



# Working List of Vaccination Survey Analyses and Software Specifications

Draft Version 2.2  
Revised February 2021



**Expanded  
Programme on  
Immunization**

This page is intentionally blank.

Acknowledgements.....	1
Introduction .....	3
Other Inputs.....	4
Vaccination Schedule Metadata .....	4
Coverage Survey Metadata.....	4
Coverage Analysis Options.....	5
Acronym List .....	6
Dose Notation .....	6
Weighted and unweighted analyses.....	6
Computational Notation .....	7
Document Revision History.....	7
All Surveys: Raw Data Listing .....	8
DATA_01 Tables of cluster level response counts .....	8
All Surveys: Describing the Survey Sample .....	9
DESC_01: Expected vs observed sample: clusters, households, & respondents.....	9
DESC_02: Response to multiple choice question (select a single response) .....	11
DESC_03: Response to multiple choice question (select all that apply).....	13
RI Survey – DATA CLEANING .....	15
RI Survey – Measures Related to Coverage .....	17
RI_COVG_01: Crude coverage .....	17
RI_COVG_02: Valid coverage .....	20
Specification to calculate dob.....	23
RI_COVG_03: Fully vaccinated .....	24
RI_COVG_04: Not vaccinated .....	25
RI_COVG_05: Clusters with suprisingly low coverage .....	26
RI Survey – Measures Related to Access .....	28
RI_ACC_01: Crude DPT1 coverage .....	28
RI Survey – Measures Related to Continuity of Services .....	29
RI_CONT_01: Dropout between two doses.....	29
RI Survey – Measures Related to quality of Services.....	31
RI_QUAL_01: Card and register availability .....	31
RI_QUAL_02: Ever had a card .....	34
A Note Regarding Continuous and Categorical Timeliness Indicators.....	35

RI_QUAL_03: Percent of DPT1 doses that were invalid.....	36
RI_QUAL_04: Percent of MCV1 doses administered before 39 weeks of age .....	37
RI_QUAL_05: Percent of DPT2 and DPT3 doses administered before 4 weeks had passed .....	39
RI_QUAL_06: Percent of valid MCV1 doses that were administered before the age of 12 months.....	41
RI_QUAL_07B: Valid coverage if there had been no missed opportunities for simultaneous vaccination .....	42
Specification for calculation of missed opportunities derived variables.....	47
RI_QUAL_08: Percent of visits with missed opportunity for simultaneous vaccination .....	52
RI_QUAL_09: Percent of children with missed opportunity for simultaneous vaccination .....	54
RI_QUAL_10: Percent of doses administered on a given date (not yet implemented).....	58
RI_QUAL_11: Percent of doses administered at a particular age, (age in days) (not yet implemented).....	59
RI_QUAL_12: Percent of later doses in a sequence administered after a particular interval.....	60
RI_QUAL_13: Percent of children who receive DPT3 by the age of 26 weeks .....	61
RI_QUAL_14: Percent of children simultaneously vaccinated where both doses were given (not yet implemented).....	62
RI_QUAL_15: Percent of children simultaneously vaccinated where both either dose was given (not yet implemented).....	63
RI_QUAL_16: Percent of children simultaneously vaccinated among children with vaccination dates (not yet implemented).....	64
RI_QUAL_17: Number of visits needed to be fully vaccinated (not yet implemented) .....	65
RI Survey – Measures Related to Providers of Services.....	66
RI_PROV_01: Percent of vaccinations performed by each source (not yet implemented).....	66
Tetanus Survey – Measures Related to Coverage .....	67
TT_COVG_01: Children born protected from neonatal tetanus.....	67
TT_COVG_02: Women protected at the time of the survey (not yet implemented).....	70
TT_COVG_03: Crude TT coverage (% who received 2+ doses during index pregnancy) (not yet implemented).....	71
TT_COVG_04: Valid dose TT coverage (not yet implemented) .....	73
Tetanus Survey – Measures Related to Access.....	74
TT_ACC_01 Women who obtained ante-natal care (ANC) (not yet implemented) .....	74
Tetanus Survey – Measures Related to Continuity of Services .....	75
TT_CONT_01: Dropout from TT1 to TT2 during the index pregnancy (not yet implemented) .....	75

Tetanus Survey – Measures Related to Quality of Services.....	77
TT_QUAL_01: Percent of TT2 doses administered before 4 weeks had passed (not yet implemented).....	77
TT_QUAL_02: % of TT3 doses administered within 26 weeks of TT2 (not yet implemented) .....	78
TT_QUAL_03: Percent of women with missed opportunity for TT1 vaccination (not yet implemented).....	79
TT_QUAL_04: Tetanus card availability (not yet implemented).....	80
Tetanus Survey – Measures Related to Providers of Services.....	81
TT_PROV_01: Percent of TT vaccinations performed by each source (not yet implemented) .....	81
TT_PROV_02: Place of delivery (not yet implemented) .....	82
Post-SIA Survey – Measures Related to Coverage.....	83
SIA_COVG_01 Crude SIA coverage .....	83
SIA_COVG_02: Crude SIA coverage where SIA dose was the first dose .....	85
SIA_COVG_03: Lifetime measles doses, by birth cohort .....	86
SIA_COVG_04: Campaign doses compared to prior number of doses received .....	89
SIA_COVG_05: Clusters with surprisingly low coverage .....	91
Post-SIA Survey – Measures Related to Quality of Services .....	93
SIA_QUAL_01: Received a campaign card .....	93
Hypothesis Testing – Testing for Differences in Coverage .....	95
CVG_DIFF_01: Differences between strata .....	95
CVG_DIFF_02: Differences between subpopulations within a stratum .....	97

This page is intentionally blank.

## ACKNOWLEDGEMENTS

This document was developed for the World Health Organization by Biostat Global Consulting in 2015. It was written by Dale Rhoda.

The organization and contents of this document have been greatly aided by reviewing documents that describe:

- a) the Coverage Survey Analysis System (COSAS)<sup>1</sup>
- b) the Pan-American Health Organization's 2014 draft description of so-called Module 6
- c) earlier versions of WHO EPI Immunization Coverage Cluster Survey guidance<sup>2</sup>
- d) the 2015 draft WHO EPI Vaccination Coverage Cluster Survey Reference Manual<sup>3</sup>
- e) the 2018 WHO Vaccination Coverage Cluster Survey Reference Manual<sup>4</sup>

---

<sup>1</sup> Management and Evaluation of Immunization Programs through the Use of Coverage Surveys and Computerized Analysis Manual for EPI Managers, Draft June, 1991.

Desvé, Giles; Havreng Jean François; Brenner, Eric. COSAS 4.3 – Programme for Analysis of Immunization Coverage Surveys, Developed by EPICENTRE for the Expanded Programme on Immunization, WHO, November, 1991.

WinCOSAS User's Manual, Draft May 2001.

<sup>2</sup> 2005 Reference Manual; 1979, 1991 and 2007 Mid-Level Manager's Manuals

<sup>3</sup>

[http://www.who.int/immunization/monitoring\\_surveillance/Vaccination\\_coverage\\_cluster\\_survey\\_with\\_annexes.pdf?ua=1](http://www.who.int/immunization/monitoring_surveillance/Vaccination_coverage_cluster_survey_with_annexes.pdf?ua=1)

<sup>4</sup> [https://www.who.int/immunization/documents/who\\_ivb\\_18.09/en/](https://www.who.int/immunization/documents/who_ivb_18.09/en/)

This page is intentionally blank.

## INTRODUCTION

This document describes some vaccination program performance measures that can be calculated from household survey data. The document is organized around three types of surveys:

1. Routine Immunization (RI) (typically examining coverage among children aged 12-23 months)
2. Tetanus (typically administered to women who have had a live birth in the past 12 months)
3. Post-SIA (typically administered directly following supplemental immunization activity (SIA or “campaign”) to assess vaccination coverage following a mass vaccination campaign); also sometimes called a Post Campaign Coverage Survey (PCCS)

Within each survey type, most of the measures are organized by vaccination program attributes that have proven useful in earlier assessments:

- Coverage (COVG)
- Access (ACC)
- Continuity of services (CONT)
- Quality of services (QUAL)
- Provider of services (PROV)

Or by the type of figure that they produce:

- Cumulative coverage curves (CCC)
- Cumulative interval curves (CIC)
- Vaccination coverage and timeliness charts (VCTC)

And several measures or types of output are described that are not specific to one type of survey:

- Raw data listing (DATA)
- Describing the survey sample (DESC)
- Testing for statistical differences in coverage (DIFF)

Some of the indicators described here have been coded into the WHO Stata software named Vaccination Coverage Quality Indicators (VCQI) (pronounced “Vicki”) but this document is intended to have a broader application than simply documenting VCQI.

Over the years, different organizations have calculated a variety of coverage survey indicators to describe vaccination program performance. Sometimes those indicators describe the same underlying programmatic construct and sometimes the indicators are identical or have much in common with those calculated by others. Sometimes there are important (or subtle) differences in how indicators are defined and calculated and best interpreted. Our goal here is to furnish a template to describe the important aspects of indicators that WHO recommends. Some elements of the template might be useful for describing indicators calculated by other individuals and organizations, too, and for communicating about similarities and differences in calculations. Finally, the template might help identify gaps where it would be useful to define additional indicators and to put new ones in context with those that already exist.

It is our expectation that this document will be updated over time.

## Introduction

### A note on strata

The updated [2018 WHO Vaccination Coverage Cluster Survey Reference Manual](#) can be used to plan a survey in a single geographic stratum, but many of the surveys conducted now include several geographic strata. This document assumes that there will be multiple strata in the dataset and that output for most of the measures will be provided in tables where the rows list the strata and the columns list different aspects of the measures (Estimated %, 95% Confidence Interval, Design Effect, N, etc.).

The remainder of this document does not mention strata at all. In every case, we assume that the measures will be calculated and reported in a separate row for each stratum. We also assume that every analysis will make use of the variables that identify the strata, so we have not listed those variables repeatedly here.

### A note on variable names

This document should be accompanied by another one entitled *Vaccination Coverage Surveys – Forms & Variable Lists (FVL) Structured for Compatibility with VCQI*. That document lists candidate questions for coverage survey forms. In many cases this document references variable names that are defined in that document. For instance, RI20 is the 20th data element from the draft list of questions for a routine immunization survey: the sex of the child whose vaccination record is being discussed. Interpreting this document will be easier if the reader also has access to the FVL document.

## OTHER INPUTS

In addition to the data from the survey itself, it is necessary to assemble at least three other kinds of input parameters to conduct coverage analyses:

1. Parameters that describe the vaccination schedule in the country
2. Parameters that describe the survey protocol
3. Parameters that describe which analyses and which outputs are desired

## VACCINATION SCHEDULE METADATA

1. List of doses and the age (or interval) at which vaccinations were scheduled in the country with the survey at the time that the cohort surveyed should have been vaccinated (often the 24 months preceding the survey, but possibly longer or shorter).
2. Provide the following parameters for the program to access:
  - a. If it is a single-dose vaccine, or the first dose in a series, indicate the dose name, youngest age at which it should be administered (days) and optionally the oldest age at which it should be administered (days)
  - b. If it is the second or later dose in a series, the name of the dose, the minimum age (days) and the minimum interval that should have elapsed since the earlier dose (days) in order for the later dose to be considered valid

## COVERAGE SURVEY METADATA

- A. Provide an indication of whether the survey protocol had data collectors seek health center records (EPI register data). There are three possible options:

- a. Health facility data were collected for every respondent (RECORDS\_SOUGHT\_FOR\_ALL)
- b. Only collected for those who did not furnish vaccination cards during the household survey (RECORDS\_SOUGHT\_IF\_NO\_CARD), or
- c. Did not seek vaccination records at the health centers at all (RECORDS\_NOT\_SOUGHT)

This distinction will inform what to do for records with dates on both the card and the register that do not agree with each other. If the survey is described as RECORDS\_SOUGHT\_IF\_NO\_CARD, but for some children the survey data include dates from both the home-based card and the health center record, then for purposes of calculating the indicators, the card will be used and the register data will be ignored (in line with advice in Section 5.4.2 of WHO's 2018 Vaccination Coverage Cluster Survey Reference Manual).

If the survey is described as RECORDS\_SOUGHT\_FOR\_ALL, the dataset might include card and register records for many respondents, and they might sometimes disagree. In the spirit of section 5.4.2, we recommend calculating the measure with both sources and reporting the results from each respondent that indicates the best news or gives the benefit of the doubt to the vaccination program. So if the card indicates that the dose was valid, while the register indicates it was not, use the data from the card for that respondent.

This document describes the logic for calculating indicators in all three circumstances: SOUGHT\_FOR\_ALL, SOUGHT\_IF\_NO\_CARD and NOT\_SOUGHT.

## COVERAGE ANALYSIS OPTIONS

Having described the vaccination schedule and the survey protocol using programming parameters, the analyst is almost ready to conduct an analysis. Different analysts may have different goals and may specify different options each time they analyze the data: Which indicators will we calculate today? Are we making tables of output? Figures, too? Which strata should be listed? And in what order? Should each indicator be calculated only once, or are we doing some what-if analysis where we calculate a single indicator several times in a single run using several sets of options?

Each indicator has its own section of the document that lists relevant analysis options.

## ACRONYM LIST

CI	Confidence interval
COSAS	Coverage Survey Analysis Software
DEFF	Design effect
DHS	USAID Demographic & Health Survey
DNK	Do not know
LCB	Lower one-sided confidence bound
HC	Health center (might be used interchangeably with “health facility”)
ICC	Intracluster correlation coefficient
MCV	Measles Containing Vaccine
MICS	UNICEF Multiple Indicator Cluster Survey
MOV	Missed opportunity for vaccination
NEFF	Effective sample size
UCB	Upper one-sided confidence bound
VCQI	Vaccination Coverage Quality Indicators (software to analyze coverage survey data)

## DOSE NOTATION

This document uses the term *dose* to mean a combination of vaccine and the number in the sequence, when applicable, so bcg, opv0, hepb, dpt1, dpt2 and dpt3 are each separate doses. In cases where an analysis is dose-specific, the document often uses the nomenclature *<dose>* to mean *substitute the appropriate dose name here*. So, when the document says to generate a variable named *got\_<dose>*, there will be a variable named *got\_bcg* when the analysis is conducted for BCG, named *got\_dpt1* when it is conducted for DPT1, etc.

## WEIGHTED AND UNWEIGHTED ANALYSES

Any template or document that defines carefully how to calculate a vaccination coverage program indicator will need to be clear whether the indicator is weighted, and if so, how the weighting is conducted. This document uses a brief “yes/no” nomenclature for whether the indicator is weighted, but if the template is used to describe other indicators, then more detail might be warranted.

Many of the analyses listed here are described as “Weighted: Yes”. Those analyses are always weighted even if there are some respondents for whom we do not have sufficient data to put them in the numerator. Valid coverage is a good example...if we do not have vaccination dates from the card or register then we cannot say that a respondent got a valid dose, but according to the conventions of indicator RI\_COVG\_02, for example, we put all respondents in the denominator so the measure is interpreted as “% of the population represented by the respondents for whom we a) had data elements required and b) found evidence that they received a valid dose of *<dose>*”.

Some of the analyses listed below are described as “Weighted: No”. These are almost exclusively analyses where only a subset of respondents will be in the denominator, so it could be confusing to draw conclusions about the overall population. In VCQI, these analyses are currently unweighted. In a future revision it might be advisable to add a capability to recognize when more than some user-specified proportion, (like 90%) of the sample is in the denominator and more than some user-specified count, (like 2) respondents in every cluster are in the denominator, in which case maybe the weighted proportion is meaningful and the user could be informed, in a footnote of the unweighted analysis, that a weighted analysis is available. The user could obtain the weighted analysis by going back and including

an option to FORCE calculation of the weighted measure. This is an idea that should be discussed with a user group, perhaps. For now, if this document says it is unweighted, it is unweighted.

## COMPUTATIONAL NOTATION

Some ideas in this document are expressed in English prose and in computer code. The computer code is often written using syntax from Stata. This is a brief key to what some of that syntax means.

==	is exactly equal to
&	Boolean AND
	Boolean OR
!	Boolean NOT
!=	is <u>not</u> equal to
missing( <i>variable</i> )	For the data row or record of interest, the value of <i>variable</i> is missing
!missing( <i>variable</i> )	For the data row or record of interest, the value of <i>variable</i> is <u>not</u> missing
bysort <i>variable</i> :	The calculation listed after the colon is calculated over and over again, once for each set of rows or records that form a group with a unique value of <i>variable</i> ; so if we said bysort gender: the calculations would be conducted once considering only rows for female respondents and once for rows considering only males
bysort <i>variable list</i> :	The calculation to the right of the colon is done once for each set of records having a unique combination of values of the variables in <i>variable list</i> . So bysort stratumID clusterID: would repeat the calculations on each unique combination of stratumID and clusterID.

## DOCUMENT REVISION HISTORY

2015-07-27	Original draft
2015-09-20	Added details on details of calculations
2015-09-21	Added additional details
2015-09-26	Added additional details and made adjustments in accordance with discussions with WHO on 9/23
2015-09-27	Draft including specifications and test plans delivered to WHO
2015-11-28	Draft to accompany VCQI Stata Software v1.00
2017-01-31	Updated so 'Derived Variables' sections use the same variable names as VCQI; updated interpretations of most indicators
2017-06-16	Added EPI logo to cover & included Acknowledgements
2019-10-09	Updated description of RI_QUAL_01 to reflect its new variety of numerators
2019-10-09	Added descriptions of new indicators SIA_COVG_04 and SIA_COVG_05. The latter is based very closely on RI_COVG_05. The former is a specifically stratified version of SIA_COVG_01.
2021-02-12	Updated description of calculations for missed opportunities for vaccination. Updated with clarifications

## ALL SURVEYS: RAW DATA LISTING

### DATA\_01 Tables of cluster level response counts

Description: Some of the forms in Annex G of the 2005 WHO Reference Manual listed cluster level results in tables of raw counts. See that document for additional detail.

This indicator has not been coded explicitly, but for RI surveys it is possible to use RI\_COVG\_05 to generate a list of clusters and the count and % of respondents in each cluster who were found to be vaccinated.

## ALL SURVEYS: DESCRIBING THE SURVEY SAMPLE

### DESC 01: Expected vs observed sample: clusters, households, & respondents

**Description:** The sample design will have employed several parameters to estimate the number of completed responses for the survey. This table will describe the expected number of clusters, households, and respondents, based on records from the design phase, and will document the actual numbers observed in the survey. It will document the number of households visited, the number of households where no one was home, the number of respondents who refused, etc.

