

South Africa - SAPRIN COVID-19  
surveillance intensive cohort April-  
December 2020: sub-dataset for  
Protective Behaviours and Secondary  
Harms paper

**SAPRIN**

Report generated on: August 13, 2021

Visit our data catalog at: <https://saprindata.samrc.ac.za/index.php>

## Overview

### Identification

---

ID NUMBER  
SAPRIN.SC19SICAD2020V1

### Version

---

VERSION DESCRIPTION  
v1: Dataset for public distribution

PRODUCTION DATE  
2021-08-13

### Overview

---

#### ABSTRACT

This study aimed to conduct active COVID-19 surveillance and estimate the health and non-health impacts of the COVID-19 epidemic in rural South Africa. In particular, the study focused on how households responded to the Covid-19 epidemic in terms of COVID-19-related knowledge and protective behaviours, health and economic impacts of NPIs, and mental health. To this end, data was collected during 2020 as part of SAPRIN's wider work to monitor and help respond to the COVID-19 epidemic. Each SAPRIN node sampled a small proportion of households in their surveillance areas (between 750 and 2250) and invited them to participate in intensive study, answering questionnaires every 2-3 weeks between April and December 2020. At each interview, a household representative responded to questions for themselves, for the household as a whole and for each household member. The data was then linked to key sociodemographic features of the responding households. This dataset contains only those variables relevant to the first academic paper published on this cohort (<https://doi.org/10.2196/26073>).

KIND OF DATA  
Event history data

UNITS OF ANALYSIS  
Interviews nested within responding households

### Scope

---

#### NOTES

The questionnaire included both household-level and individual-specific questions, the latter of which could be directly addressed by other household members if they were present. The primary respondent acted as a proxy in all other cases. COVID-19 symptom screening was included in the questionnaire, and those who satisfied the Department of Health's COVID-19 symptom criteria were referred for additional investigation, testing, and management.

#### TOPICS

Topic	Vocabulary	URI
South Africa; Self report; Prospective studies; Longitudinal studies; COVID-19; Mental health; Social welfare; Masks; Pensions		

KEYWORDS  
COVID-19, Surveillance, South Africa, Screening, Health and Demographic Surveillance System

## Coverage

### GEOGRAPHIC COVERAGE

The data are collected in three SAPRIN nodal surveillance areas in South Africa. The Agincourt site covers approximately 420 km<sup>2</sup> in Bushbuckridge district, Mpumalanga province, close to the Mozambique border. The DIMAMO site covers approximately 200 km<sup>2</sup> in Capricorn district, Limpopo Province, approximately 40 km from urban Polokwane. The AHRI site covers approximately 438 km<sup>2</sup> in the southeast of Umkhanyakude district, KwaZulu-Natal province, close to Mtubatuba town. The dataset contains within-node units separating each area into 20-40 sub-units.

### UNIVERSE

Eligibility covered all the adult (over age 18) resident population within each of the three SAPRIN nodal sites. Residence was defined as intention to sleep the majority of time at the dwelling in these areas over a four-month period

## Producers and Sponsors

### PRIMARY INVESTIGATOR(S)

Name	Affiliation
Guy Harling	Africa Health Research Institute, KwaZulu-Natal, South Africa Institute for Global Health, University College London, London, United Kingdom Medical Research Council/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 Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States Center for Population and Development Studies, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa
Francesc Xavier Gómez-Olivé	Medical Research Council/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 International Network for the Demographic Evaluation of Populations and Their Health Network, Accra, Ghana
Joseph Tlouyamma	Dikgale-Mamabolo-Mothiba Population Health Research Centre, School of Health Care Sciences, Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa Department of Computer Science, School of Mathematical and Computer Sciences, Faculty of Science and Agriculture, University of Limpopo, Mankweng, South Africa Dikgale-Mamabolo-Mothiba Population Health Research Centre, School of Health Care Sciences, Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa Department of Computer Science, School of Mathematical and Computer Sciences, Faculty of Science and Agriculture, University of Limpopo, Mankweng, South Africa
Tinofa Mutevedz	Department of Science and Innovation-Medical Research Council South African Population Research Infrastructure Network, Johannesburg, South Africa
Chodziwadziwa Whiteson Kabudula	Medical Research Council/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
Ruth Mahlako	Dikgale-Mamabolo-Mothiba Population Health Research Centre, School of Health Care Sciences, Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa
Urisha Singh	Africa Health Research Institute, KwaZulu-Natal, South Africa
Daniel Ohene-Kwofie	Medical Research Council/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
Rose Buckland	Institute for Global Health, University College London, London, United Kingdom
Pedzisai Ndagurwa	Medical Research Council/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
Dickman Gareta	Africa Health Research Institute, KwaZulu-Natal, South Africa

