VCQI

VACCINATION
COVERAGE
QUALITY
INDICATORS

Results Interpretation
Quick-Reference Guide

Draft Version 1.5
Revised February 2021

World Health Organization

Expanded Programme on Immunization

This page is intentionally blank.

This page is intentionally blank.

ACKNOWLEDGEMENTS

This page is intentionally blank.

INTRODUCTION

The Vaccination Coverage Quality Indicators (VCQI)[1] are a World Health Organization resource that define a set of measures that may be calculated using data from household surveys. VCQI documentation defines the measures and a set of freely available Stata programs calculates them and makes standard tables and figures that are suitable for copying and pasting into reports. The VCQI software also generates datasets that hold coverage results and outcome variables; these may be used to make customized tables and figures or as inputs for additional analyses.

The purpose of this guide is to help readers interpret the tables and figures that VCQI produces. It is organized by type of survey and type of indicator with Routine Immunization (RI) indicators listed first, then Maternal Tetanus (TT) surveys, and finally Supplemental Immunization Activity (SIA) surveys which are sometimes called Post-Campaign Coverage Surveys (PCCS).

INDICATORS

Most of VCQI's indicators are organized by national vaccination programme attributes that have proven useful in earlier assessments:

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

Some are organized by the type of figures that they produce:

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

Each indicator is summarized in its own brief section to describe a) whether the indicator is weighted, b) which respondents are included in the indicator (i.e., the definition of the denominator), c) what qualifies as having the outcome of interest (definition of the numerator), d) how missing or "do not know (DNK)" responses affect the calculation, and e) how to describe the estimated outcomes in plain English language sentences.

TABLES

Broadly speaking, VCQI generates three types of tables:

1. Unweighted sample proportions, where the denominator is a count of a subset of respondents and the numerator is a count of a subset of the denominator
2. Survey-weighted estimated proportions, where the denominator is the sum of weights for all respondents and the numerator is the sum of weights of the respondents with the characteristic of interest
3. Miscellaneous tables that do not fit descriptions 1 or 2

Interpreting unweighted proportions is straightforward. They are reported with two simple columns: one listing N and another listing %. The interpretation is generally of the form: "Of the N respondents

---

[1] Pronounced "Vicki"

who met the condition to be in the denominator, X% met the condition to be in the numerator of this indicator." These proportions are reported without confidence intervals or any representation of sampling uncertainty. They are simple statements about what was observed in the survey sample. It is currently a VCQI convention that when the denominator represents a subset of all survey respondents, the outcomes are described with unweighted proportions.

Interpretation of VCQI's survey-weighted proportions may be generalized to the entire population of eligible respondents. They include a point estimate, 95% confidence interval and sample size and in some cases other summary statistics. Their interpretation is generally of the form: "X% of the population who were eligible for the survey are estimated to have <had the outcome being summarized> as documented by <whatever sources of vaccination evidence your survey used>."

The document also describes additional statistics that are found in some VCQI tables, such as the design effect, intracluster correlation coefficient, and 1-sided confidence bounds.

FIGURES

VCQI can produce eight types of plots:

1. Organ pipe plots
2. Inchworm plots
3. Double-inchworm plots
4. Coverage bar charts

5. Double bar charts
6. Cumulative coverage curves
7. Cumulative interval curves
8. Vaccination coverage & timeliness charts

This guide shows examples of each and describes how to interpret them.

The document concludes with instructions on how to interpret the output from VCQI's RI date data quality report.

OTHER RESOURCES

Other helpful VCQI resources include *Getting Started with VCQI*, the *VCQI User's Guide*, the *Forms & Variable Lists (FVL) Structured for Compatibility with VCQI,* and the *Working List of Vaccination Survey Analyses and Software Specifications*. You may find the latest versions at the VCQI resources website[2].

Please send feedback on this guide to Dale Rhoda at Biostat Global Consulting Dale.Rhoda@biostatglobal.com and Carolina Danovaro at WHO danovaroc@who.int.

---

[2] http://www.biostatglobal.com/VCQI_resources.html

RI_COVG: 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 per card or recall (or register, if health centers were visited in this survey)

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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

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 per card or register

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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.

RI_COVG_03: FULLY VACCINATED

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received <u>all</u> <the doses in the list that makes up "fully vaccinated">

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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*>."

RI_COVG_04: NOT VACCINATED

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who received <u>none</u> of <the doses in the list that makes up "fully vaccinated">

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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."

RI_COVG_05: CLUSTERS WITH SURPRISINGLY LOW CRUDE 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

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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*>."

RI_ACC: 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 DPT1 / PENTA1 per card or recall (or register, if health centers were visited in this survey)

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

Note: The output for RI_ACC_01 is simply a copy of the output for RI_COVG_01 for the DPT1 (or PENTA1) dose. When it is calculated as RI_ACC_01, it may be given an interpretation that explicitly mentions access to vaccination services.

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)>*."

RI_CONT: RI SURVEY – MEASURES RELATED TO CONTINUITY OF SERVICES

RI_CONT_01: DROPOUT BETWEEN TWO CRUDE DOSES

Weighted: No

Denominator: Number of respondents who received the first dose and were age-eligible to receive the second dose before the survey

Numerator: Number of respondents who received the first dose and who were eligible but did <u>not</u> receive the second dose

Missing/DNK: If data for the first dose is missing or "do not know" (DNK), the respondent does not appear in the denominator because only those with evidence of receiving that dose are counted there. If they have evidence of vaccination for the first dose, they are in the denominator and then if evidence for the second dose is missing or DNK then they will be counted as having dropped out (did not receive the second dose) in the numerator.

Description: It is common and straightforward to calculate weighted dropout results from the crude or valid coverage tables (i.e., from RI_COVG_01 or RI_COVG_02). For any two doses, (early and later) the commonly reported dropout proportion is simply
(Cvg_early – Cvg_later) / Cvg_early.

We could write code to generate a table of that quantity, but the weights themselves may not be especially helpful for this indicator. To be consistent with other unweighted VCQI measures, if the denominator does not include all respondents, this indicator estimates and reports an unweighted proportion.

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

RI_QUAL: RI SURVEY – MEASURES RELATED TO QUALITY OF SERVICES

RI_QUAL_01: CARD AVAILABILITY

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: There may be up to 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, if the child is missing the DOB, then the date fell between the earliest possible vx date for children in this dataset 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

Missing/DNK: All respondents are in the denominator, even if the response to showing a card or register is missing or "do not know" (DNK). In the numerator, anyone with missing/DNK responses would be counted as having NOT shown a card or register.

Interpretation: To interpret columns labeled "RI Card Availability": "X% of the population eligible for the survey are estimated to have a home-based record available for interviewers to view."

To interpret the column labeled "RI Card with Dates": "X% of the population eligible for the survey are estimated to have a home-based record with 1+ vaccination dates written on it."

To interpret the column labeled "RI Card with Dates or Ticks": "X% of the population eligible for the survey are estimated to have a home-based record with 1+ dates or tick marks written on it."

To interpret the column labeled "RI Card with Only Clean Dates": "X% of the population eligible for the survey are estimated to have a home-based record with only clean dates recorded on it (where *clean* means the date fell between the child's DOB and the date of the survey (or between the earliest possible vx date and the date of the survey) and dates for dose series were in chronological order)."

To interpret columns labeled "RI Register Availability": "X% of the population eligible for the survey are estimated to have a facility-based record (*register*) available for interviewers to view."

To interpret the column labeled "RI Register with Dates": "X% of the population eligible for the survey are estimated to have a facility-based record (*register*) with 1+ vaccination dates written on it."

To interpret the column labeled "RI Register with Dates or Ticks": "X% of the population eligible for the survey are estimated to have a facility-based record with 1+ vaccination dates or tick marks written on it."

To interpret the column labeled "RI Register with Only Clean Dates": "X% of the population eligible for the survey are estimated to have a facility-based record with only clean dates recorded on it (where *clean* means the date fell between the child's DOB and the date of the survey (or between the earliest possible vx date and the date of the survey) and dates for dose series were in chronological order)."

To interpret the column labeled "RI Card or Register Availability": "X% of the population eligible for the survey are estimated to have either a home-based record (*card*) or a facility-based record (*register*), or both, available for interviewers to view."

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

Missing/DNK: All respondents are in the denominator, even if the response to having received a card is missing or "do not know" (DNK). In the numerator, anyone with missing/DNK responses would be counted as having NOT ever received a card.

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."

RI_QUAL_03: NO LONGER RECOMMENDED
We recommend that you run RI_QUAL_04 instead of RI_QUAL_03.

RI_QUAL_04: PERCENT OF DOSES ADMINISTERED BEFORE A SPECIFIED AGE
Weighted: No

Denominator: Number of respondents who had date of birth (DOB) data and a dose date, by card or register

Numerator: Number of respondents whose <dose> was given before <threshold> age (in days)

Missing/DNK: If the DOB or dose date is missing or "do not know" (DNK) then the respondent is excluded from this indicator calculation.

Description: Often used to identify % of children who received DPT1 or MCV1 too early.

Note: If the dates of vaccination on card and register disagree, and one shows that <dose> was given before <threshold> and the other shows it was given on or after <threshold>, this indicator gives the benefit of doubt by counting it as having been given on or after the threshold. That is to say that it is considered preferable for the child to have been at least <threshold> days old when they received the dose. If another assumption is appropriate, the age at vaccination per card and per register are both saved in the RI_QUAL_04 dataset and another outcome could be calculated.

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."

RI_QUAL_05: PERCENT OF SEQUENTIAL DOSES WITH AN INTERVAL THAT IS TOO BRIEF
Weighted: No

Denominator: Number of <dose2> (& <dose3>) doses administered where the date was known for that dose <u>and</u> for the preceding dose

Numerator: Number of times the later dose was administered before <threshold> days had passed from the date of the earlier dose

Missing: If the vaccination date for dose1 is missing, the dose1 to dose 2 interval cannot be calculated, so that interval is excluded from the calculation. If the date for dose 3 is missing, the dose2 to dose 3 interval cannot be calculated so that interval is excluded. If the date for dose2 is missing, neither interval can be calculated, so the respondent is excluded from the indicator.

Description: Often used to identify % of DPT2 & 3 doses administered before 28 days had passed. This indicator assumes that it is best to have the doses administered at least as far apart as <threshold> days. It differs from RI_COVG_12 in that regard, because _12 assumes it is best to have the doses administered after an interval that is shorter than <threshold> days.

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 shorter than <*threshold*> days."

RI_QUAL_06: PERCENT OF VALID <DOSE> DOSES THAT WERE ADMINISTERED BEFORE THE AGE OF 12 MONTHS
Weighted: No

Denominator: Number of children who had valid <dose>

Numerator: Number of children whose valid <dose> was received before the age of 12 months

Missing/DNK: If the DOB or dose date is missing or "do not know" (DNK) then the respondent is excluded from this indicator calculation.

Description: Often used to quantify the % of valid MCV1 doses administered before age 12 months

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

RI_QUAL_07B: VALID COVERAGE IF THERE HAD BEEN NO MISSED OPPORTUNITIES FOR SIMULTANEOUS VACCINATION (MOV)

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for all respondents who had a valid dose plus those who would have had a valid dose if they had received every dose they were eligible for at every date recorded on their card

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). This indicator does not consider a child's vaccination evidence; it only uses the list of dates (ages) when they were vaccinated. Children with no vaccination dates are not counted as those who would have had a valid dose if there were not MOVs.

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."

Note: This proportion can never be higher than the weighted % whose home-based record or health center register showed dates of vaccination. That is to say that if the dataset only records vaccination dates for respondents whose weights sum to 50% of the total weight, then this indicator can not exceed 50%.

A NOTE REGARDING RI_QUAL_08 AND RI_QUAL_09

RI_QUAL_08 and _09 all summarize MOVs in the survey dataset.