**Weighted:** No

**Denominator:** None

**Numerator:** Counts from survey design and from survey dataset

**Input variables:** Cluster ID, Household ID, Respondent ID, disposition codes from household visits, disposition codes from individual interviews, expected HH per cluster

**Table Output:** Report counts (N) and where appropriate, percentages

**Provenance:** 2015 WHO Reference Manual

To discuss

with WHO: The headings for VCQI's table are:

#### Proposed RI & TT survey summary table elements

- A. Total households in sample (expected)
- B. Total households in sample (observed)
- C. Households with information about survey eligibility from occupant
- D. Households with information about survey eligibility from neighbors
- E. Households with no information about survey eligibility

Note: We expect  $B = C + D + E$

- F. Number of eligible individuals in sample
- G. Number of respondents with completed interviews (expected)
- H. Number of respondents with completed interviews (observed)
- I. Number of eligible respondents unavailable
- J. Number of eligible respondents who refused to be interviewed
- K. Number of eligible respondents with no completed interview for another reason

Note: We expect  $F = H + I + J + K$

- L. Number of health centers visited
- M. Number of respondents whose records were sought in health centers
- N. Number of respondents whose records were found in health centers

## All Surveys: Describing the Survey Sample

Note that there may be additional elements for SIA summary tables:

### Proposed SIA summary table elements

- A. Total households in sample (expected)
  - B. Total households in sample (observed)
  - C. Households with information about survey eligibility from occupant
  - D. Households with information about survey eligibility from neighbors
  - E. Households with no information about survey eligibility
  - F. Number of eligible individuals in sample
  - G. Number of respondents with completed interviews (expected)
  - H. Number of respondents with completed interviews (observed)
  - I. Boys aged 9m-4y
  - J. Boys aged 5-9y
  - K. Boys aged 10-14y
  - L. Boys aged 15-18y
  - M. Girls aged 9m-4y
  - N. Girls aged 5-9y
  - O. Girls aged 10-14y
  - P. Girls aged 15-18y
  - Q. Number of eligible respondents with no caretaker available
  - R. Number of eligible respondents where caretaker refused to be interviewed
  - S. Number of eligible respondents with no completed interview for another reason
- 
- These categories could vary from survey to survey

**DESC\_02: Response to multiple choice question (select a single response)**

**Description:** Most surveys will include some multiple-choice questions to be summarized in the survey report. In some cases, the questions document the demographics of the survey respondents. Others are questions about the content of the survey itself. Regardless of the content of the question and responses, the options below will be available for summarizing responses and the software user will need to specify what sort of question it was, which variables hold the responses, and how to summarize them (e.g., in a weighted or unweighted fashion)

**Weighted:** Yes or No

**Denominator:** Number of respondents who answered the question (if unweighted)  
Sum of weights for all respondents (if weighted)  
Sum of weights for all respondents who answered the question  
(if the user specifies this option)

**Numerator:** Number of respondents who selected a particular choice (unweighted)  
Sum of weights for respondent who selected that choice (weighted)

**Variations:** Select the single most important response option  
Weighted % rather than unweighted  
Denominator = all respondents rather than all respondents who answered

**Input variables:** If the respondent can select at most one response, then the response can be stored in a single variable. That is the situation covered in DESC\_02.

**User inputs:** For each requested table, the user specifies

1. Whether the results are weighted or unweighted
2. Whether the denominator includes everyone or only persons who gave a valid response
3. Whether to aggregate across any responses to generate sub-totals
4. If yes, how many sub-totals
5. For each sub-total, which response options go into the sub-total
6. For each sub-total, a label for the output tables

**Derived variables:** Assess the number of response option levels, including sub-totals if applicable.

Generate a new derived indicator variable for each level, named <prefix>\_level\_*n* where *n* is an integer that goes from 1 up to the number of levels. Indicator variables take the values 0 or 1 or missing.

Set each indicator to zero.

Loop over all the indicators. Replace the indicator with 1 if the respondent selected the option represented by the indicator. If the case of sub-totals, replace the indicator with

## All Surveys: Describing the Survey Sample

1 if the respondent selected any one of the options represented by the sub-total indicator.

If the denominator includes only persons who gave valid responses, set the indicator to missing if the respondent did not give a valid response.

**Calculation:** The outcome is the survey weighted (or un-weighted) average of the indicator variable, as requested by the user.

**Interpretation:** Depends on the user's selections:

1. Unweighted, all respondents in denominator: Among the N respondents, X% selected this response option.
2. Unweighted, only valid responses in denominator: Among the N respondents who answered the question, X% selected this response option.
3. Weighted, all respondents in denominator: X% of eligible respondents in the population are estimated to be in the category of person who would select this response option.
4. Weighted, only valid responses in denominator: X% of eligible respondents in the population who would answer this question are estimated to be in the category of person who would select this response option.

**Table Output:** Report the % for each level (response option) and sub-total along with total N. Report weighted N if the user requests weighted results. Report 95% CI for each option if the calculation is weighted.

**Provenance:** COSAS & Others

**DESC\_03: Response to multiple choice question (select all that apply)**

**Description:** DESC\_03 summarizes responses to questions where the respondent can ‘select all that apply’: e.g., *What are the reasons your child has not received all the doses in the national immunization plan? (Select all that apply)* These summaries differ slightly from DESC\_02; if some respondents select numerous responses, the number of responses can exceed the sample size and the sum of % responses can exceed 100%.

**Weighted:** Yes or No

**Denominator:** Number of respondents who answered the question (if unweighted)  
Sum of weights for all respondents (if weighted)  
Sum of weights for all respondents who answered the question  
(if the user specifies this option)

**Numerator:** Number of respondents who selected a particular choice (unweighted)  
Sum of weights for respondent who selected that choice (weighted)

**Variations:** Mark all response options that apply  
Weighted % rather than unweighted  
Denominator = all respondents rather than all respondents who answered

**Input variables:** If the respondent is allowed to mention more than one response, then each response option is held in its own variable. This is the situation covered in DESC\_03.

**User inputs:** For each requested table, the user specifies

1. Whether the results are weighted or unweighted
2. Whether the denominator includes everyone or only persons who gave a valid response
3. The variables that hold the question responses
4. The value that indicates the respondent selected that response.
5. Whether to aggregate across any responses to generate sub-totals
6. If yes, how many sub-totals
7. For each sub-total, which response options go into the sub-total
8. For each sub-total, a label for the output table

**Derived variables:** Assess the number of response option levels, including sub-totals if applicable.

Generate a new derived indicator variable for each level, named <prefix>\_level\_*n* where *n* is an integer that goes from 1 up to the number of levels

Set the indicator to zero.

Loop over the indicators. Replace the indicator with 1 if the respondent selected the option represented by the indicator. If the case of sub-totals, replace the indicator with 1 if the respondent selected any one of the options represented by the sub-total indicator.

## All Surveys: Describing the Survey Sample

If the denominator includes only persons who gave valid responses, set the indicator to missing if the respondent did not give a valid response.

**Calculation:** The outcome is the weighted or un-weighted average of the indicator variable, as requested by the user.

**Interpretation:** Depends on the user's selections:

1. Unweighted, all respondents in denominator: Among the N respondents, X% selected this response option.
2. Unweighted, only valid responses in denominator: Among the N respondents who answered the question, X% selected this response option.
3. Weighted, all respondents in denominator: X% of eligible respondents in the population are estimated to be in the category of person who would select this response option.
4. Weighted, only valid responses in denominator: X% of eligible respondents in the population who would answer this question are estimated to be in the category of person who would select this response option.

**Table Output:** Report the % for each option along with total N. Also report weighted N if the user requests weighted results. Report 95% CI for each option if the calculation is weighted.

**Provenance:** COSAS & Others

## RI SURVEY – DATA CLEANING

The topic of this document is indicator definitions...not data cleaning...but of course survey data are often messy and inconsistent. In the sections that follow, we assume that the dataset has been cleaned by a process that is not described here, but whose output has the properties described below. Although electronic data capture holds the potential for prohibiting bad values, we find that organizations vary tremendously in whether the program automatic data checks, and whether they do so thoroughly. So even if the data are collected using (allegedly) smart phones or electronic tablets, there will always be a need to carefully review the data and take measures to ensure that:

1. Birth date data are clean
  - a. Each child has a single birth date that has been identified as appropriate for date-based age-at-vaccination calculations, or they do not
  - b. The birth date is on or after the earliest birth date for respondents to be eligible for this survey
  - c. The birth date is on or before the latest birth date for respondents to be eligible for this survey
  - d. The birth date is a complete (not partial) date
  - e. The date is a real date...not a nonsensical combination of numbers like February 30 or April 31
  - f. There may be several other birth date related variables...as recorded on the vaccination card, as recorded on the health center record, as remembered by the caregiver...and those variables may or may not be clean. They will not be used for date-based age-at-vaccination calculations...only the variable that holds the cleaned date will be used. In this document that variable is named 'dob'. And dob may be missing for some respondents.
  
2. Vaccination date data are clean
  - a. Each dose has two variables for recording the clean date: one for recording the date from the card and one for recording the date from a health center record
  - b. Each vaccination date occurs on or after the earliest possible birth date for eligibility in this survey
  - c. Each vaccination date occurs on or before the final date of survey fieldwork
  - d. Each vaccination date occurs on or after the child's date of birth (dob) if dob is defined
  - e. Each vaccination date is complete (not partial)
  - f. Each vaccination date is a real date
  - g. Dates for a vaccine sequence (DPT1, DPT2, DPT3) occur in order
  - h. For any child and dose, the clean vaccination date from the card may differ from the date from the EPI register, but if they are both in the dataset, they will at least both be clean dates as described here
  
3. For each dose, the dataset includes a variable to indicate that the caregiver recalls that the child had the dose. These variables are sometimes describes using the terms "recall" or "verbal history" or just "history". The variable takes (at least) four possible values: yes, no, don't know, missing.

## Routine Immunization Survey: Data Cleaning

4. For each dose, the dataset includes two “tick” variables: one to indicate partial evidence of vaccination on the card and another to indicate partial evidence of vaccination on the EPI register at the health center. Each tick variable takes one of two possible values: yes (ticked) or no (not ticked). The tick is set (yes) if there is a mark or indication that the child received the dose, but the evidence is not a complete clean date and therefore we cannot calculate the age at which the child received this dose. For any individual vaccine and source (DPT2 from card), if the date is set and clean, then the corresponding tick takes the value ‘no’. In other words, ticks are only used to represent evidence of vaccination when there is no clean date.

So to summarize, the descriptions and specifications below assume that vaccination evidence is contained in the variables listed below. For each child, any of these variables might be set and clean (as described above) or might be missing.

- a clean date of birth (dob)

For each vaccine and dose:

- a clean date from the home-based record  
(Described in the calculation instructions as: <dose>\_card\_date.)
- a clean tick from the card  
(Described in the calculation instructions as: <dose>\_card\_tick.)
- a clean date from the EPI register at the health center  
(Described in the calculation instructions as: <dose>\_register\_date.)
- a clean tick from the register  
(Described in the calculation instructions as: <dose>\_register\_tick.)
- a clean indication of caregiver recall  
(Described in the calculation instructions as <dose>\_history.)

And for the BCG vaccine there can be one additional source of information. In some survey protocols, when the child is present during the interview, the data collector looks at the child’s arm and records whether there is evidence of a BCG scar. (Item R172 in the accompanying FVL document.) The specifications below assume that this will be recorded with three values: yes, no or child not present. (Described in the calculation instructions as: bcg\_scar.)

## RI SURVEY – MEASURES RELATED TO COVERAGE

### RI COVG 01: Crude coverage

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received the vaccine dose according to [card, register, history]

Vaccines: Calculated for each dose of each vaccine

Time options: By the time of survey

Variations: By card  
By history  
By register  
By card or history (for purpose of comparison with older surveys)  
By card or history or register  
To analyze (this is the main crude coverage outcome)

Survey variables:

- Variable like RI33: Date of dose (card) [date or missing; date could be incomplete]
- Variable like RI34: No date, but tick mark on card [1 = yes; 2 = no; missing]
- Variable like RI71: Received dose (history) [1 = yes; 2 = no; 99 = DNK; missing]
- Variable like RIHC23: Date of dose (register) [date or missing; could be incomplete]
- Variable like RIHC24: No date, but tick mark in register [1=yes; 2=no; missing]
- (BCG is special in that it also has RI72: Interviewer sees evidence of scar on child; this is handled as a special case for BCG; no other vaccine has this extra question.)

Cleaned variables:

- <dose>\_card\_date
- <dose>\_card\_tick
- <dose>\_history
- <dose>\_register\_date
- <dose>\_register\_tick

Derived

variables: All coded 1 = yes; 0 = no; Each variable takes on a missing value if missing all inputs.

- Got dose according to card:  
got\_crude\_<dose>by\_card = !missing(<dose>\_card\_date) | <dose>\_card\_tick == 1
- According to history: got\_crude\_<dose>\_by\_history = <dose>\_history == 1

## Routine Immunization Survey: Measures Related to Coverage

- According to the EPI register:  
got\_crude\_<dose>\_by\_register = !missing(<dose>\_register\_date) |  
<dose>\_register\_tick == 1
  - According to card or history:  
got\_crude\_<dose>\_c\_or\_h = got\_crude\_<dose>\_c == 1 | got\_crude\_<dose>\_h == 1
  - According to card or register:  
got\_crude\_<dose>\_c\_or\_r = got\_crude\_<dose>\_c == 1 | got\_crude\_<dose>\_r == 1
  - According to card or history or register:  
got\_crude\_<dose>\_c\_or\_h\_or\_r = got\_crude\_<dose>\_c\_or\_h == 1 |  
got\_crude\_<dose>\_r == 1
- b. BCG also incorporates info about whether the data collector saw the scar. The absence of a scar does not over-ride other sorts of evidence, but presence of a scar does over-ride both card and register
- got\_crude\_bcg\_by\_scar: bcg\_scar\_history == 1
  - got\_crude\_bcg\_by\_card\_or\_history: got\_crude\_bcg\_c\_or\_h == 1 | bcg\_scar\_history == 1
  - got\_crude\_bcg\_by\_c\_or\_h\_or\_r: got\_crude\_bcg\_c\_or\_h == 1 | got\_bcg\_r == 1
- c. got <dose>\_to\_analyze: this is the definitive output used to summarize the measure and it depends:
- If RECORDS\_NOT\_SOUGHT: = got\_<dose>\_c\_or\_h
  - If RECORDS\_SOUGHT\_IF\_NO\_CARD:
    1. = \_c\_or\_h if respondent has a card
    2. = \_register == 1 | \_history == 1 if no card
    3. If BCG: also set to 1 if bcg\_scar\_history == 1
  - If RECORDS\_SOUGHT\_FOR\_ALL: - \_c\_or\_h\_or\_r

**Calculation:** This calculation is an estimated proportion (which is the same as a simple weighted average of the 0 and 1 values in the numerator). The point estimate may be obtained without taking the survey design into effect. Any estimate of uncertainty should include adjustment for the survey design.

If the respondent is in the analysis dataset, then their weight appears in the measure denominator, even if all inputs are missing. We tried to find evidence of their vaccination status, and failed, so they contribute to the denominator, but not the numerator for crude coverage.

If the respondent has a 1 for the output of interest (by card, history, register, card or history, card or history or register) then their weight appears in the measure numerator.

So the denominator is the sum of all survey weights and the numerator is the sum of weights for which evidence of coverage was found in the survey.

Table Output: Estimated %, 95% confidence interval (CI), 1-sided 95% lower confidence bound (LCB), 1-sided 95% upper confidence bound (UCB), Design Effect (DEFF), Intracluster correlation coefficient (ICC), sample size N (unweighted), population size N (weighted)

Graph: Organ pipe & Inchworm plots

Interpretation: "X% of the population who were eligible for the survey are estimated to have received <dose>, as documented by <source(s)>."

Provenance: COSAS & others

Points for discussion with WHO:

- A. MICS seems to remove observations from the calculation if there is no card and the mother is not sure whether they had the dose or not. They do not appear in either the numerator (which is appropriate) or the denominator (which I would argue is not appropriate). I recommend against this practice, but we should discuss the decision.
- B. We will want to come up with a standard policy about which method of calculation to use for CI endpoints. We recommend the so-called *modified Wilson* intervals when the sample proportion is between 0 and 1. In several reviews they have been found to have good properties: close to 95% likely to contain the true population value even when that value approaches 0% or 100% and relatively narrow intervals when compared with others that have good coverage. When the sample % is 0% or 100% we recommend calculating a so-called *modified Clopper-Pearson* interval.

## RI COVG 02: Valid coverage

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received a valid dose according to [card, register]

Vaccines: Calculated for each dose

Time options: By the time of survey, or  
By 12 months of age,

Variations: By card  
By register  
By card or register  
To analyze (this is the main valid coverage outcome)  
By age 1

The main motivation is to assess valid coverage when using dates from both cards or health center registers. One reason for reporting card and register data alone, in addition to card or register, is to show how much coverage estimates increase when the survey team goes to the effort and expense of collecting data from health centers.

Parameters: Define age at which each dose is valid  
Define minimum interval before valid subsequent dose  
Use parameters discussed at start of this document to establish the hierarchy of evidence regarding card vs. register if they disagree. (Those parameters are: RECORDS SOUGHT FOR ALL and RECORDS SOUGHT IF NO CARD.)

Survey variables:

- RI21, 22, and 23: Birthdate from caregiver [may be missing or incomplete]
- RI32 (d/m/y) Birthdate from card [may be missing or incomplete]
- RIHC22 (d/m/y) Birthdate from register [may be missing or incomplete]
  
- Variable like RI33: Date of dose (card) [date or missing]
- Variable like RIHC23: Date of dose (register) [date or missing]

Cleaned variables:

- dob
- <dose>\_card\_date
- <dose>\_register\_date

Derived variables:

- got\_valid\_<dose>\_by\_card :
  1. if a single-dose vaccine this is easy; if age was in valid window, 1; otherwise 0
  2. if a multi-dose vaccine, check the first dose date; if invalid, evaluate the next dose for validity (e.g., if DPT1 is not valid, check the DPT2 date and see if it would have been valid for DPT1); proceed until all dates have been checked and used to assign validity flags
- age\_at\_valid\_<dose>\_card: note the age at which the child received the valid dose (days)
- valid\_<dose>\_age1\_card 1 if valid dose by card was by age 1; 0 if after age 1; missing if missing age\_at\_valid\_<dose>\_card
- got\_valid\_<dose>\_by\_register : same logic as card, but using dates from the register
- age\_at\_valid\_<dose>\_register: note the age at which the child rec'd the valid dose
- valid\_<dose>\_age1\_register 1 if valid dose by card was by age 1; 0 if after age 1; missing if missing age\_at\_valid\_<dose>\_register
- got\_valid\_<dose>\_c\_or\_r :
  - = got\_valid\_<dose>\_by\_card == 1 | got\_valid\_<dose>\_by\_register == 1
- valid\_<dose>\_age1\_c\_or\_r:
  - = valid\_<dose>\_age1\_card == 1 | valid\_<dose>\_age1\_register == 1
- got\_valid\_<dose>\_to\_analyze: this is the ultimate variable used to summarize valid coverage and it considers whether RECORDS\_SOUGHT\_FOR\_ALL or RECORDS\_SOUGHT\_IF\_NO\_CARD
  1. If RECORDS\_NOT\_SOUGHT then this is equal to got\_valid\_<dose>\_by\_card
  2. If RECORDS\_SOUGHT\_FOR\_ALL then this is equal to got\_valid\_<dose>\_c\_or\_r
  3. If RECORD\_SOUGHT\_IF\_NO\_CARD then for respondents with cards, this variable is equal to got\_valid\_<dose>\_by\_card and for those without cards it is equal to got\_valid\_<dose>\_by\_register and for those without cards or register data, it is missing
- valid\_<dose>\_age1\_to\_analyze: this is the ultimate variable used to summarize valid coverage and it considers whether RECORDS\_SOUGHT\_FOR\_ALL or RECORDS\_SOUGHT\_IF\_NO\_CARD
  1. If RECORDS\_NOT\_SOUGHT then this is equal to valid\_<dose>\_age1\_card
  2. If RECORDS\_SOUGHT\_FOR\_ALL then this is equal to valid\_<dose>\_age1\_c\_or\_r
  3. If RECORD\_SOUGHT\_IF\_NO\_CARD then for respondents with cards, this variable is equal to valid\_<dose>\_age1\_card and for those without cards it is equal to valid\_<dose>\_age1\_register and for those without cards or register data, it is missing

## Routine Immunization Survey: Measures Related to Coverage

Calculation: Simple estimated proportion, as described for RI\_COVG\_01

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Organ pipe & Inchworm plots

Interpretation: "X% of the population who were eligible for the survey are estimated to have a documented record of vaccinations (<source(s)>) and to have received a valid dose of <dose>."

