negative	278/891 (31)	1.0 (0.7–1.5)
Positive	66/159 (42)	Referent

*aOR: adjusted odds ratio.

Appendix Table 6. Comparison of participants with detectable SARS-CoV-2 antibodies after the first wave (blood draw 3) and second wave (blood draw 5), by site, July 2020–April 2021, South Africa*

	Rural				Urban			
	No. (%)				No. (%)			
Age group,y	Sero+ B3	Sero+ B5	Infected wave 2	OR (95% CI	Sero+ B3	Sero+ B5	Infected wave 2	OR (95% CI)
<5	2/76 (3)	12/76 (16)	10/12 (83)	1.8 (0.3–10.6)	12/76 (16)	12/45 (27)	6/13 (46)	3.2 (0.9–11.0)
5–12	7/193 (4)	36/193 (19)	29/36 (81)	1.5 (0.5–5.1)	36/193 (19)	38/120 (32)	15/38 (39)	2.4 (1.0–5.8)
13–18	7/76 (9)	23/76 (30)	16/23 (70)	0.8 (0.2–2.9)	23/76 (30)	41/78 (53)	12/41 (29)	1.5 (0.6–3.8)
19–34	13/89 (15)	33/89 (37)	20/33 (61)	0.6 (0.2–1.7)	33/89 (37)	32/92 (35)	13/34 (38)	2.3 (0.9–5.7)
35–59	7/78 (9)	26/78 (33)	19/26 (73)	Referent	26/78 (33)	65/111 (59)	14/66 (21)	Referent
≥60	4/41 (10)	10/41 (24)	6/10 (60)	0.6 (0.1–2.6)	10/41 (24)	18/53 (34)	11/18 (61)	5.8 (1.9–17.8)

* OR: odds ratio; Sero+: seropositive.

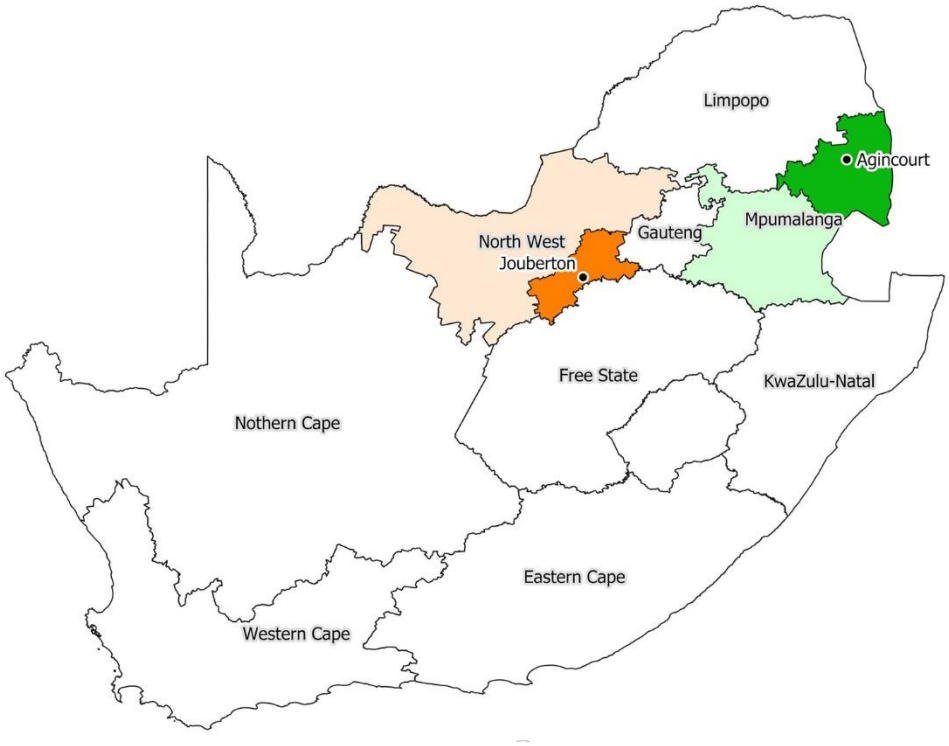
Appendix Table 7. Comparison of individuals with and without a blood draw 3 and 5 pair, South Africa*

Characteristic	Missing 3,5 pair	Univariate OR (95% CI)
Site		
Rural	115/668 (17)	1.0 (0.8–1.4)
Urban	99/598 (17)	Referent
Sex		
M	85/503 (17)	1.0 (0.7–1.3)
F	129/763 (17)	Referent
Age group, y		
<5	38/159 (24)	2.7 (1.6–4.4)
5–12	37/350 (11)	Referent
13–18	24/178 (13)	1.3 (0.8–2.3)
19–34	66/247 (27)	3.1 (2.0–4.8)
35–59	32/221 (14)	1.4 (0.9–2.4)
≥60	17/111 (15)	1.5 (0.8–2.8)
HIV status		
Negative	112/968 (12)	Referent
Positive	22/178 (12)	1.1 (0.7–1.8)

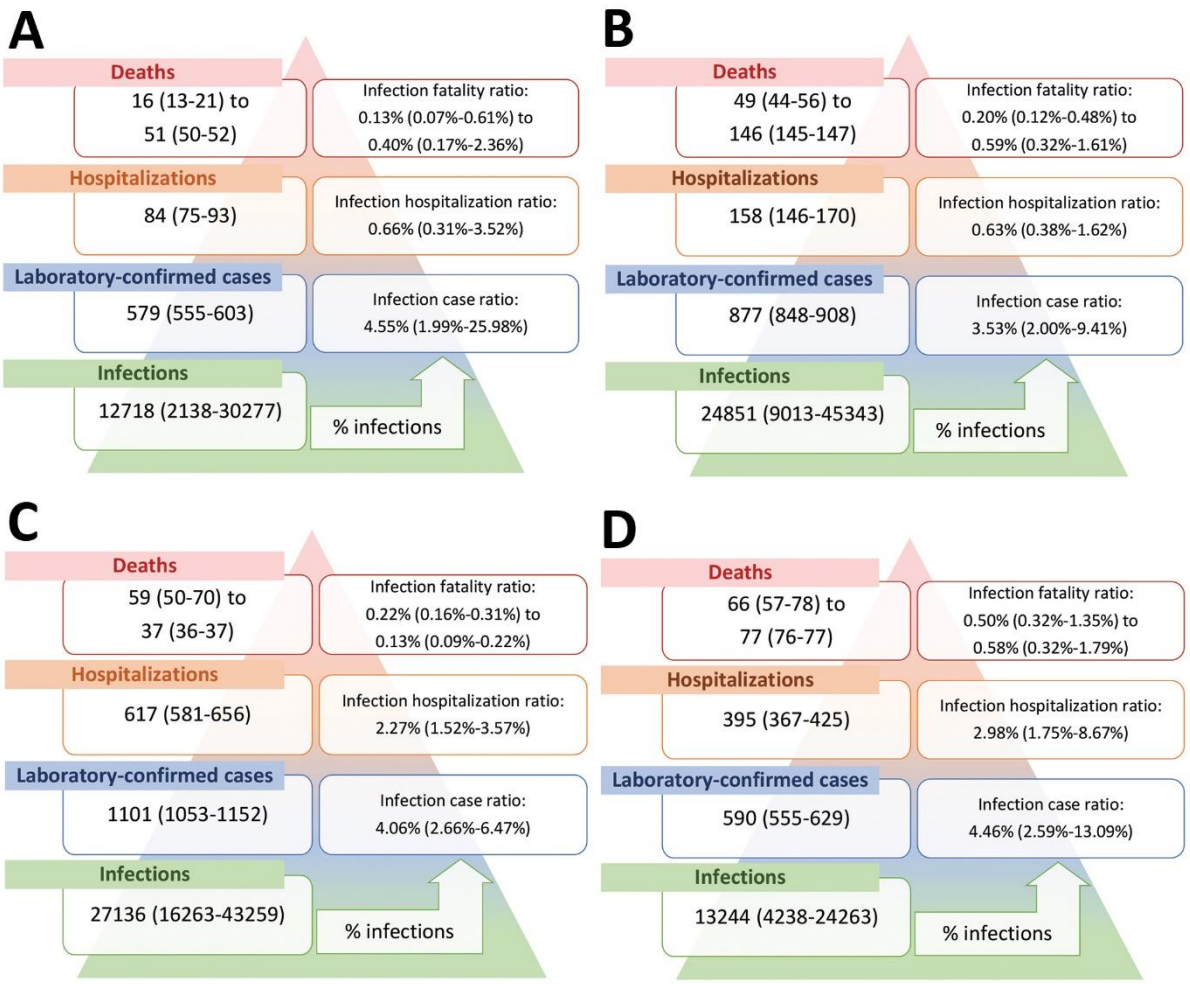
*OR: odds ratio.

Appendix Table 8. The PHIRST-C Group

Name	Affiliation
Amelia Buys	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.
Anne von Gottberg	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Cheryl Cohen	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
F. Xavier Gómez-Olivé	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Floidy Wafawanaka	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Jackie Kleynhans	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Jacques du Toit	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Jinal N. Bhiman	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Jocelyn Moyes	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Kathleen Kahn	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Kgaugelo Patricia Kgasago	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.
Limakatso Lebina	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.
Linda de Gouveia	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.
Maimuna Carrim	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.
Meredith L. McMorrow	Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa.
Mignon du Plessis	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.
Neil A. Martinson	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. Johns Hopkins University Center for TB Research, Baltimore, Maryland, United States of America.
Nicole Wolter	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Retshidisitswe Kotane	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Stefano Tempia	Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. MassGenics, Atlanta, Georgia, United States of America.
Stephen Tollman	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Science, University of the Witwatersrand.
Tumelo Moloantoa	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.



Appendix Figure 1. Location of study sites PHIRST-C, South Africa. The rural site is located in Agincourt, Ehlanzeni District (dark green), Mpumalanga Province (light green) and the urban site in Jouberton, Dr Kenneth Kaunda District (dark orange), North West Province (light orange).



Appendix Figure 2. World Health Organization population age-standardized SARS-CoV-2 infection, diagnosis, hospitalization and deaths per 100,000 population in a rural community during A) wave 1 and B) wave 2, and urban community during C) wave 1 and D) wave 2 of infections, South Africa, March 2020–March 2021.

The Vanuatu Health Program is supported by the Australian Department of Foreign Affairs and Trade Australian Aid Program. C.v.G. is a recipient of an Early Career Research Fellowship, supported by the Australian National Health and Medical Research Council.

About the Author

Dr. van Gemert is an epidemiologist and postdoctoral research fellow in the Vanuatu Health Program, Port Vila, Vanuatu, and with the Burnet Institute, Melbourne, Victoria, Australia. Her primary research interest is surveillance of infectious disease in resource-poor settings.

References

1. Vanuatu Government Office of the President. Extraordinary Gazette Numbero Special No. 3. Extension of Declaration of State of Emergency Order No. 93 of 2020 [cited 2020 Jul 13]. https://www.gov.vu/index.php/events/news/86-extension-of-the-declaration-of-the-soe-order-no-93-of-2020
2. Vanuatu Ministry of Health. Vanuatu situation report 59—23 December 2021 [cited 2022 Jan 10]. https://covid19.gov.vu/images/Situation-reports/19122021_Vanuatu_COVID19_NHEOC_SitRep_59_2.pdf
3. Yang B, Tsang TK, Wong JY, He Y, Gao H, Ho F, et al. The differential importation risks of COVID-19 from inbound travellers and the feasibility of targeted travel controls: a case study in Hong Kong. Lancet Reg Health West Pac. 2021;13:100184. https://doi.org/10.1016/j.lanwpc.2021.100184
4. Clifford S, Pearson CA, Klepac P, Van Zandvoort K, Quilty BJ, Eggo RM, et al.; CMMID COVID-19 working group. Effectiveness of interventions targeting air travellers for delaying local outbreaks of SARS-CoV-2. J Travel Med. 2020;27:taaa068. https://doi.org/10.1093/jtm/taaa068
5. World Health Organization. WHO coronavirus (COVID-19) dashboard 2022 [cited 2022 Jan 10]. https://covid19.who.int/
6. Vanuatu National Statistics Office. Statistics update: international visitor arrivals. December 2020 provisional highlights [cited 2021 Feb 10]. https://www.stats.govt.nz/information-releases/international-travel-december-2021
7. Pritchard E, Matthews PC, Stoesser N, Eyre DW, Gethings O, Vihta KD, et al. Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. Nat Med. 2021;27:1370–8. https://doi.org/10.1038/s41591-021-01410-w

Address for correspondence: Caroline van Gemert, The Burnet Institute, 85 Commercial Rd, Melbourne, VIC 3004, Australia; email: caroline.vangemert@burnet.edu.au

SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa

Jackie Kleynhans, Stefano Tempia, Nicole Wolter, Anne von Gottberg, Jinal N. Bhiman, Amelia Buys, Jocelyn Moyes, Meredith L. McMorrow, Kathleen Kahn, F. Xavier Gómez-Olivé, Stephen Tollman, Neil A. Martinson, Floidy Wafawanaka, Limakatso Lebina, Jacques D. du Toit, Waasila Jassat, Mzimasi Neti, Marieke Brauer, Cheryl Cohen, for the PHIRST-C Group[1]

Author affiliations: National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa (J. Kleynhans, S. Tempia, N. Wolter, A. von Gottberg, J.N. Bhiman, A. Buys, J. Moyes, W. Jassat, M. Neti, C. Cohen); University of the Witwatersrand, Johannesburg (J. Kleynhans, S. Tempia, N. Wolter, A. von Gottberg, J.N. Bhiman, J. Moyes, C. Cohen); US Centers for Disease Control and Prevention, Atlanta, Georgia, USA (S. Tempia, M.L. McMorrow); US Centers for Disease Control and Prevention, Pretoria, South Africa (M.L. McMorrow); MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), University of the Witwatersrand, Johannesburg (K. Kahn, F.X. Gómez-Olivé, S. Tollman, F. Wafawanaka, J. du Toit); Johns Hopkins University Center for Tuberculosis Research, Baltimore, Maryland, USA (N.A. Martinson); Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg (N.A. Martinson, L. Lebina); Africa Health Research Institute, Durban, South Africa (L. Lebina); Ampath Pathology, Pretoria (M. Brauer)

DOI: https://doi.org/10.3201/eid2805.220278

By November 2021, after the third wave of severe acute respiratory syndrome coronavirus 2 infections in South Africa, seroprevalence was 60% in a rural community and 70% in an urban community. High seroprevalence before the Omicron variant emerged may have contributed to reduced illness severity observed in the fourth wave.

South Africa has experienced 4 waves of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections, the fourth dominated by the Omicron variant of concern (*1*). Data on the proportion of the population with serologic evidence of previous infection at the time of Omicron emergence are important to contextualize the observed rapid increases and subsequent quick decline in case numbers (*1*), as well as the lower severity compared with previous variants (*2*).

[1]Additional members of the PHIRST-C group who contributed to this article are listed at the end of this article.

We previously described the seroprevalence of SARS-CoV-2 in the PHIRST-C (Prospective Household Study of SARSCoV-2, Influenza, and Respiratory Syncytial Virus Community Burden, Transmission Dynamics, and Viral Interaction) cohort in a rural and an urban community at 5 timepoints during July 2020–March 2021 (*3*). By using the same methods (Appendix, https://wwwnc.cdc.gov/EID/article/28/5/22-0278-App1.pdf), we report seroprevalence at 4 additional timepoints through November 27, 2021, spanning the third, Delta-dominated wave (Appendix Figure 1), ending the week Omicron was identified (*4*). We tested serum samples by using the Roche Elecsys Anti-SARS-CoV-2 assay (Roche Diagnostics, https://www.roche.com); we considered a cutoff index ≥1.0 an indication of prior infection. The immunoassay detects nucleocapsid (N) antibodies; thus, it does not detect postvaccination antibody responses. We obtained seroprevalence 95% credible intervals (CrIs) by using Bayesian inference with 10,000 posterior draws (*5*). We estimated the age- and sex-adjusted number of infections and age-adjusted diagnosed cases, hospitalizations, deaths, case-to-infection ratio (CIR), hospitalization-to-infection ratio (HIR), and in-hospital and excess death fatality-to-infection ratio (FIR), as described previously (*3*) (Appendix). Third-wave infections were defined as participants who had a paired blood draw (BD) from the fifth timepoint of the previous study (BD5) (collected March 22–April 11, 2021) and from the ninth timepoint of this study (BD9) (collected November 15–27, 2021) and who were seronegative at BD5 and seropositive at BD9 or seropositive at BD5 but had a ≥2-fold higher cutoff index in BD9 (because 38 possible reinfections occurred after BD5 [Appendix]). We obtained vaccination status through reviewing vaccine cards that participants kept at home. The study was approved by the University of the Witwatersrand Human Research Ethics Committee (reference no. 150808); the US Centers for Disease Control and Prevention relied on local clearance (IRB approval no. 6840).

Overall, pre–third wave (BD5) SARS-CoV-2 seroprevalence adjusted for assay sensitivity and specificity was 26% (95% CrI 22%–29%) in the rural and 41% (95% CrI 37%–45%) in the urban community. After the third wave (BD9), overall seroprevalence increased to 60% (95% CrI 56%–64%) in the rural community and 70% (95% CrI 66%–74%) in the urban community (Figure; Appendix Table 1). In both communities, the largest increase in seroprevalence was seen in children 13–18 years of age, who also had the highest seroprevalence of all ages after the third wave: 80% (95% CrI 70%–88%) in the rural community (a 49% increase) and 83% (95% CrI 73%–90%) in the urban community (a 19% increase).

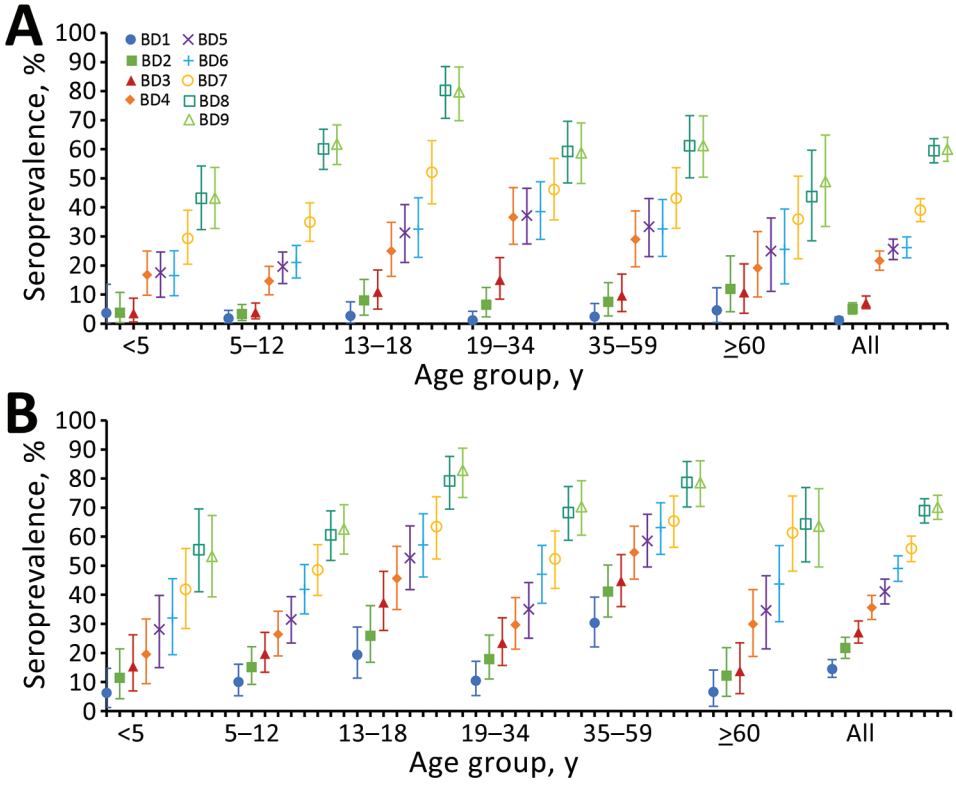


Figure. Severe acute respiratory syndrome coronavirus 2 seroprevalence at each blood collection, by age group, in a rural community (A) and urban community (B), South Africa, March 2020–November 2021. Baseline blood draw (BD1) collected July 20–September 17, 2020; second draw (BD2), September 21 – October 10, 2020; third draw (BD3), November 23–December 12, 2020; fourth draw (BD4), January 25–February 20, 2021; fifth draw (BD5), March 22–April 11, 2021; sixth draw (BD6), May 20–June 9, 2021; seventh draw (BD7), July 19–August 5, 2021; eighth draw (BD8), September 13–25, 2021; ninth draw (BD9), November 15–27, 2021. Error bars represent 95% credible intervals. Seroprevalence estimates adjusted for sensitivity and specificity of assay.

During the third wave of infections, the incidence at the rural site was 39% (95% CrI 24%–55%), resulting in a CIR of 3% (95% CI 2%–5%). HIR was 0.5% (95% CI 0.3%–0.7%) and in-hospital FIR was 0.1% (95% CI 0.1%–0.2%); excess deaths FIR was 0.5% (95% CI 0.4%–0.8%) (Figure; Appendix Figure 2).

In the urban community, the incidence during the third wave was 40% (95% CrI 26%–54%). CIR was a 5% (95% CI 4%–8%), and HIR was 2% (95% CI 2%–4%). In-hospital FIR was 0.4% (95% CI 0.3%–0.6%) and excess deaths FIR was 0.6% (95% CI 0.4%–0.9%) (Figure; Appendix Figure 2).

HIR and FIR were similar between wave 2 and 3 (Appendix Figure 3). SARS-CoV-2 vaccines became available in South Africa in February 2021, after the second wave. By the end of wave 3, only 8% (49/609) of participants were fully vaccinated (1 dose of Johnson & Johnson/Janssen or 2 doses of Pfizer-BioNTech) in the rural community and 19% (97/512) in the urban community (Appendix Table 2). Considering the overall low vaccination coverage in these communities during the study period, the similar HIR and FIR in wave 2 and 3 were likely driven by a combination of natural immunity and potentially a moderate effect attributable to vaccination.

Taken together, by the end of November 2021, just before the emergence of Omicron, the combined proportion of persons who had serologic evidence of previous infection (at any timepoint), were fully vaccinated, or both was 62% (389/631) at the rural community and 72% (411/568) at the urban community (Appendix Table 3).

After the third wave of infections in South Africa, we observed a ≥60% overall seroprevalence attributable to SARS-CoV-2 infection, ranging from 43% in rural community children <5 years of age to 83% in urban community children 13–18 years of age (Figure). CIR, HIR, and FIRs were similar between the second and third waves. Similar to our data, results from a study in Gauteng Province found seroprevalence of 56%–80% attributable to natural infection before the emergence of Omicron (*6*). The high seroprevalence before Omicron emergence may have contributed to reduced illness severity observed in the fourth wave (*2*).

Additional members of the PHIRST-C Group who contributed: Kgaugelo Patricia Kgasago, Linda de Gouveia, Maimuna Carrim, Mignon du Plessis, Retshidisitswe Kotane, and Tumelo Moloantoa.

Acknowledgments

We thank all persons participating in the study and the field teams for their hard work and dedication to the study, the laboratory teams, the PHIRST-C scientific and safety committee, the national SARS-CoV-2 National Institute for Communicable Diseases (NICD) surveillance team, and NICD Information Technology.

This work was supported by the NICD of the National Health Laboratory Service and the US Centers for Disease Control and Prevention (cooperative agreement no. 6U01IP001048-04-02 awarded to C.C.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. C.C. received grant funds from US Centers for Disease Control and Prevention, Wellcome Trust, and South African Medical Research Council. N.W. and A.v.G. received grant funds from Sanofi and the Gates Foundation.

The investigators welcome enquiries about possible collaborations and requests for access to the dataset. Data will be shared after approval of a proposal and with a signed data access agreement. Investigators interested in more details about this study, or in accessing these resources, should contact the corresponding author.

About the Author

Ms. Kleynhans is an epidemiologist in the Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases. Her primary research interests include the epidemiology of respiratory diseases like influenza and COVID-19, vaccine impact studies, and modelling of infectious disease transmission dynamics.

References

1. National Institute for Communicable Diseases. COVID-19 weekly epidemiology brief week 3. 2022 Jan 26 [cited 2022 Jan 31]. https://www.nicd.ac.za/wp-content/uploads/2022/01/COVID-19-Weekly-Epidemiology-Brief-week-3-2022.pdf
2. Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. Lancet. 2022;399:437–46. https://doi.org/10.1016/S0140-6736(22)00017-4
3. Kleynhans J, Tempia S, Wolter N, von Gottberg A, Bhiman JN, Buys A, et al.; PHIRST-C Group. PHIRST-C Group. SARS-CoV-2 seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. Emerg Infect Dis. 2021;27:3020–9. https://doi.org/10.3201/eid2712.211465
4. World Health Organization. Classification of Omicron (B.1.1.529): SARS-CoV-2 variant of concern. 2021 Nov 26 [cited 2022 Jan 5]. https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern
5. 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. eLife. 2021;10:e64206. https://doi.org/10.7554/eLife.64206

6. Madhi SA, Kwatra G, Myers JE, Jassat W, Dhar N, Mukendi CK, et al. Population immunity and Covid-19 severity with Omicron variant in South Africa. N Engl J Med. 2022;NEJMoa2119658. https://doi.org/10.1056/NEJMoa2119658

Address for correspondence: Jackie Kleynhans, Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, 1 Modderfontein Rd, Sandringham, 2192, Johannesburg, South Africa; email: jackiel@nicd.ac.za

Angiostrongylus cantonensis in a Red Ruffed Lemur at a Zoo, Louisiana, USA

Jessica Rizor, Ryan A. Yanez, Tuddow Thaiwong, Matti Kiupel

Authors affiliation: Michigan State University Veterinary Diagnostic Laboratory, Lansing, Michigan, USA

DOI: https://doi.org/10.3201/eid2805.212287

A red ruffed lemur (*Varecia rubra*) from a zoo in Louisiana, USA, was euthanized for worsening paresis. Brain and spinal cord histology identified eosinophilic meningoencephalomyelitis with intralesional adult *Angiostrongylus* sp. nematodes. PCR and sequencing confirmed *A. cantonensis* infection, indicating this parasite constitutes an emerging zoonosis in the southeastern United States.

Angiostrongylus cantonensis is a parasitic metastrongyloid nematode that has a neurotropic larval stage and is endemic throughout Southeast Asia and the Pacific Islands. The rat (*Rattus* spp.) is the main definitive host and a variety of gastropods serve as intermediate hosts. In rats, infections cause no brain damage and only some pulmonary disease in severe infections. However, in aberrant hosts, including humans and nonhuman primates, larvae cause severe eosinophilic meningoencephalitis. Clinical signs are associated with migration of the larvae and the immune response to dead or dying nematodes (*1*).

In 1987, *A. cantonensis* nematodes were detected in rats in New Orleans, Louisiana, USA (*2*); in 1995, a human case of eosinophilic meningitis was reported in North America in a child from New Orleans (*3*). *A. cantonensis* nematodes have now become endemic in the southeastern United States, as evidenced by reports of infection in a child in Texas (*4*); a horse from Mississippi (*5*); captive Geoffroy's tamarins (*Saguinus geoffroyi*) in Alabama (*6*); and several animals in Florida, including a white-handed gibbon (*Hylobates lar*), an orangutan (*Pongo pygmaeus*), a white-throated capuchin monkey (*Cebus capucinus*), a red ruffed lemur (*Varecia rubra*), and a nine-banded armadillo (*Dasypus novemcinctus*) (*7*,*8*). Ingestion of infected gastropods and paratenic hosts or unwashed contaminated vegetables are proposed routes of infection for aberrant hosts.

The International Union for Conservation of Nature lists red ruffed lemurs (*Varecia rubra*) as critically endangered (*9*). In June 2021, a 9-year-old male red ruffed lemur from a zoo in Louisiana was humanely euthanized because of hind limb paresis and a right head tilt that worsened over an 8-day period. The lemur was housed in a troop of 5 adult lemurs in an outdoor exhibit. Various snail species are common in the enclosure, but no other lemurs were clinically affected.

A necropsy performed at the Michigan State University Veterinary Diagnostic Laboratory (Lansing, Michigan, USA) identified no gross lesions. The laboratory formalin-fixed and processed the brain, the entire spinal cord, and all major organs for histopathology. Histopathologic examination revealed multiple transverse and longitudinal sections of adult nematodes within the subarachnoid space and neuropil of the cerebellum and brainstem. Nematodes were ≈50–70 μm in diameter and had a 3–4-μm thick smooth, eosinophilic cuticle and prominent lateral cords. Adult nematodes had coelomyarian musculature, and the pseudocoelom contained a reproductive tract and an intestinal tract lined by multinucleated cells with flocculent eosinophilic to brown material in the lumen (Figure). Nematodes were surrounded by hemorrhage and small numbers of eosinophils, neutrophils, macrophages, and glial cells. Several cerebellar folia were effaced by invading nematodes, hemorrhage, and inflammation. The cerebellar meninges were expanded by numerous eosinophils, fewer neutrophils, foamy macrophages, multinucleated giant cells, and lymphocytes. A representative section of thoracic spinal cord contained an identical single adult nematode in the subdural space. Another adult nematode had regionally effaced the dorsal horn in a section of lumbar spinal cord. The affected spinal cord had regional rarefaction of both gray and white

Article DOI: https://doi.org/10.3201/eid2805.220278

SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa

Appendix

Supplementary Methods

Study Population

PHIRST-C (A Prospective Household study of SARS-CoV-2, Influenza, and Respiratory Syncytial virus community burden, Transmission dynamics and viral interaction in South Africa) is a continuation from PHIRST (A Prospective Household observational cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community burden and Transmission dynamics in South Africa) (*1*). PHIRST-C is conducted in two communities with established influenza-like illness and pneumonia surveillance sites.

South Africa consists of nine provinces, which are divided into 52 districts, which form the second level of administration. Districts are further divided into municipalities. The rural site is located in the Bushbuckridge Municipality, Ehlanzeni District, Mpumalanga Province and forms part of a health and sociodemographic surveillance site (HDSS) at the Medical Research Council (MRC)/University of Witwatersrand Rural Public Health and Health Transitions Research Unit, Agincourt. During PHIRST, two of the 29 villages within the HDSS were selected according to convenience (proximity and burden of other studies within the site) each year between 2016–2018. A random selection of ≈50 households with a household size of 3 or more were approached for enrolment. The urban site is located in the Jouberton Township, Matlosana Municipality, Dr Kenneth Kaunda District, North West Province. A list of 450 global positioning system (GPS) coordinates were generated using Google Earth. Study staff approached the nearest house within 30 m of the coordinate point for enrolment. Approximately 50 households were enrolled each year during 2016–2018. At both sites, households were eligible if they consisted of 3 or more household members (sharing at least four meals a week).

All household members were approached for inclusion in the study, and households were eligible for inclusion if >80% of members consented to be enrolled.

For PHIRST-C, all households who participated in PHIRST from the urban site, and households from the 2017 and 2018 cohort in the rural site were approached for enrolment. To supplement the sample size, additional households at each site were approached using the same methods as for PHIRST. Villages in the rural site were restricted to those four used in 2017 and 2018 of PHIRST. Informed consent was obtained from participating adults or a parent/guardian for children <18 years of age. In addition to parent/guardian consent, assent was also obtained from children aged 7–17 years.

We collected baseline data and blood (blood draw [BD] 1) at enrollment (July 20–September 17, 2020) and every 2 months thereafter: BD2, September 21–October 10; BD3, November 23–December 12, 2020; BD4, January 25–February 20, 2021; BD5, March 22–April 11, 2021; BD6, May 20 – June 9, 2021; BD7, July 19 – August 5, 2021; BD8, September 13 – 25, 2021; BD9, November 15 – 27, 2021.