When interpreting those indicators, it is very important to be clear whether the analysis was done with the CRUDE option (invalid doses count) or the VALID option (early doses are ignored).

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 3rd valid dose. So if the MOV analysis does not give credit for invalid doses (specify VALID option when running VCQI) then when the child returns for the measles vaccine at age 9 months, they are considered to be eligible for a 3rd 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 invalid doses (specifies the CRUDE option), then the child is still counted as having two valid doses of DPT, but they are <u>not</u> considered eligible for a 3rd dose at the measles visit, and that visit is not considered to be a missed opportunity for DPT.

Specifying the VALID option will result in higher results for the MOV indicators. 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 they would not.

It is my (Dale Rhoda) understanding that at this time (February 2017) WHO does not formally advise countries to give additional doses in a series if the child has received the full target number of doses, but some were invalid. (The practice may vary from country to country and even within countries.) So to summarize performance of the vaccination program as it is administered, it is probably appropriate to use the CRUDE option in the analysis. But biologically, children who receive a full complement of valid doses are probably more likely to develop immunity than those who receive some or all invalid doses. So it may be informative to do the MOV analysis twice…once with the parameter set to CRUDE and again with the parameter set to VALID, and to compare the output.

RI_QUAL_08: Percent of visits with missed opportunity for simultaneous vaccination (MOV)

Weighted: No

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

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

Missing/DNK: If a child is age-eligible for a dose on a documented vaccination visit date but are missing evidence of vaccination on that date and the caregiver says they did not receive the dose, or they "do not know" (DNK) whether the child received the dose, then the child will be counted in the numerator as having experienced an MOV for that dose on that date.

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." (For example, 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.")

RI_QUAL_09: PERCENT OF CHILDREN WITH MISSED OPPORTUNITY FOR SIMULTANEOUS VACCINATION (MOV)

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

Missing/DNK: Children missing date of birth are excluded from the denominator, and children with a date of birth but no dates of vaccination are also excluded from the denominator. Children missing any evidence of receiving this <dose> or with "do not know" (DNK) response for this <dose> will be counted as having had an MOV in the numerator if they are known to have had a vaccination visit after they were age-eligible for the dose.

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

    a) experienced an MOV (for each dose)
    b) experienced an uncorrected MOV (for each dose)
       (meaning that they had not received a valid dose as of the time of the survey)
    c) experienced a corrected MOV (for each dose)
       (meaning that they received 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) for whom <u>all</u> MOVs were uncorrected,
    c) for whom <u>all</u> MOVs were corrected,
    d) for whom <u>some but not all</u> 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."

RI_QUAL_12: PERCENT OF SEQUENTIAL DOSES WITH AN INTERVAL THAT IS TOO LONG

Weighted: No

Denominator: Number of times the two doses were administered and there was a recorded date for each

Numerator: Number of times the later dose was administered after the interval (in days) in question

Missing: If either date is missing, the interval is excluded from the indicator calculation.

Description: Often used to identify % of DPT2 & DPT3 doses administered at intervals longer than 56 days. This indicator assumes that it is best to have the doses administered after an interval that is shorter than <threshold> days. It differs from RI_COVG_05 in that regard, because _05 assumes it is best to have the doses administered after an interval that is at least <threshold> days long.

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."

RI_QUAL_13: NO LONGER RECOMMENDED.
We recommend that you run RI_QUAL_04 instead of RI_QUAL_13.

RI_CCC: RI SURVEY – CUMULATIVE CURVES OF AGE AT VACCINATION AND INTERVAL BETWEEN VACCINATIONS

INTERPRETING CUMULATIVE COVERAGE CURVES

RI_CCC_01: NO LONGER RECOMMENDED

We recommend you run RI_CCC_02 instead of RI_CCC_01 (because _02 uses survey weights while _01 does not).

RI_CCC_02: CUMULATIVE COVERAGE CURVES

This figure shows an example of a weighted cumulative coverage curve:
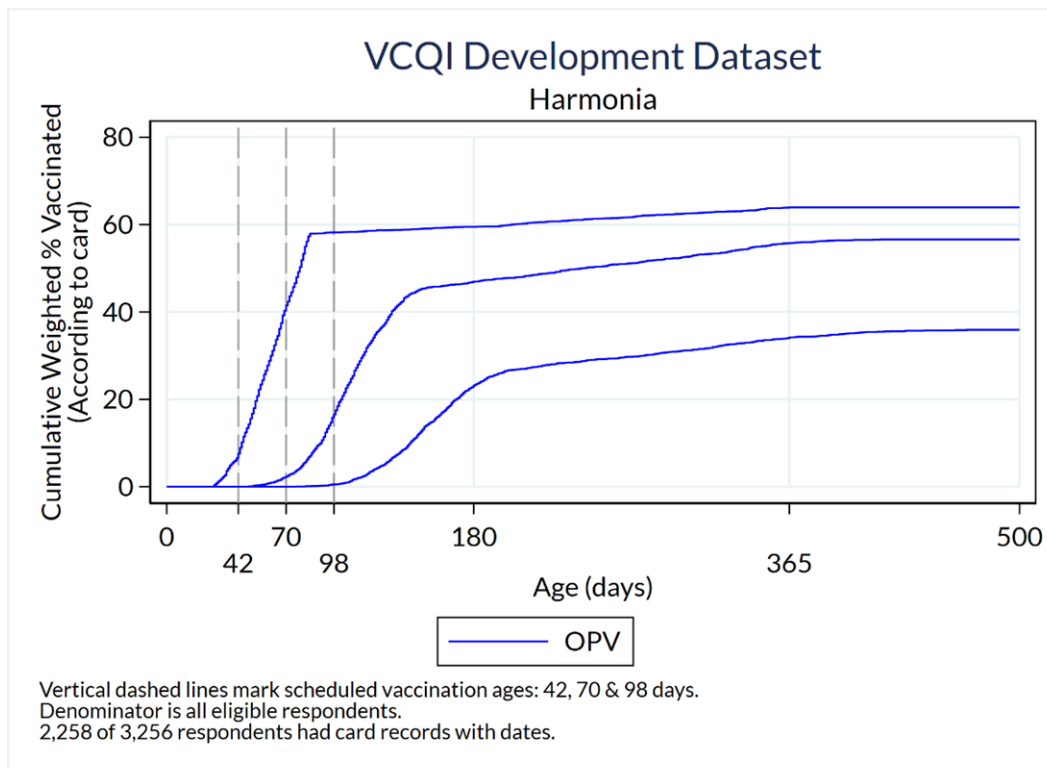
Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents who are documented as having received the dose by age X (in days)

Missing/DNK: Doses recorded with illegible dates or tick marks or whose dose date is missing altogether are not represented in these plots.

Interpretation: A point on a given cumulative coverage curve can be interpreted as: "Y% of respondents are estimated to be documented as having been vaccinated for *<dose>* by the time they were X days old."

Note that the curve does not show timing of vaccination for respondents who were indeed vaccinated but whose age at vaccination is unknown (because they are missing the date of birth or date of vaccination, or the date of vaccination is nonsensical, or partially or illegibly recorded, or because their evidence of vaccination is from caregiver recall only). To gain an appreciation for what portion of respondents have evidence of being vaccinated where the timing is unknown, the analyst should also look at the so-called Vaccination Coverage and Timeliness Charts (RI_VCTC_01).

In a perfect world, cumulative coverage curves would be a stair-step function, where 0% of respondents received the dose until the scheduled age, at which time all 100% of respondents received the dose (vertical line straight up), then horizontal line extending to the right indicating no respondent received the dose after the minimum scheduled age.

On the plots, vertical lines mark scheduled vaccination ages, so the point at which the curve crosses its associated vertical schedule line should be at 0%. If there is white space under the curve *before* it crosses the vertical line, then that percentage of respondents received the given dose early. Similarly, white space *above* the curve after it crosses the vertical line but before the plateau represents respondents who were unprotected for that given time range.

RI_CIC_01: NO LONGER RECOMMENDED

We recommend that you run RI_CIC_02 instead of _01 (because 02 incorporates survey weights, whereas _01 does not).

RI_CIC_02: CUMULATIVE INTERVAL CURVES

This figure shows an example of a weighted cumulative interval curve that summarizes the distribution of observed intra-dose intervals (in days) between receiving OPV1 and OPV2.
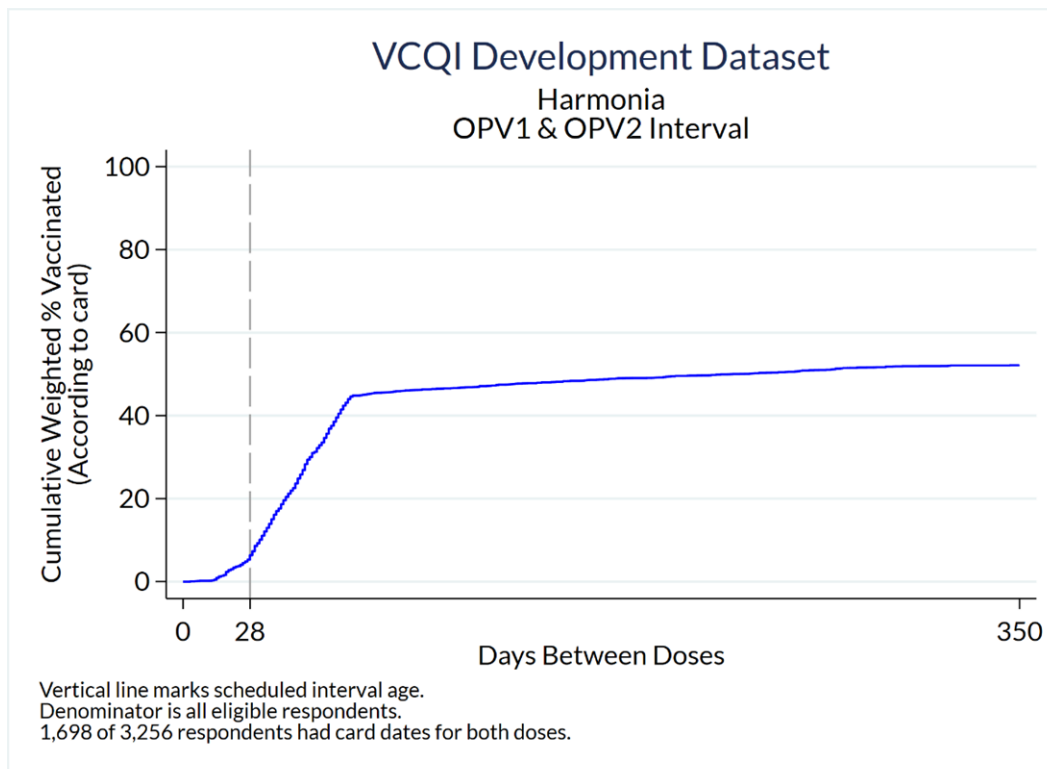
Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents whose vaccination dates indicate that the intra-dose interval was ≤ X days

Vaccines: Calculate for every two-dose antigen (interval between doses 1 and 2) and for every three-dose antigen (intervals between 1 and 2 and between 2 and 3)

Missing/DNK: Dose pairs recorded with illegible dates or tick marks or whose dose date is missing altogether are not represented in these plots.

Interpretation: A point on a given cumulative interval curve can be interpreted as: "Y% of respondents are estimated to have evidence of receiving both doses with an intra-dose interval no longer than X days."

In a perfect world, these curves would be a stair-step function, where 0% of respondents received the latter dose until the scheduled interval, at which time all the respondents received the latter dose (vertical line straight up), then a horizontal line extending to the right indicating no respondent received the latter dose after the scheduled interval.