Note: The survey report should describe what is meant by a "valid dose".

- a) The child had reached the minimum age of eligibility for this dose.
- b) If the schedule specifies a maximum age of eligibility, then the child was within the allowable age range when they received the dose.
- c) If the dose is number 2 or 3 (or higher) in a sequence, then the minimum interval had passed since receiving the earlier dose, so the child was eligible to receive the next dose.

Provenance: COSAS & others

Points for discussion with WHO:

- A. COSAS did not use later dose dates to give credit for valid doses (i.e., count DPT2 as the valid DPT1 dose). If an early dose was invalid then the later doses were not able to be considered valid. I suggest we break with that tradition and shift dates to the left, as available. Let's discuss.
- B. In the COSAS manuals it indicates that BCG is only considered to be valid if the survey data collector sees the scar on the child. Do we want to keep that important criterion, or let it drop?

## Specification to calculate dob

Inputs:

- dob d/m/y from history (may be incomplete)
- dob d/m/y from card (may be incomplete)
- dob d/m/y from register (may be incomplete)
- earliest possible dob for this survey
- latest possible dob for this survey

The goal is to calculate a birthdate to use in valid dose calculations and age-at-vaccination analyses.

For each respondent:

Generate a variable named dob and set it to missing.

1. If there is a non-missing source of m, and d, and y and if all non-missing sources agree, then construct the date from that. [Note that m could come from the card, d could come from history, and y could come from the register. The point here is that if only one month is listed in all 3 dates sources, only one day, and only one year, then combine them to make a birthdate.]

If the constructed birthdate falls before the earliest possible or after the latest possible dob for this survey then clear out the dob value; otherwise set dob equal to the constructed value.

2. Otherwise, if either the card or the history or the register has a complete set of m/d/y then construct a date from them. If more than one source has a complete set of m/d/y then construct the date for each. For each source with m/d/y, check to see that the date falls between the earliest and latest possible dob values for this survey. If there remain more than one dob, then identify the earliest of them as the dob (as this will yield the highest number of valid doses)
3. Otherwise, we do not have a good combination of m/d/y from which to construct the dob and we will not conduct valid dose analyses for this child.

Note: In VCQI, the variable is named dob\_for\_valid\_dose\_calculations, but in this document we abbreviate that with dob.

### RI COVG 03: Fully vaccinated

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received all doses in the list that makes up “fully vaccinated” in that country at that time

Time options: By the time of survey (crude or valid), or  
By 12 months of age (valid)

Variations: By any source (if crude dose analysis)  
By card or register (if valid dose analysis)

Parameters: Define list of vaccines and doses that comprise “fully vaccinated”

#### Derived

variables: This analysis uses the derived variables for the previous two measures: crude coverage and valid dose coverage. It uses a list of vaccines and doses that comprise fully vaccinated. (hereafter “the list”)

– fully\_vaccinated\_crude:

= 1 if got\_crude\_<dose>\_to\_analyze == 1 for every <dose> in the list  
= 0 otherwise

– fully\_vaccinated\_valid:

= 1 if got\_valid\_<dose>\_to\_analyze == 1 for every <dose> in the list  
= 0 otherwise

– fully\_vaccinated\_by\_age1

= 1 if got\_valid\_<dose>\_by\_age1\_to\_analyze == 1 for every <dose> in the list  
= 0 otherwise

Calculation: Simple estimated proportion, as described for RI\_COVG\_01

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Organ pipe & Inchworm plots

Interpretation: “X% of the population who were eligible for the survey are estimated to be fully vaccinated, with <either crude or valid doses> having received <list of doses to be fully vaccinated>.”

Provenance: COSAS & others

Note: Fully vaccinated by age 1 is fully vaccinated with valid doses. Be sure to emphasize this in the output.

## RI COVG 04: Not vaccinated

Weighted:	Yes
Denominator:	Sum of weights for all respondents
Numerator:	Sum of weights for respondents who received none of the doses in the list that makes up “vaccinated” in that country at that time
Time options:	By the time of survey (crude or valid), or By 12 months of age (valid)
Variations:	By any source (if crude dose analysis) By card or register (if valid dose analysis)
Parameters:	Define list of vaccines and doses that comprise “vaccinated”

**Derived variables:** This analysis uses the derived variables for the previous two measures: crude coverage and valid dose coverage. It uses a list of vaccines and doses the comprise vaccinated. (hereafter “the list”)

– not\_vaccinated\_crude:

= 1 if got\_crude\_<dose>\_to\_analyze != 1 for every <dose> in the list  
= 0 otherwise

– not\_vaccinated\_valid:

= 1 if got\_valid\_<dose>\_to\_analyze != 1 for every <dose> in the list  
= 0 otherwise

– not\_vaccinated\_by\_age1

= 1 if got\_valid\_<dose>\_by\_age1\_to\_analyze != 1 for every <dose> in the list  
= 0 otherwise

**Calculation:** Simple estimated proportion, as described for RI\_COVG\_01

**Table Output:** Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

**Graph:** Organ pipe & Inchworm plots

**Interpretation:** “X% of the population who were eligible for the survey are estimated to be un-vaccinated, having no evidence of received any <crude or valid> doses of <list of doses to be fully vaccinated> by the sources of information examined in this survey.”

**Provenance:** COSAS & others

**Note:** Not vaccinated by age 1 means not vaccinated with any valid doses by age 1. Be sure to emphasize this in the output.

## RI\_COVG\_05: Clusters with suprisingly low coverage

**Description:** This indicator identifies clusters whose coverage for one or more doses of interest is below a user-specified threshold. The user can specify a single dose or a list of several. If the cluster level coverage for any of the doses of interest is below the threshold, then the cluster will be listed as one with suprisingly low coverage.

**Weighted:** User-specifies, yes or no

**Denominator:** Count (or sum of weights) for all respondents in the cluster

**Numerator:** Count (or sum of weights) for respondents who received the dose

**Vaccines:** Calculate for select doses

**Parameters:**

- Specify which doses are under consideration: RI\_COVG\_05\_DOSE\_LIST
- Specify whether to use number of respondents or % coverage  
RI\_COVG\_05\_THRESHOLD\_TYPE = [COUNT or PERCENT]
- Specify the threshold for ‘suprisingly low’ coverage
- RI\_COVG\_05\_THRESHOLD = [integer if COUNT, and percent (0-100) if PERCENT]
- Specify whether to generate a table for every stratum and list every cluster, or whether to generate a single table that lists only the clusters that fall below the threshold.
- RI\_COVG\_05\_TABLES = [ALL\_CLUSTERS or ONLY\_LOW\_CLUSTERS]

**Input**

**Variables:** This analysis uses the stratum ID, cluster ID, survey weight and the derived variable got\_crude\_<dose>\_to\_analyze described in the section on RI\_COVG\_01.

**Derived**

**Variables** Calculate:

- cluster\_n = number of respondents in the cluster
- got\_<dose>\_count = count of persons in the cluster who rec’d a crude dose of <dose>
- got\_<dose>\_sumwt = sum of weights for those who got <dose> in cluster
- cluster\_sumwt = sum of weights for all respondents in the cluster
- got\_<dose>\_pct= weighted % of respondents who got <dose> in the cluster

Save a dataset with one row per cluster.

**Calculation:** Make a list of clusters, in order of stratum ID and cluster ID, listing count of respondents where the coverage variable == 1 and weighted % of respondents where the coverage == 1. Generate output table or tables as described below.

**Table Output:** If making tables for each stratum (ALC\_TABLES == “ALL\_CLUSTERS”) then the table name (and Excel tab name) will list the stratum ID and name; the table will list:

## Routine Immunization Survey: Measures Related to Coverage

Cluster ID & name, count of respondents in the cluster, count of respondents with got\_crude\_<dose>\_to\_analyze == 1, % covered (sum of weights in the cluster where got\_crude\_<dose>\_to\_analyze == 1 divided by sum of all weights in the cluster); highlight those whose coverage (absolute or %) falls below the specified threshold; list clusters in order of cluster ID

If making a single table that lists only the clusters with ALC across all strata, the table will list:

Stratum ID & name, Cluster ID & name, count of respondents in the cluster, absolute count of respondents with got\_<dose>\_c\_or\_h\_or\_r == 1, % covered (sum of weights in the cluster where got\_<dose>\_c\_or\_h\_or\_r == 1 divided by sum of all weights in the cluster); only list the cluster in the table if its coverage meets the user-specified ALC criterion; list the strata in order of stratum ID; list the clusters in order of cluster ID

Graph: None (But the organ pipe plots from RI\_COVG\_01 are closely related to this indicator.)

Interpretation: “Low coverage is defined here as being a cluster where fewer than <threshold> <percent or individuals> showed evidence of vaccination. The clusters highlighted in this list show evidence of low coverage for at least one of <list of doses considered>.”

Provenance: COSAS had a cluster-level coverage (or count) report

## RI SURVEY – MEASURES RELATED TO ACCESS

### RI\_ACC\_01: Crude DPT1 coverage

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for all respondents who received crude dose of DPT1

Vaccines: DPT1 (This indicator can be used to make a table re-summarizing crude coverage of any single dose.)

#### Input

Parameters: Dose name

#### Derived

variables: This measure uses the derived variable described above under 'Crude Coverage':  
got\_crude\_dpt1\_to\_analyze

Calculation: This is the same as the RI\_COVG\_01 outcome.

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (weighted), N (unweighted)

Graph: Inchworm plots – produced with the output from RI\_COVG\_01

Interpretation: "X% of the population who were eligible for the survey are estimated to have access to vaccination services because they show evidence of having received <DPT1 / PENTA1>, as documented by <source(s)>."

Provenance: COSAS & others

## RI SURVEY – MEASURES RELATED TO CONTINUITY OF SERVICES

### RI CONT 01: Dropout between two doses

Weighted: No

Description: It is straightforward to calculate the weighted dropout results from the output from RI\_COVG\_01 and RI\_COVG\_02. For any two doses, (early and later) the dropout proportion is simply  $(Cvg\_early - Cvg\_later) / Cvg\_early$ . We could write code to generate a table, but for now it is so straightforward that we are focusing on trickier measures.

Although people are accustomed to seeing the weighted estimates, the weights themselves may not be especially helpful for this indicator. It may be more meaningful, or at least easier to understand, if the software also reports the unweighted dropout proportions, where the denominator is a count of respondents who got the early dose rather than the sum of their weights.

In keeping with the consistent policy for VCQI measures, if the denominator does not include all respondents, the software estimates an unweighted proportion and reports it without uncertainty...it is representative of the sample, but may not be representative of the larger population.

Denominator: Number of respondents who received the first dose

Numerator: Number of respondents who received the first dose and did not receive the second

Vaccines: Any pair due to be administered at different ages

Time options: By the time of the survey

Survey

variables: This measure uses the got\_crude\_<dose>\_to\_analyze variables described above in the RI\_COVG\_01.

Derived

variables:

Each dropout statistic considers coverage of an earlier dose (dose1) and a later one (dose2).

Generate a new indicator variable named 'dropout\_<dose1>\_to\_<dose2>'; set it to missing.

Replace the variable with the value 0 if the respondent has evidence of receiving both doses. (got == 1 for dose1 and for dose2)

Replace the variable with the value 1 if the respondent has evidence of receiving the first, but not the second dose. (got == 1 for dose1 and 0 for dose2)

## Routine Immunization Survey: Measures Related to Continuity of Services

(To be clear: the dropout variable is missing and the respondent does not appear in the denominator if got = 0 for dose1.)

Calculation: Dropout is the unweighted average of the indicator variable.

Interpretation: "Among the <N> children who showed evidence of having received <earlier dose>, (per card or recall <or register>), <dropout>% did not show evidence of receiving <later dose>."

Table Output: Estimated %, N (unweighted)

Graph: No

Provenance: COSAS & others

## RI SURVEY – MEASURES RELATED TO QUALITY OF SERVICES

### RI QUAL 01: Card and register availability

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: There are nine numerators:

1. Sum of weights for respondents whose card (home-based record) was seen
2. Sum of weights for respondents who had card with at least one date
3. Sum of weights for respondents who had card with at least one date or tick
4. Sum of weights for respondents who had card with only clean dates  
(where *clean* means the date fell between the child's DOB and the date of the survey (or between the earliest possible vaccination date and the date of the survey) and dates for dose series were in chronological order)
5. Sum of weights for respondents whose register (facility-based record) was seen
6. Sum of weights for respondents who had register with at least one date
7. Sum of weights for respondents who had register with at least one date or tick
8. Sum of weights for respondents who had register with only clean dates
9. Sum of weights for respondents whose card or register document was seen

#### Survey

variables: RI27 (Did the interviewer see the card?)

#### Cleaned

variables:

- <dose>\_card\_date
- <dose>\_card\_tick
- <dose>\_register\_date
- <dose>\_register\_tick

#### Derived

variables:

Create new variables that count dates and ticks across all doses on the card. If RI\_RECORDS\_SOUGHT\_FOR\_ALL or RI\_RECORDS\_SOUGHT\_IF\_NO\_CARD are set to 1, the same variables are created for the register.

- <card/register>\_date\_count
- <card/register>\_tick\_count

Use them to create the additional variables.

All coded 1 = yes; 0 = no

- had\_card  
= 1 if RI27 == 1 or card\_date\_count > 0 or card\_tick\_count > 0  
= 0 otherwise

## Routine Immunization Survey: Measures Related to Quality of Services

- had\_card\_with\_dates  
= 1 if card\_date\_count > 0  
= 0 otherwise
- had\_card\_with\_dates\_or\_ticks  
= 1 if card\_date\_count > 0 or card\_tick\_count > 0  
= 0 otherwise
- had\_card\_with\_flawless\_dates  
= 1 if card\_date\_count > 0 and card\_tick\_count == 0  
= 0 otherwise

(Note that if any of the dates on the card were flawed in some way...nonsensical or too early or too late or out-of-order, then they would have been converted to ticks by the program cleanup\_dates\_and\_ticks. So if there are dates and no ticks, then all the dates on the card are free from (obvious) flaws.)

- had\_card\_or\_register  
= 1 if had\_card == 1  
= 0 otherwise

If RECORDS\_SOUGHT\_IF\_NO\_CARD or RECORDS\_SOUGHT\_FOR\_ALL are set to 1 the below variables are created:

- had\_register  
= 1 if RI27 == 1 or register\_date\_count > 0 or register\_tick\_count > 0  
= 0 otherwise
- had\_register\_with\_dates  
= 1 if register\_date\_count > 0  
= 0 otherwise
- had\_register\_with\_dates\_or\_ticks  
= 1 if register\_date\_count > 0 or register\_tick\_count > 0  
= 0 otherwise
- had\_register\_with\_flawless\_dates  
= 1 if register\_date\_count > 0 and register\_tick\_count == 0  
= 0 otherwise

Replace had\_card\_or\_register = 1 if had\_register == 1

Calculation: Estimated proportion for each “had” derived variable, similar to RI\_COVG\_01.

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Organ pipe and Inchworm plots

Interpretation: To interpret columns labeled “RI Card Availability”: “An estimated X% of the population eligible for the survey have a home-based vaccination record and can produce it for interviewers to examine.”

To interpret the column labeled “RI Card with Dates”: “An estimated X% of the population eligible for the survey have a home-based vaccination record with at least one date on it.”

To interpret the column labeled “RI Card with Dates or Ticks”: An estimated X% of the population eligible for the survey have a home-based vaccination record with at least one date or tick on it.”

To interpret the column labeled “RI Card with Only Clean Dates”: An estimated X% of the population eligible for the survey have a home-based vaccination record with dates that are free from obvious mistakes.”

Interpretation of the *register* outcomes can use similar language, substituting the phrase *facility-based vaccination record* for *home-based vaccination record*.

To interpret the column labeled “RI Card or Register Availability”: An estimated X% of the population eligible for the survey have either a home-based or a facility-based vaccination record that is available for interviewers to examine.”

Provenance: COSAS & others

RI\_QUAL\_02: Ever had a card

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who say that they ever received a card for the child

Survey

variables: RI26 (Were you ever given a card for the child? 1=yes; 2=no; 99=DNK)

Derived

variables: Calculate a new variable (ever\_had\_an\_ri\_card); set it equal to 0.  
Replace the value with 1 if the respondent ever received a card (RI26 == 1)

Calculation: Estimated proportion, calculated in a manner similar to RI\_COVG\_01.

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Organ pipe & Inchworm plots

Interpretation: "X% of the population who were eligible for the survey are estimated to have received at least one home-based record (vaccination card), even if they no longer have it or could not produce it for inspection by the survey team."

Provenance: COSAS & others

## A Note Regarding Continuous and Categorical Timeliness Indicators

In a dataset with birth dates and vaccination dates, it is possible to calculate many indicators related to *timeliness* of vaccination: how many doses are given early? On-time? Late? How often do providers give doses simultaneously (in the same visit)? How often do they miss an opportunity for simultaneous vaccination? The remainder of the indicators listed under Quality of RI Services concern timeliness.

Different programs and projects have calculated different timeliness indicators over the years and several are included in WHO's VCQI software.

Some indicators in this document are described with the example of a particular dose, but may be calculated for other doses, as well.

