South Africa

Chappell, Elizabeth, MRC Clinical Trials Unit at UCL
Herbst, Kobus, AHRI
Bärnighausen, Till, AHRI
Yapa, H. Manisha, Kirby Institute, University of New South Wales
Gareta, Dickman, AHRI

Comparing methods for estimating PCR testing coverage

Study Documentation

July 9, 2021

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	DDI.Comparing.estimates.of.testing.coverage

Table of Contents

Comparing methods for estimating PCR testing coverage

Overview

Identification	AHRI.Comparing.estimates.of.testing.coverage
Version	V1.0.0

Abstract

Early infant HIV diagnosis (EID) is critical to ensuring timely diagnosis of HIV-exposed infants, and treatment in those found to be infected. However estimates of coverage vary considerably, depending on data sources used. We used 4 methods to estimate coverage among a historical cohort of HIV-exposed infants in rural South Africa, between 2010-2016. Three datasets from this analysis are available on the repository.

· Dataset 1 (NHLS): This dataset includes data from NHLS, the National Health Laboratory Service. NHLS is responsible for laboratory testing in public facilities in South Africa, and data on all HIV DNA PCR tests conducted at the 17 healthcare facilities within the Hlabisa health sub-district (regardless of the age of the individual) between 1st June 2010 and 1st July 2017 were downloaded from the NHLS database through a secure file transfer protocol. A deterministic and probabilistic data linkage algorithm, based on first name, surname, sex, date of birth, facility at which the test was conducted and the infant's facility ID was used to identify repeat tests on the same child.

· Dataset 2 (ACDIS): This dataset includes all infants from ACDIS, AHRI's demographic surveillance database, who were born to women with HIV. Maternal HIV status was determined on the basis of either a positive HIV test result in the AHRI serosurvey, or on linkage to TIER.net, the national ART surveillance system. To estimate testing coverage, these HIV-exposed infants were linked to deduplicated PCR test data from the National Health Laboratory Service (NHLS, responsible for laboratory testing in public facilities in South Africa) from healthcare facilities in the Hlabisa health-subdistrict. Linkage was conducted using a deterministic and probabilistic linkage algorithm based on infant's first name, surname, date of birth and sex.

· Dataset 3 (MONARCH): This dataset includes data from the MONARCH trial, which evaluated the impact of a quality improvement intervention package on PMTCT processes. All children in South Africa are given a patient-held medical record at birth, called the Road-to-Health Booklet, in which information including HIV testing is recorded. As part of the trial, booklets belonging to all infants born to women receiving antenatal care at the 7 clinics in the AHRI surveillance area between July 2015 and December 2016 were photographed up to the 6-week postnatal visit. This dataset includes one record per HIV-exposed infant, with variables indicating whether a PCR test was recorded in their Road-to-Health Booklet and when their booklet was photographed.

Kind of Data	Routinely collected health record data licensed under Creative Commons Attribution 4.0 international license (CC-BY 4.0)
Unit of Analysis	Dataset 1 (NHLS): One record per infant with a PCR test in NHLS dataset Dataset 2 (ACDIS): One record per HIV-exposed infant in ACDIS Dataset 3 (MONARCH): One record per HIV-exposed infant in MONARCH

Scope & Coverage

Keywords	HIV; PCR testing; routinely collected data; early infant diagnosis
Topics	HIV; Routinely Collected Health Data
Time Period(s)	2010-2016
Countries	South Africa

Geographic Coverage
South Africa

Universe
Dataset 1 (NHLS): Infants receiving a PCR test at a healthcare facility in the Hlabisa health sub-district.

Dataset 2 (ACDIS): HIV-exposed infants living in the AHRI demographic surveillance area

Dataset 3 (MONARCH): Infants born to women with HIV receiving antenatal care at sites participating in the MONARCH trial.

Producers & Sponsors

Primary Investigator(s)	Chappell, Elizabeth, MRC Clinical Trials Unit at UCL Herbst, Kobus, AHRI Bärnighausen, Till, AHRI Yapa, H. Manisha, Kirby Institute, University of New South Wales Gareta, Dickman, AHRI
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	Wellcome Trust (WT) , Core funding SAPRIN (SAPRIN)
Other Acknowledgment(s)	Nxumalo, Siyabonga , SAPRIN Research Data Manager , Africa Health Research Institute Dube, Sweetness , Data Documentation Archivist , Africa Health Research Institute Ehlers, Eugene , Senior Software Developer , Africa Health Research Institute

Sampling

Sampling Procedure

Data Collection

Data Collection Dates	Dataset 1 (NHLS): start 2010-06-01 Dataset 1 (NHLS): end 2016-12-04 Dataset 2 (ACDIS): start 2010-06-01 Dataset 2 (ACDIS): end 2016-12-04 Dataset 3 (MONARCH): start 2015-07-01 Dataset 3 (MONARCH): end 2016-12-04

Accessibility

Access Conditions
The representative of the Receiving Organization agrees to comply with the following conditions:

1. Access to the restricted data will be limited to the Lead Researcher and other members of the research team listed in this request.
2. Copies of the restricted data or any data created on the basis of the original data will not be copied or made available to anyone other than those mentioned in this Data Access Agreement, unless formally authorized by the Data Archive.
3. The data will only be processed for the stated statistical and research purpose. They will be used for solely for reporting of aggregated information, and not for investigation of specific individuals or organizations. Data will not in any way be used for any administrative, proprietary or law enforcement purposes.
4. The Lead Researcher must state if it is their intention to match the restricted microdata with any other micro-dataset. If any matching is to take place, details must be provided of the datasets to be matched and of the reasons for the matching. Any datasets created as a result of matching will be considered to be restricted and must comply with the terms of this Data Access Agreement.

5. The Lead Researcher undertakes that no attempt will be made to identify any individual person, family, business, enterprise or organization. If such a unique disclosure is made inadvertently, no use will be made of the identity of any person or establishment discovered and full details will be reported to the Data Archive. The identification will not be revealed to any other person not included in the Data Access Agreement.
6. The Lead Researcher will implement security measures to prevent unauthorized access to licensed microdata acquired from the Data Archive. The microdata must be destroyed upon the completion of this research, unless the Data Archive obtains satisfactory guarantee that the data can be secured and provides written authorization to the Receiving Organization to retain them. Destruction of the microdata will be confirmed in writing by the Lead Researcher to the Data Archive.
7. Any books, articles, conference papers, theses, dissertations, reports, or other publications that employ data obtained from the Data Archive will cite the source of data in accordance with the citation requirement provided with the dataset.
8. An electronic copy of all reports and publications based on the requested data will be sent to the Data Archive.
9. The original collector of the data, the Data Archive, and the relevant funding agencies bear no responsibility for use of the data or for interpretations or inferences based upon such uses.
10. This agreement will come into force on the date that approval is given for access to the restricted dataset and remain in force until the completion date of the project or an earlier date if the project is completed ahead of time.
11. If there are any changes to the project specification, security arrangements, personnel or organization detailed in this application form, it is the responsibility of the Lead Researcher to seek the agreement of the Data Archive to these changes. Where there is a change to the employer organization of the Lead Researcher this will involve a new application being made and termination of the original project.
12. Breaches of the agreement will be taken seriously and the Data Archive will take action against those responsible for the lapse if willful or accidental. Failure to comply with the directions of the Data Archive will be deemed to be a major breach of the agreement and may involve recourse to legal proceedings. The Data Archive will maintain and share with partner data archives a register of those individuals and organizations which are responsible for breaching the terms of the Data Access Agreement and will impose sanctions on release of future data to these parties.

Citation Requirements

Chappell, E., Herbst, K., Baernighausen, T., Yapa, & Gareta, D. (2021). Comparing methods for estimating PCR testing coverage [Data set]. Africa Health Research Institute (AHRI). https://doi.org/10.23664/COMPARING.ESTIMATES.OF.TESTING.COVERAGE

Files Description

Dataset contains 5 file(s)

Comparing estimates of testing coverage - ACDIS dataset - WITH ID	
# Cases	2254
# Variable(s)	5

Comparing estimates of testing coverage - ACDIS dataset	
# Cases	2254
# Variable(s)	4

Comparing estimates of testing coverage - MONARCH - WITH ID	
# Cases	813
# Variable(s)	5

Comparing estimates of testing coverage - MONARCH	
# Cases	813
# Variable(s)	4

Comparing estimates of testing coverage - NHLS infants with PCR test	
# Cases	15234
# Variable(s)	3

Variables List

Dataset contains 21 variable(s)

File Comparing estimates of testing coverage - ACDIS dataset - WITH ID

#	Name	Label	Type	Format	Valid	Invalid	Question
1	acdisid	Individual internal identifier	continuous	numeric-12.0	2254	0	-
2	idnumber	-	continuous	numeric-9.0	2254	0	-
3	yob	-	discrete	numeric-43.0	2254	0	-
4	anypcr	-	discrete	numeric-11.0	2254	0	-
5	anypcr_7w	-	discrete	numeric-24.0	2254	0	-

File Comparing estimates of testing coverage - ACDIS dataset

#	Name	Label	Type	Format	Valid	Invalid	Question
1	idnumber	-	continuous	numeric-9.0	2254	0	-
2	yob	-	discrete	numeric-43.0	2254	0	-
3	anypcr	-	discrete	numeric-11.0	2254	0	-
4	anypcr_7w	-	discrete	numeric-24.0	2254	0	-

File Comparing estimates of testing coverage - MONARCH - WITH ID

#	Name	Label	Type	Format	Valid	Invalid	Question
1	BarcodeS ..	Barcoded sticker ID number (M10.....)	discrete	character-8	813	0	-
2	idnumber	-	continuous	numeric-9.0	813	0	-
3	yob	-	discrete	numeric-43.0	813	0	-
4	anypcr_7w	-	discrete	numeric-11.0	813	0	-
5	rth6wvis	-	discrete	numeric-32.0	813	0	-