The Bushbuckridge Municipality within the Ehlanzeni District is considered as predominantly rural, with main industries being agriculture and tourism (*2*). The Ehlanzeni District has a population of 1,828,738, with sixty percent of the district's population being >18 years of age. The City of Matlosana Municipality is considered to be 88.2% urban, and is located in the Dr Kenneth Kaunda District, which has a population of 797,715 (*3*). Sixty-five percent of the district population is aged >18 years.

Calculation of Case-to-infection Ratio (CIR), Hospitalization-to-infection Ratio (HIR), and Fatality-to-infection Ratio (FIR) by Wave of Infection

We calculated the age- and sex-adjusted total number of infections, age-adjusted diagnosed cases, hospitalizations, deaths, CIR (number of infections compared to diagnosed cases), HIR and in-hospital and excess death FIR as described in below equations. We defined the third wave as March 22, 2021 (week 12) to November 14, 2021 (week 45), starting with BD5, and ending the week before BD9 started. Due to differences between the sex ratios of our cohort and the district population, the infection estimates were also adjusted for sex. Data sources used in these calculations are described in the next section. Age- and sex- standardized estimates for the selected endpoints for each wave were obtained as follows:

$$Infections = \frac{\sum_i (s_i \times SAp_i)}{\sum_i (SAp_i)} \times 100{,}000$$

Where s_i is the seroprevalence in the cohort in the respective community for age and sex group i and SAp_i is the South African population for age and sex group i. Calculated for wave 3 as the number of individuals seronegative at BD5 and seropositive at BD9. Furthermore, we also included possible re-infections where the individual was already seropositive at BD5, but had a ≥2-fold higher cutoff index (COI) in BD9 compared to BD5. We therefore included infections and re-infections that occurred during the third wave. Estimates only included participants with a blood draw 5 and 9 pair, and adjusted for sensitivity and specificity of test (*4*).

The ≥2-fold increase in COI cutoff to define possible re-infections was not externally validated, but based on PCR-confirmed infections from twice-weekly nasal swab collections between March 22, 2021 (BD5) to July 19, 2021 (BD7) in the same cohort (C. Cohen et al., unpub. data, https://doi.org/10.1101/2021.07.20.21260855). Increases in COI from BD5 to BD7 ranged from 1 to 93 fold. Of the 30 individuals who had a rRT-PCR-confirmed infection between BD5 and BD7, 16 (53%, sensitivity) had a ≥2-fold increase from BD5 to BD7, and of the 312 individuals that did not have a rRT-PCR infection, 301 (96%, specificity) did not have a ≥2-fold increase. We were therefore more likely to have underestimated re-infections.

$$Diagnosed\ cases = \frac{\sum_i \left((c_i \div p_{i(d)}) \times p_{i(SA)} \right)}{\sum_i \left(p_{i(SA)} \right)} \times 100{,}000$$

Where c_i is the number of diagnosed cases (RT-PCR and antigen-based tests) from the respective district reported to the NMCSS during wave 3 (March 22, to November 14, 2021) in age group i, $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age group i.

$$Hospitalizations = \frac{\sum_i \left((h_i \div p_{i(d)}) \times p_{i(SA)} \right)}{\sum_i \left(p_{i(SA)} \right)} \times 100{,}000$$

Where h_i is the number of hospitalizations from the respective district reported to COVID-19 Sentinel Hospital Surveillance (DATCOV) during wave 3 (March 22, to November 14, 2021) (*5*) in age group i, $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age group i.

$$In-hospital\ deaths = \frac{\sum_i \left((d_i \div p_{i(d)}) \times p_{i(SA)} \right)}{\sum_i \left(p_{i(SA)} \right)} \times 100{,}000$$

Where d_i is the number of in-hospital deaths from the respective districts reported to DATCOV during wave 3 (March 22, to November 14, 2021) (*5*) in age group *i*, $p_{i(d)}$ is the district population for age group *i* and $p_{i(SA)}$ is the South African population for age group *i*.

$$Excess\ deaths = \frac{(ExD\ x\ 0.85)\ \times\ p_{(SA)}}{p_{(SA)}} \times 100{,}000$$

Where *ExD* is the rate of provincial excess deaths adjusted to the South African population reported by South African Medical Research Council during wave 3 (March 22, to November 14, 2021) (*6*), and $p_{(SA)}$ is the total South African population. According to estimates only 85% of excess deaths are attributable to COVID-19 (*6*).

For infections, hospitalizations, in-hospital (minimum) and excess (maximum) deaths, 95% CIs were calculated using the Clopper-Pearson method with a Poisson distribution.

$$Case-to-infection\ ratio\ (CIR) = \frac{Diagnosed\ cases}{Infections}\ x\ 100$$

$$Hospitalization-to-infection\ ratio\ (HIR) = \frac{Hospitalizations}{Infections}\ x\ 100$$

$$Fatality\ -to-infection\ ratio\ (FIR) = \frac{Deaths*}{Infections}\ x\ 100$$

Confidence intervals for infection ratios were calculated as ratios from the 95% confidence intervals of infection, hospitalization and death rates.

Data Sources

Population Denominators (*7*)

Population numbers for each district, by age, were obtained from the StatsSA 2021 mid-year population estimates.

Notifiable Medical Conditions Surveillance System (NMCSS) (*8*)

Reverse transcription PCR (RT-PCR) testing in South Africa to detect SARS-CoV-2 RNA started on 28 January 2020. Rapid SARS-CoV-2 antigen testing was implemented in

November 2020. All laboratories in the private and public sector performing SARS-CoV-2 tests automatically feed testing data from their lab information systems to the NMCSS, from where daily and weekly reporting on national, provincial and district SARS-CoV-2 infections are performed. No serology tests are used for this reporting. Results from antigen tests may be underestimated in the NMCSS as not all antigen test results are captured on lab information systems which feed into NMCSS. The number of cases reported to the NMCSS in the Ehlanzeni District (rural community) and Dr Kenneth Kaunda District (urban community) was used as an estimate of reported, diagnosed SARS-CoV-2 infections.

COVID-19 National Hospital Surveillance (DATCOV) (*5*)

DATCOV is a national hospital surveillance system to which all hospitals where COVID-19 admissions occurred in the public and private sector in South Africa report. The case definition includes any person admitted with a positive RT-PCR test result for SARS-CoV-2; and may include individuals for whom the main cause of hospitalization was not SARS-CoV-2. In-hospital outcome was available for all patients. The total number of hospitalizations and in-hospital deaths in each district reported to DATCOV between March 22, to November 14, 2021 was used to calculate in the HIR and minimum FIR, respectively.

South African Medical Research Council (SAMRC) Report on Weekly Deaths (*6*)

The Burden of Disease Unit at the SAMRC produces a weekly report on excess deaths in South Africa. Deaths in South Africa are registered on the Department of Home Affair's National Population Register and includes all citizens with a South African identification number. The number of excess deaths were defined as the number of all-cause deaths in that week minus the number of deaths expected for the week based on 2014–2019 trends. Excess death estimates are adjusted for incomplete reporting, and death estimates in provinces are age-standardized to the national population. Eighty-five percent of the reported excess deaths were used to calculate the maximum FIR as suggested by SAMRC.

Seroreversions

Seven individuals were seronegative at BD5, and with at least one seropositive result during BD6 and BD8, and who were seronegative again at BD9. The maximum COIs for these individuals were low, ranging from 1.09 to 34.5. This fits with an initial low antibody response and subsequent antibody titer waning. In addition, there were five individuals who were

seropositive at BD5, whose antibody titers waned below detection during BD6 to BD8, but who were seropositive again at BD9. Two of these individuals had a lower COI in BD9 than in BD5, and were not included as possible re-infections during wave 3.

Limitations

Our study is limited by a small sample size, reducing the power for accurate seroprevalence estimates in small age strata, and inclusion of only 2 geographic sites, including only households with ≥3 people, with high unemployment in selected households compared to the community unemployment rate (*9*) and therefore may not be representative of other districts and provinces in South Africa. Our definition of re-infections has not been validated elsewhere and is based on a small number of infections that occurred between BD5 and BD7, and did not consider different variants. Based on data from the same cohort, these reinfections occurred in only a small portion (3%) of the cohort (C. Cohen et al., unpub. data, https://doi.org/10.1101/2021.07.20.21260855) and would have had a negligible influence on the infection ratios. CIR, HIR, and FIR formed part of an ecologic analysis, which is inherently prone to biases. Transmission dynamics within our cohort may not be similar to the transmission dynamics within the district. Twenty percent (260/1,277) of persons did not have a BD 5+9 pair, and bias could have been introduced if the seroprevalence were different for those without a BD 5+9 blood pair. Since we did not have complete serologic results for draws between BD5 and BD9 for all individuals with a BD 5+9 pair, we did not consider individuals who may have seroreverted during wave 3 as infections or re-infections. This may have led to an underestimation of infections, and overestimation of the CIR, HIR and FIR.

References

1. Cohen C, McMorrow ML, Martinson NA, Kahn K, Treurnicht FK, Moyes J, et al.; PHIRST group. Cohort profile: a prospective household cohort study of influenza, respiratory syncytial virus and other respiratory pathogens community burden and transmission dynamics in South Africa, 2016–2018. Influenza Other Respir Viruses. 2021;15:789–803. PubMed https://doi.org/10.1111/irv.12881

2. Bushbuckridge Local Municipality. Bushbuckridge local municipality final integrated development plan, 2020/21. 2020 [cited 2021 Feb 16]. https://www.cogta.gov.za/cgta_2016/wp-content/uploads/2021/02/Bushbuckridge-Municipality.pdf

3. City of Marlosana. Matlosana spatial development framework. 2009 [cited 2021 Feb 16]. http://www.matlosana.gov.za/Documents/Policies/SPLUMA/SDF%20FRAMEWORK%202009.pdf

4. 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. eLife. 2021;10:e64206. PubMed https://doi.org/10.7554/eLife.64206

5. National Institute for Communicable Diseases. COVID-19 sentinel hospital surveillance update, week 51. 2021 Dec 28 [cited 2022 Jan 5]. https://www.nicd.ac.za/wp-content/uploads/2021/12/NICD-COVID-19-Weekly-Sentinel-Hospital-Surveillnace-update-Week-51-2021.pdf

6. South African Medical Research Council. Report on weekly deaths in South Africa, week 50. 2021 Dec 22 [cited 2022 Jan 5]. https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa

7. Statistics South Africa. Statistical release P0302. Mid-year population estimates 2021. 2021 [cited 2022 Jan 5]. http://www.statssa.gov.za/?page_id=1854&PPN=P0302

8. National Institute for Communicable Diseases. COVID-19 weekly epidemiology brief week 3. 2022 Jan 26 [cited 2022 Jan 31]. https://www.nicd.ac.za/wp-content/uploads/2022/01/COVID-19-Weekly-Epidemiology-Brief-week-3-2022.pdf

9. Kleynhans J, Tempia S, Wolter N, von Gottberg A, Bhiman JN, Buys A, et al.; PHIRST-C Group. SARS-CoV-2 seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. Emerg Infect Dis. 2021;27:3020–9. PubMed https://doi.org/10.3201/eid2712.211465

Appendix Table 1. Individuals seropositive for SARS-CoV-2 antibodies with adjusted seroprevalence and 95% credible intervals by blood draw, site and age, July 2020 – November 2021, South Africa*

Site	Age group, y	Seropositive/Tested (Seroprevalence*, 95% CrI)								
		BD1	BD2	BD3	BD4	BD5	BD6	BD7	BD8	BD9
Rural	<5	0/25 (4, 0–14)	1/49 (4, 0–14)	2/78 (4, 1–9)	14/87 (17, 10–25)	15/89 (18, 10–26)	14/88 (17, 10–25)	26/90 (29, 20–39)	33/77 (43, 32–54)	36/84 (43, 33–54)
	5–12	2/148 (2, 0–5)	5/169 (3, 0–5)	7/195 (4, 2–7)	29/202 (15, 10–20)	39/201 (20, 15–25)	42/202 (21, 16–27)	70/202 (35, 28–42)	120/200 (60, 53–67)	122/198 (62, 55–68)
	13–18	1/69 (3, 0–7)	5/72 (8, 0–7)	8/80 (11, 5–18)	20/82 (25, 16–35)	24/78 (31, 22–41)	24/75 (32, 23–43)	40/77 (52, 41–63)	59/73 (80, 71–88)	57/71 (80, 70–88)
	19–34	0/87 (1, 0–4)	5/88 (7, 0–4)	13/91 (15, 8–23)	33/91 (37, 27–47)	35/95 (37, 28–47)	34/89 (39, 29–49)	40/87 (46, 36–57)	48/81 (59, 48–70)	47/80 (59, 48–69)
	35–59	1/76 (2, 0–7)	5/77 (8, 0–7)	7/80 (10, 4–17)	23/81 (29, 20–39)	27/82 (33, 24–44)	27/84 (33, 23–43)	35/82 (43, 33–54)	49/80 (61, 50–72)	49/80 (61, 50–71)
	≥60	1/40 (5, 0–12)	4/39 (12, 0–12)	4/44 (11, 4–21)	8/45 (19, 9–32)	10/42 (25, 14–39)	10/41 (26, 14–39)	14/40 (36, 22–51)	16/37 (44, 28–60)	18/37 (49, 33–65)
	All	5/445 (1, 0–2)	25/494 (5, 0–2)	41/568 (7, 5–9)	127/588 (22, 18–25)	150/587 (26, 22–29)	151/579 (26, 23–30)	225/578 (39, 35–43)	325/548 (59, 55–64)	329/550 (60, 56–64)
Urban	<5	2/45 (6, 1–15)	5/50 (11, 4–21)	7/50 (15, 7–26)	8/44 (20, 9–32)	13/48 (28, 16–41)	14/45 (32, 19–46)	19/46 (42, 28–56)	25/45 (56, 41–70)	24/45 (53, 39–67)
	5–12	11/116 (10, 5–16)	17/116 (15, 9–22)	24/125 (20, 13–27)	32/123 (26, 19–34)	38/122 (32, 24–40)	52/125 (42, 33–50)	60/124 (49, 40–57)	75/124 (61, 52–69)	77/123 (63, 54–71)
	13–18	14/75 (19, 11–29)	19/75 (26, 17–36)	30/81 (37, 28–48)	35/77 (46, 35–57)	41/78 (53, 42–64)	44/77 (57, 46–68)	47/74 (63, 52–74)	59/74 (79, 69–88)	60/72 (83, 73–90)
	19–34	10/102 (10, 5–17)	17/98 (18, 11–26)	23/100 (23, 16–32)	29/99 (30, 21–39)	34/98 (35, 26–45)	44/94 (47, 37–57)	50/96 (52, 42–62)	65/95 (68, 59–77)	62/88 (70, 61–79)
	35–59	33/110 (30, 22–39)	45/110 (41, 32–50)	52/117 (45, 36–54)	61/112 (55, 45–64)	65/111 (59, 49–67)	70/111 (63, 54–72)	70/107 (65, 56–74)	82/104 (79, 70–86)	82/104 (79, 70–86)
	≥60	3/57 (7, 2–14)	6/55 (12, 5–22)	7/56 (14, 6–23)	16/55 (30, 19–42)	18/53 (35, 23–48)	23/53 (44, 31–57)	32/52 (61, 48–74)	33/51 (64, 51–77)	30/47 (64, 50–76)
	All	73/505 (14, 12–18)	109/504 (22, 18–25)	143/529 (27, 23–31)	181/510 (36, 31–40)	209/510 (41, 37–45)	247/505 (49, 45–53)	278/499 (56, 51–60)	339/493 (69, 65–73)	335/479 (70, 66–74)

*Adjusted for assay sensitivity and specificity. Blood draw (BD) 1, July 20–September 17, 2020; BD2, September 21–October 10, 2020; BD3, November 23–December 12, 2020; BD4, January 25–February 20, 2021; and BD5, March 22–April 11, 2021; BD6, May 20 – June 9, 2021; BD7, July 19 – August 5, 2021; BD8, September 13 – 25, 2021; BD9, November 15 – 27, 2021.

Appendix Table 2. Number and percentage of PHIRST-C participants with known vaccination status (N) vaccinated for SARS-CoV-2 by BD9 (November 15 – 27, 2021) in the rural and urban community, South Africa*

		Rural, No. (%)				Urban, No. (%)		
Age group, y	N	Not vaccinated	Partially vaccinated	Fully vaccinated	N	Not vaccinated	Partially vaccinated	Fully vaccinated
<5	93	93 (100)	0 (0)	0 (0)	49	49 (100)	0 (0)	0 (0)
5–12	205	205 (100)	0 (0)	0 (0)	125	125 (100)	0 (0)	0 (0)
13–18	83	82 (99)	1 (1)	0 (0)	78	72 (92)	3 (4)	3 (4)
19–34	101	91 (90)	2 (2)	8 (8)	96	72 (75)	8 (8)	16 (17)
35–59	85	51 (60)	9 (11)	25 (29)	111	59 (53)	5 (5)	47 (42)
≥60	42	16 (38)	10 (24)	16 (38)	53	16 (30)	6 (11)	31 (58)
All	609	538 (88)	22 (4)	49 (8)	512	393 (77)	22 (4)	97 (19)

*Partially vaccinated: one dose of Pfizer-BioNTech; Fully vaccinated: one dose of Johnson & Johnson or two doses of Pfizer-BioNTech; at time of blood collection

Appendix Table 3. Number and percentage of PHIRST-C participants with either serologic evidence of previous infection (at any draw 20 July, 2020 - November 27, 2021) or fully vaccinated by BD9 (November 15 – 27, 2021) or both in the rural and urban community, South Africa*

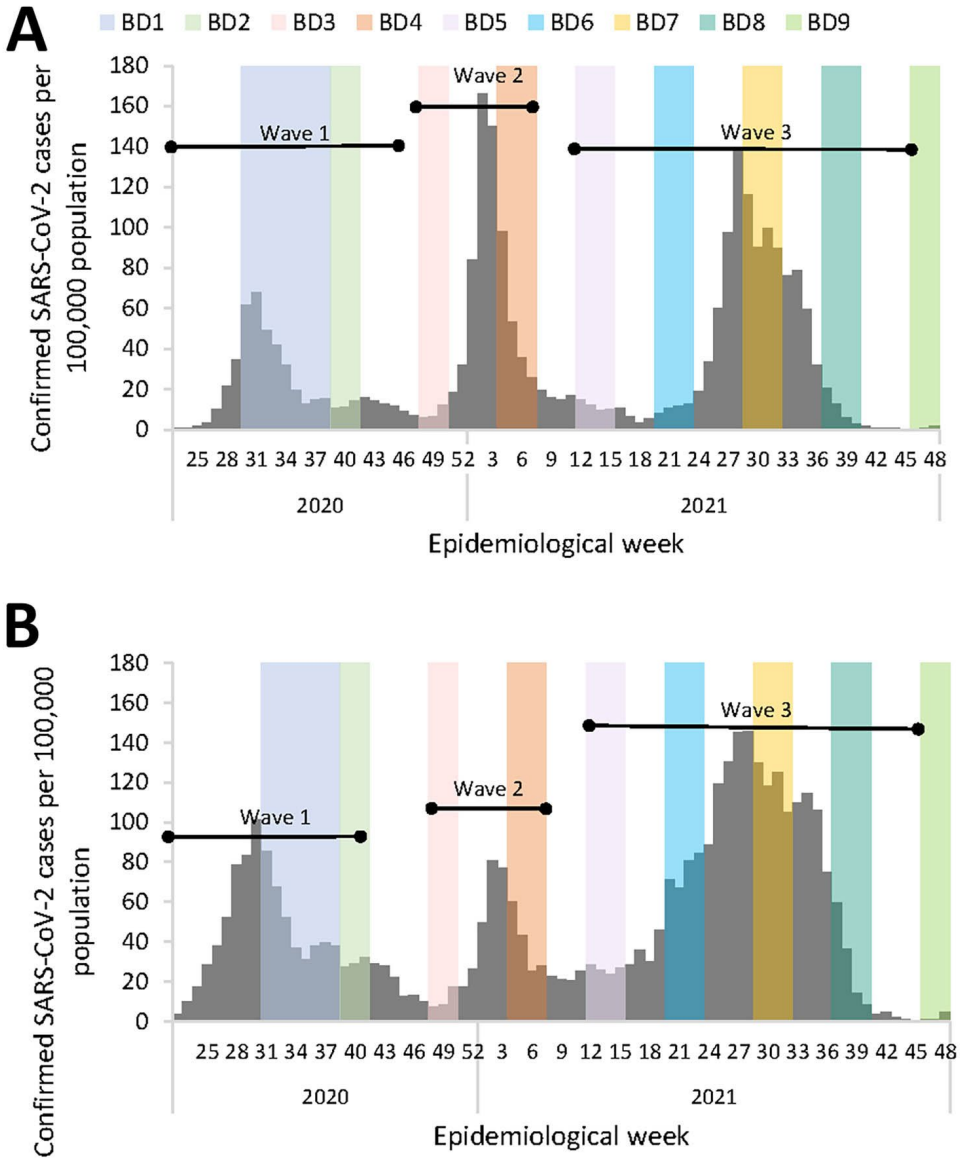
	n/N (%)	
Age group, y	Rural	Urban
<5	45/95 (47)	29/55 (53)
5–12	130/208 (63)	82/132 (62)
13–18	62/85 (73)	68/82 (83)
19–34	59/107 (55)	79/119 (66)
35–59	66/90 (73)	106/122 (87)
≥60	27/46 (59)	47/58 (81)
All	389/631 (62)	411/568 (72)

*Fully vaccinated: one dose of Johnson & Johnson or two doses of Pfizer-BioNTech; at time of blood collection

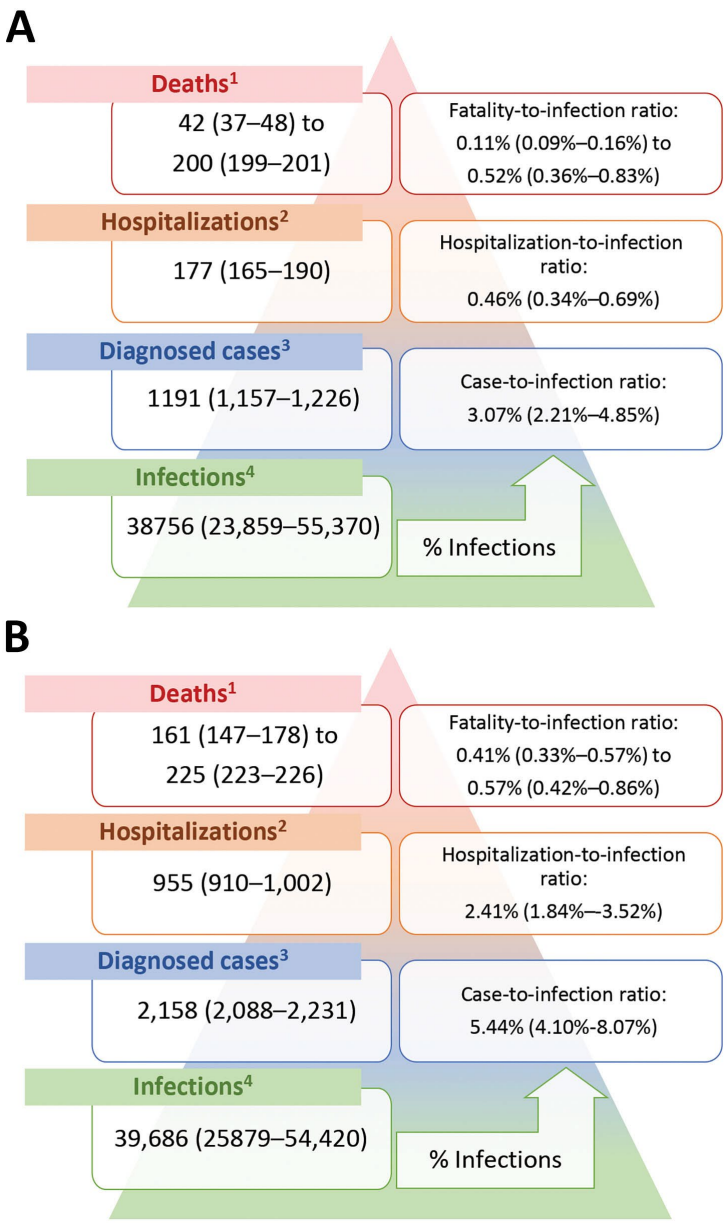
Appendix Table 4. The PHIRST-C Group

Name	Affiliation
Amelia Buys	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.
Anne von Gottberg	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Cheryl Cohen	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
F. Xavier Gómez-Olivé	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Floidy Wafawanaka	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Jackie Kleynhans	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Jacques D. du Toit	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Jinal N. Bhiman	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Jocelyn Moyes	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Kathleen Kahn	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.
Kgaugelo Patricia Kgasago	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.
Limakatso Lebina	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. Africa Health Research Institute, Durban, South Africa. Africa Health Research Institute, KwaZulu-Natal, South Africa.
Linda de Gouveia	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.

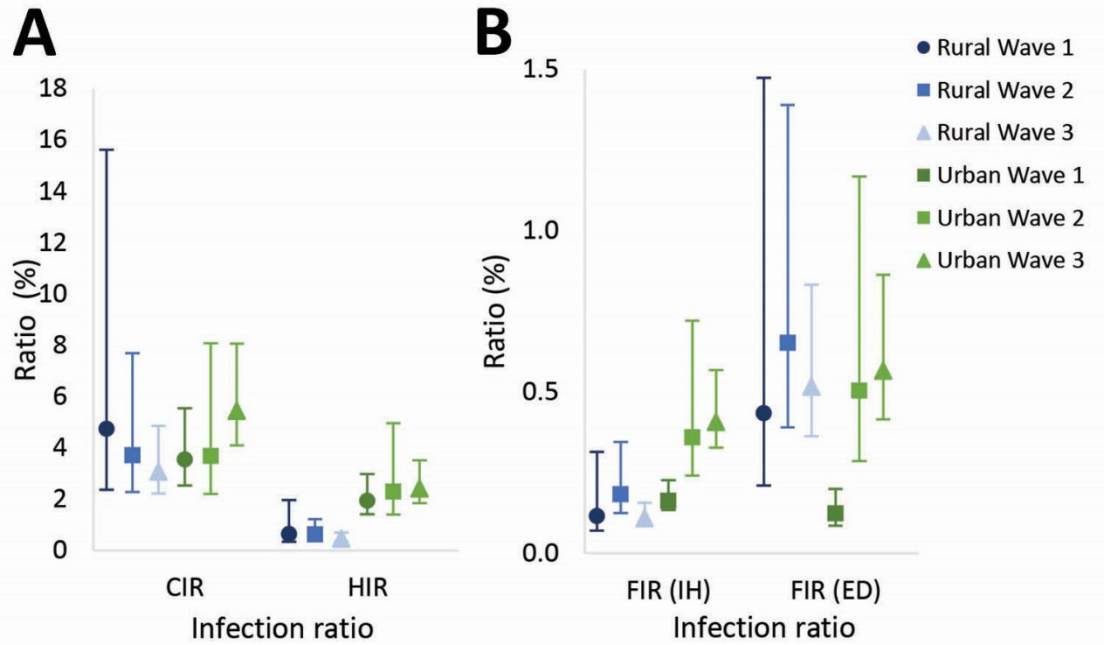
Name	Affiliation
Maimuna Carrim	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Meredith L. McMorrow	Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa.
Mignon du Plessis	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Neil A. Martinson	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. Johns Hopkins University Center for TB Research, Baltimore, Maryland, United States of America.
Nicole Wolter	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Retshidisitswe Kotane	Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Stefano Tempia	Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. MassGenics, Atlanta, Georgia, United States of America.
Stephen Tollman	MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Science, University of the Witwatersrand.
Tumelo Moloantoa	Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.



Appendix Figure 1. Timing of blood collection and district SARS-CoV-2 weekly incidence in a) rural community and b) urban community, March 2020 – November 2021, South Africa. November 23 – December 12, 2020; fourth draw (BD4) January, 25 – February 20, 2021; fifth draw (BD5) March 22 – April 11, 2021; sixth draw (BD6) May 20 – June 9, 2021; seventh draw (BD7) July 19 – August 5, 2021; eighth draw (BD8) September 13 – 25, 2021; ninth draw (BD9) November 15 – 27, 2021. Vertical lines represent 95% credible interval. Wave 1 (March 1 – November 21, 2020), 2 (November 22, 2020 – March 21, 2021) and 3 (March 22 – November 14, 2021) lines indicate period used for analysis. Laboratory-confirmed SARS-CoV-2 infections (reverse transcription polymerase chain reaction and antigen tests) as reported to the Notifiable Medical Conditions Surveillance System (NMCSS).



Appendix Figure 2. South African age-standardized SARS-CoV-2 infection, diagnosis, hospitalization and deaths per 100,000 population during wave 3 in the a) rural and b) urban community, March – November, 2021, South Africa. Standardized to South Africa mid-year population estimate for 2021. Wave 3: March 22 – November 14, 2021. [1]Minimum estimate: in-hospitalization deaths in districts based on COVID-19 Sentinel Hospital Surveillance report (DATCOV), maximum estimate: provincial excess deaths reported (South African Medical Research Council). [2]Hospitalizations in the district based on COVID-19 Sentinel Hospital Surveillance (DATCOV). [3]DiagnosedSARS-CoV-2 cases reported from districts from national SARS-CoV-2 surveillance (NMCSS). [1]Based on becoming seropositive or 2-fold increase in cutoff index (COI) between draw 5 and 9. [1]Values in bracket refer to 95% credible interval for infections and 95% confidence interval for all other estimates.



Appendix Figure 3. SARS-CoV-2 a) case-to-infection (CIR), hospitalization-to-infection (HIR) and b) in-hospital (IH) fatality-to-infection (FIR) and excess (ED) fatality-to-infection (FIR) ratios in a rural and urban community during the first, second and third wave of infections, March 2020 - November 2021, South Africa.

Clinical Infectious Diseases

MAJOR ARTICLE

Household Transmission of Severe Acute Respiratory Syndrome Coronavirus 2 From Adult Index Cases With and Without Human Immunodeficiency Virus in South Africa, 2020–2021: A Case-Ascertained, Prospective, Observational Household Transmission Study

Jackie Kleynhans,[1,2] Sibongile Walaza,[1,2] Neil A. Martinson,[3,4] Mzimasi Neti,[1] Anne von Gottberg,[1,5] Jinal N. Bhiman,[1,5] Dylan Toi,[1] Daniel G. Amoako,[1,6] Amelia Buys,[1] Kedibone Ndlangisa,[1] Nicole Wolter,[1,5] Leisha Genade,[3] Lucia Maloma,[3] Juanita Chewparsad,[3] Limakatso Lebina,[3,7] Linda de Gouveia,[1] Retshidisitswe Kotane,[1] Stefano Tempia,[1,2] and Cheryl Cohen[1,2]

[1]Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa; [2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [3]Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa; [4]Center for TB Research, Johns Hopkins University, Baltimore, Maryland, USA; [5]School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [6]School of Health Sciences, College of Health Sciences, University of KwaZulu-Natal, KwaZulu-Natal, South Africa; and [7]Africa Health Research Institute, Durban, South Africa

Background. In South Africa, 19% of adults are living with human immunodeficiency virus (HIV; LWH). Few data on the influence of HIV on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) household transmission are available.

Methods. We performed a case-ascertained, prospective household transmission study of symptomatic adult index SARS-CoV-2 cases LWH and not living with HIV (NLWH) and their contacts from October 2020 to September 2021. Households were followed up 3 times a week for 6 weeks to collect nasal swabs for SARS-CoV-2 testing. We estimated household cumulative infection risk (HCIR) and duration of SARS-CoV-2 positivity (at a cycle threshold value <30 as proxy for high viral load).

Results. HCIR was 59% (220 of 373), not differing by index HIV status (60% LWH vs 58% NLWH). HCIR increased with index case age (35–59 years: adjusted OR [aOR], 3.4; 95% CI, 1.5–7.8 and ≥60 years: aOR, 3.1; 95% CI, 1.0–10.1) compared with 18–34 years and with contacts' age, 13–17 years (aOR, 7.1; 95% CI, 1.5–33.9) and 18–34 years (aOR, 4.4; 95% CI, 1.0–18.4) compared with <5 years. Mean positivity was longer in cases LWH (adjusted hazard ratio, 0.4; 95% CI, .1–.9).