This is not a comprehensive description of everything that could be calculated regarding timeliness, but a list of indicators that have proven useful. There is certainly room to add new indicators and new types of plots to this list.

Since being introduced in 2015, VCQI has added provisions to make three types of timeliness plots: Cumulative coverage curves (CCC), cumulative interval curves (CIC), and vaccination coverage and timeliness stacked bar charts (VCTC).

Cumulative coverage curves are produced via the indicator RI\_CCC\_02 which produces only curves, no tables. For each dose it shows the continuous weighted proportion of respondents with evidence of having received the dose at an age  $\leq$  the age indicated on the horizontal axis of the figure. These curves are sometimes called *reverse Kaplan-Meier curves*. The figures also annotate the recommended age(s) of vaccination so the viewer can gauge what portion of the target population received the dose early or late.

Although the CCC characterizes much of what is known about vaccination timeliness, some people find them to contain too much detail without an simple numeric takeaway message, so VCQI also includes RI\_QUAL indicators to succinctly summarize categorical timeliness variables.

Also, the vaccination coverage and timeliness stacked bar charts (VCTC - see VCQI Results Interpretation Quick-Reference Guide for an example and interpretation) are, conceptually speaking, categorical summaries of the cumulative coverage curves. They use a stacked bar chart to visually indicate what portion of the target population has evidence of a) showing a home-based record or facility record with vaccination dates, b) receiving the dose too soon, c) receiving the dose within 28 days of when it was due, d) receiving the dose more than 28 days late but less than two months late, e) receiving the dose 2+ months late, and f) receiving the dose, but without enough data to know what age they were when they received it. The indicator (RI\_VCTC\_01) produces stacked bar charts as its main output and saves an Excel sheet documenting the % of children represented by each colored section of the stacked bar chart. The user may accept the default timeliness categories listed above or may specify a new customized set of categories.

Finally, cumulative interval curves (CIC) can be produced via the indicator RI\_CIC\_02. Like the CCC indicator, this one makes only figures – no tables. For each pair of doses in a series, the CIC shows what proportion of respondents have evidence of receiving the later dose after an interval that is  $\leq$  the number of days indicated on the horizontal axis of the figure. This makes clear what portion of respondents receive the doses too close together and what portion receive them much too far apart. In order to quantify those concepts further, VCQI can make tables using RI\_QUAL\_05 (to summarize % of dose pairs received after too brief an interval) and RI\_QUAL\_12 (to summarize % of dose pairs received with too long an interval).

RI\_QUAL\_03: Percent of DPT1 doses that were invalid

Weighted: No

Denominator: Number of respondents who had DOB data and received DPT1 with a date, by card or register

Numerator: Number of respondents whose DPT1 was invalid (given too early)

Vaccines: DPT1 (can be calculated for any other single-dose vaccine or the first in a multi-dose series; this indicator should be calculated for dose 2 or later of a multi-dose series)

Derived variables: <Omitted; See Notes below.>

Calculation: The measure is the unweighted average of the indicator variable.

Table Output: % and N (unweighted) of observations set to 0 or 1

Graph: Unweighted proportion plot

Provenance: COSAS

Interpretation: We recommend running RI\_QUAL\_04 instead of \_03. See that measure for interpretation.

Notes: This indicator assumes that it is not desirable to receive the dose early, so it sets the indicator to 0 if either \_card or \_register indicate that they received the dose after the age of eligibility.

The steps to perform calculations for RI\_QUAL\_03 and RI\_QUAL\_04 and RI\_QUAL\_13 are the same. Starting in early 2021, we recommend that users of VCQI software call RI\_QUAL\_04 three times in each RI survey analysis:

1. Once to estimate the percent of DPT1 doses that were invalid
2. Once to estimate the percent of MCV1 doses that were invalid
3. And once to estimate the percent of DPT3 doses that were given before age 6 months

**RI\_QUAL\_04: Percent of MCV1 doses administered before 39 weeks of age**

Weighted: No

Denominator: Number of respondents who had DOB data and MCV1 with a date, by card or register

Numerator: Number of respondents whose MCV1 was given before 39 weeks of age

Vaccines: MCV1

(This indicator could be calculated for any dose and threshold. It assumes that it is not desirable to receive the dose before the threshold, so sets the indicator to 0 if there is evidence that the dose was received after the threshold.)

(Note: This indicator is similar to RI\_QUAL\_03, but slightly more generic and flexible because it compares with a user-specified threshold instead of the age of vaccination eligibility and because it can be used to evaluate any dose...not just a single dose vaccine or the first dose in a multi-dose series.)

User Inputs: RI\_QUAL\_04\_DOSE\_NAME (usually MCV1)

RI\_QUAL\_04\_AGE\_THRESHOLD (in days...usually  $\text{'=(39*7)'} = 274$ )

Derived variables:

This measure uses derived variables calculated above for the valid dose analysis: dob, <dose>\_card\_date, <dose>\_register\_date, no\_card

(In the notation here, <dose> = dose name and t = threshold.)

Generate two new indicator variables: early\_<dose>\_t\_card and early\_<dose>\_t\_register

Set each to 1 if that source indicates the dose was received before the age threshold; otherwise set to 0. Set early\_<dose>\_t\_card missing if the dataset is missing dob or dose\_date\_card. Set to early\_<dose>\_t\_register missing if the dataset is missing dob or dose\_date\_register.

Generate a new indicator named got\_early\_<dose>\_t and set it to missing. (This is the final indicator to report and it considers circumstances under which RECORDS were sought. This variable does not have the to\_analyze suffix described for some earlier indicators, but it could.)

If RI\_RECORDS\_NOT\_SOUGHT, make the indicator equal to early\_<dose>\_t\_by\_card.

If RI\_RECORDS\_SOUGHT\_FOR\_ALL, make it equal to 1 if either got\_<dose>\_by\_card. Replace the indicator with \_by\_register if \_by\_card is missing or \_by\_register is 0.

If RI\_RECORDS\_SOUGHT\_IF\_NO\_CARD, make the indicator equal to \_by\_card ; replace it with \_by\_register if missing \_by\_card.

Calculation: The measure is the unweighted average of the indicator variable.

Table Output: % and N (unweighted) of observations set to 0 or 1

## Routine Immunization Survey: Measures Related to Quality of Services

Graph: None

Interpretation: "Of N respondents in the sample for whom age-at-vaccination could be calculated for <dose>, X% received it before the age of <threshold> days."

Provenance: COSAS

Notes: This indicator assumes that it is not desirable to receive the dose early, so it sets the indicator to 0 if either `_card` or `_register` indicate that they received the dose after the threshold.

**RI\_QUAL\_05: Percent of DPT2 and DPT3 doses administered before 4 weeks had passed**

Weighted: No

Denominator: Number of DPT2 & 3 doses administered where the date was known for that dose and for the preceding dose

Numerator: Number of times the DPT2 or 3 dose was administered before 4 weeks had passed from the date of the earlier dose

Vaccines: DPT (This indicator may be calculated for any 2-dose or 3-dose vaccine sequence.)

User inputs: Vaccine name & interval (days)

Survey variables:

Variables like RI39, RI47 & RI55 (dates of DPT 1/2/3 by card)

Variables like RIHC29 & RIHC37 & RIHC45 (dates of DPT 1/2/3 by register)

Cleaned

Variables: <dose1>\_card\_date, <dose2\_card\_date>, <dose3>\_card\_date  
<dose1>\_register\_date, <dose2\_register\_date>, <dose3>\_register\_date

Derived variables:

Generate a new dataset named RI\_dose\_intervals.

- Each RI respondent in the survey dataset will be represented by several rows in the new dataset.
- The new dataset will include all the stratum and cluster and household and respondent ID variables as well as four new variables: early\_dose, late\_dose, card\_interval\_days and register\_interval\_days.
- The dataset will have one row per consecutive dose interval per respondent, so each respondent will have a row where early\_dose = “DPT1” and later\_dose = “DPT2”. Each respondent will have a row where early\_dose = “DPT2” and later\_dose = “DPT3”.
- The dataset only documents intervals between consecutive doses, so each respondent does not have a row where early\_dose is “DPT1” and later\_dose is “DPT3”.

Save this interval dataset, as it may be useful for other indicators or analyses.

Drop observations for all doses except the doses of interest.

Calculate a new indicator named short\_interval\_<dose>\_<threshold>\_c and set it to 1 if the card\_interval\_days is < the threshold. Set it to 0 if the interval is >= the threshold. Set it to missing if the interval is missing.

Calculate a new indicator named short\_interval\_<dose>\_<threshold>\_r and set it to 1 if the register\_interval\_days is < the threshold. Set it to 0 if the interval is >= the threshold. Set it to missing if the interval is missing.

## Routine Immunization Survey: Measures Related to Quality of Services

Calculate a new indicator variable named `short_interval_<dose>_<threshold>`. This is the indicator that will be used to calculate the outcome. It does not have the words “to\_analyze” in its name, but it could.

If `RECORDS_NOT_SOUGHT` then set `short_interval_<dose>_<threshold>` = `short_interval_<dose>_<threshold>_c`.

If `RECORDS_SOUGHT_IF_NO_CARD` then use indicator from card (`_c`) if seen, otherwise register (`_r`).

If `RECORDS_SOUGHT_FOR_ALL` then use whichever source gives the benefit of the doubt; in this case it is considered better if the interval exceeds the threshold so give the benefit of the doubt (set `short_interval_<dose>_<threshold>` to 0) if either card or register indicate that the interval was at least as long as the threshold.

**Calculation:** The measure is the unweighted average of the indicators.

**Table Output:** % and N (unweighted denominator) (Note that N is the number of intervals for which we have dates...not the number of children...be sure to say that in a footnote)

**Interpretation:** “Of N intervals in the sample where the data include dates for both the earlier and later dose, X% of the intervals were shorter than <threshold> days.”

**Graph:** Unweighted proportion plot

**Provenance:** COSAS

**Note:** It is possible to run the indicator more than once with different length thresholds to explore the effect of a so-called “grace period”. i.e., run it with a 28 day threshold and then change the `ANALYSIS_COUNTER` and run again with a 23 day threshold to evaluate how the outcome would change if we gave a 5-day grace period and considered 23 days to be an acceptable interval.

This indicator makes a dataset that documents the length of all first-to-second dose intervals and all second-to-third dose intervals. That dataset could be useful for plotting, say, the cumulative distribution of interval lengths.

**RI\_QUAL\_06: Percent of valid MCV1 doses that were administered before the age of 12 months**

Weighted: No

Denominator: Number of children who had valid dose of MCV1 (i.e., were age-eligible when they received MCV1)

Numerator: Number of children whose valid MCV1 was received before the age of 12 months

Vaccines: MCV1 (Could be run on any dose; threshold is always 12 months for this indicator.)

Derived variables:

This measure uses variables that were calculated for earlier measures:  
got\_valid\_mcv1\_to\_analyze and valid\_mcv1\_by\_age1\_to\_analyze

Generate a new indicator variable set to missing: got\_valid\_<dose>\_before\_age1

Replace the indicator variable with 0 if got\_valid\_mcv1\_to\_analyze is 1.

Replace the indicator variable with 1 if valid\_mcv1\_by\_age1\_to\_analyze is 1.

Calculation: The measure is the unweighted average of the indicator variable.

Table Output: % and N (count of respondents where indicator is 0 or 1)

Graph: Unweighted proportion plot

Interpretation: "Of N respondents in the sample who received a valid dose of <dose>, X% were administered before the age of 1 year."

Provenance: COSAS

Note: In many countries, children are scheduled to receive the first dose of measles vaccine at age 9 months. If doses are administered in a timely manner then the vast majority of children who receive a valid dose should receive it within three months of when it is due. (It would be even better to receive it within one month or even one week of when it is due...but this traditional indicator reports the proportion of children who receive it between before the age of 1 year. Of course this indicator will not be of interest in countries where the schedule recommends that the first dose of measles vaccine be administered at age 12 months.

**RI\_QUAL\_07B: Valid coverage if there had been no missed opportunities for simultaneous vaccination**

Update: RI\_QUAL\_07B replaces the original indicator RI\_QUAL\_07.  
The updated version uses more thorough logic.

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for all respondents who could have had a valid dose based on the schedule and their visit date(s)

Vaccines: Calculate for each vaccine and dose

Input

Parameters: N/A

Input

variables: This analysis uses the following variables: respondent ID, date of birth, and visit date(s). To make the output tables, it also uses the following variables: cluster ID, stratum ID, and weights, as well as level(s) 1, 2, 3, and 4 IDs and names.

Derived

variables: The data should be in long form, meaning one row per respondent ID per visit date for that respondent. Sort the data by respondent ID and visit date.

Generate a new derived variable named `age_at_visit` as the difference between visit date and date of birth. The units of this variable is age in days.

For each respondent, generate another new derived variable named `num_days_since_last_visit` as the difference between the `age_at_visit` variable for consecutive visit dates. Make sure the first `num_days_since_last_visit` is set to 0. In Stata, the code might look something like this:

```
bysort respid: gen num_days_since_last_visit = age_at_visit[_n] - age_at_visit[_n-1]  
bysort respid: replace num_days_since_last_visit = 0 if _n==1 & !missing(age_at_visit)
```

Next, calculate whether the respondent should have received a given dose based on the country's vaccination schedule and their age in days at that visit. We will do this first for single doses, then look at doses in a series.

For single doses, generate a new derived indicator variable `got_hypo_<dose>` (where `hypo` is short for hypothetical). Its value should be 1 for the first occurrence where their age in days at the visit is greater than or equal to the schedule's minimum age in days, and 0 otherwise. For respondents who have a missing or 0 survey weight, this indicator should be set to missing. Note that for all the visit dates for a given respondent, this variable will only take on the value 1 for one of those dates, the first visit date for which the condition was met. E.g., if a respondent had 6 visit dates and the vaccination schedule assigned the minimum age in days for the birth dose BCG as 0 days old, then all 6 visit dates would be opportunities for the respondent to

receive BCG. However, only the first visit date such that the minimum age in days is met should be flagged as 1 (which would be the first visit date in this example), and all other values for this indicator should be 0 for the other visit dates (or missing if the respondent has a 0 or negative survey weight).

In Stata, the code might look something like this:

```
foreach d in `=lower("$RI_SINGLE_DOSE_LIST")' {
    bysort respid: gen got_hypo_`d' = age_at_visit >= `d'_min_age_days' if
!missing(age_at_visit) & psweight > 0 & !missing(psweight)
    bysort respid: gen got_hypo_`d'_sum = sum(got_hypo_`d')
    replace got_hypo_`d'_sum = 0 if got_hypo_`d'_sum > 1
    replace got_hypo_`d'_sum = . if psweight == 0 | missing(psweight)
    replace got_hypo_`d' = got_hypo_`d'_sum
    drop got_hypo_`d'_sum
}
```

Now consider doses in a series. For each dose in a series, a new derived indicator variable `got_hypo_<dose>` will also be created. The first dose in the series is calculated the same as single doses. However, subsequent doses in a series will need to consider the minimum age in days between doses according to the country's vaccination schedule. When there is a subsequent dose, the previous dose in that series will need to have derived a variable `num_days_since_<dose>`, which will be compared to the vaccination schedule's minimum age in days between doses. Conceptually, dose 1 in a series generates two variables:

- (1) `got_hypo_<dose>1` is an indicator variable that is 1 for the first occurrence where `age_at_visit` is equal to or greater than the minimum age in days according to the vaccination schedule, and 0 otherwise.
- (2) `num_days_since_<dose>1` is the difference between `age_at_visit` between subsequent visit dates and the visit date where the respondent got the hypothetical dose 1, and 0 otherwise. This variable will be used to generate the `got_hypo_<dose>2` indicator variable.

For example, if a child had 5 visit dates at days 0, 45, 78, 98, and 270, and the schedule called for a 3-dose series with minimum ages of 42, 70, and 98, with a minimum interval of 28 days between doses, then the child would have received the first dose at the second visit, day 45. For this respondent, `got_hypo_<dose>1` would have values 0, 1, 0, 0, 0 and `num_days_since_<dose>1` would have values 0, 0, 33, 53, and 225.

Dose 2 in a 3-dose series would generate two variables:

- (1) `got_hypo_<dose>2` is an indicator variable that is 1 for the first occurrence where `age_at_visit` is equal to or greater than the minimum age in days according to the vaccination schedule AND `num_days_since_<dose>1` is greater than or equal to the minimum interval number of days between doses 1 & 2, and 0 otherwise.
- (2) `num_days_since_<dose>2` is the difference between `age_at_visit` between subsequent visit dates and the visit date where the respondent got the hypothetical

dose 2, and 0 otherwise. This variable will be used to generate the got\_hypo\_<dose>3 indicator variable.

For our example, the child would have received the second dose at the third visit, day 78, since both the minimum age in days for dose 2 and minimum interval days between doses 1 & 2 were met. For this respondent, got\_hypo\_<dose>2 would have values 0, 0, 1, 0, 0 and num\_days\_since\_<dose>2 would have values 0, 0, 0, 20, and 192.

Dose 3 in the 3-dose series would generate one variable:

- (1) got\_hypo\_<dose>3 is an indicator variable that is 1 for the first occurrence where age\_at\_visit is equal to or greater than the minimum age in days according to the vaccination schedule AND num\_days\_since\_<dose>2 is greater than or equal to the minimum interval number of days between doses 1 & 2, and 0 otherwise.

For our example, the child would have received the third dose at the fifth visit, day 270, since both the minimum age in days for dose 2 and minimum interval days between doses 1 & 2 were met. For this respondent, got\_hypo\_<dose>3 would have values 0, 0, 0, 0, 1. Note that the last dose in the series does not need to generate the variable num\_days\_since\_<dose>X.