On the plots, vertical lines mark scheduled intervals, so the point at which the curve crosses it should be at 0%. If there is white space under the curve *before* it crosses the vertical line, then that percentage of respondents had too few days between doses (i.e., received the latter dose early). Similarly, white space *above* the curve after it crosses the vertical line but before the plateau represents respondents who experienced an interval longer than the scheduled number of interval days, and possibly were unprotected for that time period.

Notes: Two vertical lines may appear on the plot. The latter dose's minimum interval in days, based on the schedule, is used to plot one vertical line. The difference between the dose pair's minimum age in days, based on the schedule, is also calculated. If the difference is not equal to the minimum interval, then a second vertical line will appear on the plot. In this case, the interval between these two vertical lines represents a time period when a respondent could receive the latter dose, and it could be counted as early valid. Respondents having an interval period between a given dose pair that occurs before the first vertical line represents an early latter dose, whereas respondents having an interval period after the second vertical line represents a late latter dose.

For example, some countries schedule a first dose of measles rubella vaccine at age 9 months and a second dose at age 18 months. The scheduled intra-dose interval is 9 months, but the second dose would be considered valid any time after a 28-day interval. So the plot for the MCV1 – MCV2 interval might include a vertical line at 28 days and a vertical line at 270 days (9 months). Anytime the second dose is received before 28 days have passed, the latter dose is early (invalid). Anytime the second dose is received after 270 days have passed, the latter dose is late. (We use the term *late* here quite strictly. Some countries may not consider the dose to be officially *late* until one or two months have passed after the age when the dose is scheduled.)

RI_VCTC: RI SURVEY – VACCINATION COVERAGE AND TIMELINESS CHARTS

INTERPRETING VACCINATION COVERAGE AND TIMELINESS CHARTS

RI_VCTC_01: VACCINATION COVERAGE AND TIMELINESS CHARTS (VCTC)

This figure shows an example of coverage and timeliness charts for OPV1 through OPV3. The figure is a categorical representation of the same data as in the cumulative coverage curves. Each figure summarizes data for a geographic stratum. Each dose is represented with a bar. By default, timeliness is divided into four categories: too early, timely, < 2 months late, and 2+ months late and a fifth category indicates the portion whose evidence of vaccination does not allow calculation of age at vaccination.
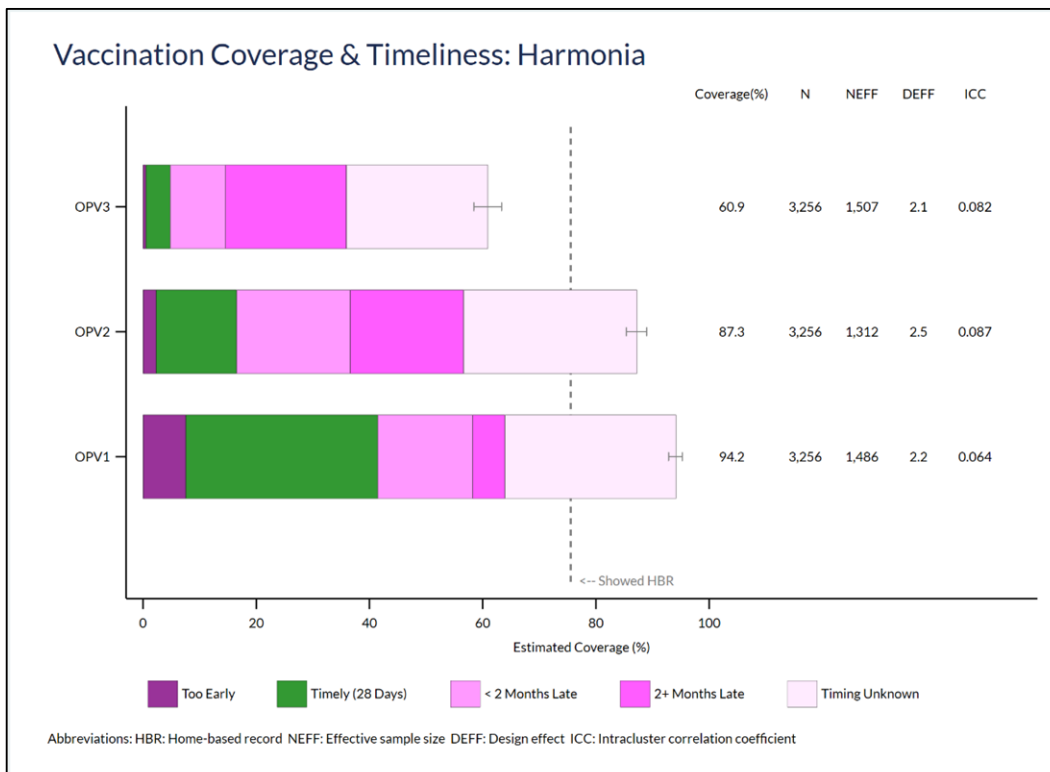
Vaccination Coverage & Timeliness: Harmonia

	Coverage(%)	N	NEFF	DEFF	ICC
OPV3	60.9	3,256	1,507	2.1	0.082
OPV2	87.3	3,256	1,312	2.5	0.087
OPV1	94.2	3,256	1,486	2.2	0.064

Legend: Too Early, Timely (28 Days), < 2 Months Late, 2+ Months Late, Timing Unknown

Abbreviations: HBR: Home-based record NEFF: Effective sample size DEFF: Design effect ICC: Intracluster correlation coefficient

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for respondents whose vaccination dates indicate the dose was given either too early (before the age it was scheduled), timely (with 28 days of the age it was scheduled), < 2 months late, or 2+ months late. The category definitions may be customized by the VCQI user.

Vaccines: User may generate the chart to show any or all doses for each stratum.

Missing/DNK: If the respondent does not have evidence of vaccination, they are not represented in any colored portion of the bar. If the evidence is from caregiver recall or an incomplete or illegible date then the child is represented in the portion labeled 'Timing Unknown' .

Interpretation: A portion of a bar may be interpreted as "<*Bar segment width*>% of respondents in <*stratum*> received <*dose*> in <*timeliness category*>." The precise numeric width of each bar portion is stored in the RI_VCTC_01 Excel tab. The precise % of respondents who showed a home-based vaccination record (HBR) is stored in the RI_QUAL_01 Excel tab.

Annotated coverage estimates from the VCTC may be used to calculate drop-out from an earlier to later dose. In the example above, 94.2% of children had evidence of receiving OPV1 and only 60.9% had evidence of receiving OPV3. Roughly a third of children who received OPV1 did not go on to receive the third dose. A common equation used to generate a weighted estimate of drop-out is 100 x (Early dose coverage – later dose coverage) / Early dose coverage. In the example above, we might say that 100 x (94.2-60.9)/94.2 = 35.4% of children with evidence of OPV1 do not have evidence of receiving OPV3.

See a later section of this document for guidance on how to interpret N, NEFF, DEFF, and ICC.

TT_COVG: TT SURVEY – MEASURES RELATED TO MATERNAL TETANUS 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 whose children were protected at birth

Missing/DNK:　Children are considered to not be born protected if they are born to women who are missing information about tetanus vaccination or who do not know how many doses they received or when the most recent dose was received

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*>."

SIA_COVG: 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 per
                <card, history, or finger mark>

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all
                sources are missing or "do not know" (DNK). In the numerator, respondents with
                missing data or DNK responses are considered to have not received the dose.

Interpretation: "X% of eligible children who were living here during the campaign are estimated to have
                been vaccinated against <campaign dose> during the recent campaign per information
                obtained *by card, by caregiver history, by finger mark>*."

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

Missing/DNK: If the coverage response is missing or "do not know" (DNK) or if the respondent does
                not know whether they received a dose before, then this dose is not considered to be
                the first dose.

Interpretation: "X% of eligible children who were living here during the campaign received their first-
                ever dose of <campaign dose> vaccine in the recent campaign."

SIA_COVG_03: LIFETIME DOSES OF <CAMPAIGN DOSE>, BY BIRTH COHORT
Weighted: Yes

Description: Each SIA will be targeted at a population of children who span several years of age.
                Each year of age is a one-year "birth cohort". In this measure, we report how each
                cohort is divided across three categories: those for whom we do not find evidence (by
                card or history or fingermark or registry) that they ever received a dose of the campaign
                vaccine, those for whom we find evidence of a single lifetime dose of that vaccine, and
                those for whom we find evidence of 2+ doses. The three categories will sum to 100%
                for each cohort. (Do not know 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 the
                   campaign vaccine

2. Sum of weights for respondents who show evidence of one lifetime dose of the campaign vaccine

3. Sum of weights for respondents who show evidence of 2+ lifetime doses of the campaign vaccine

Missing/DNK: If the number of lifetime doses is missing or "do not know" (DNK), the respondent is characterized as having never received a dose of the campaign vaccine.

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*>."

SIA_COVG_04: CAMPAIGN COVERAGE STRATIFIED BY PRIOR NUMBER OF DOSES RECEIVED
Weighted: Yes

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

Denominator: Sum of weights for all respondents

Numerator: There can be several numerators.

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

Interpretation: The output appears using several rows per stratum.

"X% of children were vaccinated during SIA."

"X% of children who received 0 doses prior to campaign were vaccinated during SIA."

"X% of children who received 1 dose prior to campaign were vaccinated during SIA."

"X% of children who received 2+ doses prior to campaign were vaccinated during SIA"

"X% of children who do not know if dose received prior to campaign were vaccinated during SIA."

"X% of children who received a dose prior to campaign, but unsure how many were vaccinated during SIA."

SIA_COVG_05: CLUSTERS WITH SURPRISINGLY LOW CAMPAIGN 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

Missing/DNK: All respondents are in the denominator, even if their coverage responses from all sources are missing or "do not know" (DNK). In the numerator, respondents with missing data or DNK responses are considered to have not received the dose.

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."

SIA_QUAL: 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

Missing/DNK: If the vaccination response is missing or "do not know" (DNK), the respondent is excluded from the denominator. If they were vaccinated, but the response regarding the card is missing or DNK then they are considered to have not received a campaign card.

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."

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

INTERPRETING OTHER STATISTICS

The earlier portions of this document describe how to interpret the main outcomes of the survey – the estimated coverage proportions. This section describes how to interpret some of the additional statistics that are reported with weighted indicators.

2-SIDED 95% CONFIDENCE INTERVAL (CI)

Reports often state that the survey team is "95% confident" that the true coverage in the target population falls within the 2-sided 95% confidence interval obtained from the sample. If the survey is believed to be free of important biases, this is an acceptable way to present results to policymakers. Strictly speaking, the confidence interval means: "If this survey were repeated, without bias, many, many times using the same target population, the same design, the same sampling frame and protocol, the same questions, and the same analysis, and if a confidence interval were calculated using the same technique for each repetition of the survey, then 95% of the intervals would indeed contain the true population coverage number".

We cannot know whether the survey is free of bias and we cannot know the magnitude of the net bias. Furthermore, we also cannot know whether the sample selected for a given survey is one of the 95% of samples that generates an interval containing the true population parameter, or whether it is one of the 5% of samples for which the entire confidence interval lies above or below the true population parameter. However, for practical purposes (and in the absence of important biases), it is acceptable to document the steps you took to minimize bias in the survey, acknowledge that bias is not quantifiable, and say if the survey were free from bias, we would be 95% confident that the true unknown population coverage figure falls within the estimated 95% CI from the survey sample.

Note that the interval that VCQI labels the 95% CI is a 2-sided interval and we consider there to be a 2.5% chance that the true probability coverage falls below the entire interval and a 2.5% chance that it falls above the entire interval.

1-SIDED 95% LOWER CONFIDENCE BOUND (LCB)

Informally, we say that in the absence of bias we are 95% confident that the true population coverage parameter falls above the LCB. The formal interpretation is similar to that described above for the 2-sided CI; when we say we are 95% confident, we mean that if the survey were repeated many times, the true parameter would fall above the LCB for 95% of those repeated surveys.