Conclusions. Index HIV status was not associated with higher HCIR, but cases LWH had longer positivity duration. Adults aged >35 years were more likely to transmit and individuals aged 13–34 to be infected SARS-CoV-2 in the household. As HIV infection may increase transmission, health services must maintain HIV testing and antiretroviral therapy initiation.

Keywords. SARS-CoV-2; COVID-19; HIV; transmission; acquisition.

By January 2022, South Africa had reported 3.6 million coronavirus disease 2019 (COVID-19) cases and 94.3 thousand deaths; the highest reported from Africa [1, 2]. South Africa experienced 4 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) waves; the first dominated by ancestral variants, followed by Beta (B.1.351), Delta (B.1.617.2), and Omicron (B.1.1.529) variant waves [3].

Although incident human immunodeficiency virus (HIV) infections and acquired immune deficiency syndrome (AIDS)-related deaths from 2010 to 2019 in South Africa declined by 53% and 61%, respectively, the burden of HIV is still high, with an estimated 19% of the adult population aged 15–49 years living with HIV (LWH); the fourth highest in sub-Saharan Africa [4]. The SARS-CoV-2 pandemic impacted several health programs, including HIV testing and care. During initial lockdowns, there was a decline in HIV testing and antiretroviral therapy initiations, which gradually returned to pre-lockdown levels in South Africa [5] and other sub-Saharan African countries [6].

Few studies have reported on the influence of HIV infection on SARS-CoV-2, with most data available from high-income countries and little evidence from sub-Saharan Africa where most people LWH reside [4]. People LWH are at greater risk for hospitalization [7–10] and death [9–13] when infected with SARS-CoV-2, but SARS-CoV-2 prevalence is similar between people LWH and people not living with HIV (NLWH) [14].

Received 11 April 2022; editorial decision 28 July 2022; published online 5 September 2022

Correspondence: J. Kleynhans, Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, 1 Modderfontein Road, Sandringham, 2192, Johannesburg, South Africa (jackiel@nicd.ac.za).

Clinical Infectious Diseases® 2023;76(3):e71–e81

© The Author(s) 2022. Published by Oxford University Press on behalf of Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1093/cid/ciac640

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

Risk for hospitalization and death increases with a decline in CD4 + T cells [8, 9, 13]. Limited data are available on the role of HIV in the transmission of SARS-CoV-2. One study showed no increase in household transmission from or acquisition of SARS-CoV-2 infection in people LWH [2]. People LWH with severe COVID-19 who are not virally suppressed shed SARS-CoV-2 for longer periods [2, 15], which could lead to increased secondary transmission.

We assessed household cumulative infection risk (HCIR), duration of SARS-CoV-2 positivity (episode duration), and serial interval in households with SARS-CoV-2 index cases LWH and NLWH from October 2020 to September 2021 during the Beta and Delta waves.

METHODS

We conducted a case-ascertained, prospective, observational, household transmission study of household contacts of symptomatic adult index SARS-CoV-2 cases LWH and NLWH at 2 sites in South Africa: Klerksdorp (North West Province) and Soweto (Gauteng Province). Planned sample size was 264 and 176 contacts from households with an index case NLWH and LWH, respectively (see the Supplementary Materials). Actual sample size was 344 and 103 household members exposed to an index case NLWH and LWH, respectively.

Screening for Index Cases

Screening procedures are detailed in the Supplementary Methods. In short, nasopharyngeal swabs were collected from clinic attendees aged ≥18 years with symptom onset ≤5 days prior to screening and tested for SARS-CoV-2 using real-time reverse-transcription polymerase chain reaction (rRT-PCR).

Household Enrollment

We approached households of individuals who tested positive for SARS-CoV-2 with symptom onset <7 days prior and no household members reporting symptoms in the 14 days prior to index screening. We enrolled households with ≥3 eligible members (sharing ≥2 meals in the same residence for ≥2 days/week) and where ≥70% of housheold members consented to participate. Households that withdrew within 10 days from index symptom start date were excluded from the analysis.

Index and Household Follow-up

We visited households 3 times a week for 6 weeks to collect nasal swabs and data on symptoms and healthcare-seeking behavior from consenting household members. At the first and last study visits, clotted blood was collected for serological testing. Follow-up started on 12 October 2020 and continued to 11 August 2021 and 28 September 2021 in Klerksdorp and Soweto, respectively.

SARS-CoV-2 Detection

Nasopharyngeal (screening) and nasal (follow-up) specimens were tested for SARS-CoV-2 genes using qualitative rRT-PCR with the Allplex 2019-nCoV kit (Seegene Inc, Seoul, South Korea). Specimens were considered positive for SARS-CoV-2 if the cycle threshold (Ct) value was <40 for any gene target.

SARS-CoV-2 Variants

We characterized the first SARS-CoV-2–positive specimen for each participant using the Allplex SARS-CoV-2 Variants I and II PCR assays (Seegene Inc, Seoul, Korea) and through full genome sequencing on the Ion Torrent Genexus platform (Thermo Fisher Scientific). We classified the infection episodes as Alpha, Beta, Delta, non-Alpha/Beta/Delta, or unknown variant (see the Supplementary Material).

Serology

We used an in-house enzyme-linked immunosorbent assay to detect antibodies against the SARS-CoV-2 spike protein [16] and nucleocapsid protein using the Roche Elecsys anti-SARS-CoV-2 assay. Individuals were considered seropositive if they tested positive on either.

Statistical Analyses

Definitions of terms used for this study are listed in Table 1. To assess factors associated with HCIR, we used logistic regression accounting for within-site and household clustering using a mixed-effects hierarchical regression model. To assess factors associated with a time-to-event analysis (serial interval and episode duration), we used a multilevel mixed-effects survival model with Weibull accelerated failure time analysis. Hazard ratios (HRs) <1 correspond to longer episode duration than observed in the reference group. Since multiple members from the same household could potentially be included in the serial interval analysis, we controlled for both site- and household-level clustering in the analysis. In the episode duration analysis, we controlled for only site clustering (1 index per household). In addition to using site to control for clustering, it was also included as a covariate in models. Episode duration was assessed at any Ct value (<40) and Ct <30 (proxy for high viral load based on virus culture studies [17]). We first assessed covariates on univariate analysis, including all with $P < .2$ in the multivariable analysis. We performed backward elimination and kept all variables with $P < .05$ in the final model, except those included a priori. We included site, SARS-CoV-2 variant, and index immune suppression related to HIV status (defined as CD4+ T-cell count <200 cells/mL) in the HCIR and episode duration models a priori irrespective of statistical significance in the multivariable model. Site was included a priori in the serial interval analysis. Due to low SARS-CoV-2 vaccination coverage in study participants (only 1 contact, Table 2), vaccination status was not included in our analyses.

Sensitivity Analysis

To assess the influence of loss to follow-up, we performed a sensitivity analysis that included only households where 65%

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

Table 1. Definitions of Terms Used for the Study

Household	A group of 3 or more people who regularly share at least 2 meals in the same residence at least 2 days per week (residential institutions excluded).
Index case	The first household member who had coronavirus disease 2019–like symptoms. We assumed that the household member screened was the index case within the household as they were the first household member to develop symptoms.
SARS-CoV-2 infection episode	At least 1 nasal swab rRT-PCR–positive for SARS-CoV-2. Individuals who seroconverted during follow-up but with no rRT-PCR–confirmed infection were not included in secondary case analyses.
SARS-CoV-2 cluster	Composed of all infections within a household within an interval between infections of ≤2 weeks including single infections within a household.
Episode duration	Duration of SARS-CoV-2 positivity. The start of symptom onset to the midpoint between the last positive swab and first negative swab. Individuals who were still SARS-CoV-2 positive on the last study visit (whether at the end of follow-up or due to early withdrawal) were right-censored for the multivariable analysis.
Serial interval	Number of days between the onset of symptoms in the index case and the onset of symptoms in the secondary case. Multivariable analyses were restricted to symptomatic secondary cases and to serial interval periods of ≤21 days as longer serial intervals could have been due to tertiary cases or secondary infections.
HCIR	The percentage of susceptible household members (based on baseline serology) who had at least 1 SARS-CoV-2–positive swab from the start of follow-up up to 2 weeks from the last SARS-CoV-2–positive swab of the index case. Considering susceptibility was important because following the second wave of SARS-CoV-2 infection in South Africa, 41% of individuals were estimated to have had previous SARS-CoV-2 infection [18]. Individuals with SARS-CoV-2 antibodies detected at baseline, but also tested positive on rRT-PCR, were included in the HCIR calculation. Individuals for whom no baseline serology was available were included in the analysis as presumed susceptible. We did not consider any secondary introductions in the household for our analysis. Households where members had SARS-CoV-2 infection with different variants of concern were excluded from the analysis.

Abbreviations: HCIR, household cumulative infection risk; rRT-PCR, real-time reverse-transcription polymerase chain reaction; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

of enrolled household members completed 65% of follow-up visits in the first 3 weeks of follow-up. To explore the effect of previous SARS-CoV-2 infection, we considered all household members irrespective of baseline serology as susceptible contacts in the HCIR analysis.

Ethics

The University of the Witwatersrand Human Research Ethics Committee approved the study protocol. Participants in follow-up received a $3.00 grocery store voucher per visit to compensate for time required for specimen collection and interview.

RESULTS

Screening, Enrollment, and Follow-up

From 2 October 2020 to 30 September 2021, we screened 1531 clinic attendees for SARS-CoV-2; 18% (277) tested positive on rRT-PCR. Of those who tested positive and met eligibility criteria for household enrollment (n = 277), 143 (52%) households were approached and 131 (92%) were enrolled. Reasons for noninclusion are shown in Figure 1. The final cohort consisted of 131 index cases and 457 household contacts (Figure 1); the median household size was 4.

Twenty-one percent (28 of 131) of index cases were LWH, and 2 index cases initially agreed but then refused HIV testing after enrollment (classified as HIV unknown during analyses). The majority (93 of 131, 71%) of index cases and contacts (265 of 457, 58%) were female (Table 2).

Of 10 584 potential study visits to individual participants, we completed 8509 (80%) visits and detected SARS-CoV-2 in 17% (1454 of 8352) of nasal swabs collected (Supplementary Figure 1).

Secondary SARS-CoV-2 Cases

We diagnosed 232 (51%) rRT-PCR–confirmed secondary cases from 457 contacts linked to 131 index cases. One-third (69 of 232) of secondary cases reported ≥1 symptom during their SARS-CoV-2 episode, reporting on average 3 symptoms (range, 1–9). The mean symptom duration was 11 days (range, 4–40). The most common symptoms reported were cough (45 of 69, 65%), headache (31 of 67, 46%), and fever (27 of 69, 39%). Five secondary cases were hospitalized (Supplementary Table 5).

Household Cumulative Infection Risk

Of 131 households, we excluded 7 (5%) from the HCIR analysis: 4 (13 contacts) had SARS-CoV-2 clusters with >1 variant detected (Supplementary Figure 2); in 3 (8 contacts), all contacts were seropositive at baseline with no rRT-PCR–confirmed SARS-CoV-2 infection during follow-up. An additional 42 contacts were excluded because they had prevalent SARS-CoV-2 antibodies at baseline and no SARS-CoV-2 detection during follow-up. We therefore included 124 of 131 (95%) index cases with 373 of 436 (86%) contacts for this analysis.

The HCIR was 59% (220 of 373) overall. The mean number of household contacts who tested positive for SARS-CoV-2 following the index episode was 2 (range, 0 to 7). On univariate analysis, HCIR was similar in households with index cases NLWH (58%, 173 of 293) and households where the index was LWH (60%; 50 of 83; odds ratio [OR], 1.0; 95% confidence interval [CI], .4–2.3).

On multivariable analysis after adjusting for site and immune suppression, factors associated with household transmission were index case aged 35–59 years (adjusted OR [aOR], 3.4;

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

Table 2. Baseline Characteristics of Severe Acute Respiratory Syndrome Coronavirus 2 Index Cases (n = 131) and Their Household Contacts (n = 457), Klerksdorp and Soweto, South Africa, September 2020–October 2021

	Overall		Index Case, n/N (%)		Household Contact, n/N (%)	
Characteristic	Index Case, n/N (%)	Household Contact, n/N (%)	Klerksdorp	Soweto	Klerksdorp	Soweto
Household characteristics						
Index case/contact	131/588 (22)	457/588 (78)	62/274 (23)	69/314 (22)	212/274 (77)	245/314 (78)
Household size						
3–5	106/131 (81)	305/457 (67)	51/62 (82)	55/69 (80)	146/212 (69)	159/245 (65)
6–10	25/131 (19)	152/457 (33)	11/62 (18)	14/69 (20)	66/212 (31)	86/245 (35)
Rooms used for sleeping						
1–2	66/117 (56)	209/401 (52)	33/57 (58)	33/60 (55)	114/192 (59)	95/209 (45)
3–4	42/117 (36)	141/401 (35)	22/57 (39)	20/60 (33)	69/192 (36)	72/209 (34)
>4	9/117 (8)	51/401 (13)	2/57 (4)	7/60 (12)	9/192 (5)	42/209 (20)
Crowding	89/131 (68)	330/457 (72)	43/62 (69)	46/69 (67)	158/212 (75)	172/245 (70)
Child aged <5 years	17/131 (13)	63/457 (14)	10/62 (16)	7/69 (10)	43/212 (20)	20/245 (8)
Household member smokes inside	30/131 (23)	93/457 (20)	18/62 (29)	12/69 (17)	51/212 (24)	42/245 (17)
Main water source inside home	92/131 (70)	321/457 (70)	38/62 (61)	54/69 (78)	124/212 (58)	197/245 (80)
Place to wash hands inside	129/131 (98)	451/457 (99)	62/62 (100)	67/69 (97)	212/212 (100)	239/245 (98)
Main cooking fuel						
Electricity	127/131 (97)	444/457 (97)	58/62 (94)	69/69 (100)	199/212 (94)	245/245 (100)
Gas/Paraffin	4/131 (3)	13/457 (3)	4/62 (6)	0/69 (0)	13/212 (6)	0/245 (0)
Monthly household income						
<US$23	5/131 (4)	17/457 (4)	4/62 (6)	1/69 (1)	15/212 (7)	2/245 (1)
US$24 to US$46	6/131 (5)	19/457 (4)	4/62 (6)	2/69 (3)	12/212 (6)	7/245 (3)
US$47 to US$93	12/131 (9)	39/457 (9)	10/62 (16)	2/69 (3)	33/212 (16)	6/245 (2)
US$94 to US$187	23/131 (18)	86/457 (19)	12/62 (19)	11/69 (16)	43/212 (20)	43/245 (18)
$US188 to US$375	21/131 (16)	68/457 (15)	7/62 (11)	14/69 (20)	23/212 (11)	45/245 (18)
US$376 to US$752	8/131 (6)	33/457 (7)	2/62 (3)	6/69 (9)	5/212 (2)	28/245 (11)
US$753 to US$1506	6/131 (5)	19/457 (4)	0/62 (0)	6/69 (9)	0/212 (0)	19/245 (8)
Refused to disclose	50/131 (38)	176/457 (39)	23/62 (37)	27/69 (39)	81/212 (38)	95/245 (39)
Individual characteristics						
Age, years						
<5	…	19/457 (4)	…	…	11/212 (5)	8/245 (3)
5–12	…	80/457 (18)	…	…	39/212 (18)	41/245 (17)
13–17	…	70/457 (15)	…	…	38/212 (18)	32/245 (13)
18–34	37/131 (28)	126/457 (28)	21/62 (34)	16/69 (23)	54/212 (25)	72/245 (29)
35–59	76/131 (58)	113/457 (25)	37/62 (60)	39/69 (57)	50/212 (24)	63/245 (26)
≥60	18/131 (14)	49/457 (11)	4/62 (6)	14/69 (20)	20/212 (9)	29/245 (12)
Sex						
Male	38/131 (29)	192/457 (42)	17/62 (27)	21/69 (30)	93/212 (44)	99/245 (40)
Female	93/131 (71)	265/457 (58)	45/62 (73)	48/69 (70)	119/212 (56)	146/245 (60)
Level of education[a]						
No schooling	6/129 (5)	15/283 (5)	3/60 (5)	3/69 (4)	1/120 (1)	14/163 (9)
Primary	4/129 (3)	21/283 (7)	2/60 (3)	2/69 (3)	14/120 (12)	7/163 (4)
Secondary	43/129 (33)	89/283 (31)	25/60 (42)	18/69 (26)	47/120 (39)	42/163 (26)
Matriculation	69/129 (53)	136/283 (48)	28/60 (47)	41/69 (59)	51/120 (43)	85/163 (52)
Post-secondary	7/129 (5)	22/283 (8)	2/60 (3)	5/69 (7)	7/120 (6)	15/163 (9)
Employment[a]						
Unemployed	55/121 (45)	159/258 (62)	25/57 (44)	30/64 (47)	68/111 (61)	91/147 (62)
Student	10/121 (8)	30/258 (12)	5/57 (9)	5/64 (8)	11/111 (10)	19/147 (13)
Employed	56/121 (46)	69/258 (27)	27/57 (47)	29/64 (45)	32/111 (29)	37/147 (25)
Smoking cigarettes[b]	17/129 (13)	66/323 (20)	11/60 (18)	6/69 (9)	35/145 (24)	31/178 (17)
Smoke indoors	5/17 (29)	15/66 (23)	5/11 (45)	0/6 (0)	14/35 (40)	1/31 (3)
HIV status						
Not living with HIV	101/131 (77)	176/457 (39)	51/62 (82)	50/69 (72)	130/212 (61)	46/245 (19)
Living with HIV	28/131 (21)	35/457 (8)	10/62 (16)	18/69 (26)	16/212 (8)	19/245 (8)
Unknown	2/131 (2)	246/457 (54)	1/62 (2)	1/69 (1)	66/212 (31)	180/245 (73)

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

Table 2. Continued

Characteristic	Overall		Index Case, n/N (%)		Household Contact, n/N (%)	
	Index Case, n/N (%)	Household Contact, n/N (%)	Klerksdorp	Soweto	Klerksdorp	Soweto
Index case living with HIV in household						
Index HIV-negative	101/131 (77)	344/457 (75)	51/62 (82)	50/69 (72)	175/212 (83)	169/245 (69)
Index HIV- positive	28/131 (21)	103/457 (23)	10/62 (16)	18/69 (26)	34/212 (16)	69/245 (28)
Index HIV unknown	2/131 (2)	10/457 (2)	1/62 (2)	1/69 (1)	3/212 (1)	7/245 (3)
CD4+ T-cell count						
Not living with HIV, not immune suppressed[c]	17/28 (61)	13/35 (37)	5/10 (50)	12/18 (67)	2/16 (13)	11/19 (58)
Living with HIV, immune suppressed[c]	4/28 (14)	2/35 (6)	1/10 (10)	3/18 (17)	1/16 (6)	1/19 (5)
Living with HIV, no CD4+ T-cell count	7/28 (25)	20/35 (57)	4/10 (40)	3/18 (17)	13/16 (81)	7/19 (37)
Underlying illness[d]	32/129 (25)	57/448 (13)	12/60 (20)	20/69 (29)	205/212 (97)	243/245 (99)
Body mass index						
Underweight	3/129 (2)	29/448 (6)	1/60 (2)	2/69 (3)	14/205 (7)	15/243 (6)
Normal	36/129 (28)	206/448 (46)	18/60 (30)	18/69 (26)	108/205 (53)	98/243 (40)
Overweight	25/129 (19)	103/448 (23)	10/60 (17)	15/69 (22)	39/205 (19)	64/243 (26)
Obese	65/129 (50)	110/448 (25)	31/60 (52)	34/69 (49)	44/205 (21)	66/243 (27)
Previous TB	7/129 (5)	9/448 (2)	2/60 (3)	5/69 (7)	6/205 (3)	3/243 (1)
Current TB	4/129 (3)	2/448 (0)	2/60 (3)	2/69 (3)	0/205 (0)	2/243 (1)
Coronavirus 2019 vaccination[e]	0/131 (0)	1/457 (0)	0/62 (0)	0/69 (0)	1/212 (0)	0/245 (0)

Abbreviations: HIV, human immunodeficiency virus; TB, tuberculosis.

[a]Individuals aged ≥18 years.

[b]Individuals aged ≥15 years.

[c]Immune suppressed defined as CD4+ T-cell count <200 cells/mL.

[d]Underlying medical conditions include self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, prematurity.

[e]At least 1 dose administered 14 days prior to enrollment.

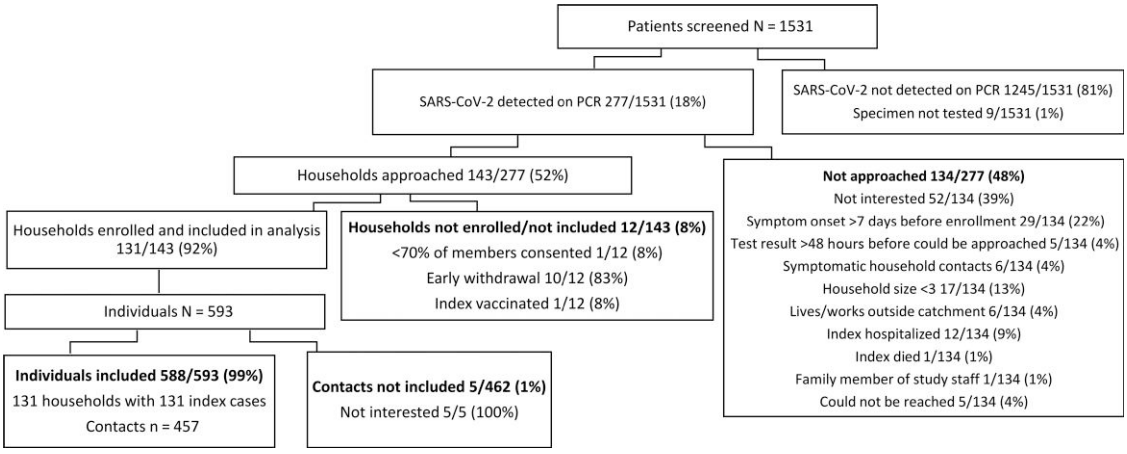


Figure 1. Adult SARS-CoV-2 index cases and household contacts enrolled, Klerksdorp and Soweto, South Africa, 2020–2021. Abbreviations: PCR, polymerase chain reaction; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

95% CI, 1.5–7.8) and ≥60 years (aOR, 3.1; 95% CI, 1.0–10.1) compared with 18–34 years; index cases with a Ct value <25 (aOR, 5.3; 95% CI, 1.6–17.6) and 25–35 (aOR, 7.5; 95% CI, 2.2–26.0) compared with Ct >35; and infection with the Delta variant compared with Beta (aOR, 4.6; 95% CI, 1.5–14.4; Figure 2, Supplementary Table 1).

Fourteen percent (17 of 124) of index cases were LWH and not immune suppressed, while 3% (4 of 124) were LWH and immune suppressed. Contacts of index cases LWH with immune suppression had higher HCIR (62%, 8 of 13) compared with contacts of index cases who were not immune suppressed (58%, 30 of 52), but this was not statistically significant on

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

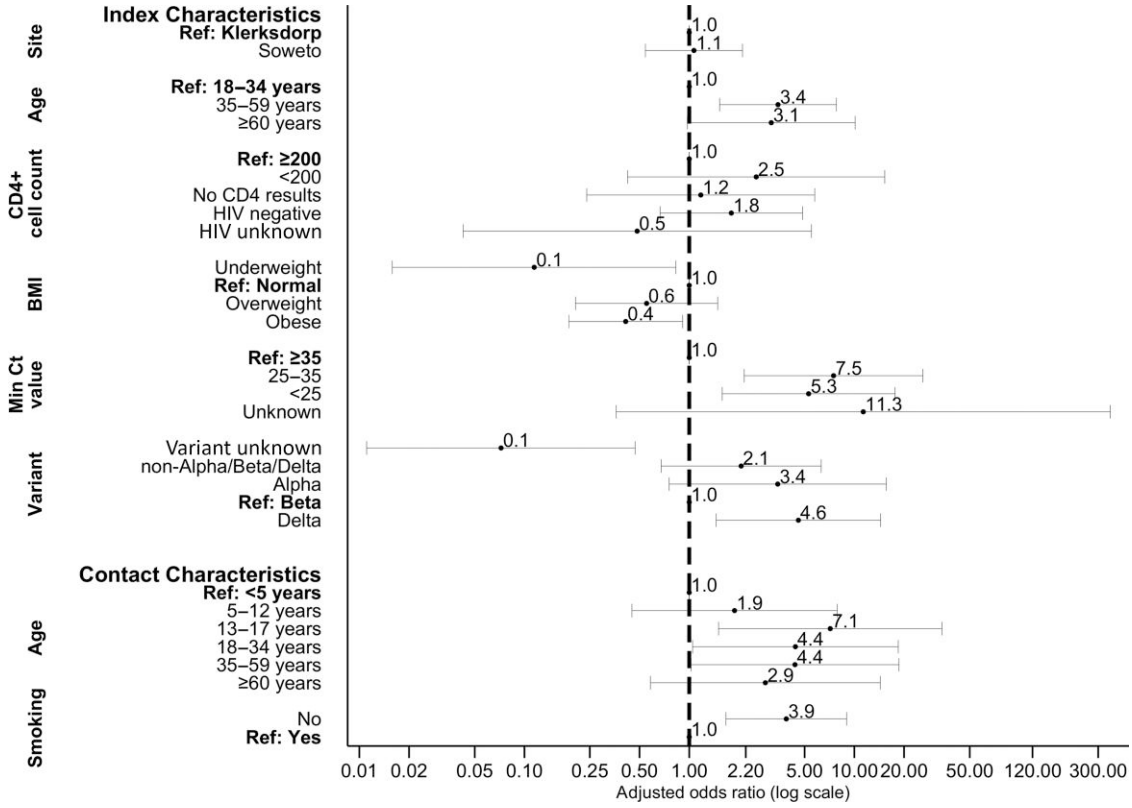


Figure 2. Factors associated with severe acute respiratory syndrome coronavirus 2 household transmission in index cases and acquisition in household contacts on multivariable logistic regression, Klerksdorp and Soweto, South Africa, 2020–2021 (n = 373). Abbreviation: HIV, human immunodeficiency virus.

multivariable analysis (aOR, 2.5; 95% CI, .4–15.3). Contact age 13–17 years (aOR, 7.1; 95% CI, 1.5–33.9) and 18–34 years (aOR, 4.4; 95% CI, 1.0–18.4) compared with <5 years and contacts not currently smoking (aOR, 3.9; 95% CI, 1.7–9.0) were associated with higher HCIR.

Episode Duration

We right-censored 6% (8 of 131) of index cases who were SARS-CoV-2–positive on their last specimen collected at the end of follow-up (n = 5) or at withdrawal from the study (n = 3). When all 131 index cases were included, the mean episode duration for index cases was 20 days (range, 3–47; Figure 3). When we excluded the 8 right-censored individuals (n = 123), the mean episode duration for index cases was 19 days (range, 3 to 45). The mean episode duration was similar for index cases NLWH (20 days; range, 3–45) compared with index cases LWH (17 days; range, 3–45; hazard ratio [HR], 0.8; 95% CI, .5–1.2).

On multivariable analysis, factors associated with longer episode duration at any Ct value, in days, were Soweto site (adjusted HR [aHR], 0.5; 95% CI, .3–.7); being aged 35–59 years (aHR, 0.4; 95% CI, .2–.6) and being aged ≥60 years (aHR, 0.2; 95% CI, .1–.5) compared with aged 18–34 years; Ct <25 (aHR, 0.6; 95% CI, .2–.9) compared with Ct >35 (Figure 4, Supplementary Table 2); and being seropositive at the end of follow-up (aHR, 0.1; 95% CI, .0–.3). Individuals infected with the Delta variant (aHR, 2.6; 95% CI, 1.4–5.0) and a nonvariant of concern (aHR, 4.0; 95% CI, 1.9–8.2) compared with Beta had shorter episode durations (Figure 4, Supplementary Table 2).

Eighty-eight (67%) index cases had ≥1 specimen with ≥1 target with Ct <30. Mean episode duration with Ct <30 was 7 days (range, 2–17). On multivariable analysis, factors associated with longer episode duration considering only specimens with at least 1 target with a Ct <30, was female sex (aHR, 0.5; 95% CI, .3–.9), LWH (aHR, 0.4; 95% CI, .1–.9), and being seropositive at the end of follow-up (aHR, 0.01; 95% CI, .001–.2; Figure 4, Supplementary Table 3).

Serial Interval

We excluded 5 index contact pairs where serial interval was >21 days; 3 with an index NLWH and 2 pairs where the index was LWH. Mean serial interval for index cases and symptomatic contact pairs included in the risk factor analysis was 6 days (range, 1–20; Figure 5). Mean serial interval for index cases NLWH was 6 days (range, 1–15); for index cases LWH it was 8 days (range, 2–20).

On multivariable analysis, pairs with contacts aged 35–59 years (aHR, 0.3; 95% CI, .1–.9) and ≥60 years (aHR, 0.2; 95%

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

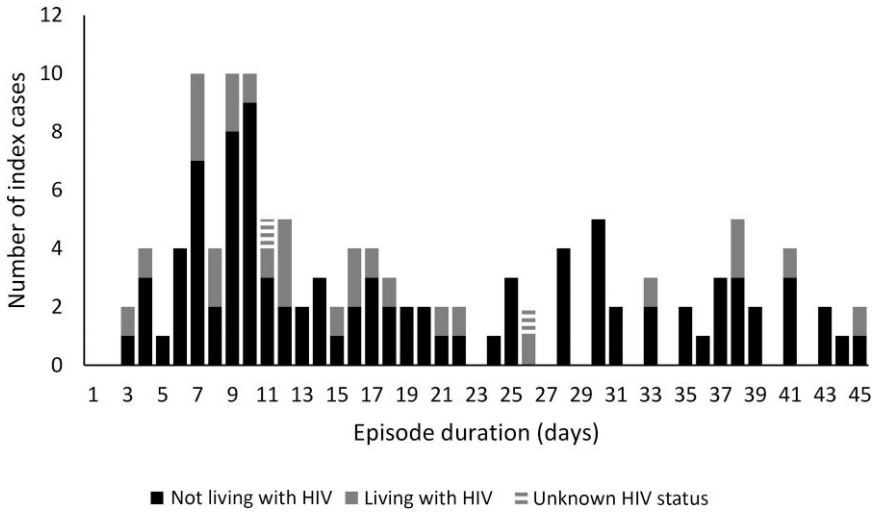


Figure 3. Severe acute respiratory syndrome coronavirus 2 episode duration in index cases living with and not living with HIV, Klerksdorp and Soweto, South Africa, 2020–2021 (n = 123). Excludes those positive at last specimen collected. Abbreviation: HIV, human immunodeficiency virus.