In Stata, for a 3-series dose vaccine, the code might look something like this:

```

foreach d in `=lower("$RI_MULTI_3_DOSE_LIST")' {
    * 1st dose in series...
    bysort respid: gen got_hypo_`d'1 = age_at_visit>=`d'1_min_age_days' ///
        if !missing(age_at_visit) & psweight>0 & !missing(psweight)
    bysort respid: gen got_hypo_`d'1_sum = sum(got_hypo_`d'1)
    gen num_days_since_`d'1_temp = num_days_since_last_visit // make this variable before
                                                                // replacing the _sum variable
    replace num_days_since_`d'1_temp = 0 if got_hypo_`d'1_sum==0 | ///
        got_hypo_`d'1_sum==1 // update new variable based on _sum values
    bysort respid: gen num_days_since_`d'1 = sum(num_days_since_`d'1_temp)
    replace got_hypo_`d'1_sum = 0 if got_hypo_`d'1_sum>1
    replace got_hypo_`d'1_sum = . if psweight==0 | missing(psweight)
    replace got_hypo_`d'1 = got_hypo_`d'1_sum
    drop got_hypo_`d'1_sum num_days_since_`d'1_temp

    * 2nd dose in series...
    bysort respid: gen got_hypo_`d'2 = age_at_visit>=`d'2_min_age_days' & ///
        num_days_since_`d'1>=`d'2_min_interval_days' if !missing(age_at_visit) & ///
        psweight>0 & !missing(psweight)
    bysort respid: gen got_hypo_`d'2_sum = sum(got_hypo_`d'2)
    gen num_days_since_`d'2_temp = num_days_since_last_visit // make this variable before
                                                                // replacing the _sum variable
    replace num_days_since_`d'2_temp = 0 if got_hypo_`d'2_sum==0 | got_hypo_`d'2_sum==1
    bysort respid: gen num_days_since_`d'2 = sum(num_days_since_`d'2_temp)
    replace got_hypo_`d'2_sum = 0 if got_hypo_`d'2_sum>1
    replace got_hypo_`d'2_sum = . if psweight==0 | missing(psweight)
    replace got_hypo_`d'2 = got_hypo_`d'2_sum
    drop got_hypo_`d'2_sum num_days_since_`d'2_temp

    * 3rd dose in series...
    bysort respid: gen got_hypo_`d'3 = age_at_visit>=`d'3_min_age_days' & ///
        num_days_since_`d'2>=`d'3_min_interval_days' if !missing(age_at_visit) & ///
        psweight>0 & !missing(psweight)
    bysort respid: gen got_hypo_`d'3_sum = sum(got_hypo_`d'3)
    replace got_hypo_`d'3_sum = 0 if got_hypo_`d'3_sum>1
    replace got_hypo_`d'3_sum = . if psweight==0 | missing(psweight)
    replace got_hypo_`d'3 = got_hypo_`d'3_sum
    drop got_hypo_`d'3_sum

    drop num_days_since_`d'1 num_days_since_`d'2
}

```

## Routine Immunization Survey: Measures Related to Quality of Services

Calculation: Weighted estimated proportion, calculated as described in RI\_COVG\_01

Table Output: Estimated %, 95% CI, N (unweighted), N (weighted)

Graph: Inchworm plots

Interpretation: “X% of the population who were eligible for the survey would have been estimated to have a documented record of vaccinations (<source(s)>) and to have received a valid dose of <dose> if there had been no missed opportunities for simultaneous vaccination.”

Provenance: COSAS & others

## Specification for calculation of missed opportunities derived variables

Calculating missed opportunities flags is conceptually straightforward but gets a little messy if the multi-dose vaccine records include not only vaccination dates, but also tick marks on the vaccination cards (as can be the case with EPI survey dataset), or doses recorded by history (as can be the case with DHS & MICS).

The following instructions document the way we calculate missed opportunities flags. There are two sets of calculations: one that gives credit only for so-called valid doses, and another set that does a more crude calculation that gives credit for doses received early. When the calculations give credit only for valid doses, then more missed opportunities will be recorded than in the crude calculations. Consider a child who receives a first dose of DPT at 4 weeks of age. The dose is invalid. Then they receive two more doses at 4 week intervals, so at 8 and 12 weeks of age. The crude dose calculation would say they received 3 doses, but the valid dose calculation would omit that early dose and say they only received two doses, so every time the child is vaccinated for another vaccine after the age of 16 weeks, the calculation would count the visit as a missed opportunity for them to receive a valid third dose of DPT.

Whether you prefer the crude or valid calculations may depend on your perspective and what you will use the numbers for.

The following calculations will produce missed opportunity flags for both crude and valid dose calculations.

Conceptually, the outline is as follows:

Make a dataset consisting only of children who have a valid date of birth and at least one vaccination date recorded. Each child may be represented by several rows in the dataset – one row for every distinct date on which they received vaccinations. Each row shows, using indicator variables, which vaccines they received on which days. Sort the dataset by child and by vaccination date. Calculate their age at each visit. Use the country's vaccination schedule and some user-inputs about grace periods to calculate which vaccine(s) the child was eligible to receive on each visit date. The missed opportunities are occasions where the child was eligible for and did not receive the vaccine. If the child received the vaccine on a later date, then the missed opportunity is said to have been "corrected". If the child never received it, then it is "uncorrected".

These calculations are straightforward for the single-dose vaccines. Those with dates can be calculated. Those with only tick marks or evidence by caregiver history should be excluded from the calculation (MOV flags that are described below are set to 0) because we don't know when they received the dose.

For multi-dose vaccines, the calculations are pretty easy if all the evidence is from dates, but becomes trickier if some doses are recorded by tick mark or history. Those are handled in a special section below.

These calculations can be conducted on a dataset from cards or registers. We do not mix data from cards and registers in the same calculations, but it is acceptable to do the calculations both ways and then pass forward the flags from one or the other or whichever flags give the benefit of the doubt or whichever show the worst-case calculations.

The MOV calculations can be calculated two ways: giving credit for valid doses only, or a crude analysis that gives credit for any dose. VCQI calculates the flags both ways, per request from WHO on 9/23/2015 and the user tells VCQI which set of outcomes to summarize in tables and figures.

Start with the RI dataset. Drop all variables except those that identify the child (RI01, RI03, RI11, RI12), the dob (use the derived variable dob), the vaccination dates, ticks, and history (recall) flags.

## Routine Immunization Survey: Measures Related to Quality of Services

1. For MOV calculation purposes, evidence from a tick mark on the card/register or from caregiver recall is equivalent: we have evidence that the child received the dose but we do not know their age on the day they received it, so we exclude that dose (and doses that come after it in a series) from MOV flag calculations. To simplify calculations, for doses with no vaccination date, and no tick, and evidence from caregiver recall, set the tick flag to 1. This allows the MOV code to ignore variables that code evidence from recall. This is simply a computational efficiency – for other indicators, the datasets should keep data from cards and registers separate from recall data.
2. Also, for MOV calculation purposes, the code would be more streamlined and efficient if we treat 2-dose antigens like 3-dose antigens, so generate faux placeholder variables to hold evidence of the third dose for 2-dose antigens. That is, if ROTA is a 2-dose antigen, the dataset will already hold variables for `rota1_card_date` and `rota2_card_date`. Generate new variables named `rota3_card_date` and `rota3_card_tick` so the code can always loop from 1 to 3 without having a lot of logic to distinguish between 2- and 3-dose antigens. Assign missing values to these new variables.
3. Transpose the dataset from its RI form or its RIHC form, where each child occupies a single row, to a new form where there is a row for every different date on which the child received a vaccination. Establish flags to indicate which vaccines the child received on each day (`got_<dose>` = 0 or 1 in each row). So each row consists of the variables that define the person, the dob, the `got_` flags, and tick flags.
4. Generate a variable called `age`, which is the difference between the date of vaccination and the date of birth, in days.
5. Generate a variable called `person`, which is a unique identifier for each child (together the variables RI01 RI03 RI11 and RI12 make unique combination; give each child their own unique value of `person`).
6. Sort the dataset by `person` and `age`.
7. Calculate new indicator variables for the single-dose vaccines:
  - a. `Credit_<dose>` indicates whether or not we would give **credit** for the dose on this date if the child received it; they must have reached the minimum age for a valid dose to get credit for it. These parameters `<dose>_min_age_days` are part of the vaccination schedule parameters supplied by the user.
    - i. Generate an indicator: `credit_<dose> = (age ≥ <dose>_min_age_days)`
    - ii. If the user has specified a crude dose analysis, replace `credit_<dose>` with 1
    - iii. If the vaccination evidence for the dose is from tick mark, replace `credit_<dose>` with 0
  - b. Generate another indicator to track whether the child is **eligible** for a valid dose if it were given on this date (if they are eligible and do not receive it, then we count it as an MOV)
    - i. `elig_<dose> = (age ≥ <dose>_min_age_days)`
    - ii. If child's age at this visit is `> <dose>_max_age_days` then replace `credit_<dose>` with 0
    - iii. If the vaccination evidence for the dose is from tick mark, replace `elig_<dose>` with 0
  - c. Generate a variable to indicate if the child **received** a dose on this date for which we give credit:  
`got_credit_<dose> = (credit_<dose> == 1 & got_<dose> == 1)`
  - d. We only care about the first time they received a valid dose, so if the flag is set more than once, only retain the first instance:
    - i. Within each person (bysort `person`): generate `cum_credit_<dose>` = running sum of valid doses rec'd to date
    - ii. Replace `got_credit_<dose> = 0` if `got_credit_<dose> == 1 & cum_credit_<dose> > 1`
    - iii. Drop the `cum_credit_<dose>` variable

Later we will drop most rows, keeping only one per person, so for the variables and flags that we want to save for later use, be sure to copy the value into all rows for this respondent.

- e. Calculate the age at which they got the dose, and populate that age in all rows for this person:
    - i. generate  $\text{dropthis} = \text{got\_credit\_<dose>} * \text{age}$   
Within each person (bysort person): generate  $\text{age\_at\_credit\_<dose>} = \text{max}(\text{dropthis})$
    - ii. Drop dropthis
  - f. Save a flag in every row indicating whether or not they got credit for a dose:
    - i. Within each person (bysort person): generate  $\text{flag\_got\_credit\_<dose>} = \text{max}(\text{got\_credit\_<dose>})$
8. Generate the same indicator variables for the multi-dose vaccines, but the only difference is in how the credit and eligible flags are set:
- a. For the first dose in the series, there is no difference... the calculation is the same as for single-dose vaccines:
    - i.  $\text{credit\_<dose>} = (\text{age} \geq \text{<dose>\_min\_age\_days})$
    - ii.  $\text{elig\_<dose>} = (\text{age} \geq \text{<dose>\_min\_age\_days})$
    - iii. if doing a crude analysis, replace  $\text{credit\_<dose>} = 1$
    - iv. Replace both  $\text{credit\_<dose>}$  and  $\text{elig\_<dose>}$  with 0 if the evidence for the dose is from a tick mark
  - b. For the 2<sup>nd</sup> dose, we only give credit and make the child eligible if they received the first dose and the age and an intra-dose interval has passed
    - i.  $\text{credit\_<dose>}2 = \text{flag\_got\_credit\_<dose>}1 == 1 \ \& \ (\text{age} > (\text{age\_at\_credit\_<dose>}1 + \text{<dose>}2\_min\_interval\_days))$
    - ii.  $\text{elig\_<dose>}2 = \text{flag\_got\_credit\_<dose>}1 == 1 \ \& \ (\text{age} > (\text{age\_at\_credit\_<dose>}1 + \text{<dose>}2\_min\_interval\_days))$
    - iii. if doing a crude analysis, we give credit even before the interval elapses, so change the credit indicator  
replace  $\text{credit\_<dose>}2 = \text{flag\_got\_credit\_<dose>}1 \ \& \ \text{age} > \text{age\_at\_credit\_<dose>}1$
    - iv. Replace both  $\text{credit\_<dose>}$  and  $\text{elig\_<dose>}$  with 0 if the evidence for either dose 1 or dose 2 is from a tick mark
  - c. For the 3<sup>rd</sup> dose, we do a similar calculation...they have to have received a second dose to be eligible for the 3<sup>rd</sup>, and we have to wait the proper interval before counting MOVs and giving credit.
    - i.  $\text{credit\_<dose>}3 = \text{flag\_got\_credit\_<dose>}2 == 1 \ \& \ (\text{age} > (\text{age\_at\_credit\_<dose>}2 + \text{<dose>}3\_min\_interval\_days))$
    - ii.  $\text{elig\_<dose>}3 = \text{flag\_got\_credit\_<dose>}2 == 1 \ \& \ (\text{age} > (\text{age\_at\_credit\_<dose>}2 + \text{<dose>}3\_min\_interval\_days))$
    - iii. if doing a crude analysis, we give credit even before the interval elapses, so change the credit indicator:  
replace  $\text{credit\_<dose>}3 = \text{flag\_got\_credit\_<dose>}2 \ \& \ \text{age} > \text{age\_at\_credit\_<dose>}2$
    - iv. Replace both  $\text{credit\_<dose>}$  and  $\text{elig\_<dose>}$  with 0 if the evidence for dose 1 or dose 2 or dose 3 is from a tick mark

Now that the preliminary flags are set for each dose, the program can proceed to calculate whether there were missed opportunities for each dose. This next set of calculations is the same for each dose, regardless of whether it is part of a series or not. Repeat these steps for every <dose>.

1. Within each person (bysort person): generate  $\text{cum\_<dose>} = \text{running sum of got\_<dose>}$
2. Generate  $\text{mov\_<dose>} = \text{elig\_<dose>} == 1 \ \& \ \text{cum\_<dose>} == 0$

## Routine Immunization Survey: Measures Related to Quality of Services

3. Within each person (bysort person): generate `cum_mov_<dose>` = running sum of `mov_<dose>`
4. Generate `cor_mov_<d>` = `cum_mov_<dose>` > 0 & `got_<dose>` == 1
5. Record in a flag in all rows for this person whether they had 1+ corrected MOVs:  
Within each person (bysort person): generate `flag_cor_mov_<dose>` = `total(cor_mov_<dose>)`  
Replace `flag_cor_mov_<dose>` = `flag_cor_mov_<dose>` > 0
6. Record in all rows the total number of MOVs  
Within each person (bysort person): generate `total_mov_<dose>` = `total(mov_<dose>)`
7. Record in a flag if they had 1+ MOVs  
Generate `flag_had_mov_<dose>` = `total_mov_<dose>` > 0
8. Set a flag if all their MOVs for this dose were uncorrected:  
`gen flag_uncor_mov = (flag_had_mov_<dose> == 1) & (flag_got_credit_<dose> == 0)`
9. Count the number of visits where the child was eligible for the dose; put it in all rows.  
Within each person (bysort person): generate `total_elig_<dose>` = `total(elig_<dose>)`

When these steps have been completed, there will be a meaningful set of missed opportunity flag variables for each respondent.

Loop over all doses and calculate a final set of derived variables:

1. `elig_for_anydose = 1` for any visit where the `elig_<dose>` flag is 1 for 1+ doses in the list
2. Set a value in all rows for the respondent:  
Within each person (bysort person): `gen total_elig_visits = count of visits where elig_for_anydose = 1`
3. `mov_for_anydose = 1` for any visit where `flag_had_mov_<dose>` is 1 for 1+ doses in the list
4. Set a value in all rows for the respondent:  
Within each person (bysort person): `gen total_mov_visits = count of visits where mov_for_anydose = 1`
5. `total_visit_movs = count of doses for which mov_<dose>` is 1 in this visit
6. Set a value in all rows for the respondent:  
Within each person (bysort person): `gen total_movs = sum of total_visit_movs across all visits`

For each respondent, drop all rows in the dataset except one and save the dataset for the purpose of later merging in the missed opportunity outcome variables. Those flags include:

1. `flag_cor_mov_<dose>` = 1 if respondent had a corrected MOV for this dose
2. `flag_uncor_mov_<dose>` = 1 if they had an uncorrected MOV
3. `flag_had_mov_<dose>` = 1 if they had 1+ MOVs for this dose
4. `total_mov_<dose>` = # of visits where respondent was eligible and did not receive this dose
5. `total_elig_<dose>` = total visits where the child was eligible for the dose (includes visit when they got credit for receiving the dose, if applicable)
6. `total_elig_visits` = total visits where child was eligible for at least one dose of something
7. `total_mov_visits` = total visits where the child had 1+ MOVs
8. `total_movs` = total MOVs across all visits

Note that a 1 for the MOV flags means we have enough data to calculate and we do find evidence of an MOV. But a 0 might mean either that we did not have the appropriate data, or that we had it and did not find evidence. If we decide later that there is an important distinction between these two cases, then we should set flags to missing (instead of 0) when we are explicitly noting that we do not have enough information to assess whether there was an MOV.

Note that the code could be arranged to calculate MOV outcome variables both ways – counting crude doses and counting only valid doses. In that case the outcome variable names would need to reflect whether they are from the crude or valid calculations. For the 2015 software we will plan to calculate two sets of flags for each child: crude and valid dose MOV flags.

**RI\_QUAL\_08: Percent of visits with missed opportunity for simultaneous vaccination**

Weighted: No

Denominator: Number of vaccination visit dates where a respondent was eligible to receive 1+ vaccinations

Numerator: Number of vaccination visit dates where a respondent did not receive all vaccinations for which they were eligible

Vaccines: Calculate for each vaccine and dose  
Calculate over all vaccines and doses  
(rate of MOVs per visit, i.e., average # of vaccines missed per visit)

Input

Parameters RI\_QUAL\_07\_VALID\_OR\_CRUDE should be set to CRUDE if the user wishes to give credit for early doses and should be set to VALID if the user does not wish to give credit for early doses.

Consider a country where DPT is scheduled to be given at 6, 10 and 14 weeks. Consider a child who received DPT at 5, 9 and 13 weeks and who received measles at 9 months of age. The child did not receive 3 valid doses of DPT...only the doses at 9 weeks and 13 weeks were valid...and they were valid for DPT1 and DPT2. The dose received at 5 weeks was an invalid dose, so the child did not receive a 3<sup>rd</sup> valid dose. So if the MOV analysis does not give credit for early doses (specify VALID for the parameter listed above) then when the child returns for the measles vaccine at age 9 months, they are considered to be eligible for the 3<sup>rd</sup> valid dose of DPT. And if they do not receive it along with measles, it is counted as a missed opportunity.

If, instead, the user gives credit for early doses (specify the CRUDE option), then the child is still counted as having two valid doses of DPT, but they are not considered eligible for a 3<sup>rd</sup> dose at the measles visit, and the measles visit is not considered to be a missed opportunity for DPT.

Specifying VALID for this parameter will result in higher results for this indicator. If the parameter is set to VALID then the child described above would be considered to have an MOV for DPT3 when they receive measles but not DPT at 9 months. If instead, the parameter is set to CRUDE then this child is not considered to be eligible for DPT at 9 months because she already had 3 DPT doses. So she is not considered to have had an MOV at the 9 month visit.

It may be informative to do the MOV analysis twice...once with the parameter set to VALID and again with the parameter set to CRUDE, and to compare the output.

Input

Variables: This analysis uses MOV outcome variables calculated in the code described in section 6.6a of this document.

Derived

Variables: This indicator does not calculate new derived variables; it simply tabulates MOV flags calculated in the code described above.

For individual doses, the calculation is the sum over all respondents of  $\text{total\_mov\_<dose>} / \text{sum over all respondents of total\_elig\_<dose>}$

For all doses combined, the calculation is the sum over all respondents of  $\text{total\_mov\_visits} / \text{sum over all respondents of total\_elig\_visits}$

Also report: sum over all respondents of  $\text{total\_movs} / \text{sum over all respondents of total\_elig\_visits}$ . (Note that this last measure to report is not a percent...it is a ratio...the average number of missed eligible doses per visit...so it should not be scaled up by 100% as is done to the other estimates to put them on the percent scale.)

**Table Output:** Report outcomes for all doses in a single very wide table, where each dose has a two columns: number of eligible visits, and % of those visits with MOV. The aggregate data over all doses show those two columns plus three additional columns which are the sum of total\_movs, the sum of all eligible visits, and the rate of total movs / total eligible visits.