Resign Gunda	Africa Health Research Institute, KwaZulu-Natal, South Africa School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa
Thobeka Mngomezulu	Africa Health Research Institute, KwaZulu-Natal, South Africa School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa Africa Health Research Institute, KwaZulu-Natal, South Africa
Siyabonga Nxumalo	Africa Health Research Institute, KwaZulu-Natal, South Africa
Emily B Wong	Africa Health Research Institute, KwaZulu-Natal, South Africa Division of Infectious Diseases, University of Alabama, Birmingham, Birmingham, AL, United States
Kathleen Kahn	Medical Research Council/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 International Network for the Demographic Evaluation of Populations and Their Health Network, Accra, Ghana
Mark J Siedner	Africa Health Research Institute, KwaZulu-Natal, South Africa Harvard Medical School and the Medical Practice Evaluation Center, Massachusetts General Hospital, Boston, MA, United States
Eric Maimela	Dikgale-Mamabolo-Mothiba Population Health Research Centre, School of Health Care Sciences, Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa Department of Public Health, School of Health Care Sciences, Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa
Stephen Tollman	Medical Research Council/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 International Network for the Demographic Evaluation of Populations and Their Health Network, Accra, Ghana
Mark Collinson	Medical Research Council/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

#### OTHER PRODUCER(S)

Name	Affiliation	Role
South African Population Research Infrastructure Network		

#### FUNDING

Name	Abbreviation	Role
Department of Science and Innovation - South African Population Research Infrastructure Network	DSI-SAPRIN	Funding for Nodes
Wellcome Trust	WT	Funding for AHRI
Wellcome Trust	WT	Fellowship for Guy Harling
Royal Society	RS	Fellowship for Guy Harling
National Institutes of Health	NIH	Funding for Emily B Wong

## Metadata Production

#### METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
South African Population Research Infrastructure Network	SAPRIN		

#### DATE OF METADATA PRODUCTION

2021-08-13

#### DDI DOCUMENT VERSION

Version 1 (August 2021)

DDI DOCUMENT ID  
DDI.SAPRIN.SC19SICAD2020V1

## Sampling

No content available

## Questionnaires

No content available

## Data Collection

### Data Collection Dates

---

<b>Start</b>	<b>End</b>	<b>Cycle</b>
2020-04-15	2020-12-24	SAPRIN

### Time Periods

---

<b>Start</b>	<b>End</b>	<b>Cycle</b>
2020-04-15		SAPRIN

## Data Processing

No content available

## Data Appraisal

No content available

## File Description

## Variable List

## SC19SICAD2020

### Content

Cases	32526
Variable(s)	42
Structure	Type: Keys: ()
Version	
Producer	
Missing Data	

## Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V43	interviewkey	Unique ID for interview	discrete	character	
V44	contact_date	Date of interview	discrete	character	
V45	source	Type of interview conducted	discrete	numeric	
V46	wave	Data collection wave	discrete	numeric	
V47	hh_intid	Unique ID for household within node	contin	numeric	
V48	llntid	Unique ID for household respondent within node	contin	numeric	
V49	Sex	Sex of household respondent	discrete	numeric	
V50	Age	Age of household respondent at interview date	contin	numeric	
V51	currently_resident		discrete	numeric	
V52	hh_current_contact	Outcome of contact attempt	discrete	numeric	
V53	hh_ar03__8	Household members are wearing face masks	discrete	numeric	
V54	hh_phq2_pos	PHQ2 score >2, possible depression	discrete	numeric	
V55	hh_gad2_pos	GAD2 score >2, possible anxiety	discrete	numeric	
V56	hseh_children	Number of children in household	discrete	numeric	
V57	hseh_adults	Number of working age adults in household	discrete	numeric	
V58	hseh_seniors	Number of people aged over 60 in household	discrete	numeric	
V59	hseh_size	Number of people in household (res + non-res)	contin	numeric	
V60	hseh_ses_quint	Quintiles of household wealth within node (1 highest)	discrete	numeric	
V61	hseh_emp_count	Number of full/parttime employed people in household	discrete	numeric	
V62	hseh_pens_count	Number of pension grant receivers in household	discrete	numeric	
V63	hseh_ogrant_count	Number of non-pension grant receivers in household	discrete	numeric	
V64	hseh_pens_bin	Any employed individual in HH at last HDSS interview	discrete	numeric	
V65	hseh_ogrant_bin	Any employed individual in the household at last HDSS interview	discrete	numeric	
V66	lockdown_level	Government lockdown level at interview date	discrete	numeric	
V67	hh_ahri_cohort	AHRI allocation cohort	discrete	numeric	
V68	node	Node data line comes from	discrete	numeric	
V69	hh_agin_cohort	Agincourt allocation cohort	discrete	numeric	
V70	hh_dimamo_cohort	Randomly selected household cohort	discrete	numeric	