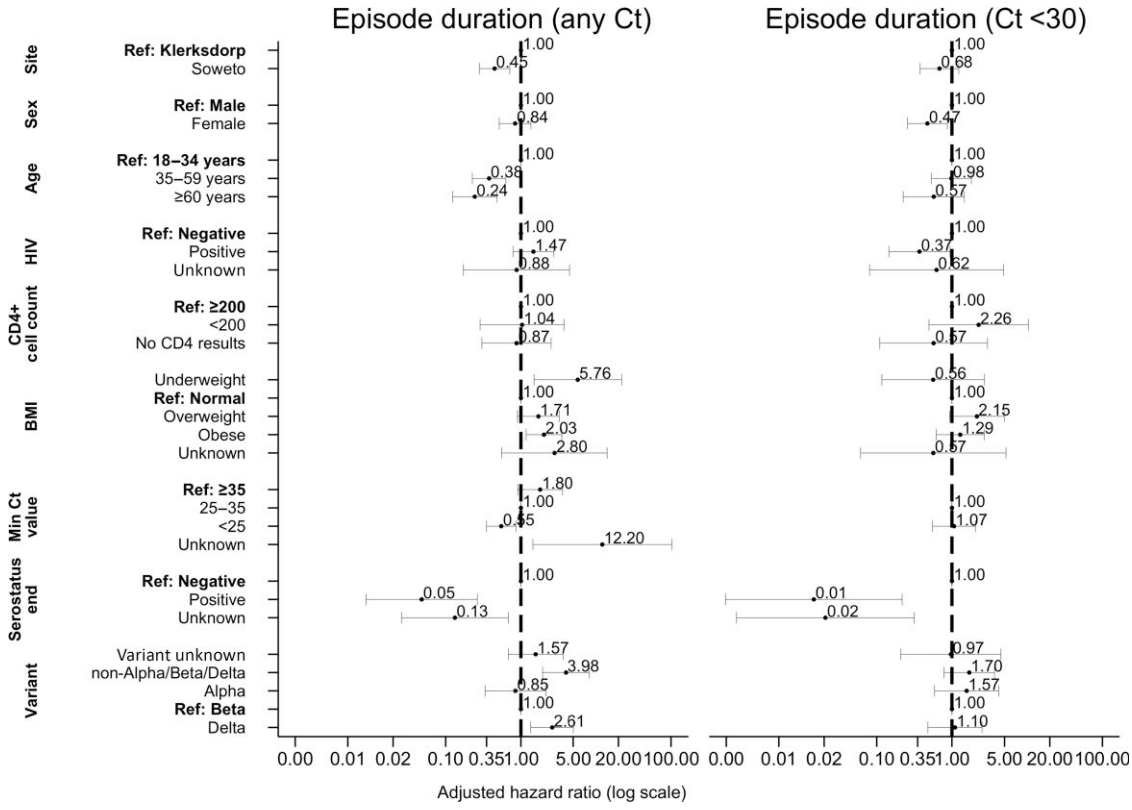


Figure 4. Factors associated with severe acute respiratory syndrome coronavirus 2 episode duration irrespective of Ct value (left), Ct <30 (right) in index cases on Weibull accelerated failure time regression, Klerksdorp and Soweto, South Africa, 2020–2021 (n = 123). Abbreviations: BMI, body mass index; Ct, cycle threshold; HIV, human immunodeficiency virus.

CI, .0–.8) compared with aged 18–34 years and where the contact was LWH (aHR, 0.1; 95% CI, .0–.8) had longer serial intervals (Figure 6, Supplementary Table 4).

Sensitivity Analysis

When individuals seropositive at baseline with no rRT-PCR–confirmed SARS-CoV-2 infection during follow-up were not

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

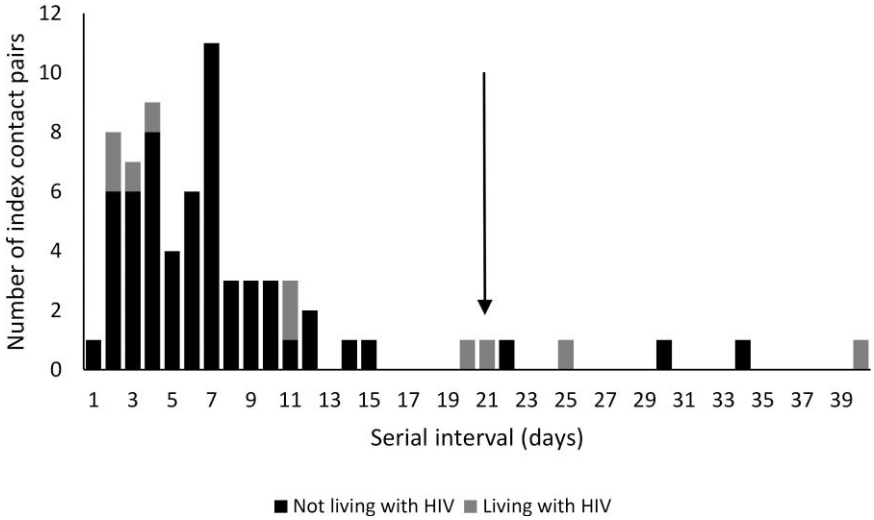


Figure 5. Interval between onset of symptoms in the index case and onset of symptoms in household contacts with severe acute respiratory syndrome coronavirus 2 (serial interval) by HIV status of the index case, Klerksdorp and Soweto, South Africa, 2020–2021 (n = 69). Arrow indicates cutoff for inclusion in analysis for factors associated with serial interval duration. Abbreviations: BMI, body mass index; Ct, cycle threshold; HIV, human immunodeficiency virus.

excluded, HCIR was 51% (220 of 436) overall. We found similar factors associated with HCIR (Supplementary Table 6). When only households where 65% of members completed 65% of visits were included, we included 112 index cases with 342 contacts. Factors associated with HCIR, episode duration, and serial interval were similar to what we observed in the main analysis (Supplementary Tables 7–9).

DISCUSSION

We performed a case-ascertained, prospective, household transmission study for SARS-CoV-2 in South Africa, including 131 index cases, 28 of whom were LWH, and 457 household contacts. We observed a 59% HCIR, HCIR being higher in households with older index cases and contacts aged 13–17 years and 18–34 years. HCIR was also higher in households with Delta-infected index cases vs Beta. The HCIR was similar in index cases LWH and NLWH. Index episode durations were longer in older individuals. Episode duration at high viral load (Ct <30) was longer in index cases LWH, and serial interval was longer in contacts LWH.

HCIR from previous studies has varied based on study design, symptom status of the index case, timing within the epidemic [19], and SARS-CoV-2 variant [20]. In our study, which included only symptomatic index cases, we estimated the HCIR at 59%, a higher estimate than the overall 37% reported from a recent meta-analysis that included 33 studies performed in 2021 and 2022 and the variant-specific 23% and 30% estimates for Beta and Delta variants, respectively [20]. In a study from Madagascar that included both symptomatic and nonsymptomatic index cases, HCIR was 39% [21]. The higher estimate seen in our study may be influenced by symptom severity, as proposed in previous studies [22], or the inclusion of only adult index cases; adult index cases result in higher HCIR [23]. In our study, households with index cases aged >35 years were 3 times more likely to result in higher HCIR compared with when index cases were aged 18–34 years. HCIR was also higher when contacts were aged 13–17 and 18–34 years compared with <5 years. Contacts aged 13–18 years were also associated with higher HCIR in the prospective household cohort study of SARS-CoV-2, influenza, and respiratory syncytial virus community burden, transmission dynamics, and viral interaction in South Africa (PHIRST-C) [2], although studies from earlier in the pandemic showed higher attack rates in elderly household members [22, 24]. This may be related to the shift in age distribution of cases from the older population to younger individuals with progression of the pandemic [2, 18, 25]. As seen previously [2, 26, 27], we also observed higher secondary attack rates where the minimum rRT-PCR Ct was lower for the index case, which could be considered a proxy for higher viral load.

We observed no difference in HCIR in households with index cases living with and without HIV. However, we observed a higher HCIR in people LWH who were immune suppressed, but this association was not statistically significant, possibly due to low numbers (n = 14) of included immunosuppressed index cases LWH. This would fit with previous studies that found that immunocompromised people LWH shed virus at low Ct values for longer [2, 15], allowing more opportunity for secondary infections, although we did not observe increased transmission in our study, possibly due to small numbers.

The mean episode duration in index cases of 19 days was higher than the 11 days reported from the household cohort

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

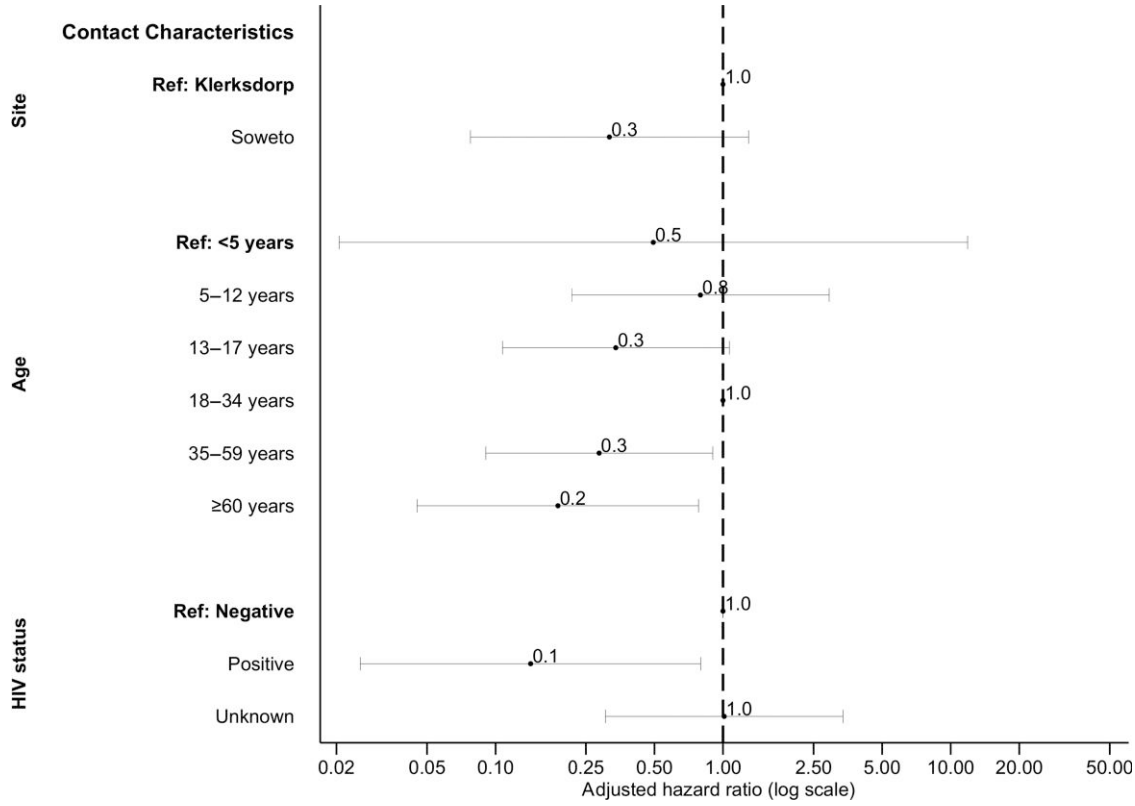


Figure 6. Factors associated with serial interval of severe acute respiratory syndrome coronavirus 2 in cases with symptomatic illness on Weibull accelerated failure time regression, Klerksdorp and Soweto, South Africa, 2020–2021 (n = 62). Abbreviation: HIV, human immunodeficiency virus.

study from South Africa [2] but similar to the 18-day estimate from a meta-analysis for viral shedding time [28]. This may be because our analysis was limited to symptomatic individuals who were shown to be associated with longer episode duration [2, 28]. Episode duration in our study was also longer than in studies of hospitalized South African patients where median episode duration was 13 days [15]. Previous studies from South Africa in the community and in hospitals found that immunocompromised people LWH shed SARS-CoV-2 for longer [2, 15]. While we did not find overall longer shedding in people LWH, when considering detection at Ct <30 (proxy for high viral load), we also observed that people LWH had longer episode durations. Longer episode durations may allow increased opportunity for viral evolution and the establishment of novel variants [29].

Previous serial interval estimates ranged from 4 to 7.5 days [2, 22], similar to our estimate of 6 days. We did not find any index characteristics related to longer serial intervals, but longer serial intervals were observed in contacts aged >35 years and those LWH. Individuals with compromised immune systems (people LWH, the elderly) may still be able to be infected with SARS-CoV-2 toward the end of the index episode when viral loads are lower. Due to their increased risk for hospitalization and death [9, 13], there should be continued support for prioritizing COVID-19 vaccination in these populations.

Our study had limitations. We assumed the first household member who presented with symptoms was the index case. If the true index cases were asymptomatic, we would have underestimated the serial interval, although HCIR estimates should not be greatly affected. Due to the delay between index screening and household enrollment, we did not have the exact date of first SARS-CoV-2 positivity in household contacts and may have also overestimated HCIR if there were multiple introductions of SARS-CoV-2 in the household. By excluding individuals who were seropositive at baseline with no SARS-CoV-2 infection during follow-up, we assumed 100% protection from previous infections, which is likely not correct, and may have overestimated HCIR. When these individuals were included, HCIR reduced by 8%. We were unable to reach the planned sample size for contacts of index cases LWH and may have been underpowered to detect some differences. By only including symptomatic index cases, our results may not be generalizable to households with asymptomatic infections. We did not consider the role of age-related contact patterns within the household on household transmission. This should be considered for future studies.

In conclusion, in 2 communities in South Africa, HCIR was higher than in previous studies [20] and not influenced by HIV status. Episode duration at high viral loads (inferred by Ct <30)

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

was increased for index cases LWH, which may lead to increased risk for secondary transmission and viral evolution. Serial interval was longer in contacts LWH. Although these findings indicate that HIV status of the index case did not affect SARS-CoV-2 transmission to household contacts, it may still play a role, especially if people LWH are not virally suppressed. Sustaining and strengthening HIV treatment and care programs should be a focus moving forward to ensure people LWH are diagnosed and virally suppressed to reduce prolonged shedding of SARS-CoV-2 and potentially reduce increased transmission, as well as their risk for hospitalization and death [9].

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Author contributions. Conception and design of study: J. K., S. W., N. A. M., A. v. G., J. N. B., M. W., L. L., S. T., and C.C. Data collection and laboratory processing: A. v. G., J. N. B., D. G. A., A. B., K. N., N. W., L. G., L. N., J. C., L. d. G., and R. K. Analysis and interpretation: J. K., S. W., S. T., and C. C. Accessed and verified underlying data: J. K., M. N., L. G., L. M., and J. C. Drafted the article: J. K. All authors critically reviewed the article and had access to all of the data reported in the study.

Acknowledgments. The authors thank all of the individuals who participated in the study; the field teams for their hard work and dedication to the study; the Centre for Respiratory Diseases and Meningitis laboratory team; and Andrew Whitelaw, June Fabian, Themba Radebe, and Thoko Sifunda from the safety committee. This study protocol was adapted from the World Health Organization (WHO) Unity Studies Early Investigation Protocol for Household Transmission, an investigation protocol for coronavirus disease 2019 infection.

Data availability. The investigators welcome enquiries about possible collaborations and requests for access to the dataset. Data will be shared after approval of a proposal and with a signed data access agreement. Investigators interested in more details about this study or in accessing these resources should contact the corresponding author.

Financial support. This research was funded by the Wellcome Trust (grant 221003/Z/20/Z) in collaboration with the Foreign, Commonwealth and Development Office, United Kingdom. For the purpose of open access, the author has applied for a CC BY-ND (Creative Commons Attribution–NonDerivative) public copyright license to any Author Accepted Manuscript version arising from this submission. A. v. G., J. K., and S. W. also report grants paid to institution in support of this work from World Health Organization Regional Office for Africa, Africa Pathogen Genomics Initiative through the African Society of Laboratory Medicine and Africa CDC, the US Centers for Disease Control and Prevention (CDC), the South African Medical Research Council, and the UK Department of Health and Social Care and managed by the Fleming Fund: SEQAFRICA project. L. L. and N. W. also report support for this work from the CDC. M. N. reports support for this work paid to institution from the CDC. N. A. M. reports support for this work paid to institution from the National Institute of Communicable Diseases.

Potential conflicts of interest. C. C. has received grant support from Sanofi Pasteur, the CDC, Wellcome Trust, Programme for Applied Technologies in Health, Bill & Melinda Gates Foundation, and South African Medical Research Council. N. W., A. v. G., J. K., and S. W. report receiving grant funds paid to institution from Sanofi Pasteur and the Bill & Melinda Gates Foundation. A. v. G. also reports an unpaid position as chairperson of the National Advisory Group on Immunisation. L. L. reports receiving grant funds from the CDC and Bill & Melinda Gates Foundation. N. M. reports receiving grant funds paid to institution from Pfizer to conduct observational studies on the burden of pneumonia in South Africa; unpaid participation on a data and safety monitoring board of a tuberculosis (TB) host-directed therapies trial and a scientific advisory board on a trial of an electronic reminder to take daily TB treatment; and an unpaid leadership or fiduciary role with the Setshaba Research Center. S. W. reports an unpaid leadership or fiduciary role for a board, society, committee, or advocacy group. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Africa CDC. Africa CDC COVID-19 Dashboard. Available at: https://africacdc.org/covid-19/. Accessed 25 January 2022.
2. Cohen C, Kleynhans J, von Gottberg A, et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020–21. Lancet Infect Dis **2022**; 22:821–34. doi:10.1016/S1473-3099(22)00069-X.
3. Wolter N, Jassat W, Walaza S, et al. Early assessment of the clinical severity of the SARS-CoV-2 Omicron variant in South Africa: a data linkage study. Lancet **2022**; 339:437–46. doi:10.1016/S0140-6736(22)00017-4.
4. UNAIDS. Available at: https://www.unaids.org/sites/default/files/media_asset/2020_aids-data-book_en.pdf. UNAIDS Data 2020. Accessed 23 January 2022.
5. Dorward J, Khubone T, Gate K, et al. The impact of the COVID-19 lockdown on HIV care in 65 South African primary care clinics: an interrupted time series analysis. Lancet HIV **2021**; 8:e158–65. doi:10.1016/S2352-3018(20)30359-3.
6. Harris TG, Jaszi E, Lamb MR, et al. Effects of the coronavirus disease 2019 pandemic on human immunodeficiency virus services: findings from 11 Sub-Saharan African countries. Clin Infect Dis **2022**; 75:e1046–53.
7. Spinelli MA, Lynch KL, Yun C, et al. SARS-CoV-2 seroprevalence, and IgG concentration and pseudovirus neutralising antibody titres after infection, compared by HIV status: a matched case-control observational study. Lancet HIV **2021**; 8: e334–41. doi:10.1016/S2352-3018(21)00072-2.
8. Ambrosioni J, Blanco JL, Reyes-Urueña JM, et al. Overview of SARS-CoV-2 infection in adults living with HIV. Lancet HIV **2021**; 8:e294–305. doi:10.1016/S2352-3018(21)00070-9.

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

9. Tesoriero JM, Swain CAE, Pierce JL, et al. COVID-19 outcomes among persons living with or without diagnosed HIV infection in New York State. JAMA Network Open **2021**; 4:e2037069. doi:10.1001/jamanetworkopen.2020.37069.
10. Danwang C, Noubiap JJ, Robert A, et al. Outcomes of patients with HIV and COVID-19 co-infection: a systematic review and meta-analysis. AIDS Res Ther **2022**; 19:3. doi:10.1186/s12981-021-00427-y.
11. Ssentongo P, Heilbrunn ES, Ssentongo AE, et al. Epidemiology and outcomes of COVID-19 in HIV-infected individuals: a systematic review and meta-analysis. Sci Rep **2021**; 11:6283. doi:10.1038/s41598-021-85359-3.
12. Dong Y, Li Z, Ding S, et al. HIV infection and risk of COVID-19 mortality: a meta-analysis. Medicine **2021**; 100:e26573. doi:10.1097/MD.0000000000026573.
13. Jassat W, Cohen C, Tempia S, et al. Risk factors for COVID-19-related in-hospital mortality in a high HIV and tuberculosis prevalence setting in South Africa: a cohort study. Lancet HIV **2021**; 8:e554–67. doi:10.1016/S2352-3018(21)00151-X.
14. Friedman EE, Devlin SA, McNulty MC, et al. SARS-CoV-2 percent positivity and risk factors among people with HIV at an urban academic medical center. PLoS One **2021**; 16:e0254994. doi:10.1371/journal.pone.0254994.
15. Meiring S, Tempia S, Bhiman JN, et al. Prolonged shedding of SARS-CoV-2 at high viral loads amongst hospitalised immunocompromised persons living with HIV, South Africa. Clin Infect Dis **2022**; 75:e144–56.
16. Wibmer CK, Ayres F, Hermanus T, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med **2021**; 27:622–5. doi:10.1038/s41591-021-01285-x.
17. Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. Euro Surveill **2020**; 25:2001483. doi:10.2807/1560-7917.ES.2020.25.32.2001483.
18. Kleynhans J, Tempia S, Wolter N, et al. SARS-CoV-2 seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. Emerg Infect Dis **2021**; 27:3020–9. doi:10.3201/eid2712.211465.
19. Hsu CY, Wang JT, Huang KC, et al. Household transmission but without the community-acquired outbreak of COVID-19 in Taiwan. J Formos Med Assoc **2021**; 120:S38–45. doi:10.1016/j.jfma.2021.04.021.
20. Madewell ZJ, Yang Y, Longini IM Jr, et al. Household secondary attack rates of SARS-CoV-2 by variant and vaccination status: an updated systematic review and meta-analysis. JAMA Network Open **2022**; 5:e229317. doi:10.1001/jamanetworkopen.2022.9317.
21. Ratovoson R, Razafimahatratra R, Randriamanantsoa L, et al. Household transmission of COVID-19 among the earliest cases in Antananarivo, Madagascar. Influenza Other Respir Viruses **2022**; 16:48–55. doi:10.1111/irv.12896.
22. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med **2020**; 382:1199–207. doi:10.1056/NEJMoa2001316.
23. Koh WC, Naing L, Chaw L, et al. What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate and associated risk factors. PLoS One **2020**; 15:e0240205. doi:10.1371/journal.pone.0240205.
24. Jing QL, Liu MJ, Zhang ZB, et al. Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. Lancet Infect Dis **2020**; 20:1141–50. doi:10.1016/S1473-3099(20)30471-0.
25. Malmgren J, Guo B, Kaplan HG. Continued proportional age shift of confirmed positive COVID-19 incidence over time to children and young adults: Washington State March–August 2020. PLoS One **2021**; 16:e0243042. doi:10.1371/journal.pone.0243042.
26. Marks M, Millat-Martinez P, Ouchi D, et al. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. Lancet Infect Dis **2021**; 21:629–36. doi:10.1016/S1473-3099(20)30985-3.
27. Cerami C, Popkin-Hall ZR, Rapp T, et al. Household transmission of SARS-CoV-2 in the United States: living density, viral load, and disproportionate impact on communities of color. Clin Infect Dis **2021**; 74:1776–85. doi:10.1093/cid/ciab701.
28. Yan D, Zhang X, Chen C, et al. Characteristics of viral shedding time in SARS-CoV-2 infections: a systematic review and meta-analysis. Front Public Health **2021**; 9:652842. doi:10.3389/fpubh.2021.652842.
29. Van Egeren D, Novokhodko A, Stoddard M, et al. Controlling long-term SARS-CoV-2 infections can slow viral evolution and reduce the risk of treatment failure. Sci Rep **2021**; 11:22630. doi:10.1038/s41598-021-02148-8.

Downloaded from https://academic.oup.com/cid/article/76/3/e71/6655792 by guest on 15 March 2023

Supplementary material:

Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study

Contents

Supplementary methods 1
Supplementary results 4
Supplementary discussion 5
Supplementary tables 6
Supplementary figures 26
References 28

Supplementary methods

Study population

We conducted the study at two sites, Klerksdorp and Soweto. Klerksdorp is located in the municipality of Matlosana in North West Province and has a population of over 385,000 people and is 115 km^2. The households of study participants in Klerksdorp include mostly single-family houses and shacks. Prevalence of HIV in the North West Province was estimated at 18% in 2020 for individuals >15 years.[1] Soweto is a township outside of Johannesburg, located in the Gauteng Province where the HIV prevalence for individuals >15 years was estimated at 15% in 2020.[1] Its population is 1.5 million and encompasses 150 km^2. Housing structures include, but are not limited to single-family houses, multi-unit dwellings, shacks within informal settlements and in the backyards of formal residences, and hostels. Clinic catchment areas were defined as Kenneth Kaunda District in North West Province and sub-district Region D of City of Johannesburg district in Gauteng Province),

Sample size

We aimed to assess a significant difference in the household cumulative infection risk (HCIR) between household contacts exposed to SARS-CoV-2 by a HIV-infected vs HIV-uninfected index case for a 95% confidence interval and 80% power. We assumed an average household size of 5 individuals, a HIV prevalence among index cases presenting with COVID-19-like-illness of 40% (based on HIV prevalence for influenza-like illness),[2] and a HCIR of 10% in household members exposed to an HIV-uninfected index case and a HCIR of 20% among household members exposed to a HIV-infected index case. The resulting total sample size for this study was 440 exposed household members: 264 exposed household members from households with an HIV-uninfected index case and 176 exposed household members from households with an index case LWH.

Screening for index cases

Screening for index cases started in October 2020 and continued September 2021, and was done at three clinics in Klerksdorp (Jouberton, Alabama and Grace Mokhomo clinics), and five in Soweto (Lillian Ngoyi, Zola, Mofolo and Chaiwelo Community Health Centres; and Mofolo clinic). Screening started in October 2020 and continued until June 2021 in Klerksdorp and September 2021 in Soweto. In July 2021, the Soweto site was reduced to two clinics.

Study staff screened clinic attendees 18 years and older on week days for suspected COVID-19 including any of the following symptoms: fever, cough, sore throat, difficulty breathing, myalgia, headaches, running nose and diarrhoea; with symptom onset ≤5 days prior to screening. Two nasopharyngeal swabs were collected from consenting eligible individuals; one of which was transported in viral or universal transport medium to the National Institute for Communicable Diseases (NICD) for SARS-CoV-2 RNA detection using real-time reverse transcription polymerase chain reaction (rRT-PCR), and the other was used for point-of-care SARS-CoV-2 testing using the rapid antigen detection kit (implemented February 2021, Standard Q COVID-19 Ag test, SD Biosensor, Chungcheongbuk-do, South Korea). All positive antigen tests were verified with rRT-PCR. Staff also completed a case investigation form to collect demographic information, symptoms, and underlying medical conditions, including HIV infection. In addition to active screening at clinics, individuals who tested positive for SARS-CoV-2 meeting study case definitions were also referred from the influenza-like illness surveillance program in Klerksdorp[2] and by contact tracing teams in the site districts as part of the national public health response to COVID-19 in South Africa. Referred individuals were included if a documented SARS-CoV-2 positive test was available.

Household enrolment

We approached households of individuals who tested positive for SARS-CoV-2, with symptom onset less than 7 days prior (up to 5 days prior to screening with another 48 hours for testing to be completed), and who reported at least two additional household members. We excluded households who resided outside the clinics' catchment areas, if household contacts reported symptoms in the 14 days prior to screening, or if the index cases were hospitalised immediately following screening). Once the head of household provided permission, all household members were approached for consenting and enrolment. Households with ≥3 eligible members, of whom ≥70% consented to participate were enrolled. To allow for sufficient possible transmission events in each enrolled household, we did not enrol households with <3 members. During enrolment we collected information on household characteristics such as household size, caretaking roles, sleeping arrangements and housing; and demographics, underlying medical conditions, education, employment, smoking, alcohol use and vaccination information. We did not collect information SARS-CoV-2 infections prior to study enrolment. Households that withdrew within 10 days from index symptom start date were excluded from the analysis. All data collected during screening and follow up visits were captured in real-time using tablets onto Research Electronic Data Capture (REDCap) databases hosted at the University of the Witwatersrand[3].

Symptoms

Symptoms recorded for all ages were self-reported or measured fever, cough, shortness of breath or difficulty breathing, nasal congestion or a runny nose, vomiting and diarrhoea. We also collected sore throat, lost sense of smell or taste, abdominal pain, muscle aches, fatigue, headache and confusion in those aged ≥5 years, and poor feeding, irritation, and lethargy for those <5 years.

HIV testing

Point of care rapid HIV enzyme-linked immunosorbent assay (ELISA) testing (WANTAI HIV 1+2Ab Rapid test, Beijing Wantai Biological, Beijing, China) was offered to all individuals at screening with an unknown HIV status, or for whom a documented negative HIV result was from > 6 months previously. For consenting household contacts HIV status was determined using rapid HIV tests if unknown or a documented negative status was not available within the previous 6 months. For children aged 10 years and younger whose mother had a documented HIV negative status during pregnancy, HIV status was considered negative. Where participants did not agree to rapid testing at point of care, but did consent to HIV testing, residual serum was tested for HIV antibodies using an ELISA assay at NICD. Data on recent CD4+ T cell count and HIV viral load was collected for

individuals living with HIV. If these data were not available, samples were collected and viral load tested by quantitative rRT-PCR (Roche Cobas Ampliprep/Cobas Taqman HIV-1 test, Roche Diagnostics, Mannheim, German) at the NICD.

SARS-CoV-2 detection

Nasopharyngeal (screening) and nasal (follow-up) specimens were collected using nylon flocked swabs and transported in viral or universal transport medium to the NICD for testing. Nasopharyngeal specimens were used during screening as this is the standard of care for SARS-CoV-2 diagnostics in South Africa due to higher sensitivity. Nasal specimens were used during follow-up due to better acceptability for participants (less discomfort than nasopharyngeal specimens) and reduced occupational transmission risk for study nurses (lower probability for aerosolisation). Total nucleic acids were extracted from 200µl of sample using a MagNA Pure 96 automated extractor and the DNA/Viral NA Small Volume v2.0 extraction kit (Roche Diagnostics, Mannheim, Germany). The E, RdRp and N SARS-CoV-2 genes were detected by rRT-PCR, using the Allplex™ 2019-nCoV kit (Seegene Inc., Seoul, South Korea). Specimens were considered positive for SARS-CoV-2 if the cycle threshold (C_t) value was below 40 for any of the gene targets.

SARS-CoV-2 variants

The first SARS-CoV-2 positive specimen for each participant was subjected to the Allplex™ SARS-CoV-2 Variants I and II PCR assays (Seegene Inc., Seoul, Korea). The Variants I assay targets the RdRp gene, HV69/70 deletion, N501Y and E484K mutations; and the Variants II assay targets the L452R, W152C, K417T and K417N mutations; allowing differentiation between the Alpha (B.1.1.7 Pango lineage), Beta (B.1.351), Delta (B.1.617.2), Gamma (P.1) variants of concern, and the US-California variant of interest (B.1.429).

In addition, the same specimen was sequenced on the Ion Torrent Genexus platform using AmpliSeq for SARS-CoV-2 (Illumina), which is based on the ARTIC SARS-CoV-2 sequencing protocol (https://www.protocols.io/view/illumina-nextera-dna-flex-library-construction-and-bhjgj4jw). Genomes were assembled using the Exatype SARS-CoV-2 pipeline (https://sars-cov-2.exatype.com/), which includes de-duplicating sequenced reads data, base-calling, de-multiplexing, removal of amplicon primer sequences prior to variant calling, mapping to a reference using Examap and finally consensus sequence generation. These consensus sequences were manually inspected and polished using Aliview v1.27 (http://ormbunkar.se/aliview/). The reference genome used throughout the assembly process was NC_045512.2 (Accession number: MN908947.3). Clade and lineage assignments were made using the online Nextclade (https://clades.nextstrain.org/) and PANGOLin (https://pangolin.cog-uk.io/) applications, which also enable identification of known variants of concern as well as novel mutations. Only specimens with a C_t value <35 on at least one gene target in the initial rRT-PCR was sequenced. If the first specimen in the episode was not successfully characterised (due to high C_t values in initial rRT-PCR or poor sample quality), a second specimen was characterised by variant PCR and sequencing.

Variants were assigned firstly based on sequencing results where sequence coverage/reads was sufficient for clade and lineage classification, followed by the variant PCR results when sequence coverage/reads was low, and then based on variants detected within the household if no sequencing or variant PCR result was available. Where no variant-specific mutations were detected in the PCR, and the RdRp gene C_t value was <35, and internal controls for both assays were detected, the episodes were classified as attributed to a non-Alpha/Beta/Delta variant. Where more than one variant was detected within one SARS-CoV-2 cluster, specimens were retested from fresh aliquots of the same sample to confirm results. Confirmed mixed variant clusters were excluded from HCIR analyses.