**Interpretation:** To interpret columns labeled "Visits with MOV for <dose>": "Respondents did not receive <dose> in X% of the N visits where they were eligible for it."

To interpret the column labeled "Visits with MOV for any dose": "Respondents did not receive all doses for which they were eligible in X% of the N visits where they were eligible for one or more doses."

To interpret the column labeled "MOVs per Visit": "On average, respondents were not given R doses for which they were eligible in each vaccination visit."

If MOVs per visit is a number smaller than 1, it may be helpful to interpret thus:

"On average, there was a missed opportunity for simultaneous vaccination in one out of every  $1/R$  visits in the survey dataset." (i.e., If the average MOVs per visit is 0.2, we might say "On average there was a missed opportunity for simultaneous vaccination in one out of every 5 visits represented in the survey dataset.")

**Provenance:** 2015 EPI cluster survey reference manual draft

**RI\_QUAL\_09: Percent of children with missed opportunity for simultaneous vaccination**

Description: This analysis identifies the number (and percent) of respondents who:

- a) experienced an MOV
- b) experienced an uncorrected MOV  
(meaning that they had not rec'd a valid dose as of the time of the survey)
- c) experienced a corrected MOV  
(meaning that they rec'd a valid dose sometime after their MOV or MOVs and before the survey)

When considering the MOVs calculated across all doses, it identifies the number (and percent) of respondents who:

- a) experienced 1+ MOVs for any doses,
- b) all MOVs were uncorrected,
- c) all MOVs were corrected,
- d) and some but not all MOVs were corrected.

Weighted: No

Denominator: Number of children with date of birth data and date of vaccination data indicating that they had 1+ visits for vaccination on days when they were eligible to receive the dose in question

Numerator: Number of children who experienced 1+ missed opportunities to be vaccinated for the dose in question

Vaccines: Calculate for each vaccine and dose  
Calculate over all vaccines and doses  
(# of children with 1+ MOV / # of children with 1+ eligible visit date in the dataset)

Input

Parameters RI\_QUAL\_07\_VALID\_OR\_CRUDE should be set to CRUDE if the user wishes to give credit for early doses and should be set to VALID if the user does not wish to give credit for early doses.

Consider a country where DPT is scheduled to be given at 6, 10 and 14 weeks. Consider a child who received DPT at 5, 9 and 13 weeks and who received measles at 9 months of age. The child did not receive 3 valid doses of DPT...only the doses at 9 weeks and 13 weeks were valid...and they were valid for DPT1 and DPT2. The dose received at 5 weeks was an invalid dose, so the child did not receive a 3<sup>rd</sup> valid dose. So if the MOV analysis does not give credit for early doses (specify VALID for the parameter listed above) then when the child returns for the measles vaccine at age 9 months, they are considered to be eligible for the 3<sup>rd</sup> valid dose of DPT. And if they do not receive it along with measles, it is counted as a missed opportunity.

If, instead, the user gives credit for early doses (specify the CRUDE option), then the child is still counted as having two valid doses of DPT, but they are not considered eligible for a 3<sup>rd</sup> dose at the measles visit, and the measles visit is not considered to be a missed opportunity for DPT.

Specifying VALID for this parameter will result in higher results for this indicator. If the parameter is set to VALID then the child described above would be considered to have an MOV for DPT3 when they receive measles but not DPT at 9 months. If instead, the parameter is set to CRUDE then this child is not considered to be eligible for DPT at 9 months because she already had 3 DPT doses. So she is not considered to have had an MOV at the 9 month visit.

It may be informative to do the MOV analysis twice...once with the parameter set to VALID and again with the parameter set to CRUDE, and to compare the output.

#### Input

Variables: This analysis uses MOV outcome variables calculated in the code described in section 6.6a of this document.

#### Derived

Variables: In the description below, the text <vc> stands for RI\_QUAL\_07\_VALID\_OR\_CRUDE and it will represent either the string 'valid' or the string 'crude'.

Start with the dataset of MOV flags merged onto the survey dataset. This dataset has one row per RI respondent.

Generate three count variables and set them all to zero: doses\_with\_mov;  
doses\_with\_uncor\_mov; doses\_with\_cor\_mov.

Generate a new derived indicator variable for each dose, child\_had\_mov\_<dose>; set it equal to flag\_had\_mov\_<dose>\_<vc>. Replace the indicator with a missing value if total\_elig\_<dose>\_<vc> is 0.

Generate a new derived indicator variable for each dose, child\_had\_uncor\_mov\_<dose> and set it equal flag\_uncor\_mov\_<dose>\_<vc>; replace it with missing if child\_had\_mov\_<dose> is not 1.

Generate a new derived indicator variable for each dose, child\_had\_cor\_mov\_<dose>; set it equal to flag\_cor\_mov\_<dose>\_<vc>. Replace the indicator with missing if child\_had\_mov\_<dose> is not 1.

For each dose:

Increment doses\_with\_mov if child\_had\_mov\_<dose> == 1

Increment doses\_with\_uncor\_mov if child\_had\_uncor\_mov\_<dose> == 1

Increment doses\_with\_cor\_mov if child\_had\_cor\_mov\_<dose> == 1

Generate a new derived indicator child\_had\_mov; set it equal to one if doses\_with\_move > 0. Set it to zero if doses\_with\_mov is zero. Set it to missing if total\_elig\_visits\_<vc> is zero.

Generate a new derived indicator variable, child\_had\_only\_uncor\_mov = 0; set it equal to 1 if doses\_with\_mov > 0 and doses\_with\_cor\_mov == 0; set it to missing if total\_elig\_visits\_<vc> == 0

## Routine Immunization Survey: Measures Related to Quality of Services

Generate a new derived indicator variable, `child_had_only_cor_mov = 0`; set it equal to 1 if `doses_with_mov > 0 & doses_with_uncor_mov == 0`; set it to missing if `total_elig_visits_<vc> == 0`

Generate a new derived indicator variable, `child_had_cor_and_uncor_mov = 0`; set it equal to 1 if `doses_with_cor_mov > 0 & doses_with_uncor_mov > 0`; set it to missing if `total_elig_visits_<vc> == 0`

Calculation: The outputs are unweighted counts and proportions, as described below.

Table Output: Report outcomes for all doses in a single very wide table, where each dose has several columns: the number of children who had at least one visit where they were eligible to receive the dose (this is the number of kids for which the indicator is either 0 or 1); the number and % of those children who had 1+ MOVs for that dose, the number and percent of children whose MOVs were uncorrected at the time of the survey, the number and percent of children whose MOVs were corrected before the survey.

The results are reported in dose-specific sub-sections of the table and the final section lists the percent of respondents with an MOV for any dose in the schedule, the number and percent for whom all MOVs were uncorrected, the number and percent for whom all MOVs were corrected, and the number and percent for whom some but not all of the MOVs were corrected.

Interpretation: To interpret columns labeled “Had MOV for <dose> %”: “Among the N children in the survey dataset who received some vaccinations on days when they were age-eligible to receive <dose>, X% of them experienced 1+ occasions where they were eligible to receive <dose> but did not receive it.”

To interpret the column labeled “MOV uncorrected for <dose> %”: “Among the N children in the survey dataset who visited vaccination services on days when they were eligible to receive <dose>, X% of them experienced uncorrected missed opportunities for vaccination with <dose>, that is, there were 1+ occasions where they were eligible to receive <dose> but did not receive it, and as of the date of the survey they still had not received it.”

To interpret the column labeled “MOV corrected for <dose> %”: “Among the N children in the survey dataset who visited vaccination services on days when they were eligible to receive <dose>, X% of them experienced corrected missed opportunities for vaccination with <dose>, that is, there were 1+ occasions where they were eligible to receive <dose> but did not receive it, but they did receive it at a later date.”

To interpret column labeled “Had MOV for any dose (%)” : “Among the N children in the survey dataset who visited vaccination services on days when they were eligible to receive any dose, X% of them experienced 1+ occasions where they did not receive all doses for which they were eligible.”

To interpret column labeled “All MOVs were uncorrected (%)” : “Among the N children in the survey dataset who experienced 1+ MOVs for any doses, X% had all of their MOVs still uncorrected at the time of the survey.”

To interpret column labeled “All MOVs were corrected (%)”: “Among the N children in the survey dataset who experienced 1+ MOVs for any doses, X% had all of their MOVs corrected by the time of the survey.”

To interpret column labeled “Some (not all) MOVs were corrected (%)”: “Among the N children in the survey dataset who experienced 1+ MOVs for any doses, X% had some but not all of their MOVs corrected by the time of the survey.”

Provenance: COSAS

RI QUAL 10: Percent of doses administered on a given date (not yet implemented)

Weighted: No

Denominator: Number of times the dose was administered, as recorded by date, in the survey dataset

Numerator: Number of times the dose was administered on the date in question

Vaccines: Calculate separately for each vaccine and dose

Input variables: Vaccination dates

Table Output: Probably best reported with a figure instead of a table

Graph: Bar chart is appropriate, to identify periods of time when no doses were given

Provenance: COSAS

Note: This could be reported as a simple count, using the numerator only, rather than a %.

**RI QUAL 11: Percent of doses administered at a particular age, (age in days) (not yet implemented)**

Weighted: No

Denominator: Number of times the dose was administered, as recorded by date, to respondents who also had their date of birth recorded in the survey dataset

Numerator: Number of times the dose was administered on the age (in days) in question

Vaccines: Calculate separately for each vaccine and dose

Input variables: DOB  
Vaccination dates

Table Output: Probably best reported with a figure instead of a table

Graph: Could be a bar chart or line graph, and could report results day-by-day or cumulative results in an inverse Kaplan-Meier curve

Provenance: COSAS

**RI\_QUAL\_12: Percent of later doses in a sequence administered after a particular interval**

Weighted: No

Denominator: Number of times the two doses were administered & both had recorded dates

Numerator: Number of times the intra-dose interval was > the threshold in question (in days)

Vaccines: Calculate separately for different dose pairs

Input variables: Clean vaccination dates

Derived

variables: For each dose pair of interest, the user inputs the names of the two doses and the interval that interests them.

For each dose pair of interest, generate three new indicator variables:

igt\_<dose1>\_<dose2>\_<threshold>\_c = 1 if the interval calculated using card dates exceeds the threshold; it is 0 if the interval is <= the threshold; it is missing if the interval is not calculable.

igt\_<dose1>\_<dose2>\_<threshold>\_r = 1 if the interval calculated using register dates exceeds the threshold; it is 0 if the interval is <= the threshold; it is missing if the interval is not calculable.

If RECORDS\_NOT\_SOUGHT, generate igt\_<dose1>\_<dose2>\_<threshold> = igt\_<dose1>\_<dose2>\_<threshold>\_c

If RECORDS\_SOUGHT\_IF\_NO\_CARD, generate igt\_<dose1>\_<dose2>\_<threshold> = igt\_<dose1>\_<dose2>\_<threshold>\_c and replace it with igt\_<dose1>\_<dose2>\_<threshold>\_r if the respondent had no card

And if RECORDS\_SOUGHT\_FOR\_ALL, generate igt\_<dose1>\_<dose2>\_<threshold> = igt\_<dose1>\_<dose2>\_<threshold>\_c and replace it with igt\_<dose1>\_<dose2>\_<threshold>\_r if \_c is missing or if \_r is 0.

This indicator is looking for instances where the interval was much longer than optimal...so it is a good thing to have the interval be shorter than the threshold. (By contrast, in RI\_QUAL\_05 it is a good thing for the interval to be er than the threshold.)

Calculation: The measure is simply the unweighted average of the indicator variable igt\_<dose1>\_<dose2>\_<threshold>.

Interpretation: "Of N intervals in the sample where the data include dates for both the earlier and later dose of <vaccine>, X% of the intervals were longer than <threshold> days."

Table Output: N and %

Graph: Unweighted proportion plot

Provenance: COSAS

RI\_QUAL\_13: Percent of children who receive DPT3 by the age of 26 weeks

Note: The code for RI\_QUAL\_13 is identical to the code for RI\_QUAL\_04...with all instances of 04 changed to 13. This is wasteful. Recommend changing RI\_QUAL\_04 to accept a list of doses and a list of thresholds and then drop the code for RI\_QUAL\_13. That way we only maintain one set of programs to do this task.

**RI QUAL 14: Percent of children simultaneously vaccinated where both doses were given (not yet implemented)**

Weighted: No

Denominator: Number of children who received both doses that have the same eligibility date in the national vaccination schedule, and who have vaccination date data for both doses

Numerator: Number of children who received the two doses on the same date

Vaccines: Calculate separately for each vaccine/dose pair

Input variables: Vaccination dates

Parameters: List of vaccine pairs where eligibility date (age) is the same in the nat'l vaccination schedule

Table Output: N and %

Provenance: PAHO Module 6

**RI QUAL 15: Percent of children simultaneously vaccinated where both either dose was given (not yet implemented)**

Weighted: No

Denominator: Number of children who received either dose that have the same eligibility date in the national vaccination schedule

Numerator: Number of children who received the two doses on the same date

Vaccines: Calculate separately for each vaccine/dose pair

Input variables: Vaccination dates

Parameters: List of vaccine pairs where eligibility date (age) is the same in the nat'l vaccination schedule

Table Output: N and %

Provenance: PAHO Module 6

**RI QUAL 16: Percent of children simultaneously vaccinated among children with vaccination dates (not yet implemented)**

Weighted: No

Denominator: Number of children with 1+ vaccination dates in the dataset

Numerator: Number of children who received the two doses on the same date

Vaccines: Calculate separately for each vaccine/dose pair

Input variables: Vaccination dates

Parameters: List of vaccine pairs where eligibility date (age) is the same in the nat'l vaccination schedule

Table Output: N and %

Provenance: PAHO Module 6

RI QUAL 17: Number of visits needed to be fully vaccinated (not yet implemented)

Weighted: No

Denominator: None

Numerator: Number of additional visits the child would need to make to be fully vaccinated

Input variables: DOB

Vaccination dates

(Or simply Got It if we assume that simultaneous valid vaccination can be performed)

Parameters: List of vaccines & doses to be fully vaccinated

Table Output: Report mean and median

Graph: Could report a histogram

Provenance: PAHO Module 6

Note: Need to be clear about whether the analysis is identical for respondents who have vaccination dates and those who do not

## RI SURVEY – MEASURES RELATED TO PROVIDERS OF SERVICES

### RI PROV\_01: Percent of vaccinations performed by each source (not yet implemented)

Weighted: No

Denominator: Number of vaccinations reported that also have source data

Numerator: Number of vaccinations reported for the particular source in question

Vaccines: Calculate for each vaccine

Input variables: Source of vaccination (probably from multiple-choice question)

Parameters: List of sources

Table Output: Unweighted % for each source, including a category for 'missing' so the total across all sources adds up to 100%.

Graph: Could use a stacked bar chart or pie graph

Provenance: COSAS

Note: The authors of the updated 2015 WHO EPI cluster survey reference manual expressed concern that respondents are not likely to be able to provide this data reliably and recommended that the question be dropped from most surveys and the measure be dropped from most survey reports.

In situations where the outcome is important to understand and where the question is deemed likely to yield high quality results, it could be inserted back into the questionnaire.

## TETANUS SURVEY – MEASURES RELATED TO COVERAGE

### TT COVG 01: Children born protected from neonatal tetanus

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for women who are protected

Variations: By card  
By history  
By card or history (this is the main outcome)

#### Survey variables:

- TT09 Start date of interview with respondent
- TT16 Age of the mother, years
- TT27 Do you have a card and may I see it. 1=seen; 2=yes, not seen; 3=no card
- TT30-35 (Dates of TT doses, by mother’s card)
- TT36-Received TT during Index pregnancy 1=yes, 2=no, 99=DNK
- TT37 (History: how many doses during index pregnancy? 1=1;2=2;3=3 or more;99=DNK)
- TT38- Received TT during previous pregnancy 1=yes, 2=no, 99=DNK
- TT39 (History: how many doses during earlier pregnancies: number; or 99=DNK)
- TT40 Received TT outside of pregnancy 1=yes, 2=no, 99=DNK
- TT41 (History: number of doses outside pregnancy: 1-6 means 1-6; 7=7+; 99=DNK)
- TT42 (History: how many years ago was your most recent dose: 0 means in the last year; a number means that number; 98 = never had one; 99=DNK)
- TTHC21-26 (Dates of TT doses, by mother’s record at the health center)

#### Derived variables:

Generate a new derived variable: `lifetime_tt_doses_by_card`; set it to 0.

- Increment `lifetime_tt_doses_by_card` if TT30 is not missing.
- Increment it again if TT31 is not missing. Again for TT32, 33, 34, and 35.
- (Now `lifetime_tt_doses_by_card` takes the values 0-6.)

Generate a new derived variable: `lifetime_tt_doses_by_history`: set it to 0

- Add TT37 to `lifetime_tt_doses_by_history` if TT37 = 1, 2, or 3
- Add TT39 to `lifetime_tt_doses_by_history` if TT39 is not missing and not 99
- Add TT41 to `lifetime_tt_doses_by_history` if TT41 is 1, 2, 3, 4, 5, 6 or 7.

Generate a new derived variable named `years_since_last_dose_c_or_h`; set it to TT42

## Tetanus Survey: Measures Related to Coverage

- If the number of years from TT09 (the current survey data collection date) back to TT30 is smaller than years\_since\_last\_dose\_c\_or\_h, then replace years\_since\_last\_dose\_c\_or\_h with that smaller number of years; Repeat for TT31-35.

Generate a new derived variable: lifetime\_tt\_doses\_by\_register; set it to 0

- Increment it if TTHC21 is not missing. Again for TT22-26.
- (Now lifetime\_tt\_doses\_by\_register takes the values 0-6.)

Generate a new derived variable named years\_since\_last\_dose\_r\_or\_h; set it to TT42.

- If the number of years from TT09 (the current survey data collection date) back to TTHC21 is smaller than years\_since\_last\_dose\_c\_or\_h, then replace years\_since\_last\_dose\_c\_or\_h with that smaller number of years; Repeat for TT31-35.

Generate a new derived indicator variable: protected\_at\_birth\_by\_card; set it to 0.

- Replace protected\_at\_birth\_by\_card with 1 if lifetime\_tt\_doses\_by\_card is  $\geq 5$
- Replace protected\_at\_birth\_by\_card with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-9 years and lifetime\_tt\_doses\_by\_card == 4.
- Replace protected\_at\_birth\_by\_card with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-4 years and lifetime\_tt\_doses\_by\_card == 3.
- Replace protected\_at\_birth\_by\_card with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-2 years and lifetime\_tt\_doses\_by\_card == 2.

Generate a new derived indicator variable: protected\_at\_birth\_by\_history; set it to 0.

