

South Africa

**Rochat, Tamsen, University of the Witwatersrand,
Manchester Metropolitan University**

Bland, Ruth, University of Glasgow

Houle, Brian, The Australian National University

Xanthe Hunt, AHRI

**Analytic dataset for violence and conduct
in the Siyakhula Cohort, 2012-2014**

Study Documentation

May 7, 2025

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	DDI.AHRI.sb-violence-01

Table of Contents

Overview	4
Scope & Coverage	4
Producers & Sponsors	4
Sampling	4
Data Collection	4
Accessibility	5
Files Description	6
sb-violence-01	6
Variables List	7
sb-violence-01	7
Variables Description	8
sb-violence-01	9

Analytic dataset for violence and conduct in the Siyakhula Cohort, 2012-2014

Overview	
Identification	AHRI.sb-violence-01
Version	V2.0.0
Abstract The Siyakhula cohort study and Vertical Transmission Study provide measures of birth weight and perinatal factors and measures of cognition and education outcomes for children aged 10-11 years in 2012-2014 in rural South Africa. The full cohort profile can be found here: Rochat, T. J., Houle, B., Stein, A., Pearson, R. M., Newell, M. L., & Bland, R. M. (2017). Cohort profile: the Siyakhula cohort, rural South Africa. <i>International journal of epidemiology</i> , 46(6), 1755-1756n. The present analytic dataset was produced for the purposes of a specific publication, focussed on violence.	
Kind of Data	Survey Data
Unit of Analysis	Individual

Scope & Coverage	
Keywords	Conduct disorder, violence, child behaviour, maternal mental health Conduct disorder, violence, child behaviour, maternal mental health
Topics	Child education, Maternal mental health, HIV, Food security, Parental discipline strategies, Child behaviour
Time Period(s)	2001-2014
Countries	South Africa
Geographic Coverage AHRI's AHRI's Population Intervention Programme Study Area (PIPSA) situated in uMkhanyakude district KwaZulu-Natal province which is approximately 200km north of Durban. The area is approximately 845km ² in size.	
Universe The cohort is comprised of children in the PIPSA who were aged 10-11 years in 2012-2014.	

Producers & Sponsors	
Primary Investigator(s)	Rochat, Tamsen, University of the Witwatersrand, Manchester Metropolitan University Bland, Ruth, University of Glasgow Houle, Brian, The Australian National University Xanthe Hunt, AHRI
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	Grand Challenges Canada (GCC) , Sponsor and Funder Wellcome Trust (WT) , Sponsor and Funder

Sampling	
Sampling Procedure Nonrandomised, prospective, intervention cohort study with re-enrollment of HIV-negative children	

Data Collection	
Data Collection Dates	start 2001-01-01

end 2006-12-04 start 2012-01-01 end 2014-12-04
--

Accessibility

Access Conditions

Access to the data requires accurate completion of the online data access application form accessible on the AHRI Data repository (<<https://data.ahri.org/>>). Data users are required to abide by the data use conditions stipulated on the application for access to the data. Failure to do so may result in their data access privileges revoked by the Data Custodian. In order to recognise the effort and intellectual contributions of AHRI investigators in producing and curating the data, users of AHRI data must acknowledge the source of the data and abide by the terms and conditions under which the data is accessed. All analytical datasets published on the AHRI Data Repository are assigned digital object identifier (DOIs) and the DOIs can be found on the Data Repository under Study Description tab - Access policy. AHRI data users are required to always cite the dataset using the DOI.

Additional checks are in place for this dataset, due to the sensitive nature of some of the outcomes. All applications to use the data will be routed to a member of the custodian team, who will evaluate applications.