Serology

We used an ELISA to detect antibodies against SARS-CoV-2 spike protein as previously described[4]. In summary, recombinant trimeric spike protein was coated at 2 μg/ml onto 96-well, high-binding plates and incubated overnight at 4°C. The plates were washed and incubated at 37 °C for 1–2 hours in blocking buffer (5% skimmed milk powder, 0.05% Tween 20, 1× PBS). Serum samples were added at 1:100 dilution, following a one-hour incubation at 37 °C and anti-human horseradish peroxidase-conjugated antibody was added for another one hour at 37°C. The signal was developed with OneStep TMB substrate (Thermo Fisher Scientific, USA) for 5 min at room temperature, followed by addition of 1 M H_2SO_4 stop solution. Absorbance at 450 nm was measured and specimens with optical density (OD)>0.4 were considered positive for anti-spike antibodies. For the nucleocapsid protein we performed the assay on the Cobas e601 instrument (Roche Diagnostics), and we considered a cut-off index (COI) >1.0 as positive for anti-nucleocapsid antibodies.

Supplementary results

Serology

At baseline, 67% (83/123) of index cases for whom serology results were available were already seropositive for SARS-CoV-2, which increased to 98% (107/109) after six weeks of follow-up. Two individuals had a rRT-PCR-confirmed SARS-CoV-2 episode, but were still seronegative for SARS-CoV-2 antibodies at the end of follow-up. The first individual who did not seroconvert was a HIV-uninfected male. Episode duration was 6 days, with only one positive swab (collected at screening) with a minimum C_t value of 38. There was no secondary transmission (0/2) detected in the household linked to the illness episode. The second was a female living with HIV, with an episode duration of 9 days and minimum C_t value of 26. There was an 80% (4/5) attack rate in the household. No HIV viral load or CD4+ T cell count were available.

Paired baseline and exit sera were available in 72% (331/457) of contacts. In those with paired sera, 39% (128/331) where already seropositive at baseline, 57% (73/128) of whom had a rRT-PCR-confirmed SARS-CoV-2 infection during follow-up. Of those initially seronegative, 47% (95/203) were seropositive after six weeks of follow-up; 80% (76/95) of whom had at least one respiratory specimen positive for SARS-CoV-2 positive specimen. The 19 contacts who seroconverted during follow-up but with no rRT-PCR-confirmed infection were not included as secondary cases in attack rate estimates or the risk factor analysis as we could not confirm if this episode was linked to the index case episode. Twenty percent (19/95) of contacts initially seronegative who had at least one respiratory specimen positive for SARS-CoV-2 positive did not seroconvert by the end of the follow-up period. Of the 91 index cases and 331 contacts with paired sera, no sero-reversions were detected.

Only one household contact was vaccinated for SARS-CoV-2 with at least one dose 14 days before enrolment. This participant was seropositive at baseline with no rRT-PCR-confirmed infection during follow up, and therefore not included in the analysis.

Mixed SARS-CoV-2 variant clusters

We identified four households where multiple SARS-CoV-2 variants were detected within one cluster. These were two Beta and Delta clusters, one Alpha and Beta cluster, and one Beta and C.1.2 cluster (Supplementary Figure 2).

Out-patient consultations, hospitalisations and deaths

The proportion of individuals infected with SARS-CoV-2 who had COVID-19-related hospitalisations, was 6% (6/131) in index cases and 1% (5/457) in contacts (Supplementary Table 5), with one death each of an index case and contact.

Supplementary discussion

Household cumulative infection risk

From our results, there were fewer secondary cases in contacts who smoked. While evidence exists that smoking may lead to poorer COVID-19 outcomes,[5] other studies have also found a negative association with smoking and SARS-CoV-2 infection.[5, 6] The latter association may be a result of confounding biases including socio-economic status and healthcare seeking behaviour,[5] or possible decreased test sensitivity[6] or could reflect a true biological interaction.

Episode duration

On multivariable analysis, index cases that were underweight (aHR 5.8, 95%CI 1.5-22.1) or obese (aHR 2.0, 95%CI 1.2-3.5) compared to a normal BMI, had shorter episode durations (Supplementary Table 2). Obesity has been shown to be associated with increased episode duration,[7] of which we observed the inverse with individuals of normal weight having longer episode durations, the reason for this is unclear.

After adjusting for confounders on multivariable analysis, individuals infected with the Delta variant had shorter episode durations than those infected with Beta, while episode duration did not differ based on variant when considering detection at a C_t<30. In contrast, households with Delta infections had higher HCIR. Longer episode duration does not necessary indicate shedding of live virus; and higher HCIR with Delta might be related to higher viral loads[8] and increased infectiousness.[9] We did observe lower minimum C_t values (proxy for higher viral load) in secondary cases infected with Delta vs Beta. The shorter episode durations with Delta infections in index cases observed could also be due to repeat infections in some index cases, which have been shown to result in shorter episode durations.[8] We did not collect information SARS-CoV-2 infections prior to study enrolment. Episode duration based on C_t<30 (as a proxy for higher viral load) was longer in females. This association of longer episode duration in females has not been observed in other studies, and warrants further investigation.

Limitations

Although we reached our overall sample size, we only managed to enrol 103 exposed household members from households with an index case LWH (59% of target). The HIV prevalence amongst screened patients were lower than anticipated (9% vs expected 40% as seen in surveillance for influenza-like illness). Furthermore, very few index cases were immunosuppressed, likely due to a high proportion of PLWH being in care. Individuals who are not on ART may be less likely to seek care and be identified as SARS-CoV-2 index cases.

We did not perform quantitative rRT-PCR to determine viral loads to validate the use of C_t<30 as a proxy for high viral load. We also did not culture samples to confirm the shedding of live virus at C_t<30, but rather based this assumption on previous literature [10].

Supplementary tables

Supplementary Table 1. Factors associated with SARS-CoV-2 household transmission in index cases and acquisition in household contacts, Klerksdorp and Soweto, South Africa, 2020-2021, (n=373).

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Index characteristics			
Site			
Klerksdorp	95/174 (54.6)	Reference	
Soweto	125/199 (62.8)	1.4 (0.7-2.8)	1.1 (0.5-2.1)
Age (years)			
18-34	54/108 (50.0)	Reference	Reference
35-59	137/218 (62.8)	1.9 (0.9-4.2)	**3.4 (1.5-7.8)**
≥60	29/47 (61.7)	1.8 (0.6-6.0)	**3.1 (1.0-10.1)**
Sex			
Male	64/111 (57.7)	0.9 (0.4-1.9)	
Female	156/262 (59.5)	Reference	
HIV			
HIV-infected	51/85 (60.0)	1.0 (0.4-2.3)	
HIV-uninfected	163/279 (58.4)	Reference	
Unknown HIV-status	6/9 (66.7)	0.8 (0.1-11.4)	
CD4+ T cell count			
HIV-infected, not immune suppressed†	30/52 (57.7)	Reference	Reference
HIV-infected, immune suppressed†	8/13 (61.5)	1.6 (0.2-13.5)	2.5 (0.4-15.3)
HIV-infected, no CD4+ T cell count	13/20 (65.0)	1.5 (0.2-8.7)	1.2 (0.2-5.8)
HIV-uninfected	163/279 (58.4)	1.2 (0.4-3.3)	1.8 (0.7-4.9)
Unknown HIV-status	6/9 (66.7)	1.0 (0.1-15.3)	0.5 (0.0-5.5)
Underlying illness‡			
No	172/293 (58.7)	0.8 (0.3-1.8)	
Yes	48/75 (64.0)	Reference	
Unknown	0/5 (0.0)	Excluded	
Body-mass index			
Underweight	3/9 (33.3)	0.2 (0.0-2.2)	**0.1 (0.0-0.8)**
Normal weight	68/108 (63.0)	Reference	Reference
Overweight	44/67 (65.7)	1.1 (0.4-3.2)	0.6 (0.2-1.5)
Obese	105/184 (57.1)	0.8 (0.4-1.7)	**0.4 (0.2-0.9)**
Unknown	0/5 (0.0)	Excluded	Excluded
Crowding			
No	56/88 (63.6)	1.5 (0.7-3.3)	
Yes	164/285 (57.5)	Reference	
Child <5 years in home			
No	190/316 (60.1)	1.8 (0.7-5.0)	
Yes	30/57 (52.6)	Reference	
Current cigarette smoke (self-reported)			
No	203/328 (61.9)	2.8 (1.0-8.0)	
Yes	17/40 (42.5)	Reference	
Unknown	0/5 (0.0)	Excluded	
Episode duration			

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
1-10 days	64/122 (52.5)	Reference	
11-21 days	53/95 (55.8)	1.5 (0.6-3.6)	
>21 days	101/151 (66.9)	**2.3 (1.0-5.3)**	
Positive at last visit	2/5 (40.0)	0.5 (0.0-8.7)	
Fever			
No	110/187 (58.8)	Reference	
Yes	110/186 (59.1)	1.1 (0.5-2.2)	
Minimum C_t value			
>35	9/39 (23.1)	Reference	Reference
25-35	86/139 (61.9)	**8.3 (2.4-28.1)**	**7.5 (2.2-26.0)**
<25	123/192 (64.1)	**8.7 (2.7-28.4)**	**5.3 (1.6-17.6)**
Unknown	2/3 (66.7)	10.2 (0.3-402.1)	11.3 (0.4-355.6)
Serostatus at follow-up end			
Seronegative	4/7 (57.1)	0.5 (0.0-7.4)	
Seropositive	182/293 (62.1)	Reference	
Unknown	34/73 (46.6)	0.4 (0.2-1.1)	
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	18/28 (64.3)	1.3 (0.4-4.2)	2.1 (0.7-6.3)
Alpha	16/20 (80.0)	3.3 (0.7-14.8)	3.4 (0.8-15.6)
Beta	144/253 (56.9)	Reference	Reference
Delta	40/49 (81.6)	**3.6 (1.2-10.7)**	**4.6 (1.5-14.4)**
Unknown	2/23 (8.7)	0.0 (0.0-0.3)	0.1 (0.0-0.5)
Contact characteristics			
Age (years)			
<5	5/19 (26.3)	Reference	Reference
5-12	40/73 (54.8)	4.2 (1.0-17.6)	1.9 (0.4-7.9)
13-17	43/60 (71.7)	**15.8 (3.4-74.0)**	**7.1 (1.5-33.9)**
18-34	56/95 (59.0)	**6.4 (1.6-26.1)**	**4.4 (1.0-18.4)**
35-59	56/90 (62.2)	**6.3 (1.5-26.4)**	4.4 (1.0-18.6)
≥60	20/36 (55.6)	**6.1 (1.2-30.2)**	2.9 (0.6-14.4)
Sex			
Male	83/160 (51.9)	Reference	
Female	137/213 (64.3)	**2.0 (1.2-3.4)**	
HIV			
HIV-uninfected	82/147 (55.8)	1.1 (0.4-3.1)	
HIV-infected	16/28 (57.1)	Reference	
HIV status unknown	122/198 (61.6)	1.3 (0.5-3.5)	
Underlying illness‡			
No	189/323 (58.5)	Reference	
Yes	30/42 (71.4)	2.6 (1.0-6.8)	
Unknown	1/8 (12.5)	0.1 (0.0-0.8)	
Body-mass index			
Underweight	15/25 (60.0)	1.0 (0.3-3.0)	
Normal weight	96/173 (55.5)	Reference	
Overweight	44/77 (57.1)	1.2 (0.6-2.4)	
Obese	64/90 (71.1)	**2.4 (1.2-4.9)**	
Unknown	1/8 (12.5)	0.1 (0.0-1.0)	

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Current cigarette smoke (self-reported)			
No	197/320 (61.6)	**2.8 (1.2-6.2)**	**3.9 (1.7-9.0)**
Yes	22/49 (44.9)	Reference	Reference
Unknown	1/4 (25.0)	0.4 (0.0-10.4)	Excluded
Sleep in same room as index			
No	150/261 (57.5)	Reference	
Yes	70/112 (62.5)	1.2 (0.7-2.1)	
Index is main caregiver			
No	183/313 (58.5)	1.2 (0.6-2.6)	
Yes	37/60 (61.7)	Reference	

OR – Odds ratio, aOR – Adjusted odds ratio. *Household cumulative infection risk (HCIR) excludes household members seropositive at baseline with no SARS-CoV-2 detected on rRT-PCR during follow-up and households with mixed SARS-CoV-2 variant clusters. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes $p<0.05$.

Supplementary Table 2. Factors associated with SARS-CoV-2 episode duration (irrespective of C_t value) in index cases, Klerksdorp and Soweto, South Africa, 2020-2021, (n=123). HR – Hazard ratio*

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Site			
Klerksdorp	17.0±11.6 (4.0-45.0)	Reference	Reference
Soweto	21.2±12.6 (3.0-45.0)	**0.6 (0.5-0.9)**	**0.4 (0.3-0.7)**
Age (years)			
18-34	16.5±11.3 (4.0-41.0)	Reference	Reference
35-59	18.6±12.6 (3.0-45.0)	0.7 (0.5-1.1)	**0.4 (0.2-0.6)**
≥60	27.3±10.0 (6.0-43.0)	**0.5 (0.3-0.8)**	**0.2 (0.1-0.5)**
Sex			
Male	18.5±13.2 (3.0-45.0)	Reference	Reference
Female	19.3±11.9 (3.0-45.0)	1.1 (0.7-1.6)	0.8 (0.5-0.0)
HIV			
HIV-infected	17.2±12.0 (3.0-45.0)	0.8 (0.5-1.2)	1.5 (0.8-2.7)
HIV-uninfected	19.6±12.4 (3.0-45.0)	Reference	Reference
Unknown HIV-status	18.5±10.6 (11.0-26.0)	1.3 (0.3-5.6)	0.9 (0.2-4.5)
CD4+ T cell count			
HIV-infected, not immune suppressed†	18.8±12.8 (3.0-45.0)	Reference	Reference
HIV-infected, immune suppressed†	14.1±9.0 (7.0-33.0)	1.4 (0.5-4.3)	1.0 (0.3-3.7)
HIV-infected, no CD4+ T cell count	14.1±9.0 (7.0-33.0)	1.9 (0.8-4.7)	0.9 (0.3-2.5)
HIV-uninfected	19.6±12.4 (3.0-45.0)	1.0 (0.6-1.6)	Omitted
Unknown HIV-status	18.5±10.6 (11.0-26.0)	1.4 (0.3-6.2)	Omitted
Underlying illness‡			
No	19.1±12.4 (3.0-45.0)	Reference	
Yes	19.8±12.2 (3.0-41.0)	0.8 (0.5-1.3)	
Unknown	9.0±0.0 (9.0-9.0)	4.0 (1.0-17.0)	
Body-mass index			
Underweight	13.0±4.6 (9.0-18.0)	**3.0 (0.9-10.1)**	**5.8 (1.5-22.1)**
Normal weight	20.1±12.9 (6.0-45.0)	Reference	Reference
Overweight	19.9±14.5 (3.0-43.0)	1.0 (0.6-1.7)	1.7 (0.9-3.2)
Obese	18.8±11.3 (4.0-44.0)	1.2 (0.8-1.9)	**2.0 (1.2-3.5)**
Unknown	9.0±0.0 (9.0-9.0)	4.6 (1.1-19.9)	2.8 (0.6-14.2)
Crowding			
No	18.0±11.7 (3.0-43.0)	Reference	
Yes	19.6±12.5 (3.0-45.0)	0.8 (0.5-1.2)	
Child <5 years in home			
No	19.5±12.2 (3.0-45.0)	0.7 (0.4-1.3)	
Yes	16.6±12.5 (4.0-35.0)	Reference	
Current cigarette smoke (self-reported)			
No	19.9±12.5 (3.0-45.0)	0.6 (0.3-1.1)	
Yes	15.3±10.3 (3.0-37.0)	Reference	
Unknown	9.0±0.0 (9.0-9.0)	2.8 (0.6-12.5)	
Fever			
No	15.8±10.5 (3.0-45.0)	Reference	
Yes	22.4±13.1 (3.0-45.0)	**0.7 (0.5-0.9)**	
Minimum C_t value			
>35	10.8±7.0 (3.0-26.0)	**0.4 (0.2-0.8)**	1.8 (0.9-3.6)

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
25-35	15.2±11.8 (3.0-45.0)	Reference	Reference
<25	24.1±11.5 (9.0-43.0)	**0.3 (0.1-0.5)**	**0.5 (0.3-0.9)**
Unknown	8.0±0 (8.0-8.0)	2.2 (0.3-17.0)	12.2 (1.4-103.2)
Serostatus at follow-up end			
Seronegative	7.5±2.1 (6.0-9.0)	Reference	Reference
Seropositive	20.6±12.4 (3.0-45.0)	**0.1 (0.0-0.5)**	**0.0 (0.0-0.3)**
Unknown	13.5±10.2 (3.0-38.0)	0.2 (0.1-1.0)	0.1 (0.0-0.7)
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	13.6±9.8 (4.0-38.0)	2.4 (1.3-4.6)	**4.0 (1.9-8.2)**
Alpha	21.4±15.1 (7.0-41.0)	0.9 (0.4-1.9)	0.8 (0.3-2.1)
Beta	20.4±12.9 (3.0-45.0)	Reference	Reference
Delta	20.1±9.9 (4.0-41.0)	1.2 (0.7-2.1)	**2.6 (1.4-5.0)**
Unknown	11.4±7.0 (5.0-28.0)	3.0 (1.5-5.9)	1.6 (0.7-3.6)

*Estimated using Weibull accelerated failure time regression adjusted for clustering by site and household. Hazard ratio <1 corresponds to longer episode duration than reference group used. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes $p<0.05$.

Supplementary Table 3. Factors associated with SARS-CoV-2 episode duration (C_t value<30) in index cases, Klerksdorp and Soweto, South Africa, 2020-2021, (n=123). HR – Hazard ratio*

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Site			
Klerksdorp	6.9±2.2 (3.0-13.0)	Reference	Reference
Soweto	7.0±2.6 (2.0-17.0)	0.6 (0.4-1.0)	0.7 (0.4-1.2)
Age (years)			
18-34	6.6±2.8 (3.0-17.0)	Reference	Reference
35-59	7.0±2.2 (2.0-13.0)	1.0 (0.6-1.6)	1.0 (0.5-1.8)
≥60	7.3±2.5 (4.0-12.0)	**0.3 (0.1-0.6)**	0.6 (0.2-1.5)
Sex			
Male	6.7±2.4 (3.0-12.0)	Reference	Reference
Female	7.0±2.5 (2.0-17.0)	0.7 (0.4-1.1)	**0.5 (0.3-0.9)**
HIV			
HIV-infected	7.3±2.8 (2.0-12.0)	**0.4 (0.2-0.7)**	**0.4 (0.1-0.9)**
HIV-uninfected	6.8±2.3 (3.0-17.0)	Reference	Reference
Unknown HIV-status	9.0±0 (9.0-9.0)	0.6 (0.1-4.4)	0.6 (0.1-4.9)
CD4+ T cell count			
HIV-infected, not immune suppressed†	7.2±3.1 (2.0-12.0)	Reference	Reference
HIV-infected, immune suppressed†	7.3±2.1 (5.0-9.0)	**4.1 (1.1-10.9)**	2.3 (0.5-10.3)
HIV-infected, no CD4+ T cell count	7.5±3.3 (4.0-12.0)	3.3 (1.0-9.3)	0.6 (0.1-3.0)
HIV-uninfected	6.8±2.3 (3.0-17.0)	**4.4 (2.1-21.6)**	Omitted
Unknown HIV-status	9.0±0 (9.0-9.0)	2.7 (0.3-0.0)	Omitted
Underlying illness‡			
No	7.0±2.6 (2.0-17.0)	Reference	
Yes	6.6±2.0 (3.0-11.0)	1.5 (0.9-2.5)	
Unknown	9.0±0 (9.0-9.0)	0.7 (0.1-5.5)	
Body-mass index			
Underweight	10.5±2.1 (9.0-12.0)	0.8 (0.2-3.6)	0.6 (0.1-2.7)
Normal weight	6.7±1.7 (4.0-10.0)	Reference	Reference
Overweight	6.1±2.2 (3.0-12.0)	**2.7 (1.3-5.3)**	2.1 (0.9-5.0)
Obese	7.2±2.7 (2.0-17.0)	1.7 (1.0-3.0)	1.3 (0.6-2.7)
Unknown	9.0±0 (9.0-9.0)	1.2 (0.2-9.4)	0.6 (0.1-5.2)
Crowding			
No	6.6±2.2 (3.0-13.0)	Reference	
Yes	7.1±2.5 (2.0-17.0)	0.7 (0.4-1.0)	
Child <5 years in home			
No	6.9±2.5 (2.0-17.0)	0.8 (0.4-1.6)	
Yes	6.8±1.2 (5.0-9.0)	Reference	
Current cigarette smoke (self-reported)			
No	6.8±2.3 (2.0-13.0)	1.2 (0.6-2.3)	
Yes	7.8±3.6 (5.0-17.0)	Reference	
Unknown	9.0±0 (9.0-9.0)	0.8 (0.1-6.5)	
Fever			
No	6.3±2.4 (2.0-13.0)	Reference	
Yes	7.4±2.4 (3.0-17.0)	**1.1 (0.7-1.7)**	
Minimum C_t value			

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
>35	No observations	No observations	No observations
25-35	11.1±6.5 (3.0-31.0)	Reference	Reference
<25	24.1±11.5 (9.0-43.0)	**0.7 (0.4-1.2)**	1.1 (0.6-2.1)
Unknown	No observations	No observations	No observations
Serostatus at follow-up end			
Seronegative	4.0±0 (4.0-4.0)	Reference	Reference
Seropositive	7.0±2.5 (2.0-17.0)	**0.2 (0.0-1.2)**	**0.01 (0.01-0.2)**
Unknown	6.6±2.1 (3.0-10.0)	0.2 (0.0-1.9)	**0.02 (0.01-0.3)**
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	6.8±1.6 (5.0-10.0)	1.3 (0.7-2.7)	1.7 (0.8-3.7)
Alpha	6.0±2.2 (3.0-9.0)	1.6 (0.7-3.8)	1.6 (0.6-4.2)
Beta	7.2±2.6 (3.0-17.0)	Reference	Reference
Delta	6.4±2.1 (3.0-11.0)	**0.4 (0.2-0.8)**	1.1 (0.5-2.5)
Unknown	6.0±5.7 (2.0-10.0)	1.1 (0.3-4.5)	1.0 (0.2-4.5)

*Estimated using Weibull accelerated failure time regression adjusted for clustering by site and household. Hazard ratio <1 corresponds to longer episode duration than reference group used. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Omitted categories of co-variates were co-linear in multivariable model. Bold face font denotes p<0.05.

Supplementary Table 4. Factors associated with serial interval of SARS-CoV-2 in cases with symptomatic illness, Klerksdorp and Soweto, South Africa, 2020-2021, (n=62). HR – Hazard ratio*

Characteristic	Serial interval, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Index characteristics			
Site			
Klerksdorp	5.8±4.1 (1.0-21.0)	Reference	
Soweto	7.5±4.0 (2.0-20.0)	0.3 (0.1-1.3)	0.3 (0.1-1.3)
Age (years)			
18-34	6.6±4.6 (1.0-15.0)	Reference	
35-59	6.4±3.6 (2.0-21.0)	0.6 (0.1-3.0)	
≥60	7.0±6.4 (2.0-20.0)	0.9 (0.1-7.9)	
Sex			
Male	6.7±4.6 (2.0-21.0)	0.9 (0.2-3.9)	
Female	6.5±3.9 (1.0-20.0)	Reference	
HIV			
HIV-infected	9.3±7.9 (2.0-21.0)	0.2 (0.0-1.4)	
HIV-uninfected	6.1±3.2 (1.0-15.0)	Reference	
Underlying illness‡			
No	6.9±4.4 (1.0-21.0)	Reference	
Yes	17.0±5.5 (2.0-14.0)	1.5 (0.4-6.1)	
Body-mass index			
Normal weight	7.2±5.8 (2.0-21.0)	0.7 (0.1-4.8)	
Overweight	6.3±3.7 (2.0-15.0)	Reference	
Obese	6.4±3.7 (1.0-20.0)	0.7 (0.1-3.5)	
Crowding			
No	5.9±4.1 (1.0-15.0)	Reference	
Yes	6.8±4.1 (2.0-21.0)	0.8 (0.2-3.3)	
Child <5 years in home			
No	6.5±4.2 (1.0-21.0)	0.4 (0.0-3.5)	
Yes	6.9±3.6 (2.0-12.0)	Reference	
Current cigarette smoke (self-reported)			
No	6.7±4.2 (1.0-21.0)	0.2 (0.0-5.9)	
Yes	4.3±0.5 (4.0-5.0)	Reference	
Episode duration			
1-10 days	6.8±4.0 (2.0-15.0)	2.0 (0.3-12.3)	
11-21 days	7.4±4.6 (1.0-21.0)	Reference	
>21 days	5.9±4.0 (2.0-20.0)	3.9 (0.6-24.0)	
Unknown			
Fever			
No	7.0±4.7 (1.0-20.0)	1.0 (0.3-4.0)	
Yes	6.2±3.7 (2.0-21.0)	Reference	
Minimum C_t value			
25-35	6.3±5.5 (1.0-21.0)	0.5 (0.1-2.1)	
<25	6.7±3.1 (2.0-15.0)	Reference	
Serostatus at end			
Seronegative	6.1±3.0 (2.0-14.0)	Reference	
Seropositive	6.9±4.6 (1.0-21.0)	0.6 (0.1-2.5)	

Characteristic	Serial interval, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Unknown	5.0±2.0 (3.0-7.0)	1.7 (0.1-22.8)	
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	8.2±2.3 (6.0-12.0)	0.2 (0.0-1.0)	
Alpha	3.6±1.2 (2.0-5.0)	7.1 (0.7-73.2)	
Beta	6.1±4.6 (1.0-21.0)	Reference	
Delta	8.2±4.1 (2.0-15.0)	0.4 (0.1-2.7)	
Contact characteristics			
Age (years)			
<5	5.5±0.7 (5.0-6.0)	1.9 (0.1-26.0)	0.5 (0.0-11.9)
5-12	5.4±3.4 (1.0-10.0)	0.9 (0.2-3.4)	0.8 (0.2-2.9)
13-17	6.3±3.5 (2.0-12.0)	0.4 (0.1-1.3)	0.3 (0.1-1.1)
18-34	4.4±2.6 (2.0-9.0)	Reference	Reference
35-59	7.5±4.7 (2.0-20.0)	**0.3 (0.1-1.1)**	**0.3 (0.1-0.9)**
≥60	8.3±4.8 (3.0-21.0)	**0.3 (0.1-1.0)**	**0.2 (0.0-0.8)**
Sex			
Male	6.2±4.6 (1.0-20.0)	Reference	
Female	6.7±3.8 (2.0-21.0)	0.8 (0.3-1.9)	
HIV			
HIV-uninfected	5.4±4.0 (1.0-21.0)	Reference	Reference
HIV-infected	11.1±5.6 (3.0-20.0)	**0.2 (0.0-1.0)**	**0.1 (0.0-0.8)**
Unknown	6.6±2.8 (2.0-12.0)	0.7 (0.2-2.3)	1.0 (0.3-3.4)
Underlying illness‡			
No	6.3±4.3 (1.0-21.0)	2.6 (0.7-9.4)	
Yes	7.6±3.3 (2.0-14.0)	Reference	
Body-mass index			
Underweight	6.5±5.7 (3.0-15.0)	0.3 (0.1-1.8)	
Normal weight	6.7±4.4 (1.0-20.0)	0.5 (0.1-1.6)	
Overweight	5.9±3.4 (2.0-14.0)	Reference	
Obese	6.7±4.2 (2.0-21.0)	0.5 (0.1-1.6)	
Current cigarette smoke (self-reported)			
No	6.3±3.7 (1.0-21.0)	Reference	
Yes	7.6±5.9 (3.0-20.0)	0.5 (0.1-1.9)	
Serostatus at baseline			
Seronegative	6.6±3.4 (1.0-15.0)	1.3 (0.5-3.2)	
Seropositive	6.1±4.3 (2.0-21.0)	Reference	
Unknown	8.5±7.9 (2.0-20.0)	0.6 (0.1-4.4)	

*Estimated using Weibull accelerated failure time regression adjusted for clustering by site and household. Hazard ratio <1 corresponds to prolonged serial interval compared to reference group. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity.

Supplementary Table 5. COVID-19 related hospitalisations (n=11) and deaths (n=2), Klerksdorp and Soweto, South Africa, 2020-2021.

Index/ Contact	Age group	Comorbidities	Outcome	Length of admission (days)	Admission diagnosis	Infecting variant
Index	35-59	Hypertension	Discharged alive	6	Shortness of breath	Beta
Contact	≥60	None	Discharged alive	15	Pneumonia	Beta
Contact	13-17	Congenital heart disease	Discharged alive	2	Shortness of breath	Alpha
Index	≥60	None	Discharged alive	3	Shortness of breath	Beta
Index	≥60	Hypertension, diabetes	Discharged alive	13	Shortness of breath	Beta
Index	≥60	None	Discharged alive	4	Shortness of breath	Alpha
Index	35-59	Hypertension	Died	30	Lower respiratory tract infection	Unknown
Index	35-59	None	Discharged alive	3	Shortness of breath	Delta
Contact	≥60	Asthma	Died	8	Pneumonia	Delta
Contact	≥60	None	Discharged alive	11	Not available	non-Alpha/Beta /Delta
Contact	≥60	None	Discharged alive	12	Shortness of breath	Beta

Supplementary Table 6. Factors associated with SARS-CoV-2 household transmission in index cases and acquisition in household contacts in sensitivity analysis when including individuals seropositive at baseline with no rRT-PCR infection during follow-up as susceptible, Klerksdorp and Soweto, South Africa, 2020-2021, (n=444).