- Replace protected\_at\_birth\_by\_history with 1 if lifetime\_tt\_doses\_by\_history is  $\geq 5$
- Replace protected\_at\_birth\_by\_history with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-9 years and lifetime\_tt\_doses\_by\_history == 4.
- Replace protected\_at\_birth\_by\_history with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-4 years and lifetime\_tt\_doses\_by\_history == 3.
- Replace protected\_at\_birth\_by\_history with 1 if years\_since\_last\_dose\_c\_or\_h is between or equal to 0-2 years and lifetime\_tt\_doses\_by\_history == 2.

Generate a new derived indicator variable: protected\_at\_birth\_by\_c\_or\_h; set it to 0.

- Replace protected\_at\_birth\_by\_c\_or\_h with 1 if either protected\_at\_birth\_by\_card is 1 or protected\_by\_birth\_by\_history is 1.

Generate a new derived indicator variable: protected\_at\_birth\_by\_register; set it to 0.

- Replace protected\_at\_birth\_by\_register with 1 if lifetime\_tt\_doses\_by\_register is  $\geq 5$
- Replace protected\_at\_birth\_by\_register with 1 if years\_since\_last\_dose\_r\_or\_h is between or equal to 0-9 years and lifetime\_tt\_doses\_by\_register == 4.

- Replace protected\_at\_birth\_by\_register with 1 if years\_since\_last\_dose\_r\_or\_h is between or equal to 0-4 years and lifetime\_tt\_doses\_by\_register == 3.
- Replace protected\_at\_birth\_by\_register with 1 if years\_since\_last\_dose\_r\_or\_h is between or equal to 0-2 years and lifetime\_tt\_doses\_by\_register == 2.

Gen a new derived indicator named protected\_at\_birth\_c\_or\_h ; set it to 0.

- Replace it with 1 if either protected\_at\_birth\_by\_card or protected\_at\_birth\_by\_history is 1.

Generate a new derived indicator named protected\_at\_birth\_by\_c\_or\_h\_or\_r; set it to 0.

- Replace it with 1 if either protected\_at\_birth\_c\_or\_h or protected\_at\_birth\_by\_register is 1.

Generate a new indicator named protected\_at\_birth\_to\_analyze; set it to protected\_at\_birth\_c\_or\_h.

- If TT\_RECORDS\_SOUGHT\_IF\_NO\_CARD is 1; set the new indicator = protected\_at\_birth\_c\_or\_h\_or\_r if no card was available (TT27 is not = 1)
- If TT\_RECORDS\_SOUGHT\_FOR\_ALL is 1, set the new indicator to protected\_at\_birth\_c\_or\_h\_or\_r.

Calculation: This calculation is an estimated proportion, as described in RI\_COVG\_01.

Table Output: For the main outcome: protected\_at\_birth\_to\_analyze: Show

Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (weighted), N (unweighted)

If requested generate a table that shows protection by card (estimated % and 95% CI) alongside protection by history (estimated % and 95% CI) and protection by card or history (all the statistics listed above.)

Graph: Organ pipe plots & inchworm plots

Interpretation: “X% of babies born in the 12 months preceding the survey are estimated to have been protected at birth from neonatal tetanus, according to evidence given from [maternal vaccination card and/or maternal recall of their vaccination history and/or health center records of maternal vaccinations].”

Provenance: COSAS & others

Notes: The commonly quoted timeframe of protection codified here was taken from Table 1 in the document available at:

[http://who.int/reproductivehealth/publications/maternal\\_perinatal\\_health/immunization\\_tetanus.pdf](http://who.int/reproductivehealth/publications/maternal_perinatal_health/immunization_tetanus.pdf)

TT\_COVG\_02: Women protected at the time of the survey (not yet implemented)

Items to  
discuss with  
WHO

- A. The COSAS documentation has much more detail on what it calculates for the RI survey...and relatively little detail for the TT survey. It is not clear to us what this means and how it differs from 8.1.

TT\_COVG\_03: Crude TT coverage (% who received 2+ doses during index pregnancy) (not yet implemented)

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who report having received 2+ doses during the index pregnancy

Time options: During pregnancy

Variations: By card  
By history  
By register  
By card or history (for purpose of comparison with older surveys)  
By card or history or register (this is the main crude coverage outcome)

Survey variables:

- TT17 (date of birth in the index pregnancy)
- TT27 (have a card for mother's vaccination record: 1=yes, seen; 2=yes, not seen; 3=no; 99 = DNK)
- TT30-35 (Dates of TT doses, by mother's card)
- TTHC21-TTHC26 (Dates of TT doses, by register)
- TT36 (History: rec'd doses during index pregnancy: 1=yes; 2=no; 99=DNK)
- TT37 (History: how many doses during index pregnancy? 1=1;2=2;3=3 or more;99=DNK)

Derived variables:

Generate five new derived indicator variables, and set them to zero:

got\_2plus\_preg\_tt\_c  
got\_2plus\_preg\_tt\_h  
got\_2plus\_preg\_tt\_r  
got\_2plus\_preg\_tt\_c\_or\_h  
got\_2plus\_preg\_tt\_c\_or\_h\_or\_r

Replace got\_2plus\_preg\_tt\_h with 1 if TT37 is 2 or 3.

Generate a new derived variable: start\_of\_index\_pregnancy. Set it to be 300 days before TT17 (allowing for delivery up to 1 month late)

Evidence from card:

Generate a new derived variable: tt\_dates\_in\_index\_pregnancy\_c. Set it to zero.

## Tetanus Survey: Measures Related to Coverage

Increment the variable if the date in TT30 falls on or between start\_of\_index\_pregnancy and TT17

Increment again if TT31 falls in that interval.

Increment again for each of the dates TT32 to TT35 if they fall in that interval.

Replace got\_2plus\_preg\_tt\_c with 1 if tt\_dates\_in\_index\_pregnancy\_c is  $\geq 2$ .

Evidence from register:

Generate a new derived variable: tt\_dates\_in\_index\_pregnancy\_r. Set it to zero.

Increment the variable if the date in TTHC21 falls on or between start\_of\_index\_pregnancy and TT17

Increment again if TTHC22 falls in that interval.

Increment again for each of the dates TTHC23 – TTHC26 if they fall in that interval.

Replace got\_2plus\_preg\_tt\_r with 1 if tt\_dates\_in\_index\_pregnancy\_r is  $\geq 2$ .

Replace got\_2plus\_preg\_tt\_c\_or\_h with 1 if got\_2plus\_preg\_tt\_c == 1 or got\_2plus\_preg\_tt\_h == 1

Replace got\_2plus\_preg\_tt\_c\_or\_h\_or\_r with 1 if got\_2plus\_preg\_tt\_c\_or\_h == 1 or got\_2plus\_preg\_tt\_r == 1

Calculation: Estimated proportion, as described in RI\_COVG\_01

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Inchworm plots

Provenance: COSAS & others

Points for discussion with WHO:

- A. The COSAS documentation has much more detail on what it calculates for the RI survey...and relatively little detail for the TT survey. One can imagine other indicators here, as well. Proportion who received 0, 1, or 2+ doses during the index pregnancy. Proportion who received  $n+$  lifetime doses. Let's discuss what else we want to show for crude TT coverage.

TT\_COVG\_04: Valid dose TT coverage (not yet implemented)

Points to  
discuss with  
WHO:

- A. What does the phrase “valid coverage” mean in the context of maternal TT doses? Does it mean that there was at least a 4 week interval between doses during the pregnancy? Or between all subsequent doses? Is this something we want to calculate and report?

## TETANUS SURVEY – MEASURES RELATED TO ACCESS

### TT\_ACC\_01 Women who obtained ante-natal care (ANC) (not yet implemented)

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for women who report having had ANC

Survey

variables: TT18 (did you receive ANC? 1=yes; 2=no; 99=DNK)

Derived

variables: Generate a new derived indicator variable (received\_anc); equal to 0.  
Replace the variable with 1 if TT18 is 1.

Calculation:

Estimate the weighted average of received\_anc. When estimating uncertainty, use an estimation procedure that accounts for the sample design, like the Clopper-Pearson confidence interval.

Interpretation:

“X% of women in the population who had a live birth in the 12 months preceding the survey are estimated to have received some kind of ante-natal care during that pregnancy.”

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Inchworm plots

Provenance: COSAS & others

Topics to  
discuss with  
WHO:

A. Some may wish to generate indicators for women who had 2+ or 3+ visits.

## TETANUS SURVEY – MEASURES RELATED TO CONTINUITY OF SERVICES

### TT\_CONT\_01: Dropout from TT1 to TT2 during the index pregnancy (not yet implemented)

Weighted: No

Denominator: Number of respondents who received the first dose

Numerator: Number of respondents in the denominator who did not receive the second dose

Vaccines: First and second TT dose during index pregnancy

Derived variables:

This measure uses the got\_2plus\_preg\_tt\_c\_or\_h\_or\_r variables described above in the RI\_COVG\_01.

Generate a new derived indicator variable got\_1plus\_preg\_tt\_c\_or\_h\_or\_r; set it to 0.

Replace the indicator with 1 if any of the following are true:

- d) TT37 is 1 or 2 or 3 (true by history)
- e) tt\_dates\_in\_index\_pregnancy\_c is > 0
- f) tt\_dates\_in\_index\_pregnancy\_r is > 0

Generate a new derived indicator variable dropout\_tt\_in\_index\_pregnancy; set it to missing.

Replace the indicator with 1 if got\_1plus\_preg\_tt\_c\_or\_h\_or\_r == 1.

Replace the indicator with 0 if got\_2plus\_preg\_tt\_c\_or\_h\_or\_r == 1.

Calculation: Dropout is the (unweighted) average of the indicator variable.

Interpretation:

Among mothers in the survey sample who received a tetanus vaccination during the index pregnancy, X% did not go on to receive a second dose during that pregnancy.

Table Output: % (not weighted, so do not report CI), N (unweighted)

Graph: No

Provenance: COSAS & others

Topics to discuss with WHO:

## Tetanus Survey: Measures Related to Continuity of Services

- A. Of course if the first dose in the index pregnancy was the 5<sup>th</sup> or 6<sup>th</sup> lifetime dose, then their care provider might have decided that they were protected, and didn't need a second dose during pregnancy, right? Should we handle that case in a special manner?

TETANUS SURVEY – MEASURES RELATED TO QUALITY OF SERVICES

TT\_QUAL\_01: Percent of TT2 doses administered before 4 weeks had passed (not yet implemented)

Weighted: No

Denominator: Number of TT2 doses administered where the date was known for that dose and for the preceding dose

Numerator: Number of times the TT2 dose was administered before 4 weeks had passed from the date of the earlier dose

Vaccines: TT

Input variables: Vaccination dates

Parameters: Grace period under 28 days, if any

Table Output: N and %

Graph: None

Provenance: COSAS

TT\_QUAL\_02: % of TT3 doses administered within 26 weeks of TT2 (not yet implemented)

Weighted: No

Denominator: Number of respondents who had dates for TT2 and TT3

Numerator: Number of respondents whose TT3 was given within 26 weeks of TT2

Vaccines: TT

Input variables: Vaccination dates

Table Output: N and %

Graph: None

Provenance: COSAS

Note: (I'm not familiar with this one...ask Pierre / Felicity / RobertS about it.)

TT\_QUAL\_03: Percent of women with missed opportunity for TT1 vaccination (not yet implemented)

Weighted: No

Denominator: Number of respondents who had not had TT1 before index pregnancy and who received ANC

Numerator: Number of respondents who had still not had TT1 when the baby was born

Input variables: Birth date of child  
TT1 date

Table Output: N and %

Provenance: COSAS

Note: (Double-check this one with Pierre / Felicity / RobertS)

## TT\_QUAL\_04: Tetanus card availability (not yet implemented)

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who show a card with 1+ vaccination dates on it

### Survey

variables: TT27 (Mother has a vaccination card for herself: 1=seen; 2=yes, but not seen; 3=no)  
TT30-TT35 (date of TT1-TT6 doses, from mother's card if seen)

### Derived

variables: Calculate a new variable (showed\_tt\_card\_with\_dates); set it equal to 0.  
Replace the value with 1 if the respondent showed a card (TT27 == 1)  
and the card had at least one vaccination date recorded on it.  
(!missing(TT30) | !missing(TT31) | !missing(and so on...))

### Calculation:

Estimate the weighted average of showed\_tt\_card\_with\_dates. When estimating uncertainty, use an estimation procedure that accounts for the sample design, like the Clopper-Pearson confidence interval.

(In some cases the user may wish to also calculate the weighted proportion of respondents who showed cards that did not have any dates on them or the proportion who showed cards regardless of whether the card listed any dates. This can be accomplished by calculating an additional derived variable to capture that idea. We will not do that in this early version of the software.)

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Inchworm plots

Provenance: COSAS & others

Points to  
discuss with  
WHO:

- A. Should we only count women in the numerator if their card shows 1+ tetanus vaccination dates?

## TETANUS SURVEY – MEASURES RELATED TO PROVIDERS OF SERVICES

### TT PROV 01: Percent of TT vaccinations performed by each source (not yet implemented)

Weighted: No

Denominator: Number of TT vaccinations reported that also have source data

Numerator: Number of TT vaccinations reported for the particular source in question

Input variables: Source of vaccination (probably from multiple-choice question)

Parameters: List of sources

Table Output: Unweighted % for each source, including a category for 'missing' so the total across all sources adds up to 100%.

Graph: Could use a stacked bar chart or pie graph

Provenance: COSAS

Note: The authors of the updated 2015 WHO EPI cluster survey reference manual expressed concern that respondents are not likely to be able to provide this data reliably and recommended that the question be dropped from most surveys and the measure be dropped from most survey reports.

In situations where the outcome is important to understand and where the question is deemed likely to yield high quality results, it could easily be inserted back into the questionnaire.

**TT PROV 02: Place of delivery (not yet implemented)**

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who selected a particular choice

Survey

variables: TT22 (Where did you deliver your baby? Question gives 5 categories and Other.)

Derived

variables:

Generate seven new indicator variables delivery\_location\_1 thru delivery\_location\_7; set all to 0.

Replace the value of delivery\_location\_1 with 1 if TT22 = 1. Replace the value of delivery\_location\_2 with 1 if TT22 =2. Replace the value of delivery\_location\_3 with 1 if TT22 = 3. And so on up through 6.

Replace the value of delivery\_location\_7 with 1 if TT22 is missing.

Calculation:

There are seven weighted averages to calculate here. Each is a simple weighted average of a 0/1 indicator variable.

Table Output:

Report the estimated % and 95% CI for each option along with total N (weighted) and total N (unweighted)

(Note often with weighted averages we also report LCB, UCB, DEFF and ICC, but they are not reported here. This could be changed, of course, but we don't think they will be of interest.)

Provenance: COSAS & Others

## POST-SIA SURVEY – MEASURES RELATED TO COVERAGE

### SIA\_COVG\_01 Crude SIA coverage

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received the vaccine dose according to [card, history, or finger mark]

Vaccines: Calculated for each dose

Time options: During the campaign

Variations: By card  
By history  
By finger mark  
By card or history or finger mark (this is the main coverage outcome)

#### Input

Parameters SIA\_FINGERMARKS\_SOUGHT is 1 if the survey accepts fingermarks as evidence of vaccination. (Captured in question SIA22 in the FVL document.)

#### Survey variables:

SIA20 (Was the child vaccinated in the campaign? 1=yes, by card; 2 = yes, by history; 3=no; 99=DNK)

SIA22 (Receive a finger mark? 1=yes, seen; 2 = yes, unseen; 3=no; 99=DNK)

This measure assumes that observations where SIA17 = 2 or 99 have already been dropped or had their weights set to zero.

#### Derived

variables: Generate a new derived indicator variable: got\_sia\_dose\_by\_card; set it to 0

- Replace the indicator with 1 if SIA20 is 1

Generate a new derived indicator variable: got\_sia\_dose\_by\_history; set it to 0

- Replace the indicator with 1 if SIA20 is 2

If SIA\_FINGERMARKS\_SOUGHT == 1

- Generate a new derived indicator variable: got\_sia\_dose\_by\_fingermark; set it to 0
- Replace the indicator with 1 if SIA22 is 1

Generate a new derived indicator variable: got\_campaign\_dose; set it to 0.

- Replace with 1 if the card, history, or fingermark indicators are set to 1.

Calculation: The outcome is an estimated proportion, similar to RI\_COVG\_01.

## Post-SIA Survey: Measures Related to Coverage

Interpretation: “X% of eligible children who were living here during the campaign are estimated to have been vaccinated against [measles] during the recent campaign per information obtained [by card, by caregiver history, by finger mark].”

Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Graph: Inchworm plots

Provenance: COSAS & others

SIA COVG 02: Crude SIA coverage where SIA dose was the first dose

- Weighted: Yes
- Denominator: Sum of weights for all respondents
- Numerator: Sum of weights for respondents who received the SIA dose and had never received a dose before
- Vaccines: Calculated for each dose
- Time options: During the campaign
- Input variables: Weight  
 Got it during the campaign  
 Never had it before the campaign
- Survey variables: SIA20 (Did they get a dose in the campaign? 1 or 2=yes; 3=no; 99-DNK)  
 SIA27 (Had they had a dose before? 1 or 2=yes; 3=no; 99 = DNK)  
 This measure assumes that observations where SIA17 = 2 or 99 have already been dropped or had their weights set to zero.
- Derived variables: Generate a new derived indicator variable : sia\_is\_first\_measles\_dose; set it to 0.  
 Replace the indicator with a 1 if SIA20 is 1 or 2 and SIA27 is 3
- Calculation:  
 This calculation is a simple weighted average. The point estimate may be obtained without taking the survey design into effect. Any estimate of uncertainty should include adjustment for the survey design.
- Interpretation: “X% of eligible children who were living here during the campaign received their first-ever dose of [measles] vaccine in the recent campaign.”
- Table Output: Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)
- Graph: Inchworm plots & organ pipe plots (any clusters with extraordinarily high coverage might warrant attention)
- Provenance: COSAS & others

### SIA COVG 03: Lifetime measles doses, by birth cohort

Weighted: Yes

Description:

Each SIA will be targeted at a population of kids who span several years of age. Each year of age is considered to be a so-called one-year “birth cohort”. In this measure we report how each cohort that was targeted by the campaign is divided across three categories: those for whom we do not find evidence (by card or history or registry) that they ever received a dose of MCV (0 doses); those for whom we find evidence of a single lifetime dose of MCV (1 dose); and those for whom we find evidence of 2+ doses (2+ doses). The three categories will sum to 100% for each cohort. (DNK is not evidence and is treated as a zero.)