File Comparing estimates of testing coverage - MONARCH

#	Name	Label	Type	Format	Valid	Invalid	Question
1	idnumber	-	continuous	numeric-9.0	813	0	-
2	yob	-	discrete	numeric-43.0	813	0	-
3	anypcr_7w	-	discrete	numeric-11.0	813	0	-
4	rth6wvis	-	discrete	numeric-32.0	813	0	-

File Comparing estimates of testing coverage - NHLS infants with PCR test

#	Name	Label	Type	Format	Valid	Invalid	Question
1	idnumber	1 id	continuous	numeric-10.0	15234	0	-
2	yob	-	discrete	numeric-43.0	15234	0	-
3	anypcr_7w	-	discrete	numeric-24.0	15234	0	-

Variables Description

Dataset contains 21 variable(s)

File : Comparing estimates of testing coverage - ACDIS dataset - WITH ID

acdisid: Individual internal identifier

Information	[Type= continuous] [Format=numeric] [Range= 153612-179129] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-] [Mean=168989.54 /-] [StdDev=6446.752 /-]

idnumber

Information	[Type= continuous] [Format=numeric] [Range= 1-2254] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-] [Mean=1127.5 /-] [StdDev=650.818 /-]

yob

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	2010	192	8.5%
2	2011	386	17.1%
3	2012	342	15.2%
4	2013	349	15.5%
5	2014	437	19.4%
6	2015 - before introduction of birth testing	121	5.4%
7	2015 - after introduction of birth testing	240	10.6%
8	2016	187	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.

anypcr

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test	1289	57.2%
1	PCR test	965	42.8%

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.

anypcr_7w

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test by 7w of age	1690	75.0%
1	PCR test by 7w of age	564	25.0%

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.

File : Comparing estimates of testing coverage - ACDIS dataset

idnumber

Information	[Type= continuous] [Format=numeric] [Range= 1-2254] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-] [Mean=1127.5 /-] [StdDev=650.818 /-]

yob

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	2010	192	8.5%
2	2011	386	17.1%
3	2012	342	15.2%
4	2013	349	15.5%
5	2014	437	19.4%
6	2015 - before introduction of birth testing	121	5.4%
7	2015 - after introduction of birth testing	240	10.6%
8	2016	187	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.

anypcr

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test	1289	57.2%
1	PCR test	965	42.8%

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.

anypcr_7w

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=2254 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test by 7w of age	1690	75.0%
1	PCR test by 7w of age	564	25.0%

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.

File : Comparing estimates of testing coverage - MONARCH - WITH ID

BarcodeStickerId: Barcoded sticker ID number (M10.....)

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

idnumber

Information	[Type= continuous] [Format=numeric] [Range= 1-813] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-] [Mean=407 /-] [StdDev=234.837 /-]

yob

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	2010	0	
2	2011	0	
3	2012	0	
4	2013	0	
5	2014	0	
6	2015 - before introduction of birth testing	0	
7	2015 - after introduction of birth testing	265	32.6%
8	2016	548	67.4%

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.

anypcr_7w

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test	435	53.5%
1	PCR test	378	46.5%

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.

rth6wvis

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Booklet not seen at 6 week visit	523	64.3%
1	Booklet seen at 6 week visit	290	35.7%

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.

File : Comparing estimates of testing coverage - MONARCH

idnumber

Information	[Type= continuous] [Format=numeric] [Range= 1-813] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-] [Mean=407 /-] [StdDev=234.837 /-]

yob

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	2010	0	
2	2011	0	
3	2012	0	
4	2013	0	
5	2014	0	
6	2015 - before introduction of birth testing	0	
7	2015 - after introduction of birth testing	265	32.6%
8	2016	548	67.4%

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.

anypcr_7w

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test	435	53.5%
1	PCR test	378	46.5%

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.

rth6wvis

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=813 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Booklet not seen at 6 week visit	523	64.3%
1	Booklet seen at 6 week visit	290	35.7%

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.

File : Comparing estimates of testing coverage - NHLS infants with PCR test

idnumber: 1 id

Information	[Type= continuous] [Format=numeric] [Range= 1-17268] [Missing=*]
Statistics [NW/ W]	[Valid=15234 /-] [Invalid=0 /-] [Mean=8599.716 /-] [StdDev=4985.859 /-]

yob

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=15234 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	2010	1044	6.9%
2	2011	2084	13.7%
3	2012	2234	14.7%
4	2013	2173	14.3%
5	2014	2373	15.6%
6	2015 - before introduction of birth testing	582	3.8%
7	2015 - after introduction of birth testing	2221	14.6%
8	2016	2523	16.6%

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.

anypcr_7w

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=15234 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No PCR test by 7w of age	6698	44.0%
1	PCR test by 7w of age	8536	56.0%

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.