Citation Requirements

Rochat, T., Bland, R., Houle, B., & Xanthe, H. (2025). Analytic dataset for violence and conduct in the Siyakhula Cohort, 2012-2014 [Data set]. Africa Health Research Institute. DOI:<https://doi.org/10.23664/AHRI.SB-VIOLENCE-01>

Files Description

Dataset contains 1 file(s)

sb-violence-01	
# Cases	1526
# Variable(s)	27
<u>Notes</u> sb-violence-01.dta Public use file prepared	

Variables List

Dataset contains 27 variable(s)

File sb-violence-01							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	ssbstudyid	SSB StudyId	continuous	numeric-12.0	1526	0	-
2	mumacdis ..	ACDIS_DSID or VTS Study Number of the mother	discrete	character-6	1526	0	-
3	hhadults ..	Number of resident adults	discrete	numeric-10.0	1523	3	-
4	hkids_res	Number of resident children	discrete	numeric-10.0	1524	2	-
5	male	Male child indicator	discrete	numeric-10.0	1526	0	-
6	in_vts	VTS intervention exposure	discrete	numeric-10.0	1526	0	-
7	repeated ..	Repeated grade indicator	discrete	numeric-10.0	1526	0	-
8	cagecat	Child age indicator	discrete	numeric-10.0	1526	0	-
9	mum_pcg	Primary caregiver	discrete	numeric-10.0	1526	0	-
10	anxiety	Clinical anxiety	discrete	numeric-9.0	1526	0	-
11	depression	Depression clinical	discrete	numeric-9.0	1526	0	-
12	hiv_expo ..	Child's HIV exposure	discrete	numeric-25.0	1524	2	-
13	crime	Crime exposure	discrete	numeric-9.0	1526	0	-
14	childgrant	Child support grant indicator	discrete	numeric-9.0	1522	4	-
15	foodinse ..	Food insecurity indicator	discrete	numeric-27.0	1513	13	-
16	mumrelat ..	Mother's relationship status	discrete	numeric-15.0	1507	19	-
17	pcg_empl ..	Caregiver employed indicator	discrete	numeric-9.0	1524	2	-
18	pcg_edu	Caregiver education	discrete	numeric-18.0	1505	21	-
19	pcg_age	Caregiver age	discrete	numeric-10.0	1503	23	-
20	psipsub ..	Parental distress indicator	discrete	numeric-9.0	1526	0	-
21	psipcdis ..	Relationship dysfunction indicator	discrete	numeric-9.0	1526	0	-
22	psidsub ..	Difficult child indicator	discrete	numeric-9.0	1526	0	-
23	mumsmack	Mother smacks child indicator	discrete	numeric-18.0	1523	3	-
24	physdisc ..	Reports physical discipline indicator	discrete	numeric-9.0	1520	6	-
25	ebf	Exclusive breastfeeding indicator	discrete	numeric-9.0	1526	0	-
26	violent ..	Violent and antisocial indicator	discrete	numeric-22.0	1526	0	-
27	dadinvol ..	Father figure involved indicator	discrete	numeric-9.0	1522	4	-

Variables Description

Dataset contains 27 variable(s)

File : sb-violence-01

ssbstudyid: SSB StudyId

Information [Type= continuous] [Format=numeric] [Range= 14-2074] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-] [Mean=1061.322 /-] [StdDev=600.259 /-]

mumacdisorvts_id: ACDIS_DSID or VTS Study Number of the mother

Information [Type= discrete] [Format=character] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

hhadults_res: Number of resident adults

Information [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]

Statistics [NW/ W] [Valid=1523 /-] [Invalid=3 /-]

Value	Label	Cases	Percentage
1	0-3	921	60.5%
2	4+	602	39.5%
Sysmiss		3	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

hhkids_res: Number of resident children

Information [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]

Statistics [NW/ W] [Valid=1524 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
1	0-3	576	37.8%
2	4+	948	62.2%
Sysmiss		2	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

male: Male child indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	866	56.7%
1	Yes	660	43.3%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

in_vts: VTS intervention exposure

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	627	41.1%
1	Yes	899	58.9%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

repeatedgrade: Repeated grade indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	907	59.4%