Denominator: Sum of weights for all respondents

Numerator: There are three numerators:

1. Sum of weights for respondents who report never having received a dose of MCV
2. Sum of weights for respondents who show evidence of one lifetime dose of MCV
3. Sum of weights for respondents who show evidence of 2+ lifetime doses of MCV

Vaccines: Calculated for each dose

Time options: During the campaign

Survey

variables: SIA20 (Did they get a dose in the campaign? 1 or 2=yes; 3=no; 99-DNK)  
SIA27 (Had they had a dose before? 1 or 2=yes; 3=no; 99 = DNK)  
SIA28 (Date of first RI MCV dose, by card)  
SIA29 (Got first RI MCV dose, according to tick on card)  
SIA30 (Date of second RI MCV dose, by card)  
SIA31 (Got second RI MCV dose, according to tick on card)  
SIA32 (Date of first SIA MCV dose, by SIA card)  
SIA33 (Date of second SIA MCV dose, by SIA card)  
SIA11 (Household ID )  
SIA12 (Individual number of the eligible child )

Merged

variables: HM29 (Completed years of age for the child in question)

Derived

variables: Generate a new derived variable : lifetime\_mcv\_doses; set it to 0.  
Increment the variable if SIA20 is 1 or 2.  
Increment the variable if SIA28 is not missing or SIA29 is 1  
Increment the variable if SIA30 is not missing or SIA31 is 1  
Increment the variable if SIA32 is not missing  
Increment the variable if SIA33 is not missing

Consider adding a question to the survey to ascertain, by history, how many doses they had received before...this could be SIA34: According to the documented sources you have shown, [name] has received (1, 2, 3, 4, or 5) doses of MCV. Have they received any other doses of MCV, not counting one from the recent campaign and not counting those that are on their vaccination card or prior SIA cards? That turns out to be a little complicated...so let's talk about it with WHO.

Increment the variable if SIA27 is 1 and (SIA28, 29, 30, 31, 32, and 33 are all missing)  
(This is the case where the caregiver says they had a dose before, but they don't provide a card to prove it; give them credit for a single dose. Maybe in the future we will change the questionnaire to ask how many doses the child had before.)

Generate a new derived indicator variable: lifetime\_mcv\_zero; set it to zero.  
Replace the indicator with 1 if the value of lifetime\_mcv\_doses is zero.

Generate a new derived indicator variable; lifetime\_mcv\_one; set it to zero.  
Replace the indicator with 1 if the value of lifetime\_mcv\_doses is one.

Generate a new derived indicator variable; lifetime\_mcv\_two\_plus; set it to zero.  
Replace the indicator with 1 if the value of lifetime\_mcv\_doses is  $\geq 2$ .

This measure assumes that observations where SIA17 = 2 or 99 have already been dropped or had their weights set to zero.

**Calculation:** Identify the minimum and maximum values of HM29 in the SIA dataset; this represents the earliest birth cohort (could be as low as zero completed years of age) and the oldest birth cohort represented in the dataset (could possibly be a few years above the top end of the age of eligibility for the SIA in cases where survey designers are assessing how many kids-too-old were vaccinated in the campaign).

For each integer value of age-in-years-completed between the minimum and the maximum observed in the dataset, the calculation is three simple weighted averages of the indicator variables lifetime\_mcv\_zero, lifetime\_mcv\_one and lifetime\_mcv\_two\_plus, restricted to children with the value of HM29 in question. The point estimates may be obtained without taking the survey design into effect. Any estimate of uncertainty should include adjustment for the survey design.

So if the SIA and the post-SIA survey targeted children under the age of five, then the calculation will be conducted for kids aged 0, 1, 2, 3, and 4 years of completed age.

**Interpretation:** "X% of children in the age cohort who had completed Y years gave verbal or documented indication of having received [0, 1, or 2+] lifetime doses of [the campaign vaccine]."

**Table Output:** Estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted)

Listed for all three outcomes in each row (0 first, then 1, then 2+) and listing a new row for each age cohort within each survey stratum.

**Graph:** None for now

**Provenance:** WHO request, 2015

Aspects to discuss with WHO

## Post-SIA Survey: Measures Related to Coverage

- A. Consider adding a question to the survey to ascertain, by history, how many doses they had received before...this could be SIA34: According to the documented sources you have shown, [name] has received (1, 2, 3, 4, or 5) doses of MCV. Have they received any other doses of MCV, not counting one from the recent campaign and not counting those that are on their vaccination card or prior SIA cards? That turns out to be a little complicated...so let's talk about it with WHO.
- B. For kids aged 12-23m, we may also have evidence from the RI dataset and the RIHC dataset...shall we merge those in with the SIA dataset to include those data for the kids who might have MCV1 or MCV2 in the register but not have a card available (or not have them recorded on the card)? If so, update the spec here to say that we merge in the RI dataset(s)

## SIA COVG 04: Campaign doses compared to prior number of doses received

Description: This indicator shows the campaign coverage stratified by the prior number of doses received.

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: There are up to ten numerators, each of the below for those vaccinated during campaign and not vaccinated during campaign

1. Sum of weights for all respondents who received the campaign dose and had received 0 doses prior to campaign
2. Sum of weights for all respondents who received the campaign dose and had received 1 dose prior to campaign
3. Sum of weights for all respondents who received the campaign dose and had received 2+ doses prior to campaign
4. Sum of weights for all respondents who received the campaign dose and did not know if they had received a dose prior to campaign
5. Sum of weights for all respondents who received the campaign dose and had received at least one dose prior to campaign, but not sure how many

Vaccines: Calculated for campaign dose

Input

Variables: This analysis uses the stratum ID, cluster ID, survey weight and the derived variable got\_sia\_dose described in the section on SIA\_COVG\_01.

Input

Parameters PRIOR\_SIA\_DOSE\_MAX is set to SINGLE if the questionnaire was worded to capture simply whether prior doses had been received, but not to quantify them. (e.g., via a question like SIA27). The parameter is set to PLURAL if the questionnaire included questions to allow recording of a number of prior doses. If the parameter is set to SINGLE then the stratified output table has levels: Zero prior doses and 1+ prior doses and Do not know. If the parameter is set to PLURAL then the table reports levels: Zero, 1, 2+, Unknown and Doses received, but unsure how many.

Survey

variables:

SIA20 (Was the child vaccinated in the campaign? 1=yes, by card; 2 = yes, by history; 3=no; 99=DNK)

SIA22 (Receive a finger mark? 1=yes, seen; 2 = yes, unseen; 3=no; 99=DNK)

SIA27(Was the child vaccinated prior to the campaign 1=yes, by card; 2 = yes, by history; 3=no; 99=DNK)

SIA28/SIA29 = Date and tick for first vaccination

SIA30/SIA31 = Date and tick for second vaccination

SIA32/SIA33 = Dates for first and second campaigns prior to current campaign

## Post-SIA Survey: Measures Related to Coverage

This measure assumes that observations where SIA17 = 2 or 99 have already been dropped or had their weights set to zero.

### Derived

#### variables:

Generate a new derived indicator variable: doses\_prior\_to\_sia; set it to 0

- Foreach date or tick provided in variables SIA28-SIA33 increase by 1.
- Maximum number of prior doses documented in VCQI-compatible SIA dataset is currently 4 (2 RI plus 2 SIA) Note: Perhaps we should consider changing that to accommodate situations where caregivers know of more than 4 doses, or know of more than two prior SIA doses.

**Calculation:** This indicator operates in a manner different from any other currently (2019) in VCQI. If the user has specified some level 4 stratifiers, this indicator sets those aside briefly and instead substitutes its own homemade level 4 stratifier that documents the number of prior doses received. Then when the indicator is finished, it copies the user's level 4 stratifiers back into place so they will be used for other indicators.

This implies that this indicator is not compatible with other user-specified level 4 stratifiers.

**Interpretation:** "X% of eligible children who were living here during the campaign, and who had received Y prior doses, are estimated to have been vaccinated against [measles] during the recent campaign per information obtained [by card, by caregiver history, by finger mark]."

**Table Output:** Estimated %, 95% CI, Weighted N who were vaccinated during the campaign, Weighted N of the denominator

**Graph:** None. There is code in place to make inchworm plots, but in practice they turn out to be inconveniently tall if there are a large number of strata and the value label is very long if the input parameter is PLURAL, so the plots have been commented out (turned off) for now.

**Provenance:** This table was proposed in 2018 by Felicity Cutts.

## SIA COVG 05: Clusters with suprisingly low coverage

Description: This indicator identifies clusters whose coverage campaign dose is below a user-specified threshold.

Weighted: User-specifies, yes or no

Denominator: Count (or sum of weights) for all respondents in the cluster

Numerator: Count (or sum of weights) for respondents who received the dose

Parameters:

- Specify whether to use number of respondents or % coverage  
SIA\_COVG\_05\_THRESHOLD\_TYPE = [COUNT or PERCENT]
- Specify the threshold for ‘suprisingly low’ coverage
- SIA\_COVG\_05\_THRESHOLD = [integer if COUNT, and percent (0-100) if PERCENT]
- Specify whether to generate a table for every stratum and list every cluster, or whether to generate a single table that lists only the clusters that fall below the threshold.
- SIA\_COVG\_05\_TABLES = [ALL\_CLUSTERS or ONLY\_LOW\_CLUSTERS]

Input

Variables: This analysis uses the stratum ID, cluster ID, survey weight and the derived variable got\_sia\_dose described in the section on SIA\_COVG\_01.

Derived

Variables

Calculate:

- cluster\_n = number of respondents in the cluster
- cluster\_sumwt = sum of weights for all respondents in the cluster
- got\_sia\_count = count of persons in the cluster who rec’d dose
- got\_sia\_sumwt = sum of weights for all respondents in cluster that rec’d dose
- got\_sia\_pct= weighted % of respondents who got <dose> in the cluster

Save a dataset with one row per cluster.

Calculation: Make a list of clusters, in order of stratum ID and cluster ID, listing count of respondents where the coverage variable == 1 and weighted % of respondents where the coverage == 1. Generate output table or tables as described below.

Table Output: If making tables for each stratum (SIA\_COVG\_05\_TABLES == “ALL\_CLUSTERS”) then the table name (and Excel tab name) will list the stratum ID and name; the table will list:

Cluster ID & name, count of respondents in the cluster, count of respondents with got\_sia\_dose == 1, % covered (sum of weights in the cluster where got\_sia\_dose == 1 divided by sum of all weights in the cluster); highlight those whose coverage (absolute or %) falls below the specified threshold; list clusters in order of cluster ID.

## Post-SIA Survey: Measures Related to Coverage

If making a single table that lists only the clusters with surprisingly low coverage across all strata, the table will list:

Stratum ID & name, Cluster ID & name, count of respondents in the cluster, absolute count of respondents with `got_sia_dose == 1`, % covered (sum of weights in the cluster where `got_sia_dose == 1` divided by sum of all weights in the cluster); only list the cluster in the table if its coverage meets the user-specified low coverage criterion; list the strata in order of stratum ID; list the clusters in order of cluster ID

Graph: None (But the organ pipe plots from `SIA_COVG_01` are closely related to this indicator.)

Interpretation: “Low coverage is defined here as being a cluster where fewer than <threshold> <percent or individuals> showed evidence of vaccination. The clusters highlighted in this list show evidence of low coverage.”

Provenance: COSAS had a cluster-level coverage (or count) report

## POST-SIA SURVEY – MEASURES RELATED TO QUALITY OF SERVICES

### SIA\_QUAL\_01: Received a campaign card

Weighted: No

Denominator: Number of respondents who were vaccinated in the campaign

Numerator: There are three numerators:  
 Number of vaccinated respondents whose card was seen by survey data collectors  
 Number of vaccinated respondents who reported having a card, but it was not seen  
 Number of vaccinated respondents who either showed a card or reported receiving one

Time options: During the campaign

#### Survey

variables: SIA20 (rec'd vaccine during campaign: 1 (by card) 2 (by history) 3 (no) 99 (DNK))  
 SIA21 (rec'd vaccination card in the campaign 1 (seen) 2 (history) 3 (no) 99 (DNK))  
 This measure assumes that observations where SIA17 = 2 or 99 have already been dropped or had their weights set to zero.

#### Derived

#### variables:

Generate a new derived variable indicator equal to missing: campaign\_card\_seen

- Replace the indicator with 1 if SIA21 == 1 & (SIA20 == 1 | SIA20 == 2)
- Replace the indicator with 0 if SIA21 != 1 & (SIA20 == 1 | SIA20 == 2)

Generate a new derived variable indicator equal to missing: campaign\_card\_unseen

- Replace the indicator with 1 if SIA21 == 2 & (SIA20 == 1 | SIA20 == 2)
- Replace the indicator with 0 if SIA21 != 2 & (SIA20 == 1 | SIA20 == 2)

Generate a new derived indicator variable equal to missing: got\_campaign\_card

- Replace the indicator with 1 if they got a card and were vaccinated (SIA21==1 | SIA21 ==2) & (SIA20 == 1 | SIA20 == 2)
- Replace the indicator with 0 if they did not get a card, or do not know, but were vaccinated (SIA21 !=1 & SIA21 != 2) & (SIA20 == 1 | SIA20 == 2)

Calculation: There are three measures: % who had card (seen), % who had card (unseen) and the total of the two: % who had card. The measures are unweighted averages of the indicator variables.

Interpretation: “Among the N children who were vaccinated in the campaign, X% demonstrated that they received a card.”

“Among the N children who were vaccinated in the campaign, X% reported having received a campaign card, but did not show it.”

Post-SIA Survey: Measures Related to Quality of Services

“Among the N children who were vaccinated in the campaign, X% either demonstrated that they received, or reported having received a campaign card.”

Table Output: % and N (count of those who were vaccinated in campaign)

Graph: None

Provenance: COSAS & others

## HYPOTHESIS TESTING – TESTING FOR DIFFERENCES IN COVERAGE

### CVG\_DIFF\_01: Differences between strata

**Description:** This indicator allows the user to test the hypothesis that coverage is the same between two strata, i.e., two provinces or two districts. The user specifies the strata of interest and the variable being tested and then runs the test. Results are written to a database and optionally to a spreadsheet.

**Weighted:** Yes

**Null**

**hypothesis:** Underlying population level coverage is the same in the two strata being tested

**Alternative**

**hypothesis:** Coverage in the first stratum is not equal to coverage in the second.

**Survey**

**variables:** Any variable in the RI, SIA, or TT dataset that takes only values of 0, 1, or missing.

**Inputs:**

The inputs are a set of global macros which are defined before the first test and then re-defined for subsequent tests.

CVG_DIFF_01_STRATUM_LEVEL	2 or 3
CVG_DIFF_01_ANALYSIS_COUNTER	Set to whatever value was used in the analysis that generated the dataset – usually 1.
CVG_DIFF_01_ID_OR_NAME	Set to ID or NAME to specify whether the user will identify the two strata using their IDs or their names, both of which must match those in the appropriate name dataset. So if the hypothesis is between strata at level 2, then the IDs or the NAMES specified below must match the IDs and NAMES in the LEVEL2_NAME_DATASET
CVG_DIFF_01_STRATUM_ID1	ID of stratum 1 to test (if ID_OR_NAME is ID)
CVG_DIFF_01_STRATUM_ID2	ID of stratum 2 to test (if ID_OR_NAME is ID)
CVG_DIFF_01_STRATUM_NAME1	Name of stratum 1 (if ID_OR_NAME is NAME)
CVG_DIFF_01_STRATUM_NAME2	Name of stratum 2 (if ID_OR_NAME is NAME)
CVG_DIFF_01_INDICATOR	Name of the indicator that generated the variable to test. i.e., RI_COVG_01 or TT_COVG_01
CVG_DIFF_01_VARIABLE	Name of the coverage variable to be tested. i.e., got_crude_penta3_by_card or protected_at_birth_to_analyze

**Calculation:** The p-value for the hypothesis test is from the Rao-Scott adjusted chi-square for survey variables.

## Hypothesis Testing

Interpretation: “The probability of observing two strata of this size with sample proportions that differ by this much or more if the underlying coverage were the same is equal to the p-value.”

Table Output: Stratum level, stratum 1 ID & name, stratum 2 ID & name, variable tested, unweighted N and weighted N, coverage and 95% CI in stratum 1 & stratum 2, difference in coverage, degrees of freedom for the test, 95% CI for the difference, and Rao-Scott p-value for the test.

Graph: None

Provenance: 2015 WHO Reference Manual

**CVG DIFF 02: Differences between subpopulations within a stratum**

**Description:** This indicator allows the user to test the hypothesis that coverage is the same between two subpopulations, i.e., urban vs. rural, male vs. female, literate caregiver vs. illiterate, received ante-natal care vs. did not. The user specifies the stratum of interest and the two subpopulations and the variable being tested and then runs the test. Results are written to a database and optionally to a spreadsheet.

**Weighted:** Yes

**Null**

**hypothesis:** Underlying population coverage is the same in the two subpopulations being tested

**Alternative**

**hypothesis:** Coverage in the first subpopulation is not equal to coverage in the second.

**Survey**

**variables:** Any variable in the RI, SIA, or TT dataset that takes only values of 0, 1, or missing.

**Inputs:**

The inputs are a set of global macros which are defined before the first test and then re-defined for subsequent tests.

CVG_DIFF_02_STRATUM_LEVEL	1 or 2 or 3
CVG_DIFF_02_ANALYSIS_COUNTER	Set to whatever value was used in the analysis that generated the dataset – usually 1.
CVG_DIFF_02_ID_OR_NAME	Set to ID or NAME to specify whether the user will identify the two strata using their IDs or their names, both of which must match those in the appropriate name dataset. So if the hypothesis is between strata at level 2, then the IDs or the NAMES specified below must match the IDs and NAMES in the LEVEL2_NAME_DATASET
CVG_DIFF_02_STRATUM_ID	ID of stratum to test (if ID_OR_NAME is ID)
CVG_DIFF_02_STRATUM_NAME	Name of stratum (if ID_OR_NAME is NAME)
CVG_DIFF_02_INDICATOR	Name of the indicator that generated the variable to test. i.e., RI_COVG_01 or TT_COVG_01
CVG_DIFF_02_SUBPOP_VARIABLE	Variable that holds the levels of the subpopulation (i.e., urban_cluster, sex, caregiver_literate, etc.)
CVG_DIFF_02_SUBPOP_ID1	Level of first population (must be an integer)
CVG_DIFF_02_SUBPOP_ID2	Level of second population (must be an integer)
CVG_DIFF_02_VARIABLE	Name of the coverage variable to be tested. i.e., got_crude_penta3_by_card or protected_at_birth_to_analyze

**Calculation:** The p-value for the hypothesis test is from the Rao-Scott adjusted chi-square for survey variables.

## Hypothesis Testing

Interpretation: “The probability of observing two subpopulations this size with sample proportions that differ by this much or more if the underlying coverage were the same is equal to the p-value.”

Table Output: Stratum level, stratum ID & name, subpopulation variable, subpopulation id & name for subpopulations 1 & 2; unweighted and weighted N, variable tested, coverage and 95% CI in stratum 1 & stratum 2, difference in coverage, degrees of freedom for the test, 95% CI for the difference, and Rao-Scott p-value for the test.

Graph: None

Provenance: 2015 WHO Reference Manual