1-SIDED 95% UPPER CONFIDENCE BOUND (UCB)

Informally, we say that in the absence of bias we are 95% confident that the true population coverage parameter falls below the UCB. Formally, the interpretation involves many repeated surveys in a manner like that of the LCB and 95% CI described above.

DESIGN EFFECT (DEFF)

When outcomes are estimated using data from a complex sample, they often have a variance (degree of uncertainty due to sampling) that is larger than would be observed with a simple random sample using the same sample size. The ratio of the observed variance to the variance that would have been observed with the same outcome estimate from a simple random sample is known as the design effect. The design effect is useful for planning subsequent surveys. VCQI constrains estimates of design effect to be ≥ 1.

EFFECTIVE SAMPLE SIZE (NEFF)

NEFF is the number of respondents from a simple random sample who would produce a result as precise (i.e. a confidence interval as narrow) as the observed outcome. Because VCQI constrains the DEFF to be ≥ 1, NEFF is constrained to be ≤ N. VCQI's vaccination coverage and timeliness charts (VCTC) have an option to annotate the plot with values of NEFF. VCQI calculates NEFF in the manner recommended by Dean & Pagano, 2015[3]. Because the calculation adjusts for degrees of freedom in the complex sample, it is not always true that NEFF = N / DEFF, as might be indicated in some survey literature. When the degrees of freedom are large, that relationship does hold in the VCQI output.

INTRACLUSTER CORRELATION COEFFICIENT (ICC)

The ICC is another quantity that is useful for planning subsequent surveys. It measures the correlation of the outcome within clusters in the sample. ICC ranges from small negative values to 1 and often falls between 0 and 1. An ICC that is negative means the outcome is spatially homogeneous. In the limit of homogeneity where coverage in every cluster is equal, the ICC will equal -1 / (average number of respondents per cluster). At the other extreme, an ICC of 1 means that the outcome is perfectly correlated with which cluster a respondent occupies – in some clusters every respondent had the outcome and in the other clusters, no respondent had the outcome of interest. Higher ICCs lead to higher values of design effect.

N (UNWEIGHTED)

N is the number of respondents in the denominator of the measure.

N (WEIGHTED)

N is the sum of the weights of the respondents in the denominator of the measure.

---

[3] Dean, Natalie, and Marcello Pagano. "Evaluating confidence interval methods for binomial proportions in clustered surveys." *Journal of Survey Statistics and Methodology* 3, no. 4 (2015): 484-503.

INTERPRETING ORGAN PIPE PLOTS

Many VCQI indicators make organ pipe plots – bar charts that show cluster-level coverage of a binary indicator for a single geographic stratum. Organ pipe plots are described briefly in Section 6.12 of the 2018 WHO Vaccination Coverage Cluster Survey Reference Manual[4]. They were the topic of a presentation at the 2018 Stata conference. An updated copy of Mary Prier's slides from that presentation are available online[5].
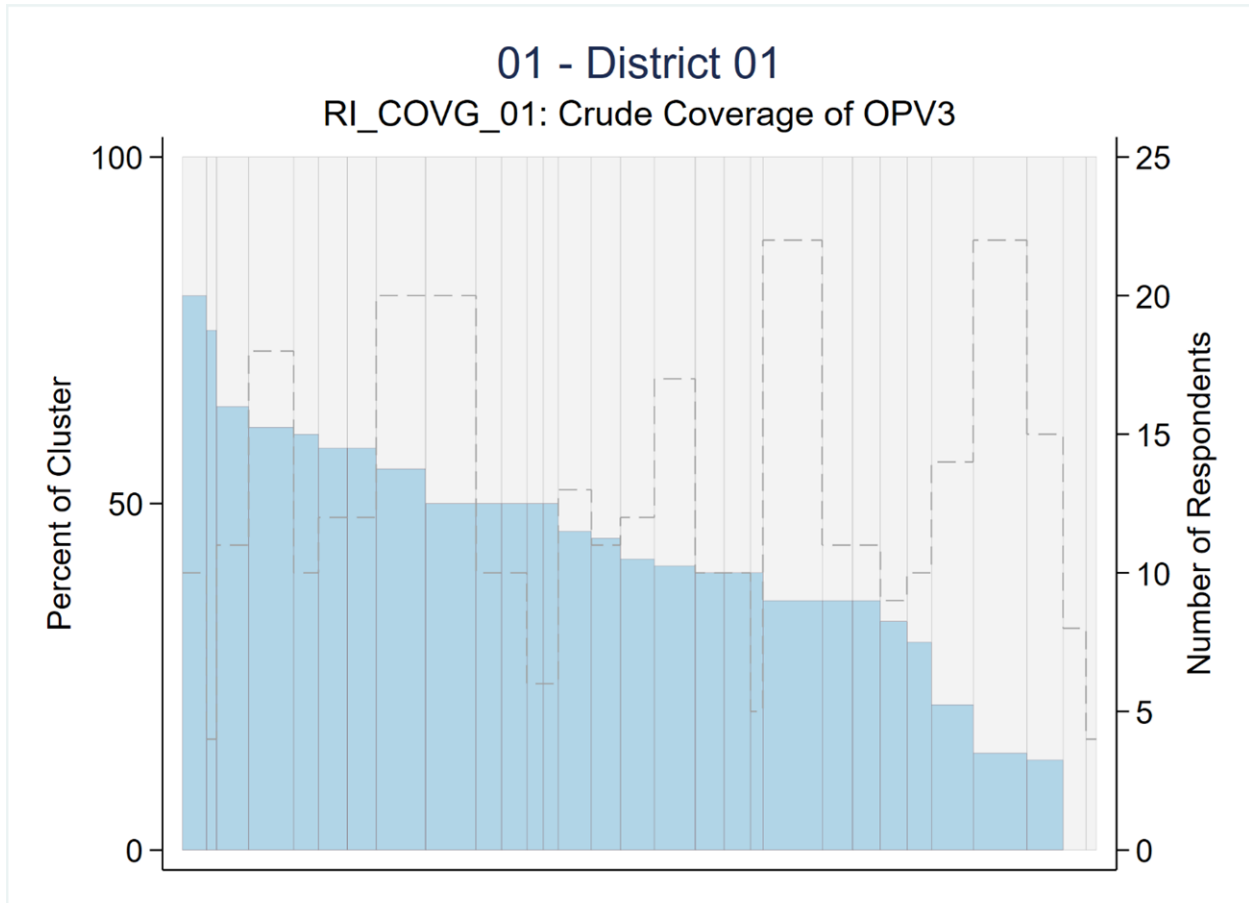
Each plot shows data for a single geographic stratum. This example is for District 01. Each plot shows data for a single binary outcome. This one is for crude coverage – whether the child had any evidence at all of having received OPV3. The entire area of the figure represents 100% of the eligible respondents. The shaded area of the plot represents the weighted proportion of respondents who have the outcome

---

[4] https://www.who.int/immunization/documents/who_ivb_18.09/en/
Note that there is a typographical error in Figure 4 in early copies of the manual. It indicated that the intracluster correlation coefficient (ICC) for the top-left sub-plot is 0, but the correct value is - 1/15. The comparable figure in the PowerPoint slides has been updated and is correct.
[5] https://github.com/BiostatGlobalConsulting/organ-pipe-plots/blob/master/opplot_presentation.pptx

of interest. The numeric weighted coverage estimate appears in the Excel tab for RI_COVG_01 but does not appear on the plot.

Each bar or column represents a single cluster. The width of the bar is proportional to the sum of the weights of respondents in that cluster. The shaded portion of the bar represents respondents with the outcome of interest. The weighted portion of respondents in the cluster with the outcome may be read off the left-side numeric scale that runs from 0 to 100%. The unshaded portion of each bar represents respondents who do NOT have the outcome. Organ pipe plots always sort the clusters so those with the highest cluster-level coverage appear at the left side and those with the lowest coverage appear at the right. In the plot above, two clusters at the far right do not have any shaded portion. None of the interviews in those two clusters yielded vaccination evidence for OPV3.

A faint gray dashed line runs across the figure indicating the number of respondents in each cluster. The scale for the number of respondents is at the right side of the plot. In this example, the scale runs from 0-25 and it looks like two clusters have as few as N=4 respondents and two have as many as N=22.

The main takeaway messages from an organ pipe plot concern:

a) What proportion of respondents in the stratum are estimated to have the outcome?
(This is the overall shaded fraction of the plot.)
b) Are there any clusters where no respondents have the outcome of interest?
(These are columns at the right side of the plot with no shaded portion.)
If yes, and if the number of respondents there is reasonably high, then it may be warranted to follow up on the location of those clusters to learn why no one has the outcome. One helpful assessment can be to examine the responses in those clusters to the question: Why hasn't your child received all the vaccinations they should have? Then compare responses in those clusters to those from the other clusters.
c) Is the coverage relatively homogeneous across clusters? That is to say, is the shaded part of each bar relatively the same height or do we observe bars at left that are much higher than those at right? This heterogeneity is an indication of spatial heterogeneity (or *spottiness*) in the outcome. The more homogeneous the outcome, the lower the ICC and the more heterogeneous the outcome, the higher the ICC.

If the user turned on the VCQI option to VCQI_SAVE_OP_PLOT_DATA then each plot will have an accompanying Stata dataset that lists the following quantities for each cluster:

1) Bar number (counts from left-to-right and ranges from 1 up to the number of clusters in the stratum)
2) Stratum ID
3) Stratum Name
4) Cluster ID
5) Cluster Name
6) Number of respondents
7) Bar width (on scale of 0-100%)
8) Cumulative sum of weights from the far left side of the plot through this cluster (ranges from 0-100%)
9) Shaded portion of the bar (ranges from 0-100%)

10) Name of the corresponding organ pipe plot image file
11) Name of the dataset file
12) VCQI indicator that made the plot
13) Value of VCQI global ANALYSIS_COUNTER at the time the plot was made

The order of the rows in the dataset is the same as the left-to-right order of the bars in the plot. The plots contain minimal numeric detail, by design, so the user can focus on overall coverage and its hetero- or homogeneity. All the quantitative details are contained in the corresponding dataset which may be found in the PLOTS_OP sub-folder of the VCQI output folder.