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Index characteristics			
Site			
Klerksdorp	95/208 (45.7)	Reference	
Soweto	125/236 (53.0)	1.4 (0.7-2.7)	1.2 (0.6-2.4)
Age (years)			
18-34	54/131 (41.2)	Reference	Reference
35-59	137/257 (53.3)	1.9 (0.8-4.2)	**3.5 (1.5-8.6)**
≥60	29/56 (51.8)	1.4 (0.5-4.6)	2.5 (0.7-8.3)
Sex			
Male	64/133 (48.1)	0.9 (0.4-2.0)	
Female	156/311 (50.2)	Reference	
HIV			
HIV-infected	51/103 (49.5)	1.0 (0.4-2.3)	
HIV-uninfected	163/331 (49.2)	Reference	
Unknown HIV-status	6/10 (60.0)	0.9 (0.1-14.3)	
CD4+ T cell count			
HIV-infected, not immune suppressed†	30/60 (50.0)	Reference	Reference
HIV-infected, immune suppressed†	8/18 (44.4)	1.4 (0.2-11.2)	1.8 (0.2-12.5)
HIV-infected, no CD4+ T cell count	13/25 (52.0)	1.0 (0.2-5.6)	0.9 (0.2-5.0)
HIV-uninfected	163/331 (49.2)	1.1 (0.4-3.0)	1.6 (0.5-4.6)
Unknown HIV-status	6/10 (60.0)	1.0 (0.1-17.4)	0.5 (0.0-8.0)
Underlying illness‡			
No	218/432 (50.5)	3.3 (0.3-41.6)	
Yes	2/7 (28.6)	Reference	
Unknown	0/5 (0.0)	Excluded	
Body-mass index			
Underweight	3/10 (30.0)	0.3 (0.0-2.7)	**0.1 (0.0-1.2)**
Normal weight	68/125 (54.4)	Reference	Reference
Overweight	44/86 (51.2)	0.9 (0.3-2.4)	0.4 (0.1-1.1)
Obese	105/218 (48.2)	0.7 (0.3-1.7)	**0.4 (0.2-1.0)**
Unknown	0/5 (0.0)	Excluded	Excluded
Crowding			
No	56/116 (48.3)	1.0 (0.5-2.2)	
Yes	164/328 (50.0)	Reference	
Child <5 years in home			
No	190/381 (49.9)	1.4 (0.5-3.9)	
Yes	30/63 (47.6)	Reference	
Current cigarette smoke (self-reported)			
No	203/391 (51.9)	2.6 (0.9-7.6)	
Yes	17/48 (35.4)	Reference	
Unknown	0/5 (0.0)	Excluded	

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Episode duration			
1-10 days	64/147 (43.5)	Reference	
11-21 days	53/115 (46.1)	1.3 (0.5-3.2)	
>21 days	101/176 (57.4)	**2.3 (1.0-5.3)**	
Positive at last visit	2/6 (33.3)	0.6 (0.0-10.7)	
Fever			
No	110/218 (50.5)	Reference	
Yes	110/226 (48.7)	1.2 (0.6-2.3)	
Minimum C_t value			
>35	9/45 (20.0)	Reference	Reference
25-35	86/174 (49.4)	**6.3 (1.8-22.4)**	**5.1 (1.3-20.1)**
<25	123/222 (55.4)	**8.2 (2.4-28.4)**	**5.4 (1.4-20.3)**
Unknown	2/3 (66.7)	14.6 (0.3-748.6)	11.7 (0.2-594.6)
Serostatus at follow-up end			
Seronegative	4/7 (57.1)	0.7 (0.0-12.3)	
Seropositive	182/343 (53.1)	Reference	
Unknown	34/94 (36.2)	0.4 (0.1-0.9)	
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	18/32 (56.3)	1.4 (0.4-4.6)	1.9 (0.6-6.2)
Alpha	16/25 (64.0)	1.8 (0.4-7.6)	2.3 (0.5-10.0)
Beta	144/302 (47.7)	Reference	Reference
Delta	40/58 (69.0)	**2.6 (0.9-7.5)**	3.0 (1.0-9.4)
Unknown	2/27 (7.4)	0.0 (0.0-0.3)	0.1 (0.0-0.5)
Contact characteristics			
Age (years)			
<5	5/19 (26.3)	Reference	Reference
5-12	40/80 (50.0)	3.3 (0.8-13.6)	1.8 (0.4-7.9)
13-17	43/67 (64.2)	**9.7 (2.2-42.9)**	**6.1 (1.3-29.0)**
18-34	56/124 (45.2)	**3.7 (0.9-14.8)**	2.8 (0.6-12.0)
35-59	56/109 (51.4)	**3.8 (0.9-15.6)**	3.0 (0.7-13.3)
≥60	20/45 (44.4)	**3.4 (0.7-15.8)**	2.1 (0.4-10.3)
Sex			
Male	83/188 (44.2)	Reference	
Female	137/256 (53.5)	**1.7 (1.1-2.8)**	
HIV			
HIV-uninfected	82/170 (48.2)	1.3 (0.5-3.4)	
HIV-infected	16/35 (45.7)	Reference	
HIV status unknown	122/239 (51.1)	1.1 (0.5-2.9)	
Underlying illness‡			
No	213/425 (50.1)	Reference	
Yes	6/10 (60.0)	1.2 (0.2-6.6)	
Unknown	1/9 (11.1)	0.1 (0.0-0.7)	
Body-mass index			
Underweight	15/29 (51.7)	1.0 (0.4-2.8)	
Normal weight	96/199 (48.2)	Reference	
Overweight	44/100 (44.0)	0.9 (0.5-1.7)	

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Obese	64/107 (59.8)	**1.9 (1.0-3.6)**	
Unknown	1/9 (11.1)	0.1 (0.0-0.9)	
Current cigarette smoke (self-reported)			
No	197/377 (52.3)	**2.7 (1.3-5.8)**	**3.1 (1.4-6.9)**
Yes	22/62 (35.5)	Reference	Reference
Unknown	1/5 (20.0)	0.3 (0.0-6.0)	Excluded
Sleep in same room as index			
No	150/316 (47.5)	Reference	
Yes	70/128 (54.7)	1.4 (0.8-2.3)	
Index is main caregiver			
No	183/375 (48.8)	1.1 (0.5-2.2)	
Yes	37/69 (53.6)	Reference	

OR – Odds ratio, aOR – Adjusted odds ratio. *Household cumulative infection risk (HCIR) excludes household with mixed SARS-CoV-2 variant clusters. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes $p<0.05$.

Supplementary Table 7. Factors associated with SARS-CoV-2 household transmission in index cases and acquisition in household contacts sensitivity analysis including only households where 65% of household members completed 65% of follow-up during the first three weeks, Klerksdorp and Soweto, South Africa, 2020-2021, (n=342).

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
Index characteristics			
Site			
Klerksdorp	92/174 (52.9)	Reference	
Soweto	108/168 (64.3)	1.7 (0.8-3.7)	1.2 (0.6-2.7)
Age (years)			
18-34	48/94 (51.1)	Reference	Reference
35-59	124/198 (62.6)	1.7 (0.7-4.3)	**4.2 (1.6-11.3)**
≥60	28/50 (56.0)	1.3 (0.3-4.9)	2.2 (0.6-7.9)
Sex			
Male	59/103 (57.3)	0.9 (0.4-2.1)	
Female	141/239 (59.0)	Reference	
HIV			
HIV-infected	42/75 (56.0)	0.7 (0.3-1.9)	
HIV-uninfected	152/258 (58.9)	Reference	
Unknown HIV-status	6/9 (66.7)	0.7 (0.0-12.6)	
CD4+ T cell count			
HIV-infected, immune not suppressed†	26/47 (55.3)	Reference	Reference
HIV-infected, immune suppressed†	6/10 (60.0)	2.0 (0.1-26.9)	2.1 (0.2-18.3)
HIV-infected, no CD4+ T cell count	10/18 (55.6)	0.9 (0.1-6.5)	1.1 (0.2-5.8)
HIV-uninfected	152/258 (58.9)	1.5 (0.5-4.9)	2.5 (0.8-7.7)
Unknown HIV-status	6/9 (66.7)	1.0 (0.0-22.3)	0.5 (0.0-7.0)
Underlying illness¶			
No	154/261 (59.0)	1.0 (0.4-2.7)	
Yes	46/79 (58.2)	Reference	
Unknown	0/2 (0.0)	Excluded	
Body-mass index			
Underweight	2/7 (28.6)	0.2 (0.0-3.1)	**0.0 (0.0-0.6)**
Normal weight	61/100 (61.0)	Reference	Reference
Overweight	40/58 (69.0)	1.5 (0.5-5.1)	0.6 (0.2-2.0)
Obese	97/175 (55.4)	0.9 (0.3-2.1)	**0.4 (0.1-0.9)**
Unknown	0/2 (0.0)	Excluded	Excluded
Crowding			
No	48/74 (64.9)	1.6 (0.6-4.1)	
Yes	152/268 (56.7)	Reference	
Child <5 years in home			
No	173/288 (60.1)	2.2 (0.7-6.8)	
Yes	27/54 (50.0)	Reference	
Current cigarette smoke (self-reported)			
No	186/305 (61.0)	3.2 (0.9-11.0)	
Yes	14/35 (40.0)	Reference	
Unknown	0/2 (0.0)	Excluded	
Episode duration			
1-10 days	58/116 (50.0)	Reference	

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
11-21 days	48/84 (57.1)	1.8 (0.7-5.2)	
>21 days	92/136 (67.7)	**2.9 (1.1-7.3)**	
Positive at last visit	2/5 (40.0)	0.5 (0.0-11.3)	
Fever			
No	103/176 (58.5)	Reference	
Yes	97/166 (58.4)	1.1 (0.5-2.4)	
Minimum C_t value			
>35	7/36 (19.4)	Reference	Reference
25-35	79/136 (58.1)	**9.5 (2.4-37.5)**	**7.9 (2.0-32.3)**
<25	112/166 (67.5)	**15.0 (3.8-58.6)**	**5.9 (1.5-23.3)**
Unknown	2/4 (50.0)	4.6 (0.2-113.5)	8.0 (0.3-186.0)
Serostatus at follow-up end			
Seronegative	4/7 (57.1)	0.4 (0.0-7.0)	
Seropositive	174/284 (61.3)	Reference	
Unknown	22/51 (43.1)	0.3 (0.1-1.0)	
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	16/25 (64.0)	1.2 (0.3-4.5)	1.9 (0.5-6.5)
Alpha	13/17 (76.5)	2.3 (0.4-12.6)	2.4 (0.5-12.8)
Beta	130/227 (57.3)	Reference	Reference
Delta	39/47 (83.0)	**4.1 (1.2-13.9)**	**5.2 (1.4-19.9)**
Unknown	2/26 (7.7)	0.0 (0.0-0.2)	0.0 (0.0-0.3)
Contact characteristics			
Age (years)			
<5	5/16 (31.3)	Reference	Reference
5-12	37/71 (52.1)	2.8 (0.6-13.2)	1.5 (0.3-6.5)
13-17	41/58 (70.7)	**11.7 (2.3-60.4)**	**6.5 (1.3-32.6)**
18-34	47/79 (59.5)	**5.0 (1.1-23.3)**	**4.3 (0.9-20.0)**
35-59	53/87 (60.9)	**4.3 (0.9-19.7)**	3.7 (0.8-17.2)
≥60	17/31 (54.8)	**4.2 (0.7-24.0)**	2.5 (0.4-13.7)
Sex			
Male	75/149 (50.3)	Reference	
Female	125/193 (64.8)	**2.3 (1.3-4.2)**	
HIV			
HIV-uninfected	80/146 (54.8)	1.4 (0.4-4.4)	
HIV-infected	14/26 (53.9)	Reference	
HIV status unknown	106/170 (62.4)	1.6 (0.5-4.8)	
Underlying illness¶			
No	171/299 (57.2)	Reference	
Yes	28/39 (71.8)	3.2 (1.1-9.0)	
Unknown	1/4 (25.0)	0.2 (0.0-5.4)	
Body-mass index			
Underweight	13/24 (54.2)	0.9 (0.3-2.7)	
Normal weight	88/166 (53.0)	Reference	
Overweight	42/72 (58.3)	1.4 (0.7-3.0)	
Obese	56/76 (73.7)	**3.9 (1.7-9.0)**	
Unknown	1/4 (25.0)	0.2 (0.0-6.7)	
Current cigarette smoke (self-reported)			

Characteristic	HCIR*, n/N (%)	Univariate, OR (95%CI)	Multivariable, aOR (95%CI)
No	179/292 (61.3)	**3.2 (1.4-7.8)**	**4.2 (1.7-10.4)**
Yes	20/47 (42.6)	Reference	Reference
Unknown	1/3 (33.3)	0.6 (0.0-23.7)	Excluded
Sleep in same room as index			
No	135/242 (55.8)	Reference	
Yes	65/100 (65.0)	1.5 (0.8-2.8)	
Index is main caregiver			
No	163/280 (58.2)	1.2 (0.5-2.7)	
Yes	37/62 (59.7)	Reference	

OR – Odds ratio, aOR – Adjusted odds ratio. *Household cumulative infection risk (HCIR) excludes household members seropositive at baseline with no SARS-CoV-2 detected on rRT-PCR during follow-up and households with mixed SARS-CoV-2 variant clusters. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes $p<0.05$.

Supplementary Table 8. Factors associated with SARS-CoV-2 episode duration in index cases sensitivity analysis including only households where 65% of household members completed 65% of follow-up during the first three weeks, Klerksdorp and Soweto, South Africa, 2020-2021, (n=117). HR – Hazard ratio*

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Site			
Klerksdorp	17.1±12.3 (4.0-45.0)	Reference	Reference
Soweto	22.2±13.4 (3.0-47.0)	**0.7 (0.5-1.0)**	**0.5 (0.3-0.8)**
Age (years)			
18-34	14.7±10.7 (4.0-41.0)	Reference	Reference
35-59	19.8±13.6 (3.0-47.0)	0.6 (0.4-0.9)	**0.3 (0.2-0.5)**
≥60	27.9±10.7 (6.0-44.0)	**0.4 (0.2-0.7)**	**0.2 (0.1-0.4)**
Sex			
Male	20.6±14.4 (3.0-47.0)	Reference	
Female	19.2±12.5 (3.0-46.0)	1.2 (0.8-1.8)	
HIV			
HIV-infected	20.0±13.0 (3.0-45.0)	Reference	
HIV-uninfected	19.6±13.2 (3.0-47.0)	1.0 (0.6-1.5)	
Unknown HIV-status	18.5±10.6 (11.0-26.0)	1.3 (0.3-5.6)	
CD4+ T cell count			
HIV-infected, immune not suppressed†	21.7±14.1 (3.0-45.0)	Reference	Reference
HIV-infected, immune suppressed†	20.7±15.5 (8.0-38.0)	1.2 (0.3-4.2)	0.8 (0.2-3.6)
HIV-infected, no CD4+ T cell count	15.3±9.3 (9.0-33.0)	1.9 (0.7-5.0)	0.8 (0.3-2.5)
HIV-uninfected	19.6±13.2 (3.0-47.0)	1.1 (0.6-2.0)	0.8 (0.4-1.6)
Unknown HIV-status	18.5±10.6 (11.0-26.0)	1.5 (0.4-6.8)	0.8 (0.2-4.2)
Underlying illness¶			
No	19.4±13.0 (3.0-45.0)	Reference	
Yes	20.7±13.4 (3.0-47.0)	0.8 (0.5-1.3)	
Unknown	9.0±0 (9.0-9.0)	4.2 (1.0-17.8)	
Body-mass index			
Underweight	13.5±6.4 (9.0-18.0)	**2.7 (0.6-11.5)**	3.5 (0.7-18.2)
Normal weight	21.6±14.0 (6.0-45.0)	Reference	Reference
Overweight	20.5±15.6 (3.0-47.0)	1.1 (0.6-1.8)	**2.0 (1.0-3.9)**
Obese	18.6±11.6 (4.0-46.0)	1.3 (0.8-2.0)	**2.2 (1.3-3.8)**
Unknown	9.0±0 (9.0-9.0)	4.5 (0.6-34.0)	1.3 (0.2-11.3)
Crowding			
No	17.6±11.9 (3.0-43.0)	Reference	
Yes	20.5±13.5 (3.0-47.0)	0.8 (0.5-1.2)	
Child <5 years in home			
No	20.5±13.1 (3.0-47.0)	0.6 (0.3-1.1)	
Yes	13.9±11.3 (4.0-35.0)	Reference	
Current cigarette smoke (self-reported)			
No	20.4±13.2 (3.0-47.0)	0.6 (0.3-1.1)	
Yes	14.7±11.3 (3.0-37.0)	Reference	
Unknown	9.0±0 (9.0-9.0)	2.7 (0.3-20.4)	
Fever			
No	17.1±11.9 (3.0-46.0)	Reference	

Characteristic	Episode duration, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Yes	22.3±13.8 (3.0-47.0)	**0.6 (0.4-0.9)**	
Minimum C_t value			
>35	11.3±7.0 (3.0-26.0)	Reference	Reference
25-35	16.2±13.4 (3.0-47.0)	**0.5 (0.2-0.9)**	0.5 (0.3-1.1)
<25	24.6±12.1 (7.0-44.0)	**0.3 (0.1-0.5)**	**0.3 (0.1-0.6)**
Unknown	8.0±0 (8.0-8.0)	2.2 (0.3-17.7)	6.2 (0.7-53.4)
Serostatus at follow-up end			
Seronegative	7.5±2.1 (6.0-9.0)	Reference	Reference
Seropositive	21.1±13.3 (3.0-47.0)	**0.1 (0.0-0.5)**	**0.0 (0.0-0.2)**
Unknown	13.2±9.7 (3.0-38.0)	0.3 (0.1-1.2)	**0.1 (0.0-0.7)**
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	14.6±9.8 (7.0-38.0)	2.1 (1.1-4.2)	**3.3 (1.5-6.9)**
Alpha	23.8±14.9 (7.0-41.0)	0.7 (0.3-1.8)	0.6 (0.2-1.5)
Beta	20.8±13.8 (3.0-47.0)	Reference	Reference
Delta	21.8±11.8 (4.0-44.0)	1.2 (0.6-2.2)	**2.2 (1.1-4.5)**
Unknown	10.9±6.9 (5.0-28.0)	2.8 (1.4-5.6)	1.4 (0.6-3.2)

*Estimated using Weibull accelerated failure time regression adjusted for clustering by site and household. Hazard ratio <1 corresponds to longer episode duration than reference group used. Excluded 8 index cases positive at last specimen collected. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes $p<0.05$.

Supplementary Table 9. Factors associated with serial interval of SARS-CoV-2 in cases with symptomatic illness sensitivity analysis including only households where 65% of household members completed 65% of follow-up during the first three weeks, Klerksdorp and Soweto, South Africa, 2020-2021, (n=60). HR – Hazard ratio*

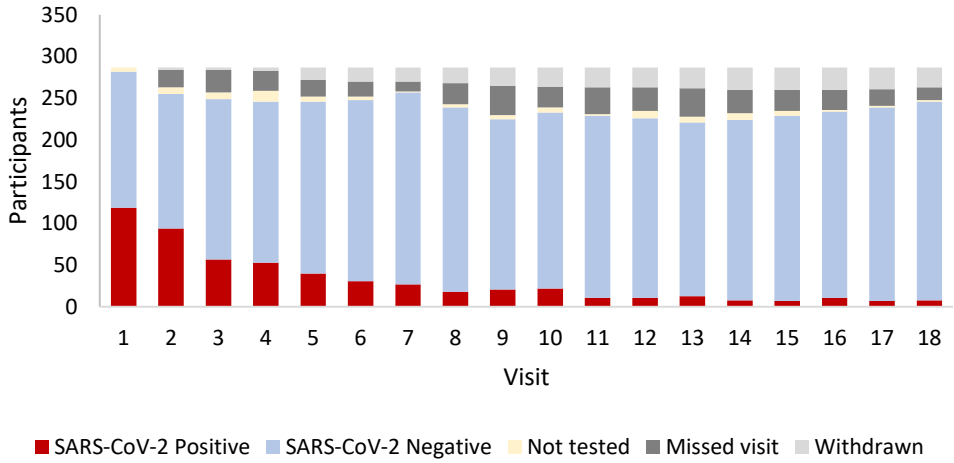
Characteristic	Serial interval, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Index characteristics			
Site			
Klerksdorp	5.9±4.1 (1.0-21.0)	Reference	
Soweto	7.5±4.0 (2.0-20.0)	0.4 (0.1-1.5)	0.3 (0.1-1.4)
Age (years)			
18-34	6.6±4.6 (1.0-15.0)	Reference	
35-59	6.6±3.6 (2.0-21.0)	0.5 (0.1-2.7)	
≥60	7.0±6.4 (2.0-20.0)	0.9 (0.1-7.7)	
Sex			
Male	6.6±4.7 (2.0-21.0)	1.3 (0.3-5.6)	
Female	6.7±3.9 (1.0-20.0)	Reference	
HIV			
HIV-infected	11.7±7.6 (3.0-21.0)	0.1 (0.0-0.6)	
HIV-uninfected	6.1±3.2 (1.0-15.0)	Reference	
Underlying illness‡			
No	7.1±4.4 (1.0-21.0)	Reference	
Yes	17.0±5.5 (2.0-14.0)	1.7 (0.4-6.8)	
Body-mass index			
Normal weight	7.2±5.8 (2.0-21.0)	1.3 (0.2-8.4)	
Overweight	6.9±3.5 (2.0-15.0)	Reference	
Obese	6.3±3.7 (1.0-20.0)	1.3 (0.2-7.0)	
Crowding			
No	5.8±4.1 (1.0-15.0)	Reference	
Yes	7.1±4.1 (2.0-21.0)	0.6 (0.1-2.3)	
Child <5 years in home			
No	6.6±4.2 (1.0-21.0)	0.4 (0.0-3.3)	
Yes	6.9±3.6 (2.0-12.0)	Reference	
Current cigarette smoke (self-reported)			
No	6.8±4.2 (1.0-21.0)	0.2 (0.0-5.1)	
Yes	4.3±0.5 (4.0-5.0)	Reference	
Episode duration			
1-10 days	7.2±3.9 (2.0-15.0)	1.5 (0.2-9.1)	
11-21 days	7.4±4.6 (1.0-21.0)	Reference	
>21 days	5.9±4.0 (2.0-20.0)	3.7 (0.6-21.6)	
Unknown			
Fever			
No	6.9±4.8 (1.0-20.0)	1.5 (0.4-5.7)	
Yes	6.5±3.7 (2.0-21.0)	Reference	
Minimum C_t value			
25-35	6.6±5.7 (1.0-21.0)	0.6 (0.2-2.5)	
<25	6.7±3.1 (2.0-15.0)	Reference	
Serostatus at end			

Characteristic	Serial interval, days, Mean ± SD (Range)	Univariate, HR* (95%CI)	Multivariable, aHR (95%CI)
Seronegative	6.1±3.0 (2.0-14.0)	Reference	
Seropositive	7.1±4.7 (1.0-21.0)	0.5 (0.1-2.2)	
Unknown	5.0±2.0 (3.0-7.0)	1.7 (0.1-23.5)	
SARS-CoV-2 variant			
non-Alpha/Beta/Delta	11.0±8.2 (6.0-12.0)	0.2 (0.0-0.9)	
Alpha	4.2±0.8 (3.0-5.0)	2.5 (0.2-30.5)	
Beta	6.0±4.7 (1.0-21.0)	Reference	
Delta	8.2±4.1 (2.0-15.0)	0.4 (0.0-2.5)	
Contact characteristics			
Age (years)			
<5	5.5±0.7 (5.0-6.0)	1.9 (0.1-25.3)	0.5 (0.0-12.1)
5-12	6.0±3.3 (1.0-10.0)	0.8 (0.2-3.6)	0.7 (0.2-3.0)
13-17	6.3±3.5 (2.0-12.0)	0.4 (0.1-1.4)	0.3 (0.1-1.1)
18-34	4.6±2.6 (2.0-9.0)	Reference	Reference
35-59	7.5±4.7 (2.0-20.0)	**0.4 (0.1-1.2)**	**0.3 (0.1-1.0)**
≥60	8.2±5.0 (3.0-21.0)	0.3 (0.1-1.2)	**0.2 (0.1-1.0)**
Sex			
Male	6.2±4.6 (1.0-20.0)	Reference	
Female	6.9±3.8 (2.0-21.0)	0.8 (0.3-1.9)	
HIV			
HIV-uninfected	5.5±4.0 (1.0-21.0)	Reference	Reference
HIV-infected	11.1±5.6 (3.0-20.0)	**0.2 (0.0-0.9)**	**0.2 (0.0-0.8)**
Unknown	6.7±2.7 (2.0-12.0)	0.7 (0.2-2.4)	1.1 (0.3-4.2)
Underlying illness‡			
No	6.4±4.3 (1.0-21.0)	2.1 (0.6-7.6)	
Yes	7.5±3.5 (2.0-14.0)	Reference	
Body-mass index			
Underweight	6.5±5.7 (3.0-15.0)	0.3 (0.0-1.6)	
Normal weight	6.7±4.4 (1.0-20.0)	0.4 (0.1-1.4)	
Overweight	5.7±3.4 (2.0-14.0)	Reference	
Obese	7.1±4.1 (2.0-21.0)	0.4 (0.1-1.3)	
Current cigarette smoke (self-reported)			
No	6.4±3.7 (1.0-21.0)	Reference	
Yes	7.6±5.9 (3.0-20.0)	0.6 (0.2-2.0)	
Serostatus at baseline			
Seronegative	6.5±3.4 (1.0-15.0)	1.3 (0.5-3.3)	
Seropositive	6.3±4.3 (2.0-21.0)	Reference	
Unknown	10.7±8.1 (5.0-20.0)	0.1 (0.0-1.8)	

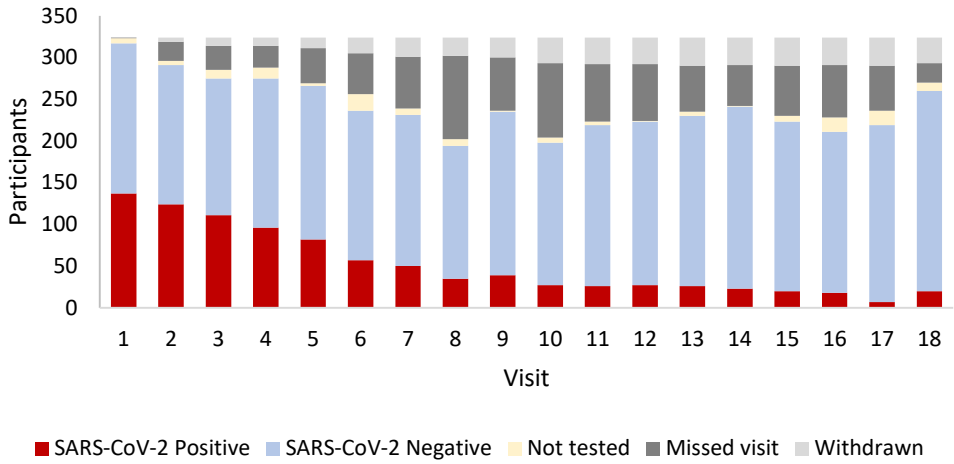
*Estimated using Weibull accelerated failure time regression adjusted for clustering by site and household. Hazard ratio <1 corresponds to prolonged serial interval. †Immune suppressed defined as CD4+ T cell count <200 cells/ml. ‡Underlying illness: self-reported history of diabetes, hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity. Bold face font denotes p<0.05.

Supplementary figures

a)

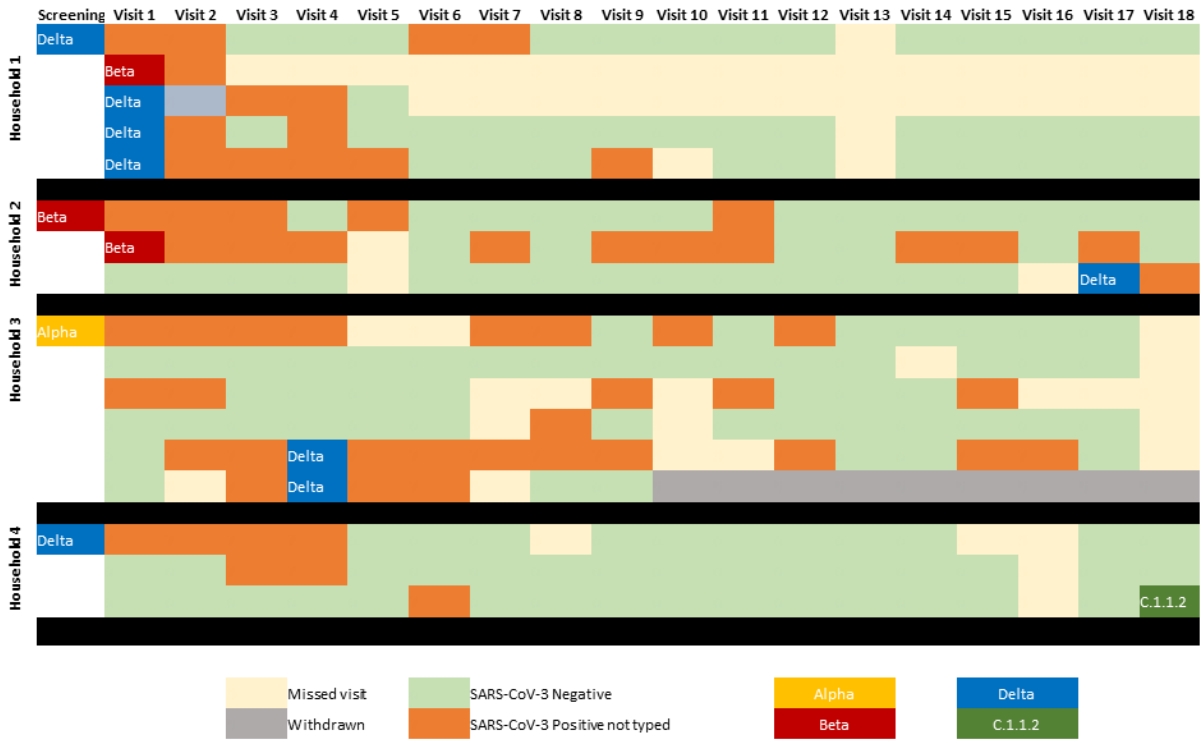


b)



Supplementary Figure 1. SARS-CoV-2 positivity by follow-up visit for index cases and contacts*, a) Klerksdorp (n=287) and b) Soweto (n=325), South Africa, 2020-2021.

*Not tested: sample collected but not tested due to lost in transit, leakage or labelling error. Missed visit: visit not conducted as individual was not available.



304

305 **Supplementary Figure 2.** Mixed SARS-CoV-2 clusters detected during follow-up, Klerksdorp and
306 Soweto, South Africa, 2020-2021. Black line denotes household separation

307

References

1. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. South Afr J HIV Med **2017**; 18(1): 694.
2. Tempia S, Walaza S, Moyes J, et al. Attributable Fraction of Influenza Virus Detection to Mild and Severe Respiratory Illnesses in HIV-Infected and HIV-Uninfected Patients, South Africa, 2012-2016. Emerg Infect Dis **2017**; 23(7): 1124-32.
3. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform **2019**; 95: 103208-18.
4. Wibmer CK, Ayres F, Hermanus T, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med **2021**; 27(4): 622-5.
5. Shastri MD, Shukla SD, Chong WC, et al. Smoking and COVID-19: What we know so far. Respir Med **2021**; 176: 106237-.
6. Vallarta-Robledo JR, Sandoval JL, Baggio S, et al. Negative Association Between Smoking and Positive SARS-CoV-2 Testing: Results From a Swiss Outpatient Sample Population. Frontiers in Public Health **2021**; 9.
7. Meiring S, Tempia S, Bhiman JN, et al. Prolonged shedding of SARS-CoV-2 at high viral loads amongst hospitalised immunocompromised persons living with HIV, South Africa. Clin Infect Dis **2022**.
8. Sun K, Tempia S, Kleynhans J, et al. Persistence of SARS-CoV-2 immunity, Omicron's footprints, and projections of epidemic resurgences in South African population cohorts. medRxiv **2022**: 2022.02.11.22270854.
9. Ong SWX, Chiew CJ, Ang LW, et al. Clinical and Virological Features of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Variants of Concern: A Retrospective Cohort Study Comparing B.1.1.7 (Alpha), B.1.351 (Beta), and B.1.617.2 (Delta). Clin Infect Dis **2021**.
10. Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. Euro Surveill **2020**; 25(32): 2001483.

Association of close-range contact patterns with SARS-CoV-2 household transmission, South Africa, 2020-2021

Author list

Jackie Kleynhans, MPH (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Lorenzo Dall'Amico, PhD (ISI Foundation, Turin, Italy.)

Laetitia Gauvin, PhD (ISI Foundation, Turin, Italy.)

Michele Tizzoni, PhD (ISI Foundation, Turin, Italy.)

Lucia Maloma, BA(Hons) (Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.)

Sibongile Walaza, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Neil A. Martinson, MD (Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. Johns Hopkins University Center for TB Research, Baltimore, Maryland, United States of America.)

Anne von Gottberg, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Nicole Wolter, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Mvuyo Makhasi, MSc (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Cheryl Cohen, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Ciro Cattuto, PhD (ISI Foundation, Turin, Italy. University of Turin, Department of Informatics, Turin, Italy.)

Stefano Tempia, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

For the SA-S-HTS Group

Keywords: SARS-CoV-2, transmission, household, contacts

Abstract (234/250)

Background

Households are an important location for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission, especially during periods where travel and work was restricted to essential services. We aimed to assess the association of close-range contact patterns with SARS-CoV-2 transmission.

Methods

We deployed proximity sensors for two weeks to measure face-to-face interactions between household members after SARS-CoV-2 was identified in the household, in South Africa, 2020 - 2021. We calculated duration, frequency and average duration of close range proximity events with SARS-CoV-2 index cases. We assessed the association of contact parameters with SARS-CoV-2 transmission using mixed effects logistic regression accounting for index and household member characteristics.