V71	contact_month	Month of interview	discrete	numeric
V72	hh_tm01_any	Any household member has left home in past 7 days	discrete	numeric
V73	nodewave	Combination of node (100s) and study wave	contin	numeric
V74	hh_maxedu	Summarized highest educational attainment in household	discrete	numeric
V75	hh_success_bin	Outcome of contact attempt successful (binary)	discrete	numeric
V76	nodemonth	Combination of node (100s) and month of 2020	contin	numeric
V77	hh_xx01_bin	Respondent feels they know enough or more about COVID-19 pandemic	discrete	numeric
V78	hh_hv01_bin	Anyone visited respondents home in past 7 days	discrete	numeric
V79	hh_hi01_bin	Any household member unable to access healthcare in past 7 days	discrete	numeric
V80	hh_hi04_bin	COVID has affected household's ability to earn money in past 7 days	discrete	numeric
V81	hh_hs01_bin	Any daily medicine missed in past week	discrete	numeric
V82	hh_ar03_out	Household members are avoiding going out/to work	discrete	numeric
V83	hh_ar03_trans	Household members are avoiding public transport	discrete	numeric
V84	hh_grpAgeSex6	Six age-sex categories of hh primary respondent	discrete	numeric



## Unique ID for interview (interviewkey)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 28601
Format: character	Invalid: 0
Width: 11	

## Date of interview (contact\_date)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 22927
Format: character	Minimum: NaN
Width: 11	Maximum: NaN

## Type of interview conducted (source)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 32526
Format: numeric	Invalid: 0
Width: 12	
Decimals: 0	
Range: 1-3	

## Data collection wave (wave)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 32526
Format: numeric	Invalid: 0
Width: 12	
Decimals: 0	
Range: 1-14	

## Unique ID for household within node (hh\_intid)

File: SC19SICAD2020

### Overview

Type: Continuous	Valid cases: 32517
Format: numeric	Invalid: 9
Width: 12	Minimum: 2003
Decimals: 0	Maximum: 139824099001
Range: 2003-139824099001	Mean: 639565531.5
	Standard deviation: 3717366467.5

## Unique ID for household respondent within node (lIntId)

File: SC19SICAD2020

### Overview

Type: Continuous  
Format: numeric  
Width: 12  
Decimals: 0  
Range: 8003-2797886002

Valid cases: 22927  
Invalid: 9599  
Minimum: 8003  
Maximum: 2797886002  
Mean: 611974653.7  
Standard deviation: 836200771.7

## Sex of household respondent (Sex)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 12  
Decimals: 0  
Range: -1-2

Valid cases: 22927  
Invalid: 9599

## Age of household respondent at interview date (Age)

File: SC19SICAD2020

### Overview

Type: Continuous  
Format: numeric  
Width: 12  
Decimals: 0  
Range: 0-954

Valid cases: 22927  
Invalid: 9599  
Minimum: 0  
Maximum: 954  
Mean: 32.7  
Standard deviation: 23.9

## (currently\_resident)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 12  
Decimals: 0  
Range: 1-3

Valid cases: 22927  
Invalid: 9599

## Outcome of contact attempt (hh\_current\_contact)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 25  
Decimals: 0  
Range: 1-93

Valid cases: 32526  
Invalid: 0

## Household members are wearing face masks (hh\_ar03\_\_8)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 12  
Decimals: 0  
Range: 0-1

Valid cases: 22627  
Invalid: 9899

## PHQ2 score >2, possible depression (hh\_phq2\_pos) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22957  
Invalid: 9569

## GAD2 score >2, possible anxiety (hh\_gad2\_pos) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22957  
Invalid: 9569

## Number of children in household (hseh\_children) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-14

Valid cases: 23095  
Invalid: 9431

## Number of working age adults in household (hseh\_adults) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-17

Valid cases: 23095  
Invalid: 9431

## Number of people aged over 60 in household (hseh\_seniors) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-3

Valid cases: 23095  
Invalid: 9431

## Number of people in household (res + non-res) (hseh\_size) File: SC19SICAD2020

### Overview

Type: Continuous  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-26

Valid cases: 23095  
Invalid: 9431  
Minimum: 0  
Maximum: 26  
Mean: 5.7  
Standard deviation: 3.4