INTERPRETING INCHWORM PLOTS

Weighted VCQI indicators produce inchworm plots to summarize estimated proportions. These plots are described at some length in Annex M of the 2018 WHO Vaccination Coverage Cluster Survey Reference Manual. The figure below shows three very useful 95% confidence intervals for an outcome where the estimated proportion happens to be 50.0%:
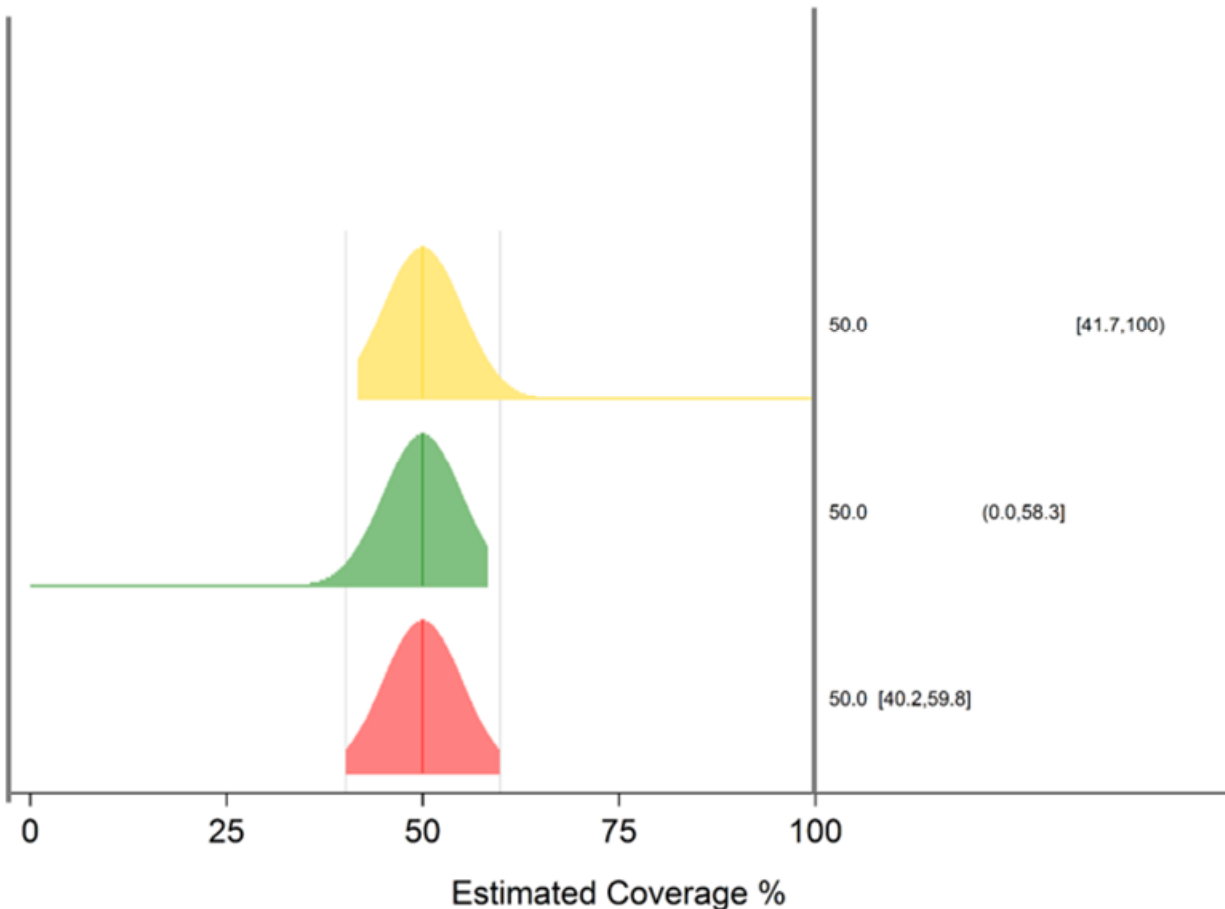
The red shape (bottom) represents the traditional 2-sided 95% confidence interval (CI), where 2.5% of the confidence or probability has been clipped off the left side and 2.5% of probability has been clipped off the right side of the distribution. In this case the base of the distribution represents the 95% confidence interval, extending from 40.2% to 59.8%. The relative height of the distribution from left-to-right represents the relative degree of our confidence that the population parameter lies at that x-value. The distribution is constructed by stacking confidence intervals. The base is the 95% confidence interval. Stacked atop it is the 94% CI. Then the 93% CI. All the way up to the 1% CI.

The green shape (middle) represents the interval defined by 0% at the left edge and the 1-sided 95% upper confidence bound (UCB). In this case, 5% of confidence or probability has been clipped from the right side of the distribution. The true population parameter is just as likely to fall inside the green distribution as it is inside the red. The red interval extends from 0% up to 58.3%.

The yellow shape (top) is the complement to the green one. It extends from the 1-sided 95% lower confidence bound (LCB) up to 100%. Five percent of confidence or probability has been clipped from

the left side. The interval runs from 41.7% to 100%. Again, in the absence of any net bias for this outcome, the true population parameter is as likely to fall in the yellow interval as the green or as the red.

Note that the LCB and UCB are always closer to the point estimate than the upper- and lower-limits of the 2-sided CI.

Rather than plot three distributions for every estimated outcome, VCQI shows the bottom distribution – the one that covers the 2-sided CI and then annotates the distribution with faint tick markings where the 1-sided LCB and UCB would fall, as demonstrated in the figure below. The vertical placement of the tick marks does not carry any meaning. It is only their x-coordinate (along the coverage axis) that is meaningful.

The figure below shows an inchworm with the LCB and UCB tick marks and all three of the notable 95% CIs in the text at right.
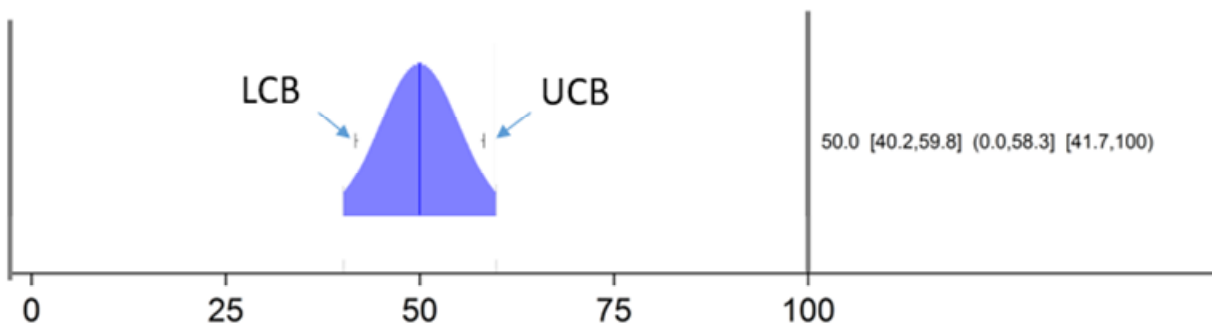
By default, the text at the right side of the plots shows the point estimate and 2-sided CI. The user may ask for other intervals as described in Annex C of the *VCQI User's Guide*.

The most common interpretation of an outcome from an inchworm plot would be: "<Point estimate> percent of eligible respondents are estimated to have the <outcome being summarized>. The 2-sided 95% confidence interval extends from <lower 2-sided CI confidence limit> to <upper limit>%."

If the survey protocol has a goal of using the LCB and UCB to classify coverage, then interpretation might proceed as described in Annex N of the 2018 WHO reference manual: "We are 95% confident that coverage falls above the LCB and 95% confident that coverage falls below the UCB. If an important programmatic coverage threshold falls between the LCB and UCB then this survey does not have an adequate sample for us to be 95% confident whether coverage is above or below that threshold."

VCQI inchworm plot shapes are colored according to which level of stratum they represent. Level 1 strata usually have a shaded bar behind the results to draw the viewer's eye and help set those results of from those of levels 2, 3 or 4. The user may change the colors as described in Annex C of the *VCQI User's Guide*.

The figure below shows a typical VCQI inchworm plot. Distributions for levels 1 and 2 are dark blue while those for level 3 strata are lighter. The level 1 stratum has a shaded band behind it to draw the reader's eye to those national level results. In this case level 3 strata are nested within level 2 strata, so all the districts for the Southern Province appear in the top half of the figure and those for the Northern Province appear in the bottom half of the figure. Within each province, the districts are sorted bottom-to-top by coverage (low-to-high). Rather than have strata sorted by outcomes, the user may also request that the strata appear in inchworm plots in the same order they appear in VCQI tables. The user may also request that outcomes be sorted bottom-to-top from highest-to-lowest estimated coverage. See Annex C of the *VCQI User's Guide* to learn how to customize inchworm plots.

TT - Protected at Birth

Stratum	Estimate
District 05	82.2% (77.7,86.0)
District 02	82.2% (75.9,87.1)
District 01	82.1% (77.4,86.0)
Southern Province	81.8% (79.6,83.8)
District 03	81.6% (74.4,87.1)
District 04	81.1% (76.0,85.3)
Harmonia	81.5% (79.9,83.0)
District 10	85.8% (81.6,89.1)
District 07	84.0% (79.3,87.7)
District 06	83.7% (78.3,88.0)
District 09	81.9% (76.9,85.9)
Northern Province	81.2% (78.7,83.4)
District 08	77.5% (72.0,82.1)

Estimated Coverage %

Text at right: Point Estimate (2-sided 95% Confidence Interval)

INTERPRETING DOUBLE INCHWORM PLOTS

Several VCQI indicators produce outcomes that are meant to be compared with those from other indicators. For example, it makes sense to compare valid coverage with crude coverage (RI_COVG_02 and RI_COVG_01), or the proportion that had a valid dose with the proportion who would have had a valid dose if there had been no missed opportunities for simultaneous vaccination (RI_COVG_02 and RI_COVG_07B). These indicators produce single inchworm plots for their outcomes and conclude by producing so-called *double inchworm plots* where each row shows two inchworm shapes. One is shaded and represents the new outcome. The other has a gray outline and is hollow, and that usually represents the previously calculated comparator indicator. The figure below shows how crude coverage (hollow shapes) compares with valid coverage by age 1 (colored shapes). The text at right lists the point estimates for the two outcomes. If the reader wants to also see the 95% confidence intervals, the VCQI user may request that those be listed on the figure (via the VCQI_DOUBLE_IWPLOT_CITEXT input parameter – see Annex C of the *VCQI User's Guide*) or may find them in the appropriate Excel output tabs. VCQI suppresses the LCB and UCB tick marks for double-inchworm plots because they can be very cluttery when the distributions are close together. As with single inchworm plots, the user has control over whether strata are listed in order of observed outcome or in order that the strata appear in tables. If sorted by outcome, the user has control over whether the figures sort bottom-to-top from high-to-low or from low-to-high outcomes. See Annex C of the *VCQI User's Guide*.

RI - Valid Coverage by Age 1 of MCV1

Southern Province	41.3 | 67.0
District 01	35.6 | 58.7
District 02	41.5 | 72.1
District 03	46.1 | 72.9
District 04	35.6 | 58.4
District 05	52.2 | 81.7
Northern Province	57.1 | 89.9
District 06	55.7 | 86.9
District 07	55.4 | 84.5
District 08	57.2 | 91.4
District 09	62.4 | 96.2
District 10	51.8 | 89.9
Harmonia	49.3 | 78.5
0: Rural	47.6 | 74.8
1: Urban	50.3 | 81.0