Results

We included 342 individuals (89 SARS-CoV-2 index cases and 253 household members). On multivariable analysis, factors associated with SARS-CoV-2 acquisition were index cases with minimum C_t value <30 (aOR 9.8 95%CI 1.3-71.8) vs >35, being infected with the Delta variant (aOR 3.9 95%CI 1.0-14.9) vs Beta, contacts aged 13-17 years (aOR 8.0 95%CI 1.0-62.2) vs <5 years and female contacts (aOR 2.4 95%CI 1.1-4.8). No contact parameters were associated with acquisition (aOR 1.0 95%CI 1.0-1.0) for all three of duration, frequency and average duration.

Conclusion

We did not find an association between close-range proximity events and SARS-CoV-2 household transmission. It may be that droplet-mediated transmission during close-proximity contacts play a smaller role than airborne transmission of household SARS-CoV-2, due to high contact rates in households or study limitations.

Introduction

South Africa has experienced five waves of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, with over 4 million laboratory-confirmed cases by August 2022 [1]. The true burden is highly underestimated, since based on seroprevalence data, after the third wave of infection, 43% to 83% of the 59.5 million South African inhabitants had already been infected, varying by age and setting [2,3].

SARS-CoV-2 transmission is mainly via the respiratory route, with droplet-mediated transmission thought to be the most important but airborne transmission also occurs [4,5]. Infection from contaminated surfaces has also been described [4]. Although infection risk is highest from symptomatic individuals [6], with the most infectious period one day before symptom onset [4], asymptomatic individuals can still transmit SARS-CoV-2 [7,8]. Households are a focal point for SARS-CoV-2 transmission [9,10], especially during peaks of non-pharmaceutical intervention (NPI) restrictions, when movement outside of the household was limited [9]. Transmission within households can in turn lead to spill over to the community [11].

Prior to the widespread availability of SARS-CoV-2 vaccines, most countries relied on NPIs to reduce the transmission of the virus, including wearing face masks, social and physical distancing. While mobility and contact survey data showed that the implementation of NPIs led to a reduction in community contacts [9,12] and in turn opportunity for infection, it is still unknown what the role of contact patterns are in the transmission of SARS-CoV-2 in the household. Most analysis relating contact patterns and SARS-CoV-2 transmission done to date has been based on low resolution data collected from contact tracing [13], mobility data [9] and contact surveys [12]. To obtain high-resolution contact data, devices broadcasting and receiving radio frequency waves can be used to measure the frequency and duration of close-proximity contacts. This has been used previously to collect contact data in among others, schools [14], workplaces [15], hospitals [16] and households [17], which can, in turn, be used for modelling disease transmission. Specifically, for SARS-CoV-2 so far, high-resolution contact data were collected on cruise ships to identify areas of high contact, and to investigate the usefulness of NPIs [18].

Understanding the drivers of SARS-CoV-2 transmission in the household, especially contact patterns, can help inform NPIs for future SARS-CoV-2 resurgences and potentially future emerging pathogens with pandemic potential. We aimed to assess the association of household close-range contact patterns with the transmission of SARS-CoV-2 in the household using proximity sensors deployed after the identification of SARS-CoV-2 in the household.

Methods

Screening, enrolment and follow-up

We nested a contact study within a case-ascertained, prospective, household transmission study for SARS-CoV-2, implemented in two urban communities in South Africa, Klerksdorp (North West Province) and Soweto (Gauteng Province) from October 2020 through September 2021. Methods for the main study have been reported previously [19]. In short, symptomatic adults (aged ≥18 years, symptom onset ≤5 days prior) consulting at clinics were screened for SARS-CoV-2 with real-time reverse transcription polymerase chain reaction (rRT-PCR) on nasopharyngeal swabs. We enrolled household contacts of SARS-CoV-2 infected individuals identified through screening (presumptive index) with ≥2 household contacts of whom none reported symptoms prior to index case onset. We visited enrolled households three times a week to collect nasal swabs and data on symptoms and

healthcare seeking. At enrolment household characteristics (household size, number of rooms used for sleeping, smoking inside the household and household income) and individual characteristics (demographics, education, employment, smoking, HIV infection, underlying illness, if SARS-CoV-2 index case was main caregiver, or sleeping in same room as index case) were collected. Nasopharyngeal (screening) and nasal swabs (follow-up) were tested for SARS-CoV-2 on rRT-PCR using the Allplex™ 2019-nCoV kit (Seegene Inc., Seoul, South Korea) and the first positive of each infection episode was characterised using the Allplex™ SARS-CoV-2 Variants I and II PCR assays (Seegene Inc., Seoul, Korea) and through whole genome sequencing on the Ion Torrent Genexus platform (Thermo Fisher Scientific, USA). We classified the infection episodes as Alpha, Beta, Delta, non-Alpha/Beta/Delta or unknown variant where we were unable to classify the sample as a variant of concern due to primary testing done elsewhere, low viral load or poor sequence quality. Households with multiple SARS-CoV-2 variants circulating at the same time (mixed clusters) were excluded from the analysis. We also collected serum at the first and final household visit for serological testing, using an in-house ELISA to detect antibodies against SARS-CoV-2 spike protein [20] and nucleocapsid protein using Roche Elecsys anti-SARS-CoV-2 assay. Individuals were considered seropositive if they tested positive on either assay. Individuals sero-positive at the start of follow-up with no rRT-PCR confirmed SARS-CoV-2 infection during follow-up were excluded from the analysis as they may have been protected from infection [21].

Contact pattern measurements

At the first or second visit during follow-up, we deployed wearable radio frequency (RF) proximity sensors [15] for two weeks to measure close-range interactions (<1.5 meters) between household members. The proximity sensors exchange low-power radio packets in the ISM (Industrial, Scientific and Medical) radio band. Exchange of packets and Received Signal Strength Indicator (RSSI), suitably thresholded, are used to assess proximity between the devices. A contact interval between two devices is defined as a sequence of consecutive 20-second intervals within which at least one radio packet was exchanged. Each sensor had a unique hardware identifier that was linked to participant study identifiers. Sensors were worn in a PVC pouch either pinned to clothing on the chest, or on a lanyard around the neck based on participant preference. We requested participants to wear the device while at home, to store them separately from other household member sensors at night, and to complete a log sheet every day for the periods the sensors were put on and taken off. During each household visit during the sensor deployment period, field workers confirmed sensors were worn. A deployment log was completed for each household to link the sensor identifier to the participant identifier and to log the date and time sensors were deployed and collected. After sensor collection, batteries were removed to prevent further package exchange between sensors. Sensors were transported to the study office where each sensor was connected to a computer and data downloaded.

Data analysis

We assumed the first individual with COVID-19 compatible symptoms in the household (individual screened at clinic) was the index case. Any household member testing positive for SARS-CoV-2 within the two weeks from the last positive result for the index case was considered a secondary SARS-CoV-2 case. Contact event data were cleaned using an automated pipeline. We excluded any close-range proximity events outside of the deployment period and that occurred during a 5-minute time slice that the accelerometer did not detect any movement of the sensor. Due to a technical

error, some sensors at the Klerksdorp site did not have a valid time stamp and needed additional processing to align the time series of close-range proximity events. This was achieved by computing, for each pair of tags X and Y, the temporal shift that maximizes the correlation between the time series of the number of packets per unit time transmitted by X and received by Y, and the reciprocal time series of the number of packets per unit time transmitted by Y and received by X (operation that can be efficiently carried out working in the frequency domain via Fourier transformation). This allowed us to build a temporal alignment graph between sensors and – as long as there was at least one sensor with a valid timestamp in the household – to use such graph to propagate the valid timestamp to all other sensors, thus recovering global temporal alignment. For the analysis, we only considered close-range proximity events that occurred one day after deployment and one day before collection.

We assessed three contact parameters: 1) duration (median daily cumulative time in contact in seconds), 2) frequency (median daily number of contacts with the index/infected individuals over the deployment period) and 3) average duration (cumulative time in contact divided by the cumulative number of close-range proximity events over full deployment period). Median values were preferred over mean values due to the rightly skewed data, and the different number of days with measured contact data for each household after data cleaning. We assessed contact parameters in two ways: 1) median number of close-range proximity events with the presumptive index case and 2) median number of close-range proximity events with all SARS-CoV-2 infected household members (as confirmed by rRT-PCR). The latter assessment was to take into account that the transmission could have been from any of the infected household members, and not necessarily the index case, or that the index case was misclassified.

We constructed contact matrices by combining the median duration and frequency of close-range proximity events for all participants between each age group, respectively. To normalize the matrix based on number of participants, we divided the cumulative contact duration and frequency by the total number of individuals in the two age groups being investigated in each cell.

We assessed the association of contact parameters with SARS-CoV-2 household transmission using the Wilcoxon rank-sum test (considering $p<0.05$ as significant) and through logistic regression controlling for individual characteristics associated with transmission. To assess factors associated with SARS-CoV-2 household transmission, we performed logistic regression with a mixed effects hierarchical regression model to account for household- and site-level clustering. For the analysis with a defined index case (i.e. investigating close-range proximity events with all presumptive index cases, first person with COVID-19 symptoms), we included only household contacts with their SARS-CoV-2 infection status as the outcome, assessing both index (transmission) and contact (acquisition) characteristics. For the analysis with no defined index case (i.e. investigating close-range proximity events with all SARS-CoV-2 infected household members), we included all enrolled household members (originally considered presumptive index and household contacts), assessing only their own characteristics. For the analysis with close-range proximity events with all SARS-CoV-2 infected household members, we included an offset term in the model to account for the number of SARS-CoV-2 infected members in contact with (number of nodes). We first built the model using individual characteristics to assess factors associated with SARS-CoV-2 transmission (excluding contact parameters). We included age and SARS-CoV-2 variant *a priori*, and assessed other co-variates on univariate analysis, keeping those with $p>0.2$ in the multivariable analysis. For which we then performed backwards elimination, keeping only those with $p<0.05$, and comparing each subsequent model to the previous using a likelihood ratio test. Finally, we generated three separate models for

each analysis (index and infected household members), including each contact parameter to the final model to assess the association with transmission.

Ethics

The study protocol was approved by University of the Witwatersrand Human Research Ethics Committee (Reference M2008114). Participants in follow-up received grocery store vouchers of USD 3 per visit to compensate for time required for specimen collection and interview, and an additional voucher once proximity sensors were returned with no visible damage.

Results

We screened 1,531 individuals and identified 277 (18%) positive for SARS-CoV-2, of which 127 (46%) were enrolled and included in the household cumulative infection risk analysis [19], with 373 household contacts. After data cleaning, we had contact data for 89 (70%) index cases and 253 (69%) household contacts (Supplementary Figure 1). Seventy-three individuals (15%) were excluded due to non-compliance, where no contacts were logged, or sensors was stationary for the period based on accelerometer data. We were more likely to have contact data for individuals from the Soweto site, from larger households, and with no household member reporting smoking indoors (Supplementary Table 1). The median number of household members included in the analysis was 4 (inter quartile range [IQR] 3-5), with a median of 3 (IQR 1-4) SARS-CoV-2 cases per household and a median of 67% (IQR 50-100%) of household members infected (including index cases). Sixty-six percent (227/342) of individuals included in the analysis lived in a household with 3-5 members, and 49% (168/342) lived in a home with only 1-2 rooms used for sleeping, with 78% (265/342) living in households where crowding was reported (>2 people per sleeping room, Supplementary Table 1).

The overall median duration of daily close-range proximity events was 1,240 seconds (21 minutes, IQR 543-2785 seconds), with a 40 second (IQR 33-52 seconds) average duration per contact event, and a median of 28 (IQR 13-63) close-range proximity events per day amongst household members (Figure 1, Supplementary Table 2). The highest median daily contact duration was observed between 5-12 year olds (Figure 2 A), who also had the highest frequency of contacts with all age groups (Figure 2 D). In Klerksdorp we observed a higher median daily close-range proximity duration and frequency in children aged 5-12 and 13-17 years (Figure 2 B, E), whereas in Soweto we observed higher close-range proximity duration and frequency between adults aged 35-59 and younger adults aged 18-34 (Figure 2 C, F).

We did not find any association between any of the contact parameters (either with the index case or all SARS-CoV-2 infected household members) and SARS-CoV-2 infection in the household using the Wilcoxon rank-sum test (p-values ranging 0.14-0.92, Table 1).

When assessing factors associated with SARS-CoV-2 transmission from presumptive index cases and acquisition in household members, a longer average contact duration was associated with a higher risk for SARS-CoV-2 transmission (OR 1.75, 95%CI 1.0 to 3.0) on univariate analysis, but this was no longer significant on multivariable analysis (Table 2). Sleeping in the same room as the index case was also not associated with transmission (OR 0.98, 95%CI 0.49 to 1.94). Furthermore, on multivariable analysis, factors significantly associated with higher SARS-CoV-2 transmission and acquisition was index case minimum C_t value <30 (aOR 9.8 95%CI 1.3-71.8) compared to >35, being infected with the Delta variant (aOR 3.9 95%CI 1.0-14.9) compared to Beta, contacts aged 13-17 years (aOR 8.0 95%CI 1.0-62.2) compared to <5 years and female contacts (aOR 2.4 95%CI 1.1-4.8). Neither close-range proximity duration (aOR 1.0 95%CI 1.0-1.0), frequency (aOR 1.0 95%CI 1.0-1.0)

or average close-range proximity duration (aOR 1.0 95%CI 1.0-1.0) with the index case were associated with acquisition (Table 2).

When not considering transmission from the presumptive index case, but rather all SARS-CoV-2 infected household members, factors significantly associated with SARS-CoV-2 acquisition on multivariable analysis were contacts aged 13-17 years (aOR 10.1 95%CI 1.1-87.3) and 35-59 years (aOR 8.5 95%CI 1.0-70.4) compared to <5 years, being obese (aOR 5.2 95%CI 1.9-14.7) compared to normal weight, and not currently smoking (aOR 3.1 95%CI 1.0-9.6). Neither close-range proximity duration (aOR 1.0 95%CI 1.0-1.0), frequency (aOR 1.0 95%CI 1.0-1.0) or average close-range proximity duration (aOR 1.0 95%CI 1.0-1.0) with SARS-CoV-2 infected household members was associated with acquisition (Table 3).

Discussion

In this case-ascertained, prospective household transmission study we did not find an association between the duration and frequency of close-range proximity events with SARS-CoV-2 infected household members with transmission in the household.

High-resolution contact patterns have been previously used to investigate influenza virus transmission routes in a hospital setting [22], and contact surveys to show the association between contacts, locations and influenza infection [23], the high correlation between pneumococcal infection risk and contact behaviour [24], and in the context of tuberculosis transmission where contact with adults are more important than contact with children [25]. To our knowledge, there are few data available on the direct association of close-range proximity events and SARS-CoV-2, and none making use of high-resolution contact data. During contact tracing efforts early in the pandemic in Singapore, it was found that sharing a bedroom with an index case and speaking to the index case for more 30 minutes or longer increased the risk for infection [26]. We did not see similar results when assessing sharing a bedroom with the index case, and this may be due to the already high level of crowding in included households. Although we observed an increase in infection risk with higher average contact durations with the index case on univariate analysis, this association was no longer seen when adjusting for age and other index and contact factors associated with transmission/acquisition.

There are several possible reasons why we did not observe an association with close-range proximity events and SARS-CoV-2 transmission, these can be classified as related to transmission dynamics or study limitations. One possibility is that along with droplet-mediated transmission during close-proximity contacts, airborne [4,5], and to a lesser extent, fomite-mediated transmission [4] may also play a role in the transmission of SARS-CoV-2 in the household. More evidence is becoming available showing that aerosol transmission may be a more important transmission route for SARS-CoV-2 than initially anticipated, especially so in poorly ventilated indoor environments[5,27]. Households in these communities do not have central air-conditioning or heating [28], and during the winter months ventilation may by poorer than in summer, although we did not measure this. Furthermore, sensors only measure face-to-face interactions, and if individuals were close to each other but not directly facing one another for extended durations, we would not have measured this, although sharing of the same air may have occurred. The ventilation within households should be considered in future studies, as this can be a target for intervention strategies to reduce secondary transmission. The high level of interaction in relatively crowded South African households may already be above the threshold for transmission risk, with host-characteristics like index viral load and contact age being more important to determine infection risk in this context. It is of interest that close-range proximity patterns within the household did not fully account for the differences in transmission based on age;

with teenagers and adults experiencing the highest infection risk, but children aged 5-17 years having the highest contacts.

Our study had limitations both in design and execution. Due to the nature of the case-ascertained study design, we would have missed the period when the index case was most infectious, just before symptom onset [4], and the close-range contact patterns measured during the study may have been different after the household members were aware of the index SARS-CoV-2 case (leading to reduced contact), and again once secondary cases were informed of their infection status (leading to increased contact). We also did not collect any information on possible NPI usage in the households, like wearing of masks. A study from South Africa showed that individuals staying at home were less likely to wear a mask [29], but these data were not ascertained during a time when a household member was infected with SARS-CoV-2. We also did not consider where contacts took place (indoors or outdoors), which relates to ventilation and may have influenced transmission. We may have also misclassified the true index case if they were asymptomatic, and did not consider tertiary transmission chains in the index-directed analysis. To adjust for possible misclassification, we performed a grouped assessment investigating close-range proximity events with all SARS-CoV-2 infected household members. This grouped analysis may also have diluted possible associations with the true infector. We also did not consider multiple introductions within the household, although we did exclude households with more than one SARS-CoV-2 variant detected. During the peaks of waves of infection in South Africa, one variant was responsible for the majority of the infections [30], and the additional introductions within the household were likely to have been the same as the initial variant. Higher resolution sequencing data may be useful to more accurately identify chains of transmission within the household. Combining contact data with clinical and virological/bacteriological data have been shown to be useful to reconstruct transmission networks [31], and we will consider this for future analyses. Our measurement of close-range contact patterns was also limited by compliance, as during the cleaning process we identified 73 sensors that were not worn, based on accelerometer data. We also had limited data in some households, where some individuals did not consent to the contact aspect for the study, or where we were unable to retrieve data due to hardware failure, lost or damaged tags.

In conclusion, we did not observe an association with close-proximity contacts and SARS-CoV-2 transmission in the household. A case-ascertained, prospective household transmission study may not be well suited to investigate this question. A possible other study design to consider is randomly selected prospective household cohorts, but deployment of sensors for extended periods of time may be logistically challenging and lead to participant fatigue and households in a cohort may not experience infection episodes unless the community attack rate is very high. High-resolution contacts in other settings like schools or workplaces where contacts are less frequent could be useful to identify the type of contact events that may lead to SARS-CoV-2 transmission. It may be that aerosol transmission plays a more important role than droplet-mediated transmission, which would make ventilation within households can also be an important consideration for future studies, and increased ventilation can be a method to reduce secondary transmission in households. Nevertheless, our study provides high-resolution household contact data that can be used to parametrise future transmission models.

Competing interests

CC has received grant support from Sanofi Pasteur, US Centers for Disease Control and Prevention, Welcome Trust, Programme for Applied Technologies in Health (PATH), Bill & Melinda Gates Foundation and South African Medical Research Council (SA-MRC). NW and AvG report receiving grant funds from Sanofi Pasteur and Bill & Melinda Gates Foundation. NM reports receiving grant

funds from Pfizer. All other authors have no competing/conflict of interest. LD, LG, MT and CCa acknowledge support from the Lagrange Project of ISI Foundation funded by CRT Foundation. LD and CCa acknowledge support from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 101003688 (EpiPose) and No. 101016233 (PERISCOPE).

Author contributions

Conception and design of study: JK, LD, LG, MT, SW, NM, AvG, NW, MM, CCo, CCa, ST

Data collection and laboratory processing: LM, JK, SW, LM, NM, AvG, NW

Analysis and interpretation: JK, LD, LG, MT, SW, CCo, CCa, ST

Accessed and verified underlying data: JK, LD, LG, MT, CCa

Drafted the Article: JK

All authors critically reviewed the Article. All authors had access to all the data reported in the study.

Data availability

The investigators welcome enquiries about possible collaborations and requests for access to the dataset. Data will be shared after approval of a proposal and with a signed data access agreement. Investigators interested in more details about this study, or in accessing these resources, should contact the corresponding author.

Other members of the South African SARS-CoV-2 Household Transmission Study (SA-S-HTS) Group

Amelia Buys, ND (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Daniel G. Amoako, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Health Sciences, College of Health Sciences, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.)

Dylan Toi, MD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Jinal N. Bhiman, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.)

Juanita Chewparsad, MSc (Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.)

Kedibone Ndlangisa, PhD (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Leisha Genade, MPH (Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa.)

Limakatso Lebina, PhD (Perinatal HIV Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. Africa Health Research Institute (AHRI), Durban, South Africa)

Linda de Gouveia, ND (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Mzimasi Neti, MPH (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Retshidisitswe Kotane, BCMP (Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa.)

Tables

Table 1. Association of contact parameters with SARS-CoV-2 household acquisition* using the Wilcoxon rank-sum test, Klerksdorp and Soweto, South Africa, September 2020 – October 2021

Contact parameter	p-value
Median daily contact duration with index	0.71
Median daily contact frequency with index	0.92
Median daily average contact duration with index	0.14
Median daily contact duration with infected household members	0.40
Median daily contact frequency with infected household members	0.42
Median daily average contact duration with infected household members	0.69

*Outcome investigated: testing positive for SARS-CoV-2

Table 2. Factors associated with SARS-CoV-2 household transmission from index cases and acquisition in household contacts (contact parameters with index case), Klerksdorp and Soweto, South Africa, 2020-2021, (n=253)

	SARS-CoV-2 infection[1]		Univariate analysis	Multivariable analysis (no contact parameter)	Multivariable analysis (including contact parameter)		
	Negative, N = 100	Positive, N = 153	OR (95%CI)	aOR (95%CI)	Duration aOR (95%CI)	Frequency aOR (95%CI)	Average duration aOR (95%CI)
Index Characteristics							
Site							
Klerksdorp	48 (45%)	59 (55%)	Reference				
Soweto	52 (36%)	94 (64%)	1.79 (0.76 to 4.24)				
Age (years)							
18-34	36 (47%)	40 (53%)	Reference	Reference	Reference	Reference	Reference
35-59	49 (36%)	89 (64%)	2.08 (0.79 to 5.48)	2.08 (0.76 to 5.66)	2.08 (0.75 to 5.76)	2.25 (0.80 to 6.36)	2.13 (0.78 to 5.82)
≥60	15 (38%)	24 (62%)	1.83 (0.48 to 6.89)	1.99 (0.49 to 8.01)	1.98 (0.48 to 8.13)	2.20 (0.52 to 9.25)	2.03 (0.50 to 8.22)
Minimum Ct value							
>35	17 (81%)	4 (19%)	Reference	Reference	Reference	Reference	Reference
30-35	23 (46%)	27 (54%)	7.42 (1.22 to 45.10)	6.90 (0.76 to 62.84)	6.74 (0.74 to 61.68)	7.48 (0.79 to 71.24)	6.52 (0.72 to 58.94)
<30	59 (33%)	120 (67%)	15.84 (2.96 to 84.67)	9.77 (1.34 to 71.35)	**9.79 (1.33 to 71.81)**	**10.67 (1.40 to 81.48)**	**9.39 (1.30 to 67.98)**
Unknown	1 (33%)	2 (67%)	15.26 (0.32 to 728.79)	7.28 (0.12 to 428.62)	6.92 (0.12 to 405.98)	7.94 (0.13 to 493.09)	6.05 (0.11 to 348.40)
SARS-CoV-2 variant							
non-Alpha/Beta/Delta	5 (36%)	9 (64%)	1.25 (0.23 to 6.91)	2.24 (0.35 to 14.54)	2.12 (0.33 to 13.72)	2.06 (0.31 to 13.58)	2.02 (0.31 to 13.03)
Alpha	2 (15%)	11 (85%)	4.87 (0.62 to 38.60)	5.13 (0.54 to 48.67)	5.69 (0.53 to 61.60)	4.78 (0.49 to 46.53)	5.85 (0.56 to 60.50)
Beta	71 (41%)	101 (59%)	Reference	Reference	Reference	Reference	Reference

Delta	8 (21%)	30 (79%)	3.28 (0.96 to 11.24)	3.83 (1.00 to 14.62)	**3.88 (1.01 to 14.92)**	**3.98 (1.02 to 15.56)**	**3.82 (1.00 to 14.62)**
Variant Unknown	14 (88%)	2 (12%)	0.06 (0.01 to 0.42)	0.15 (0.02 to 1.58)	0.16 (0.02 to 1.67)	0.15 (0.01 to 1.59)	0.16 (0.02 to 1.59)
Contact characteristics							
Age (years)							
<5	8 (73%)	3 (27%)	Reference	Reference	Reference	Reference	Reference
5-12	21 (44%)	27 (56%)	4.25 (0.65 to 27.89)	1.57 (0.24 to 10.25)	1.85 (0.27 to 12.41)	1.84 (0.27 to 12.51)	1.80 (0.27 to 12.08)
13-17	11 (26%)	31 (74%)	15.12 (2.07 to 110.46)	6.96 (0.95 to 50.96)	**7.97 (1.02 to 62.21)**	**8.99 (1.13 to 71.67)**	**8.14 (1.08 to 61.16)**
18-34	23 (38%)	38 (62%)	5.62 (0.90 to 35.11)	2.68 (0.42 to 17.11)	3.06 (0.46 to 20.56)	3.35 (0.49 to 22.80)	3.08 (0.47 to 20.13)
35-59	27 (40%)	41 (60%)	4.36 (0.68 to 27.83)	1.74 (0.27 to 11.34)	1.99 (0.29 to 13.52)	2.19 (0.32 to 15.09)	2.00 (0.30 to 13.23)
≥60	10 (43%)	13 (57%)	4.72 (0.60 to 37.47)	1.79 (0.23 to 14.14)	2.06 (0.25 to 17.33)	2.34 (0.27 to 20.14)	2.06 (0.25 to 16.72)
Sex							
Male	53 (49%)	56 (51%)	Reference	Reference	Reference	Reference	Reference
Female	47 (33%)	97 (67%)	2.54 (1.28 to 5.06)	2.42 (1.19 to 4.90)	**2.43 (1.19 to 4.97)**	**2.33 (1.14 to 4.79)**	**2.42 (1.19 to 4.92)**
Sleep in same room as index							
No	69 (40%)	103 (60%)	Reference				
Yes	31 (38%)	50 (62%)	0.98 (0.49 to 1.94)				
Median Daily Contact Duration With Index (Seconds)	305 (102, 685)	230 (100, 700)	1.00 (1.00 to 1.00)		1.00 (1.00 to 1.00)		
Median Daily Contact Frequency With Index	6 (3, 15)	6 (3, 14)	1.00 (0.98 to 1.02)			1.01 (0.99 to 1.03)	
Median Daily Average Contact Duration With Index (Seconds)	33 (28, 41)	33 (26, 42)	**1.75 (1.02 to 3.02)**				1.00 (0.99 to 1.01)

[1] n (%); Median (inter quartile range). aOR: adjusted odds ratio. Significant associations on multivariable analysis in bold face.

Factors investigated but not found significant on multivariable analysis: index sex, HIV status, underlying conditions, body mass index, current smoking, episode duration, serostatus at follow-up end; contact HIV status, underlying conditions, body mass index, current smoking, cared for by index.

Table 3. Factors associated with SARS-CoV-2 acquisition within in the household (contact parameters with SARS-CoV-2 infected household members), Klerksdorp and Soweto, South Africa, 2020-2021, (n=342)

	SARS-CoV-2 infection[1]		Univariate analysis	Multivariable analysis (no contact parameter)	Multivariable analysis (including contact parameter)		
	Negative, N = 100	Positive, N = 242	OR (95%CI)	aOR (95%CI)	Duration aOR (95%CI)	Frequency aOR (95%CI)	Average duration aOR (95%CI)
Site							
Klerksdorp	48 (33%)	96 (67%)	Reference				
Soweto	52 (26%)	146 (74%)	1.78 (0.62 to 5.09)				
Age (years)							
<5	8 (73%)	3 (27%)	Reference	Reference	Reference	Reference	Reference
5-12	21 (44%)	27 (56%)	4.05 (0.55 to 30.01)	1.95 (0.24 to 15.82)	2.29 (0.27 to 19.36)	2.28 (0.27 to 18.93)	2.09 (0.24 to 18.23)
13-17	11 (26%)	31 (74%)	17.66 (2.21 to 141.22)	**10.10 (1.17 to 87.27)**	**14.22 (1.53 to 132.15)**	**13.13 (1.44 to 119.41)**	**13.24 (1.41 to 124.29)**
18-34	23 (27%)	63 (73%)	15.74 (2.25 to 109.94)	7.78 (0.98 to 61.76)	**10.89 (1.28 to 92.68)**	**10.16 (1.22 to 84.96)**	**8.83 (1.04 to 74.97)**
35-59	27 (23%)	90 (77%)	19.51 (2.81 to 135.57)	**8.54 (1.04 to 70.37)**	**11.88 (1.35 to 104.30)**	**11.29 (1.31 to 97.56)**	**9.11 (1.03 to 80.58)**
≥60	10 (26%)	28 (74%)	19.98 (2.42 to 164.74)	7.35 (0.76 to 71.46)	**10.44 (1.01 to 108.14)**	9.84 (0.97 to 100.11)	8.24 (0.79 to 85.97)
Body mass index							
Normal weight	51 (36%)	90 (64%)	Reference				
Underweight	7 (35%)	13 (65%)	0.75 (0.19 to 3.00)	0.76 (0.17 to 3.36)	0.70 (0.16 to 3.13)	0.74 (0.17 to 3.29)	0.72 (0.16 to 3.31)
Overweight	22 (32%)	47 (68%)	1.74 (0.74 to 4.08)	1.32 (0.53 to 3.30)	1.28 (0.51 to 3.21)	1.30 (0.52 to 3.24)	1.28 (0.50 to 3.25)

Obese	17 (16%)	90 (84%)	8.72 (3.55 to 21.41)	**5.22 (1.86 to 14.71)**	**5.27 (1.85 to 15.00)**	**5.13 (1.82 to 14.46)**	**4.82 (1.69 to 13.76)**
Unknown	3 (60%)	2 (40%)	0.50 (0.02 to 12.92)	NA	NA	NA	NA
Current smoking							
Yes	18 (42%)	25 (58%)	Reference	Reference	Reference	Reference	Reference
No	81 (27%)	215 (73%)	3.31 (1.28 to 8.55)	**3.12 (1.01 to 9.59)**	3.02 (0.97 to 9.35)	**3.08 (1.00 to 9.49)**	3.05 (0.97 to 9.54)
Unknown	1 (33%)	2 (67%)	11.30 (0.15 to 835.38)	NA	NA	NA	NA
SARS-CoV-2 variant							
non-Alpha/Beta/Delta	5 (26%)	14 (74%)	1.52 (0.15 to 14.88)	1.51 (0.14 to 16.69)	1.28 (0.11 to 14.50)	1.22 (0.11 to 13.06)	1.70 (0.13 to 21.89)
Alpha	2 (12%)	15 (88%)	4.54 (0.29 to 70.12)	6.11 (0.34 to 109.45)	4.78 (0.25 to 92.24)	4.67 (0.27 to 80.37)	44.92 (1.07 to 1,881.29)
Beta	71 (31%)	161 (69%)	Reference	Reference	Reference	Reference	Reference
Delta	8 (16%)	42 (84%)	2.98 (0.56 to 15.91)	3.07 (0.52 to 18.05)	2.66 (0.44 to 16.07)	2.68 (0.46 to 15.45)	3.64 (0.55 to 23.83)
Variant Unknown	14 (58%)	10 (42%)	0.54 (0.09 to 3.30)	0.53 (0.08 to 3.59)	0.50 (0.07 to 3.48)	0.49 (0.07 to 3.23)	0.45 (0.06 to 3.43)
Median daily contact duration with infected household members (seconds)	520 (215, 1,278)	680 (162, 1,750)	1.00 (1.00 to 1.00)		1.00 (1.00 to 1.00)		
Median daily contact frequency with infected household members	14 (5, 26)	16 (4, 37)	1.00 (0.99 to 1.00)			1.00 (0.99 to 1.01)	
Median daily average contact duration with infected household members (seconds)	35 (30, 44)	38 (28, 47)	**0.97 (0.95 to 0.99)**				0.99 (0.98 to 1.00)

[1] n (%); Median (inter quartile range). aOR: adjusted odds ratio. Significant associations on multivariable analysis in bold face.
Factors investigated but not found significant on multivariable analysis: sex, HIV status, underlying conditions.