File : sb-violence-01

repeatedgrade: Repeated grade indicator

Value	Label	Cases	Percentage
1	Yes	619	40.6%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

cagecat: Child age indicator

Information	[Type= discrete] [Format=numeric] [Range= 7-11] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
7		59	3.9%
8		182	11.9%
9		596	39.1%
10		562	36.8%
11	11	127	8.3%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

mum_pcg: Primary caregiver

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Other	240	15.7%
1	Mother	1286	84.3%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

anxiety: Clinical anxiety

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	1476	96.7%
1	Yes	50	3.3%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

depression: Depression clinical

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	1439	94.3%
1	Yes	87	5.7%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

hiv_exposure: Child's HIV exposure

Information	[Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]
Statistics [NW/ W]	[Valid=1524 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
0	Unexposed	774	50.8%
1	Exposed	472	31.0%
2	Affected	278	18.2%

File : sb-violence-01

hiv_exposure: Child's HIV exposure

Value	Label	Cases	Percentage
Sysmiss		2	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

crime: Crime exposure

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	1340	87.8%
1	Yes	186	12.2%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

childgrant: Child support grant indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1522 /-] [Invalid=4 /-]

Value	Label	Cases	Percentage
0	No	87	5.7%
1	Yes	1435	94.3%
Sysmiss		4	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

foodinsecurity: Food insecurity indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]
Statistics [NW/ W]	[Valid=1513 /-] [Invalid=13 /-]

Value	Label	Cases	Percentage
0	No	775	51.2%
1	At least once in last 4 wks	738	48.8%
2	At least once/wk for 4 wks	0	
Sysmiss		13	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

mumrelationship: Mother's relationship status

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=1507 /-] [Invalid=19 /-]

Value	Label	Cases	Percentage
1	Single	246	16.3%
2	With bio father	687	45.6%
3	New partner	574	38.1%
Sysmiss		19	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

pcg_employed: Caregiver employed indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1524 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
0	No	1124	73.8%

File : sb-violence-01

pcg_employed: Caregiver employed indicator

Value	Label	Cases	Percentage
1	Yes	400	26.2%
Sysmiss		2	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

pcg_edu: Caregiver education

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=1505 /-] [Invalid=21 /-]

Value	Label	Cases	Percentage
0	None	92	6.1%
1	Primary	530	35.2%
2	Secondary	517	34.4%
3	Matriculation/post	366	24.3%
Sysmiss		21	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

pcg_age: Caregiver age

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=1503 /-] [Invalid=23 /-]

Value	Label	Cases	Percentage
1	30 or less	359	23.9%
2	31-40	586	39.0%
3	41+	558	37.1%
Sysmiss		23	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

psipdsubscale_90: Parental distress indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	1403	91.9%
1	Yes	123	8.1%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

psipedisubscale_90: Relationship dysfunction indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	1377	90.2%
1	Yes	149	9.8%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

psidsubscale_90: Difficult child indicator

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1526 /-] [Invalid=0 /-]

File : sb-violence-01

psidcsubscale_90: Difficult child indicator

Value	Label	Cases	Percentage
0	No	1393	91.3%
1	Yes	133	8.7%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

mumsmack: Mother smacks child indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1523 /-] [Invalid=3 /-]

Value	Label	Cases	Percentage
0	<=1 smack per week	633	41.6%
1	>1 smack per week	890	58.4%
Sysmiss		3	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

physdiscipline: Reports physical discipline indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1520 /-] [Invalid=6 /-]

Value	Label	Cases	Percentage
0	No	1025	67.4%
1	Yes	495	32.6%
Sysmiss		6	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ebf: Exclusive breastfeeding indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	<6 months	276	18.1%
1	6+ months	1250	81.9%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

violent_antisocial: Violent and antisocial indicator

Information [Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]

Statistics [NW/ W] [Valid=1526 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	None	1212	79.4%
1	Violent	174	11.4%
2	Violent and antisocial	140	9.2%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

dadinvolved: Father figure involved indicator

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=1522 /-] [Invalid=4 /-]

Value	Label	Cases	Percentage
0	No	958	62.9%
1	Yes	564	37.1%

File : sb-violence-01

dadinvolved: Father figure involved indicator

Value	Label	Cases	Percentage
Sysmiss		4	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.