Estimated Coverage %
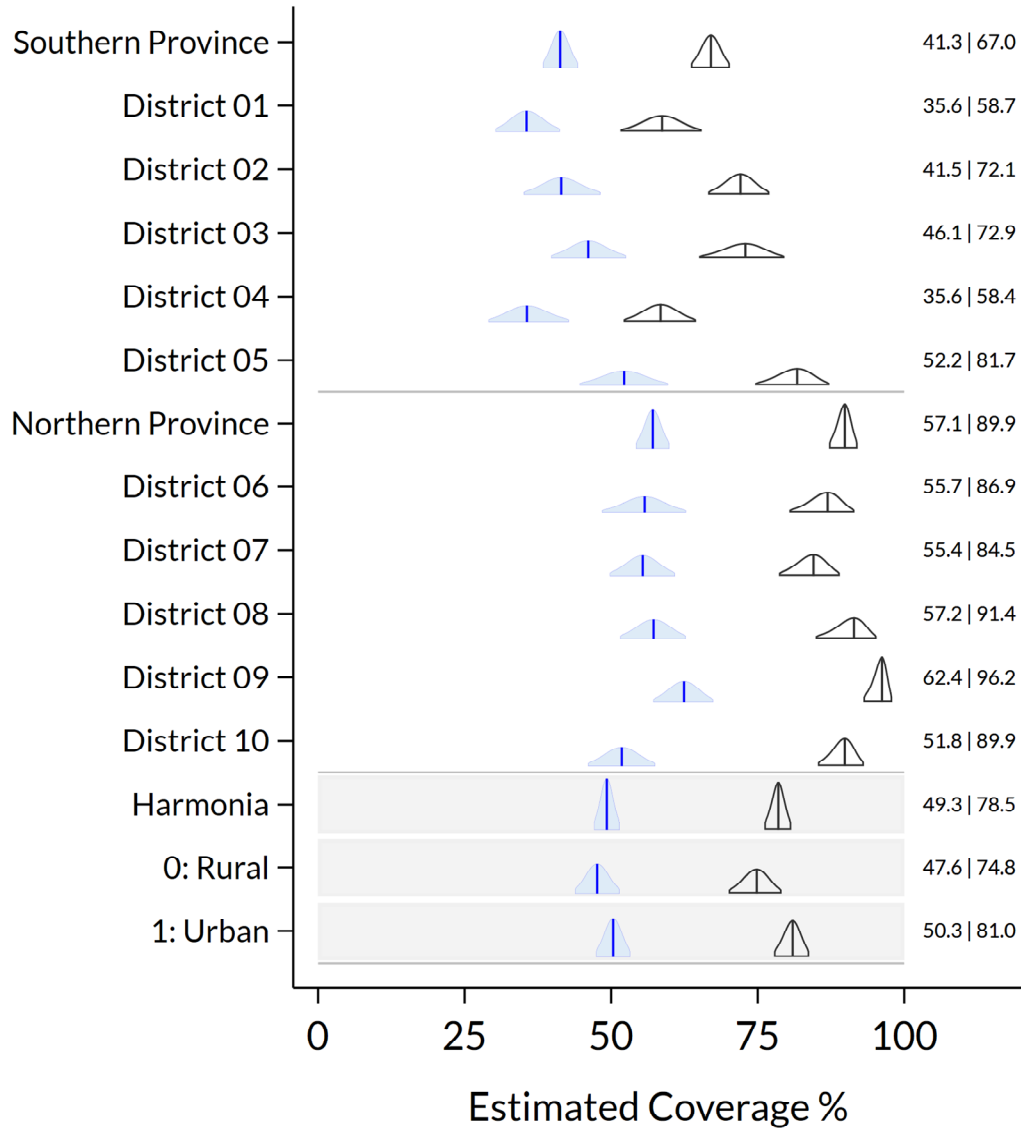
Text at right: Point Estimates from colored and from gray hollow distributions
Gray hollow shape is crude coverage; colored shape is valid coverage

Interpretation can be as simple as "While the estimated proportion of respondents with <outcome 1> in <stratum of interest> was X1%, for <outcome 2> it was X2%" Recall that if the two 95% confidence intervals do not overlap (as is the case for every row of the figure above) then we can be 95% confident that the population parameter for one outcome is higher than the other. If the two distributions do overlap slightly then we cannot draw a confident conclusion using an eyeball test: it would be necessary to calculate a formal p-value. This is possible using VCQI's COVG_DIFF indicators. See the *VCQI User's Guide*.

INTERPRETING COVERAGE BAR CHARTS

The VCQI user can request (via the input parameter named IWPLOT_SHOWBARS) that VCQI make bar charts instead of inchworm plots. In that case, every inchworm plot will be replaced with a horizontal bar chart with a simple 1-dimensional representation of the 2-sided 95% confidence interval, as shown below. For single bar charts with one row per stratum, the 2-sided 95% confidence interval is shown with a horizontal line and the line is punctuated with small vertical tick marks at the value of the 1-sided 95% LCB and UCB.

TT - Protected at Birth

Stratum	Estimate
District 05	82.2% (77.7,86.0)
District 02	82.2% (75.9,87.1)
District 01	82.1% (77.4,86.0)
Southern Province	81.8% (79.6,83.8)
District 03	81.6% (74.4,87.1)
District 04	81.1% (76.0,85.3)
Harmonia	81.5% (79.9,83.0)
District 10	85.8% (81.6,89.1)
District 07	84.0% (79.3,87.7)
District 06	83.7% (78.3,88.0)
District 09	81.9% (76.9,85.9)
Northern Province	81.2% (78.7,83.4)
District 08	77.5% (72.0,82.1)

Estimated Coverage %
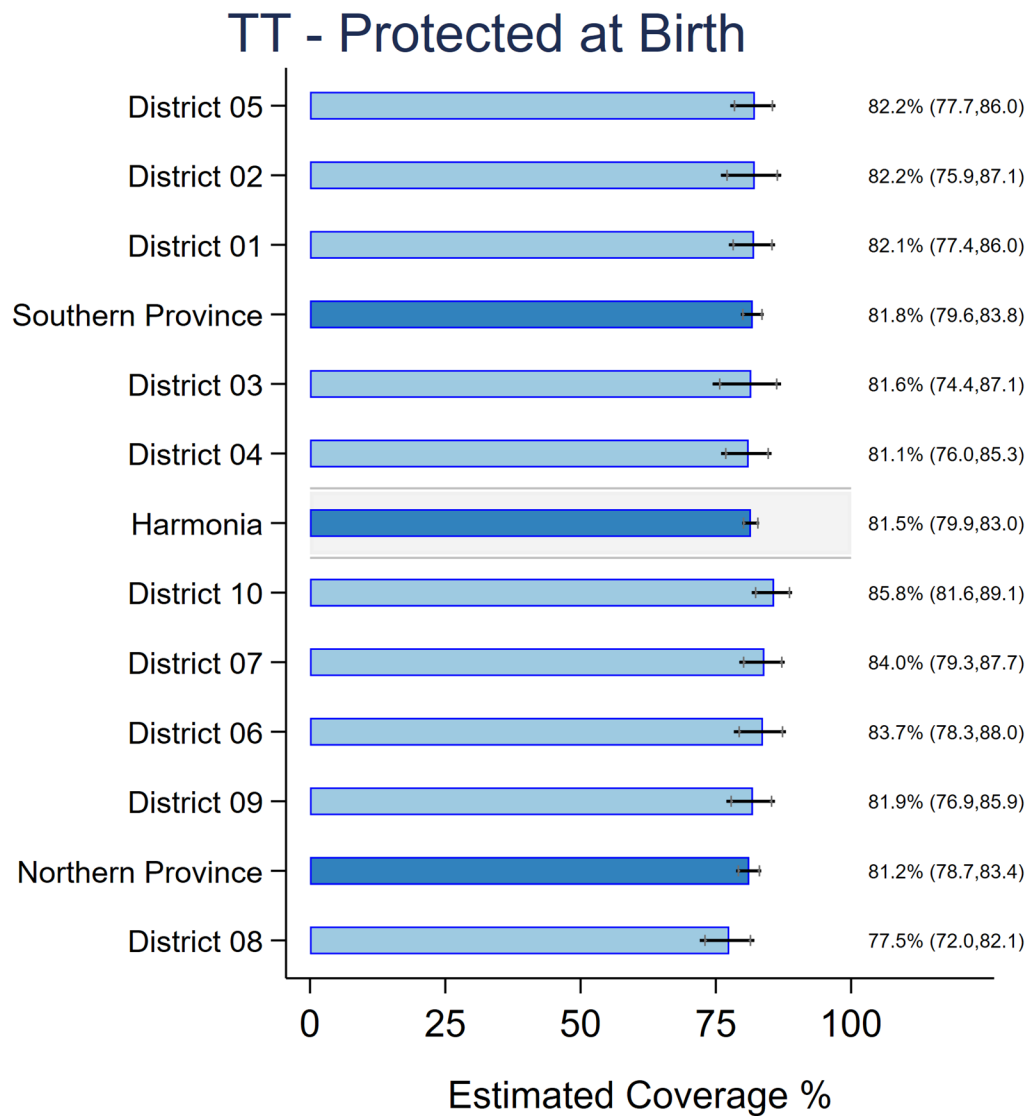
Text at right: Point Estimate (2-sided 95% Confidence Interval)

INTERPRETING DOUBLE BAR CHARTS

When the user requests bar charts, VCQI produces double bar charts instead of double inchworm plots. The newer outcome is typically shown in shaded bars and the contrasting previous comparator outcome is shown with gray unfilled bars. Each bar shows a 2-sided confidence interval with a horizontal line and 1-sided LCB and UCB ticks from the single bar charts are omitted in double bar charts. See *VCQI User's Guide* Annex C for guidance on customizing these figures.

RI - Would have Valid MCV1 if no MOVs (%)

District	Value
District 09	88.0 | 62.4
Northern Province	82.1 | 57.1
District 08	81.8 | 57.2
District 06	81.4 | 55.7
District 07	80.2 | 55.4
District 10	77.5 | 51.8
Harmonia	75.9 | 49.3
District 03	77.0 | 46.1
District 05	76.9 | 52.2
District 02	72.3 | 41.5
Southern Province	69.7 | 41.3
District 04	66.4 | 35.6
District 01	62.5 | 35.6

Estimated Coverage %
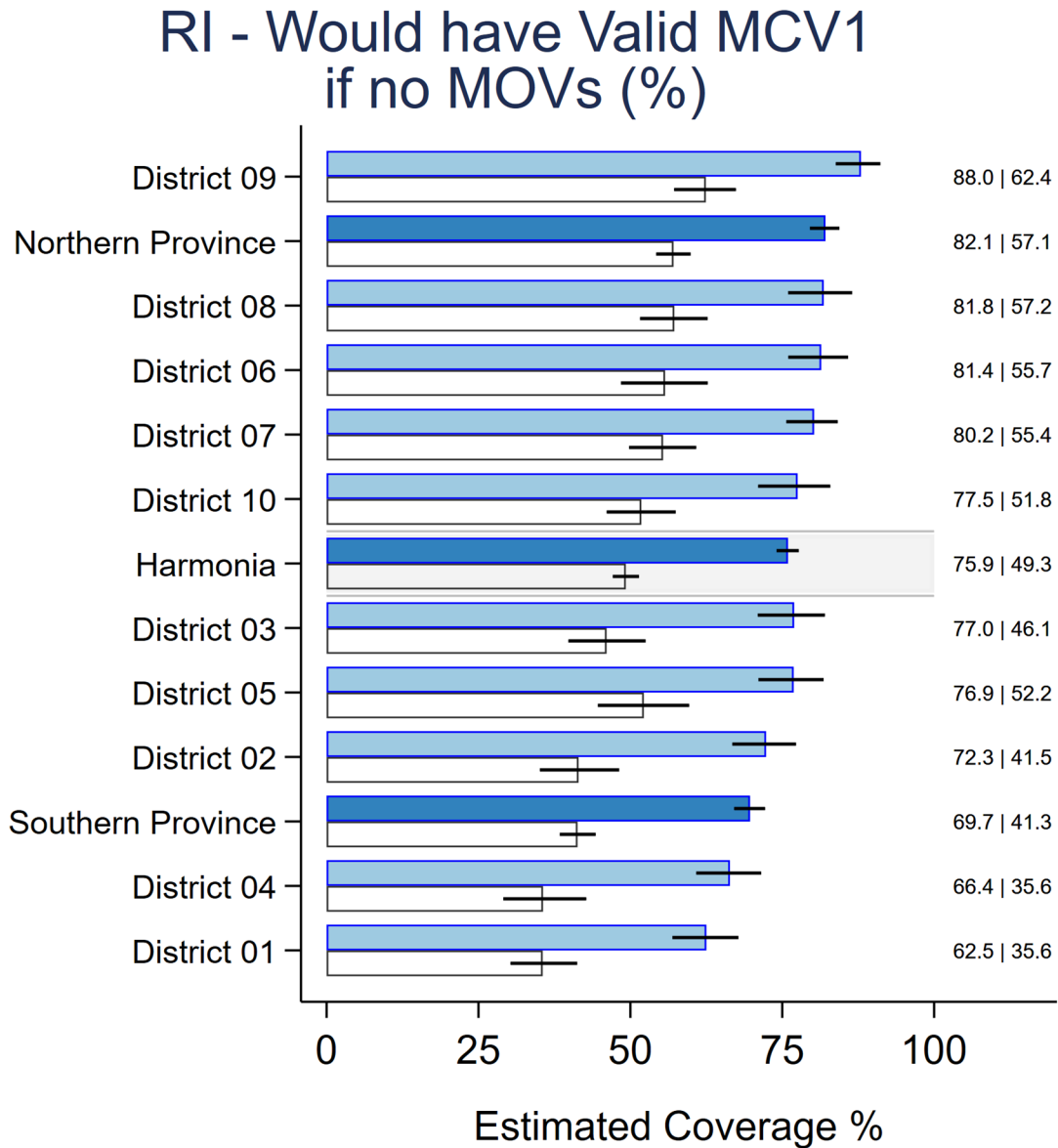
Text at right: Point estimates from colored and from gray hollow bars

Gray hollow bar shows valid coverage; colored bar shows valid coverage if no MOVs

INTERPRETING UNWEIGHTED PROPORTION PLOTS

Many of VCQI's unweighted indicators make so-called *unweighted proportion plots* that simply show the estimated proportion using a small triangle. There is no estimate of a confidence interval for these unweighted outcomes. The figure lists the estimated proportion and the value of N at the right side of the plot. Unweighted proportion plots use a symbol coloring scheme that is similar to inchworm plots: Level 1 usually has a gray stripe behind the results to draw the viewer's eye. The triangle for each stratum is the same size as every other triangle. Again, the user may use the PLOT_OUTCOMES_IN_TABLE_ORDER parameter to request that strata be listed in the same order as in VCQI tables. Or the user may use the SORT_PLOT_LOW_TO_HIGH parameter to say whether strata are sorted from low-to-high or high-to-low outcome values. See Annex C of the *VCQI User's Guide*.