Figures

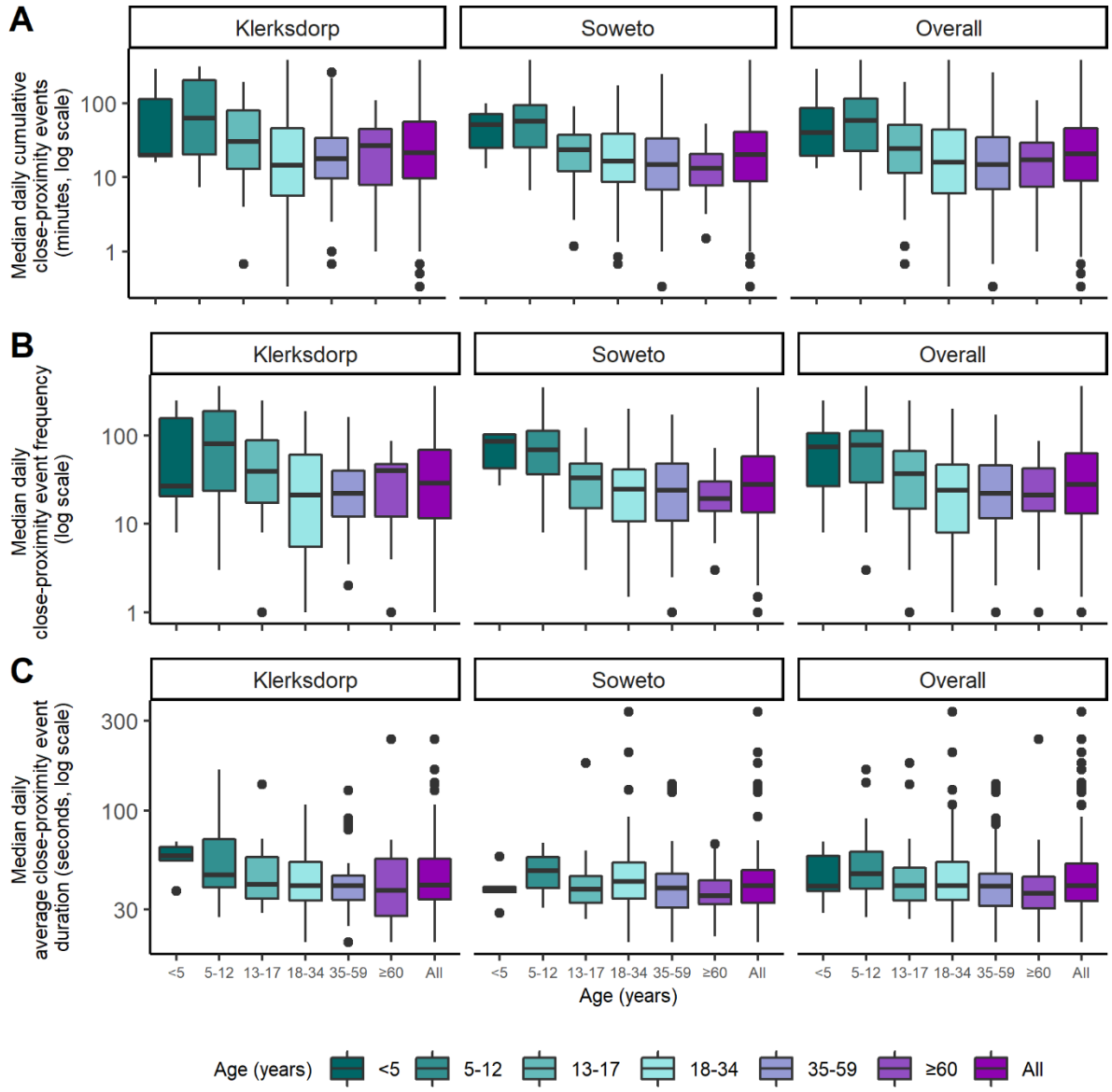


Figure 1. Contact parameters by age group (year) and site, Klerksdorp and Soweto, South Africa, September 2020 – October 2021. Horizontal line represents the median, box represents the 25th and 75th percentile, whiskers represent 1st and 99th percentile, circles indicate outliers.

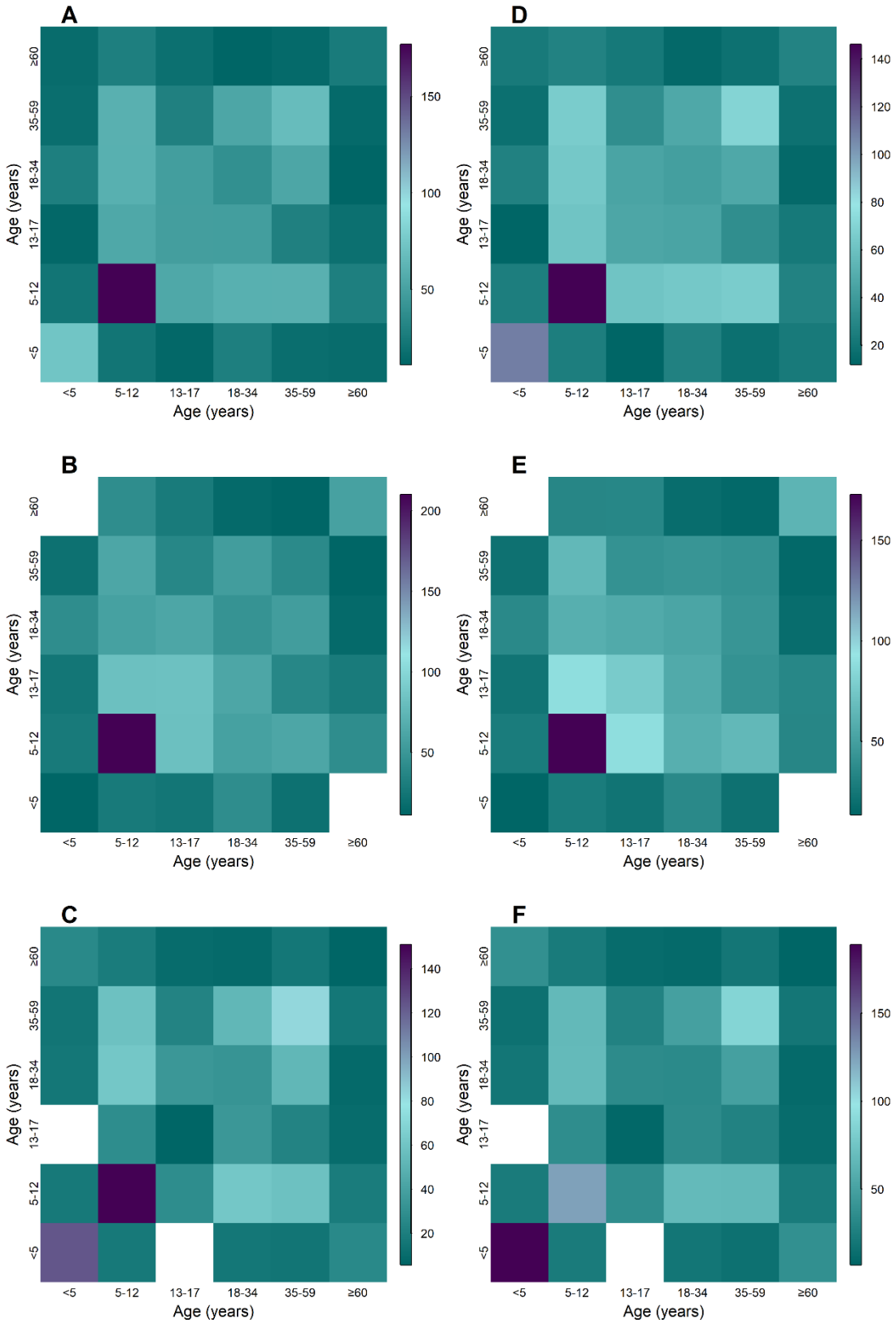


Figure 2. Aged-based contact matrices based close-proximity event duration (A-C) and frequency (D-E) for entire deployment period overall (A, D) Klerksdorp (B,D), and Soweto (C,F), September 2020 – October 2021. Teal denotes lowest value, purple highest, white no data for age group combination

References

1 National Institute for Comminicable Diseases. *COVID-19 Weekly Epidemiology Brief*, <https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-epidemiological-brief/> (2022).

2 Kleynhans, J. *et al.* SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa. *Emerging Infectious Disease journal* **28**, 1055, doi:10.3201/eid2805.220278 (2022).

3 Cable, R. *et al.* Estimates of prevalence of anti-SARS-CoV-2 antibodies among blood donors in eight provinces of South Africa in November 2021. *Res Sq*, doi:10.21203/rs.3.rs-1359658/v1 (2022).

4 Meyerowitz, E. A., Richterman, A., Gandhi, R. T. & Sax, P. E. Transmission of SARS-CoV-2: A Review of Viral, Host, and Environmental Factors. *Ann. Intern. Med.* **174**, 69-79, doi:10.7326/M20-5008 (2020).

5 Wang, C. C. *et al.* Airborne transmission of respiratory viruses. *Science* **373**, eabd9149, doi:doi:10.1126/science.abd9149 (2021).

6 Madewell, Z. J., Yang, Y., Longini, I. M., Jr., Halloran, M. E. & Dean, N. E. Household Transmission of SARS-CoV-2: A Systematic Review and Meta-analysis. *JAMA network open* **3**, e2031756-e2031756, doi:10.1001/jamanetworkopen.2020.31756 (2020).

7 Liu, T. *et al.* Transmission dynamics of 2019 novel coronavirus (2019-nCoV). *bioRxiv*, 2020.2001.2025.919787, doi:10.1101/2020.01.25.919787 (2020).

8 Cohen, C. *et al.* SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020-21. *The Lancet Infectious Diseases* **22**, 821-834, doi:10.1016/S1473-3099(22)00069-X (2022).

9 Aleta, A. *et al.* Quantifying the importance and location of SARS-CoV-2 transmission events in large metropolitan areas. *Proceedings of the National Academy of Sciences* **119**, e2112182119, doi:doi:10.1073/pnas.2112182119 (2022).

10 Hsu, C.-Y. *et al.* Household transmission but without the community-acquired outbreak of COVID-19 in Taiwan. *J. Formos. Med. Assoc.* **120**, S38-S45, doi:https://doi.org/10.1016/j.jfma.2021.04.021 (2021).

11 Nande, A., Adlam, B., Sheen, J., Levy, M. Z. & Hill, A. L. Dynamics of COVID-19 under social distancing measures are driven by transmission network structure. *PLoS Comput. Biol.* **17**, e1008684, doi:10.1371/journal.pcbi.1008684 (2021).

12 Liu, C. Y. *et al.* Rapid Review of Social Contact Patterns During the COVID-19 Pandemic. *Epidemiology* **32**, 781-791, doi:10.1097/ede.0000000000001412 (2021).

13 McAloon, C. G. *et al.* Numbers of close contacts of individuals infected with SARS-CoV-2 and their association with government intervention strategies. *BMC Public Health* **21**, 2238, doi:10.1186/s12889-021-12318-y (2021).

14 Salathe, M. *et al.* A high-resolution human contact network for infectious disease transmission. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 22020-22025, doi:10.1073/pnas.1009094108 (2010).

15 Cattuto, C. *et al.* Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* **5**, e11596-e11596, doi:10.1371/journal.pone.0011596 (2010).

16 Voirin, N. *et al.* Combining high-resolution contact data with virological data to investigate influenza transmission in a tertiary care hospital. *Infect. Control Hosp. Epidemiol.* **36**, 254-260, doi:10.1017/ice.2014.53 (2015).

17 Kiti, M. C. *et al.* Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors. *EPJ data science* **5**, 21, doi:10.1140/epjds/s13688-016-0084-2 (2016).

18 Pung, R. *et al.* Using high-resolution contact networks to evaluate SARS-CoV-2 transmission and control in large-scale multi-day events. *Nature communications* **13**, 1956, doi:10.1038/s41467-022-29522-y (2022).

19 Kleynhans, J. *et al.* Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study. *Clin. Infect. Dis.*, doi:10.1093/cid/ciac640 (2022).
20 Wibmer, C. K. *et al.* SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* **27**, 622-625, doi:10.1038/s41591-021-01285-x (2021).
21 Torresi, J., Edeling, M. A., Nolan, T. & Godfrey, D. I. A Complementary Union of SARS-CoV-2 Natural and Vaccine Induced Immune Responses. *Front. Immunol.* **13**, 914167, doi:10.3389/fimmu.2022.914167 (2022).
22 Whitney, C. G. Changing Epidemiology of Pneumococcal Disease in the Era of Conjugate Vaccines. *Curr Epidemiol Rep* **3**, 125-135, doi:10.1007/s40471-016-0077-5 (2016).
23 Kwok, K. O. *et al.* Social contacts and the locations in which they occur as risk factors for influenza infection. *Proc R Soc Lond B Biol Sci* **281**, 20140709, doi:10.1098/rspb.2014.0709 (2014).
24 Qian, G. *et al.* Association of pneumococcal carriage in infants with the risk of carriage among their contacts in Nha Trang, Vietnam: A nested cross-sectional survey. *PLoS Med.* **19**, e1004016, doi:10.1371/journal.pmed.1004016 (2022).
25 Dodd, P. J. *et al.* Age- and Sex-Specific Social Contact Patterns and Incidence of Mycobacterium tuberculosis Infection. *Am. J. Epidemiol.* **183**, 156-166, doi:10.1093/aje/kwv160 (2016).
26 Ng, O. T. *et al.* SARS-CoV-2 seroprevalence and transmission risk factors among high-risk close contacts: a retrospective cohort study. *The Lancet Infectious Diseases* **21**, 333-343, doi:10.1016/S1473-3099(20)30833-1 (2021).
27 Duval, D. *et al.* Long distance airborne transmission of SARS-CoV-2: rapid systematic review. *BMJ* **377**, e068743, doi:10.1136/bmj-2021-068743 (2022).
28 Mathee, A. *et al.* Housing Quality in a Rural and an Urban Settlement in South Africa. *Int. J. Environ. Res. Public Health* **18**, 2240, doi:10.3390/ijerph18052240 (2021).
29 Burger, R., Christian, C., English, R., Maughan-Brown, B. & Rossouw, L. Predictors of mask-wearing during the advent of the COVID-19 pandemic: Evidence from South Africa. *Transl. Behav. Med.* **12**, doi:10.1093/tbm/ibab132 (2021).
30 National Institute for Comminicable Diseases. *Tracking SARS-CoV-2 Variants*, <https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/sars-cov-2-genomic-surveillance-update/> (2022).
31 Campbell, F., Cori, A., Ferguson, N. & Jombart, T. Bayesian inference of transmission chains using timing of symptoms, pathogen genomes and contact data. *PLoS Comput. Biol.* **15**, e1006930, doi:10.1371/journal.pcbi.1006930 (2019).

Supplementary materials: Association of close-range contact patterns with SARS-CoV-2 household transmission, South Africa, 2020-2021

Kleynhans J, Dall'Amico L, Gauvin L, Tizzoni M, Maloma L, Walaza S, Martinson NA, Wolter N, Makhasi M, Cohen C, Cattuto C, Tempia S, for the SA-S-HTS Group

Supplementary Table 1. Baseline characteristics of SARS-CoV-2 index cases (n=127) and their household contacts (n=373) included in the household cumulative infection risk study and included in the contact study, Klerksdorp and Soweto, South Africa, September 2020 – October 2021.

	Overall n=500	No contact data n=158	Included in contact analysis n=342	p-value
Site				
Klerksdorp	236 (47.2)	92 (58.2)	144 (42.1)	0.001
Soweto	264 (52.8)	66 (41.8)	198 (57.9)	
Index				
Index	127 (25.4)	38 (24.1)	89 (26.0)	0.718
Contact	373 (74.6)	120 (75.9)	253 (74.0)	
Household size				
3-5	350 (70.0)	123 (77.8)	227 (66.4)	0.012
6-10	150 (30.0)	35 (22.2)	115 (33.6)	
Rooms used for sleeping				
1-2	245 (49.0)	77 (48.7)	168 (49.1)	0.557
2-4	157 (31.4)	55 (34.8)	102 (29.8)	
>4	51 (10.2)	13 (8.2)	38 (11.1)	
Unknown	47 (9.4)	13 (8.2)	34 (9.9)	
Crowding				
No	127 (25.4)	50 (31.6)	77 (22.5)	0.038
Yes	373 (74.6)	108 (68.4)	265 (77.5)	
Child <5 years				
No	426 (85.2)	137 (86.7)	289 (84.5)	0.610
Yes	74 (14.8)	21 (13.3)	53 (15.5)	
HH member smokes inside				
No	403 (80.6)	116 (73.4)	287 (83.9)	0.008
Yes	97 (19.4)	42 (26.6)	55 (16.1)	
Main water source inside home				
Water taps inside home	353 (70.6)	121 (76.6)	232 (67.8)	0.059
Water tap outside Home/tanker	147 (29.4)	37 (23.4)	110 (32.2)	
Main cooking fuel				
Electricity	483 (96.6)	153 (96.8)	330 (96.5)	1.000
Gas/Paraffin	17 (3.4)	5 (3.2)	12 (3.5)	
Monthly household income				
0-R800	42 (8.4)	20 (12.7)	22 (6.4)	0.098
R801-1600	42 (8.4)	17 (10.8)	25 (7.3)	
R1601-3200	91 (18.2)	26 (16.5)	65 (19.0)	
R3201-6400	78 (15.6)	21 (13.3)	57 (16.7)	
R6401-12800	36 (7.2)	12 (7.6)	24 (7.0)	
>R12801	20 (4.0)	9 (5.7)	11 (3.2)	

Refused	191 (38.2)	53 (33.5)	138 (40.4)	
Age (years)				
<5	19 (3.8)	8 (5.1)	11 (3.2)	0.739
5-12	73 (14.6)	25 (15.8)	48 (14.0)	
13-17	60 (12.0)	18 (11.4)	42 (12.3)	
18-34	131 (26.2)	45 (28.5)	86 (25.1)	
35-59	165 (33.0)	48 (30.4)	117 (34.2)	
≥60	52 (10.4)	14 (8.9)	38 (11.1)	
Sex				
Male	197 (39.4)	60 (38.0)	137 (40.1)	0.730
Female	303 (60.6)	98 (62.0)	205 (59.9)	
Level of education*				
No schooling / kindergarten	18 (3.6)	5 (3.2)	13 (3.8)	0.963
Primary	23 (4.6)	8 (5.1)	15 (4.4)	
Secondary	112 (22.4)	33 (20.9)	79 (23.1)	
Matriculation	170 (34.0)	52 (32.9)	118 (34.5)	
Post-secondary	20 (4.0)	7 (4.4)	13 (3.8)	
Unknown	157 (31.4)	53 (33.5)	104 (30.4)	
Employment*				
Unemployed	171 (34.2)	52 (32.9)	119 (34.8)	0.732
Student	33 (6.6)	8 (5.1)	25 (7.3)	
Employed	111 (22.2)	36 (22.8)	75 (21.9)	
Unknown	185 (37.0)	62 (39.2)	123 (36.0)	
Smoking cigarettes**				
Yes	66 (13.2)	23 (14.6)	43 (12.6)	0.503
No	428 (85.6)	132 (83.5)	296 (86.5)	
Unknown	6 (1.2)	3 (1.9)	3 (0.9)	
HIV status				
Negative	244 (48.8)	88 (55.7)	156 (45.6)	0.095
Positive	56 (11.2)	17 (10.8)	39 (11.4)	
Unknown	200 (40.0)	53 (33.5)	147 (43.0)	
Underlying illness***				
No	418 (83.6)	129 (81.6)	289 (84.5)	0.412
Yes	72 (14.4)	24 (15.2)	48 (14.0)	
Unknown	10 (2.0)	5 (3.2)	5 (1.5)	
Body-mass index				
Normal weight	208 (41.6)	67 (42.4)	141 (41.2)	0.756
Underweight	28 (5.6)	8 (5.1)	20 (5.8)	
Overweight	101 (20.2)	32 (20.3)	69 (20.2)	
Obese	153 (30.6)	46 (29.1)	107 (31.3)	
Unknown	10 (2.0)	5 (3.2)	5 (1.5)	
SARS-CoV-2 infection				
Negative	153 (30.6)	53 (33.5)	100 (29.2)	0.621
Positive (index)	127 (25.4)	38 (24.1)	89 (26.0)	
Positive (not index)	220 (44.0)	67 (42.4)	153 (44.7)	

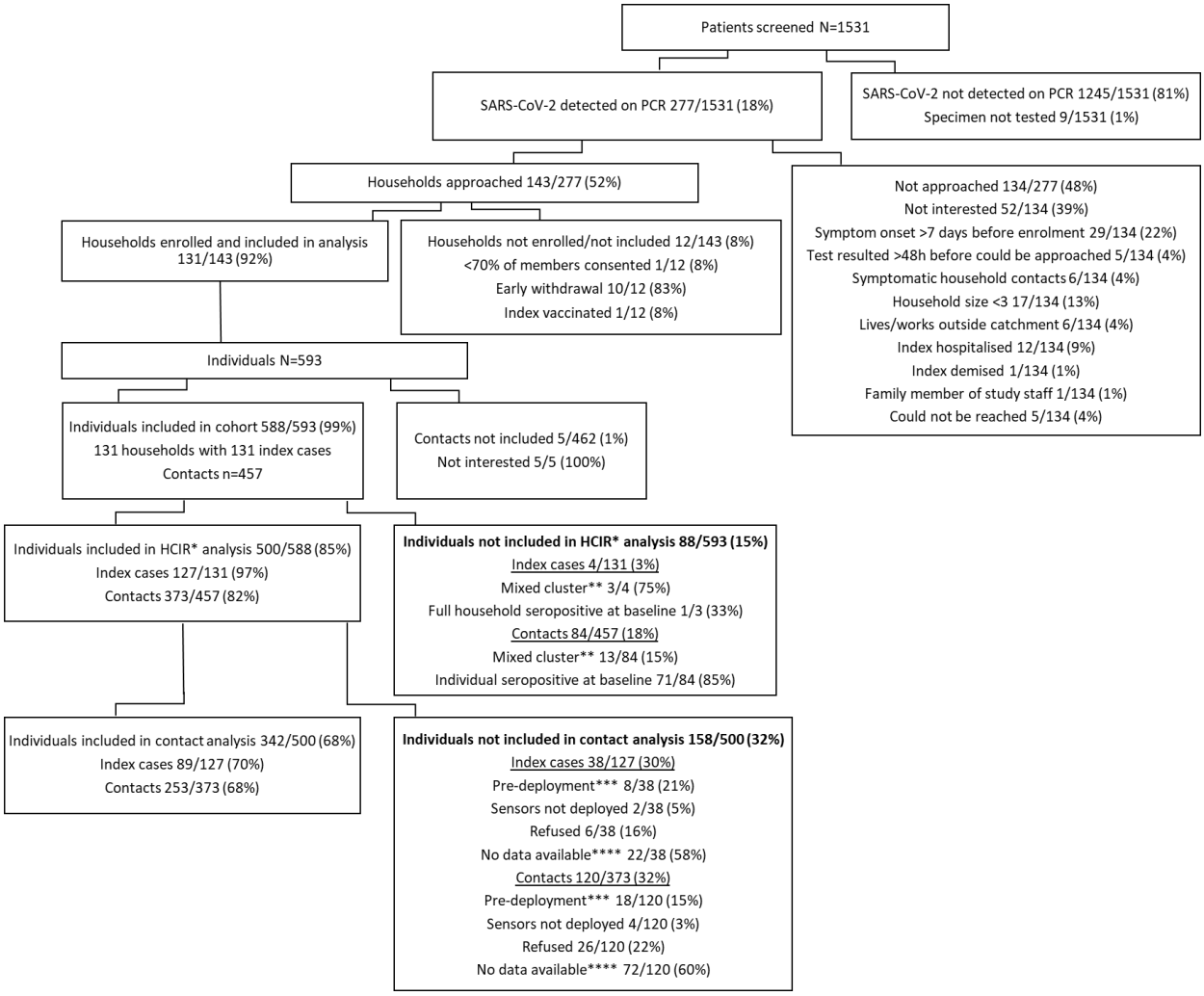
Values in headers indicate number of individuals. p-values calculated using Chi-squared test. *For individuals ≥18 years old. **For individuals ≥15 years old. ***Self-reported history of diabetes,

hypertension, asthma, lung disease, heart disease, stroke, spinal cord injury, epilepsy, cancer, liver disease, renal disease, pre-maturity.

Supplementary Table 2. Contact parameters by age group (year) and site, Klerksdorp and Soweto, South Africa, September 2020 – October 2021

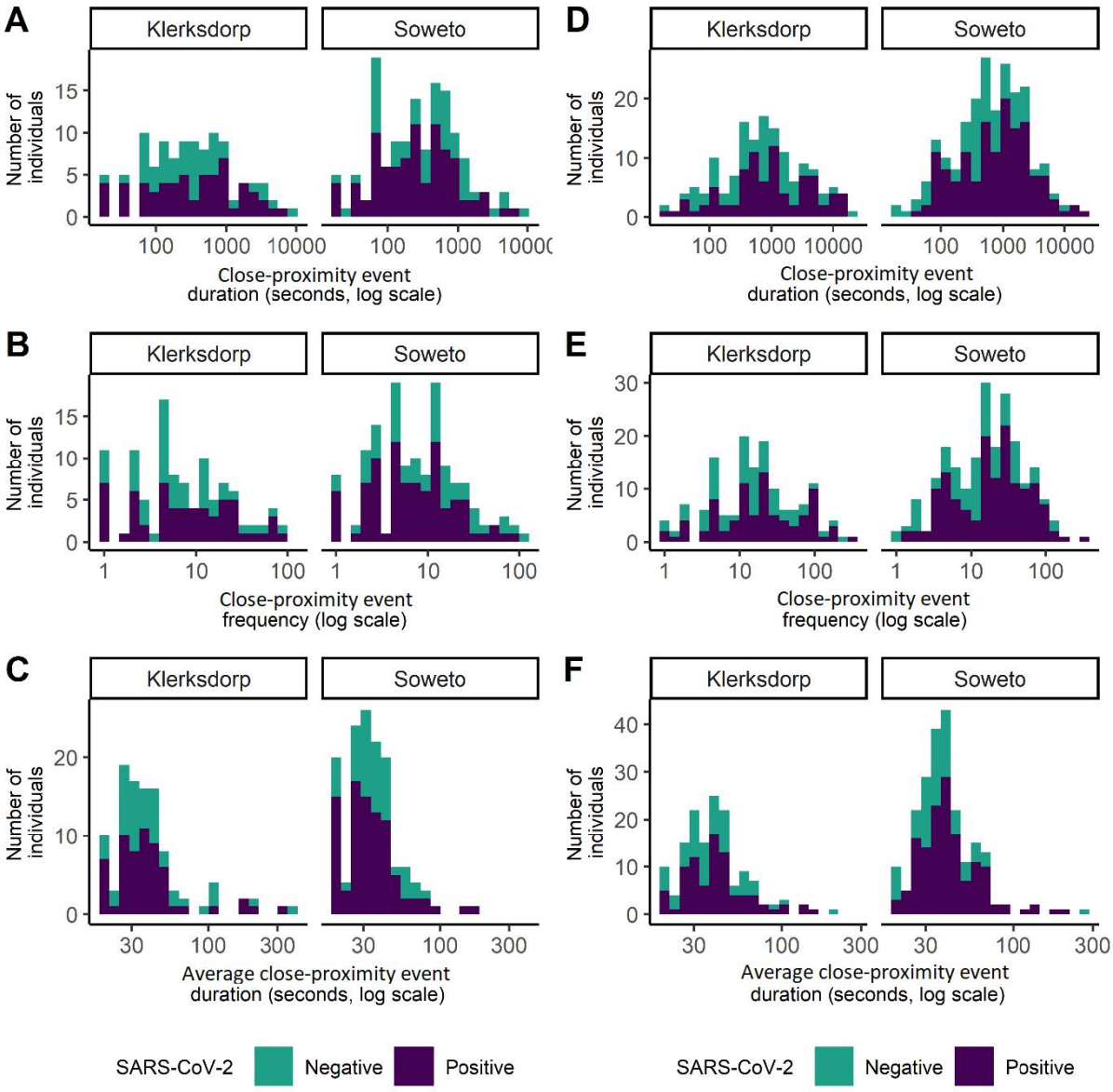
	n	Median daily contact duration (seconds, median (IQR))	Median daily contact frequency (median (IQR))	Median daily average contact duration* (seconds, median (IQR))
Overall	342	1240 (543, 2785)	28.00 (13, 63)	40.00 (33, 52)
<5	11	2440 (1175, 5230)	74 (27, 105)	40 (38, 57)
5-12	48	3510 (1382, 6975)	78 (29, 112)	46 (39, 60)
13-17	42	1485 (710, 3100)	37 (15, 66)	40 (33, 50)
18-34	86	970 (368, 2665)	24 (8, 46)	40 (34, 53)
35-59	117	900 (420, 2080)	22 (12, 46)	40 (31, 46)
≥60	38	1030 (450, 1765)	21 (14, 42)	36 (30, 45)
Klerksdorp				
Overall	144	1295 (580, 3385)	29 (12, 69)	40 (34, 55)
<5	5	1200 (1150, 6820)	26 (20, 156)	57 (54, 64)
5-12	21	3780 (1230, 12430)	80 (24, 188)	45 (39, 71)
13-17	23	1840 (805, 4855)	40 (18, 88)	41 (34, 57)
18-34	39	880 (340, 2795)	21 (6, 60)	40 (33, 53)
35-59	42	1075 (580, 2065)	22 (12, 40)	40 (33, 45)
≥60	14	1630 (520, 2732)	40 (13, 47)	38 (28, 56)
Soweto				
Overall	198	1220 (532, 2498)	28 (14, 58)	40 (32, 48)
<5	6	3150 (1570, 4295)	86 (45, 103)	38 (37, 39)
5-12	27	3480 (1530, 5750)	68 (36, 112)	48 (39, 57)
13-17	19	1410 (765, 2240)	33 (15, 48)	38 (32, 45)
18-34	47	1000 (520, 2355)	24 (11, 41)	42 (34, 53)
35-59	75	890 (410, 2030)	24 (11, 48)	39 (31, 46)
≥60	24	815 (470, 1240)	19 (14, 30)	35 (32, 43)

*Cumulative close-range proximity events duration divided by cumulative contact frequency



Supplementary Figure 1. Participants screened, enrolled, included in household transmission study and included in contact analysis, Klerksdorp and Soweto, South Africa, 2020-2021.

*Not included in HCIR analysis in study main manuscript (Kleynhans J, Walaza S, Martinson NA, et al. Household transmission of SARS-CoV-2 from adult index cases living with and without HIV in South Africa, 2020-2021: a case-ascertained, prospective observational household transmission study. Clin Infect Dis 2022.), **More than one SARS-CoV-2 variant detected in household, ***Household completed first week of follow-up before proximity sensors were deployed in field, ****No data available due to: Non-compliance based on lack of contacts logged or stationary accelerometer (n=73), Tag hardware failure (n=12), Tag lost/damaged (n=9)



Supplementary Figure 2. Distribution of A) median daily close-proximity event duration B) median daily close-proximity event frequency, D) average close-proximity event duration with index case, D) median daily close-proximity event duration E) median daily close-proximity event frequency, F) average close-proximity event duration with SARS-CoV-2 infected household members, by SARS-CoV-2 infection and site.