## Quintiles of household wealth within node (1 highest) (hseh\_ses Quint) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 8  
Decimals: 0  
Range: 1-5

Valid cases: 26929  
Invalid: 5597

## Number of full/parttime employed people in household (hseh\_emp\_count) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-10

Valid cases: 31116  
Invalid: 1410

## Number of pension grant receivers in household (hseh\_pens\_count) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-3

Valid cases: 31116  
Invalid: 1410

## Number of non-pension grant receivers in household (hseh\_ogrant\_count) File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 31116
Format: numeric	Invalid: 1410
Width: 9	
Decimals: 0	
Range: 0-16	

Any employed individual in HH at last HDSS interview (hseh\_pens\_bin)  
File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 31385
Format: numeric	Invalid: 1141
Width: 9	
Decimals: 0	
Range: 0-1	

Any employed individual in the household at last HDSS interview (hseh\_ogrant\_bin)  
File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 31385
Format: numeric	Invalid: 1141
Width: 9	
Decimals: 0	
Range: 0-1	

Government lockdown level at interview date (lockdown\_level)  
File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 31385
Format: numeric	Invalid: 1141
Width: 23	
Decimals: 0	
Range: 1-8	

AHRI allocation cohort (hh\_ahri\_cohort)  
File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 10469
Format: numeric	Invalid: 22057
Width: 11	
Decimals: 0	
Range: 1-7	

Node data line comes from (node)  
File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 32526
Format: numeric	Invalid: 0
Width: 9	
Decimals: 0	
Range: 1-3	

## Agincourt allocation cohort (hh\_agin\_cohort)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 12392
Format: numeric	Invalid: 20134
Width: 11	
Decimals: 0	
Range: 0-6	

## Randomly selected household cohort (hh\_dimamo\_cohort)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 9634
Format: numeric	Invalid: 22892
Width: 12	
Decimals: 0	
Range: 1-7	

## Month of interview (contact\_month)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 31385
Format: numeric	Invalid: 1141
Width: 9	
Decimals: 0	
Range: 1-12	

## Any household member has left home in past 7 days (hh\_tm01\_any)

File: SC19SICAD2020

### Overview

Type: Discrete	Valid cases: 22836
Format: numeric	Invalid: 9690
Width: 9	
Decimals: 0	
Range: 0-1	

## Combination of node (100s) and study wave (nodewave)

File: SC19SICAD2020

### Overview

Type: Continuous  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 101-311

Valid cases: 32526  
Invalid: 0  
Minimum: 101  
Maximum: 311  
Mean: 203.7  
Standard deviation: 78.1

## Summarized highest educational attainment in household (hh\_maxedu) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 19  
Decimals: 0  
Range: 1-4

Valid cases: 26477  
Invalid: 6049

## Outcome of contact attempt successful (binary) (hh\_success\_bin) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 32526  
Invalid: 0

## Combination of node (100s) and month of 2020 (nodemonth) File: SC19SICAD2020

### Overview

Type: Continuous  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 104-312

Valid cases: 31385  
Invalid: 1141  
Minimum: 104  
Maximum: 312  
Mean: 202.5  
Standard deviation: 78

## Respondent feels they know enough or more about COVID-19 pandemic (hh\_xx01\_bin) File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22999  
Invalid: 9527

## Anyone visited respondents home in past 7 days (hh\_hv01\_bin)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22989  
Invalid: 9537

## Any household member unable to access healthcare in past 7 days (hh\_hi01\_bin)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22976  
Invalid: 9550

## COVID has affected household's ability to earn money in past 7 days (hh\_hi04\_bin)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22959  
Invalid: 9567

## Any daily medicine missed in past week (hh\_hs01\_bin)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22984  
Invalid: 9542

## Household members are avoiding going out/to work (hh\_ar03\_out)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22627  
Invalid: 9899

## Household members are avoiding public transport (hh\_ar03\_trans)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 9  
Decimals: 0  
Range: 0-1

Valid cases: 22627  
Invalid: 9899

## Six age-sex categories of hh primary respondent (hh\_grpAgeSex6)

File: SC19SICAD2020

### Overview

Type: Discrete  
Format: numeric  
Width: 12  
Decimals: 0  
Range: 1-13  
Invalid: 101

Valid cases: 22458  
Invalid: 10068