The figure below shows, out of the total vaccination visits where the respondent was age eligible for MCV1 (N), the proportion who did not receive it (%). See the sections above for the individual indicators to learn how to describe the interpretation in a meaningful sentence.

RI - Visits with MOV for MCV1

Stratum	Sample Proportion %	
Harmonia	△	46.7% N= 3,006
Southern Province	△	51.3% N= 1,458
District 01	△	55.0% N= 278
District 02	△	53.2% N= 331
District 03	△	50.3% N= 306
District 04	△	54.3% N= 232
District 05	△	44.7% N= 311
Northern Province	△	42.3% N= 1,548
District 06	△	44.4% N= 304
District 07	△	43.6% N= 330
District 08	△	41.2% N= 291
District 09	△	41.1% N= 365
District 10	△	41.1% N= 258
Harmonia	△	46.7% N= 3,006
0: Rural	△	48.8% N= 1,147
1: Urban	△	45.3% N= 1,859

Sample Proportion %

Text at right: Unweighted sample proportion (%) and N

ROUTINE IMMUNIZATION DATE DATA QUALITY REPORT

The user may request an Excel data quality report that quantifies the number (and %) of dates in the dataset that are concerning for one reason or another. This annex describes how to request the data quality report, and how to interpret its contents.

HOW TO REQUEST THE RI DATE DATA QUALITY REPORT

Block D of the control program includes this section:

```
* For RI analysis, there is an optional report on data quality
* Set this global to 1 to generate that report
* It appears in its own separate Excel file

vcqi_global VCQI_REPORT_DATA_QUALITY        0
```

If the global is set to 0, VCQI skips the report, if it is set to 1, VCQI produces a date data quality report in a separate Excel file named <VCQI_ANALYSIS_NAME>_dates_ticks.xlsx. (Recall that the VCQI_ANALYSIS_NAME is set by the user in Block B of the control program.)

The data quality report uses some of the user-inputs: the Block D list of doses to be included in the analysis, the indication of whether vaccination evidence was sought at health centers, and the earliest and latest valid dates of vaccination for children in this dataset.

The data quality holds worksheets of five different kinds. Each is described briefly here with examples and additional detail below.

1. DOB: The child's date of birth (DOB) can be provided in up to three (possibly contradictory) ways: via the caregiver's recall (history), via the home-based vaccination record (card), or via the facility-based vaccination register (register). For each respondent, VCQI examines all three sources and identifies which DOB to use for all its timeliness calculations. The first tab of the data quality report summarizes how many DOBs were specified using the three sets of input variables, and which methods were used to successfully assign the DOB for analysis purposes for which portions of the children.
2. Dose dates: The report includes a separate worksheet for each dose in the dataset. Each worksheet summarizes the number of respondents who provided vaccination evidence for that dose via each of the three sources (history, card, register) and documents the % of dates that were incomplete, nonsensical, or outside the allowable range. In cases where evidence comes from more than one source, the sheet also summarizes the % whose evident is concordant across sources, and if discordant, how so.
3. Dose series: Each 2- and 3-dose series has its own worksheet to document the number (and %) of children where dates in the series are either out-of-order or identical.
4. Dose summary: The worksheet named 'Dose Summary' summarizes the number of dates provided across all three sources and what portion of them had data quality problems. It also briefly summarizes the evidence of concordance across sources for each dose.
5. Dates VCQI changed to ticks: VCQI engages in a data cleaning step where problematic dates are removed from the dataset and replaced with tick mark evidence. The child still receives credit for receiving a crude dose but VCQI refrains from analyzing anything about the timeliness of doses with dates that are out of range for this survey or dates that are out-of-order within a dose series. The final tab documents the number of dose dates that were changed to ticks after failing one or more data quality checks.

DOB Worksheet

The DOB worksheet answers the following questions:

1. How many (and what %) respondents provided DOB via card, register, and history (recall)?
2. How many DOBs were full dates, and how many were partially missing?
3. Among the full dates, how many were nonsensical (not real dates)?
4. How many were too early to be correct? Too late?
5. How many fell within the proper range for this survey?
6. Among children who had DOB data from more than one source, how many had dates that disagreed?
7. When VCQI selects a DOB to use for timeliness calculations, how many respondents have one selected?
8. How many have a DOB selected using the *earliest birthdate* method? The *single birthdate* method?

DOB: Present	Card n	Card %	Card denom	Register n	Register %	Register denom	History n	History %	History denom	group
How many do not have data from source?	783	24%	3256	0	0%	3256	0	0%	3256	P1
How many have data from source?	2473	76%	3256	3256	100%	3256	3256	100%	3256	P1
How many missing?	0	0%	2473	0	0%	3256	0	0%	3256	P2
How many partially missing?	0	0%	2473	0	0%	3256	0	0%	3256	P2
How many full dates?	2473	100%	2473	3256	100%	3256	3256	100%	3256	P2
How many full, but nonsensical dates?	0	0%	2473	0	0%	3256	0	0%	3256	P3
How many full and sensible dates?	2473	100%	2473	3256	100%	3256	3256	100%	3256	P3
How many sensible dates prior to earliest possible date?	0	0%	2473	0	0%	3256	0	0%	3256	P4
How many sensible dates past survey date?	0	0%	2473	0	0%	3256	0	0%	3256	P4
How many sensible dates within proper range?	2473	100%	2473	3256	100%	3256	3256	100%	3256	P4

Worksheets in the data quality report have several characteristics in common:

- The far right column is named 'group' and it indicates which rows use the same denominator. In the example above:
    - The first two rows (group P1) use the same denominator: All respondents
    - The next three rows (group P2) use the same denominator: Those with data from this source
    - The next two rows (group P3) use the same denominator: Those with full dates
    - The final three rows (group P4) use the same denominator: Those with full and sensible dates
- Within each denominator group, the figures in the n column must sum to the denominator and the figures in the % column must sum to 100%.
- Non-zero numbers in cells that are shaded pink indicate a problem with data quality or concordance of vaccination evidence.

Interpreting the DOB worksheet

- The first two rows show the number and % of respondents who did and did not provide DOB information via card, register, and recall (history).
- The next three rows show the number and % of respondents whose DOB data are missing, partially missing, or include full dates.
- The next two rows show the number and % of full dates that are sensible (correspond to a real calendar date) and are nonsensical (do not correspond to a real date, e.g., February 30 or September 31).
- The final three rows document the number and % of sensible dates that are too early (fall prior to the earliest possible valid date for this dataset), too late (fall after the final survey interview was conducted) or fall within the appropriate range.

The DOB: Present table is followed by a table that summarizes DOB discordance between pairs of sources.

Date discordance: Full date DOBs disagree	n	%	denom
Card/register?	0	0%	2473
Card/history?	0	0%	2473
Register/history?	0	0%	3256

Note: If the pink cells hold non-zero numbers, that is an indication of discordance between sources.
Each respondent is assigned a DOB to be used in valid dose calculations and vaccination timeliness calculations. VCQI uses two methods to assign that DOB. If all the DOB sources agree (are concordant) or if only one of the sources yields a full triplet of month, day, and year, then VCQI simply uses that so-called *single birthdate* as the DOB for timeliness calculations.

In some cases the DOBs will not agree. Perhaps the caregiver lists the DOB as January 21, 2013 and the home-based record (card) lists it as January 12, 2013. In the case of discordant dates, VCQI assigns the earliest of the candidate dates (January 12, 2013 in this case) to be the DOB for timeliness calculations. (Technically, VCQI assigns the earliest of the candidate dates that falls in the range of valid dates for this survey. If the earliest possible vaccination date for this survey were January 15, 2013, then VCQI would assign January 21, 2013 as the DOB for timeliness calculations for the respondent in question. The range of valid dates is specified by the user in Block D of the control program.) This method of assignment is called the *earliest birthdate* method.

The discordance table is followed by a table describing the number and proportion of respondents who have a DOB assigned and the number who have it assigned via the *earliest birthdate* versus the *single birthdate* method.

Assigned for timeliness calculations	n	%	denom
How many DOBs not assigned?	0	0%	3256
How many DOBs assigned?	3256	100%	3256
How many assigned using the earliest birthdate?	0	0%	3256
How many assigned using the single birthdate?	3256	100%	3256

Note: If card, register & history yield 2+ birthdates then VCQI assigned the earliest plausible birthdate for timeliness calculations.

DOSE WORKSHEET

The dose worksheet answers the following questions:

1. How many (and what %) respondents showed evidence of receiving this dose from card, register, and recall (history)?
2. Among evidence in the form of dates, how many were full dates and how many were partially missing?
3. Among full dates, how many were sensible and how many were nonsensical?
4. Among sensible dates, how many were too early, too late, and within the proper range?
5. Among respondents who furnished evidence for this dose from more than one source:
   a. How many children had discordant evidence (some sources said they got it while other sources did not)
   b. How many children had multiple sources confirm that they received the dose?
   c. Agreed with perfect concordance?
   d. Agreed with various small disagreements about the details of the date?

The rows in the <DOSE>: Present table are similar to those in the DOB Present table, with the exception of the final row. This row indicates the number (and %) of respondents whose evidence was from a simple tick mark for card or register, and the number and % whose caregiver said they received the dose (History).

The <DOSE>: Present table is followed by up to three tables that document concordance and discordance of evidence for this dose across card, register, and recall (history). Again, data in pink cells indicates a data quality problem and rows within a single group have values of n that sum to the denominator and values of % that sum to 100%.

BCG: Present	Card			Register			History			group
	n	%	denom	n	%	denom	n	%	denom	
How many do not have data from source?	783	24%	3256	0	0%	3256	0	0%	3256	P1
How many have data from source?	2473	76%	3256	3256	100%	3256	3256	100%	3256	P1
How many missing?	880	36%	2473	854	26%	3256	0	0%	0	P2
How many partially missing?	0	0%	2473	0	0%	3256	0	0%	0	P2
How many full dates?	1593	64%	2473	2402	74%	3256	0	0%	0	P2
How many full, but nonsensical dates?	0	0%	1593	0	0%	2402	0	0%	0	P3
How many full and sensible dates?	1593	100%	1593	2402	100%	2402	0	0%	0	P3
How many sensible dates prior to earliest possible date?	41	3%	1593	16	1%	2402	0	0%	0	P4
How many sensible dates past survey date?	0	0%	1593	0	0%	2402	0	0%	0	P4
How many sensible dates within proper range?	1552	97%	1593	2386	99%	2402	0	0%	0	P4
Tick mark: Yes (for Card & Register) or History: Yes (for History)	155	6%	2473	56	2%	3256	1940	60%	3256	

Note: BCG Scar data is excluded from these tables

Note: Nonsensical dates refer to participants who had all three date components that did not result a calendar date... eg 2/29/2012, 6/31/2014, 14/2/0214 etc.

Note: Sensible dates refer to participants who had all three date components that resulted in a true calendar date... eg 2/28/2012, 6/1/2014 etc.

Note: If the pink cells hold non-zero numbers, that is an indication of incomplete or nonsensical dates.

Note: Within each group, the numbers in the n column must sum up to the denominator and the % numbers add up to 100%. The rows within each group are mutually exclusive and exhaustive.

BCG: Concordance of evidence				
Card & Register	n	%	denom	group
1. Have neither card or register	0	0%	3256	CR1
2. Have card but no register	0	0%	3256	CR1
3. Have register but no card	783	24%	3256	CR1
4. Have both card and register	2473	76%	3256	CR1
4.1. Card&Register agree: did not get it	631	26%	2473	CR2
4.2. Card&Register agree: got it	1748	71%	2473	CR2
4.2.1. Card&Register agree: perfect agreement on date or tick	1564	89%	1748	CR3
4.2.2. Card&Register agree: but some discordance about evidence	184	11%	1748	CR3
4.2.2.1. both have full dates that disagree	74	40%	184	CR4
4.2.2.2. both have partial dates that disagree	0	0%	184	CR4
4.2.2.3. one full and one partial date	0	0%	184	CR4
4.2.2.4. one date (full or partial) and one tick	110	60%	184	CR4
4.3. Card&Register disagree:One got it, Other did not	94	4%	2473	CR2
4.3.1 Card got it, Register did not	0	0%	94	CR5
4.3.2 Card did not get it, Register did	94	100%	94	CR5

Note: If the pink cells hold non-zero numbers, that is an indication of discordance between sources.

Card & History	n	%	denom	group
1. Have neither card or history	0	0%	3256	CH1
2. Have card but no history	0	0%	3256	CH1
3. Have history but no card	783	24%	3256	CH1
4. Have both card and history	2473	76%	3256	CH1
4.1. Card&History agree: did not get it	542	22%	2473	CH2
4.2. Card&History agree: got it	1276	52%	2473	CH2
4.3. Card&History one got it, other not	655	26%	2473	CH2
4.3.1 Card got it, history did not	472	72%	655	CH3
4.3.2 History got it, card did not	183	28%	655	CH3

Note: If the pink cells hold non-zero numbers, that is an indication of discordance between sources.

Register & History	n	%	denom	group
1. Have neither reg or history	0	0%	3256	RH1
2. Have reg but no history	0	0%	3256	RH1
3. Have history but no reg	0	0%	3256	RH1
4. Have both reg and history	3256	100%	3256	RH1
4.1. Register&History agree: did not get it	660	20%	3256	RH2
4.2. Register&History agree: got it	1802	55%	3256	RH2
4.3. Register&History one got it, other not	794	24%	3256	RH2
4.3.1 Register got it, history did not	656	83%	794	RH3
4.3.2 History got it, register did not	138	17%	794	RH3

Note: If the pink cells hold non-zero numbers, that is an indication of discordance between sources.

DOSE SERIES WORKSHEET

Data in the dose series worksheet addresses these questions:

1. How many (and what %) respondents who have dates for 2 or 3 doses on the card or register have dates that
   a. Are the same for an earlier and later dose?
   b. Show the later dose occurring before the earlier dose?
   c. Show the doses occurring in the correct order?

As with the other sheets, non-zero numbers in pink cells indicate a problem with data quality and in rows in the same group, values in the n column must sum to the denominator and values in the % column sum to 100%.

PENTA Series: Quality of date order for 2 or more doses	Card			Register			
Penta with only first and second doses in series	n	%	denom	n	%	denom	group
How many where only the 1st and 2nd dose in series are within proper range?	690	28%	2473	787	24%	3256	P1
How many with only 1st and 2nd doses in series within range have the same date?	0	0%	690	0	0%	787	P2
How many with only 1st and 2nd doses in series within range have different dates?	690	100%	690	787	100%	787	P2
How many with only 1st and 2nd doses in series with different dates have 2nd dose before 1st?	5	1%	690	5	1%	787	P3
How many with only 1st and 2nd dose in series with different dates are in correct chronological order?	685	99%	690	782	99%	787	P3

Penta with only 1st and 3rd doses in series	n	%	denom	n	%	denom	group
How many where only the 1st and 3rd dose in series are within proper range?	87	4%	2473	44	1%	3256	P4
How many with only 1st and 3rd doses in series within range have the same date?	0	0%	87	0	0%	44	P5
How many with only 1st and 3rd doses in series within range have different dates?	87	100%	87	44	100%	44	P5
How many with only 1st and 3rd doses in series with different dates have 2nd dose before 1st?	0	0%	87	0	0%	44	P6
How many with only 1st and 3rd dose in series with different dates are in correct chronological order?	87	100%	87	44	100%	44	P6

Penta with only 2nd and 3rd doses in series	n	%	denom	n	%	denom	group
How many where only the 2nd and 3rd dose in series are within proper range?	63	3%	2473	27	1%	3256	P7
How many with only 2nd and 3rd doses in series within range have the same date?	0	0%	63	1	4%	27	P8
How many with only 2nd and 3rd doses in series within range have different dates?	63	100%	63	26	96%	27	P8
How many with only 2nd and 3rd doses in series with different dates have 2nd dose before 1st?	0	0%	63	0	0%	26	P9
How many with only 2nd and 3rd dose in series with different dates are in correct chronological order?	63	100%	63	26	100%	26	P9

47

Penta with all 3 doses in series	n	%	denom	n	%	denom	group
How many where all 3 doses in series are within proper range?	1015	41%	2473	1876	58%	3256	P10
How many where all 3 doses in series are within range and have the same date?	0	0%	1015	0	0%	1876	P11
How many where all 3 doses in series are within range and 1st and 2nd dose have the same date?	0	0%	1015	0	0%	1876	P11
How many where all 3 doses in series are within range and 1st and 3rd dose have the same date?	0	0%	1015	0	0%	1876	P11
How many where all 3 doses in series are within range and 2nd and 3rd dose have the same date?	0	0%	1015	0	0%	1876	P11
How many where all 3 doses in series are within range and have different dates?	1015	100%	1015	1876	100%	1876	P11
How many where all 3 doses in series are within range and have different dates and 2nd dose is before 1st?	16	2%	1015	23	1%	1876	P12
How many where all 3 doses in series are within range and have different dates and 3rd dose is before 2nd?	11	1%	1015	22	1%	1876	P12
How many where all 3 doses in series are within range and have different dates and 3rd dose is before 1st?	13	1%	1015	16	1%	1876	P12
How many where all 3 doses in series are within range with different dates and in chronological order?	988	97%	1015	1831	98%	1876	P12

DOSE SUMMARY WORKSHEET

A single Dose Summary worksheet is the final (right-most) worksheet in the date data quality report. This is the single table that gives an overview of date data quality for vaccination evidence. Data in the Dose Summary address the following questions:

1. Aggregating across all doses, how many possible doses have evidence from card, register, and recall (history)?
2. How many doses (and what %) have evidence from dates, from tick marks, and from caregiver recall?
3. What portion of the dates are full and what portion are partial dates?
4. Among full dates, how many are sensible vs. nonsensical?
5. Among sensible dates, how many fall too early, too late, and within the proper range?
6. Among respondents with evidence from more than one source, for how many is the evidence concordant? Discordant?

As with the other worksheets, non-zero numbers in pink cells indicate problems with data quality or discordance. Within each group, the numbers in the n column must sum to the denominator and those in the % column must sum to 100%.

Doses: Present	Card			Register			History			
	n	%	denom	n	%	denom	n	%	denom	group
How many do not have data from source?	783	24%	3256	0	0%	3256	0	0%	3256	P1
How many have data from source?	2473	76%	3256	3256	100%	3256	3256	100%	3256	P1
How many missing?	15379	35%	44514	13989	24%	58608	0		0	P2
How many partially missing?	0	0%	44514	0	0%	58608	0		0	P2
How many full dates?	29135	65%	44514	44619	76%	58608	0		0	P2
How many full, but nonsensical dates?	0	0%	29135	0	0%	44619	0		0	P3
How many full and sensible dates?	29135	100%	29135	44619	100%	44619	0		0	P3
How many sensible dates prior to earliest possible date?	132	0%	29135	60	0%	44619	0		0	P4
How many sensible dates past survey date?	40	0%	29135	57	0%	44619	0		0	P4
How many sensible dates within proper range?	28963	99%	29135	44502	100%	44619	0		0	P4
Tick mark: Yes (for Card & Register) or History: Yes (for History)	3100	7%	44514	930	2%	58608	30645	52%	58608	

In this table, n=2,473 respondents showed cards that held 29,135 vaccination dates. They were all full and sensible dates, but 132 were earlier than the earliest proper date for this dataset and 40 were later than the latest proper date. Most (28,963) fell in the proper range for this survey. Furthermore, 3,100 doses on cards were recorded with a tick mark instead of a date.

According to register records, n=3,256 respondents provided data on 44,619 full and sensible dates. A subset of 60 dates were too early and 57 were too late and the rest fell within the proper date range for this survey. There were 930 tick marks on the register records.

According to caregiver recall, 30,645 doses were received by the children studied in this survey.

Beneath the Doses: Present table is a large table summarizing high-level concordance and discordance for each dose. To see the details of date discordance between card and register it is necessary to visit the worksheet for that specific dose.

All Dates: Concordance of evidence								
Dose Numbers: Comparison between sources								
BCG	**Card & Register**			**Card & History**			**Register & History**	
	Card Yes	Card No		Card Yes	Card No		Register Yes	Register No
Register Yes	1748	94	History Yes	1276	183	History Yes	1802	138
Register No	0	631	History No	472	542	History No	656	660
BCG Scar data is excluded from these tables.								
Note: If the pink cells hold non-zero numbers, that is an indication of discordance between sources.								

In order to furnish a reasonably clean dataset to the code that calculates VCQI indicators, VCQI replaces partial dates, too early dates, too late dates, and dates in a series that are equal or out-of-order. In each case, the problematic date evidence is replaced with a tick mark. The user may consult the VCQI log to learn how many vaccination evidence dates were converted to ticks. The log comments that summarize tick mark substitutions appear with the keyword 'Data'.

If the vaccination evidence includes photographs of cards or register records, the data quality report might serve as the start of a process of data cleaning. If the report indicates a concerningly high number of records with incomplete or nonsensical dates, or dates that are too early or late, then it may be worthwhile to make a list of respondents who show those problems and pull the photos of their cards or register records and decide whether it is possible to clean up the input data before conducting the final analytic run of VCQI code

DATES VCQI CHANGED TO TICKS WORKSHEET

The final worksheet lists each dose and documents the number of dates that were changed to ticks and why. If the survey captured photos of home-based vaccination records or if data were originally collected on paper forms then it may be worthwhile to check the original data and learn whether any vaccination dates were entered incorrectly into the survey database. Follow instructions in Annex G of the *VCQI User's Guide* If so, make corrections to the RI or RIHC dataset and run VCQI again. It is a best practice for the survey report to document the number of dates that VCQI changed to ticks. You may copy the tables from this worksheet into the survey report.

Number of Dates VCQI Changed to Tick by Dose and Reason	Partial Dates	Nonsensical Dates	Card Dates Before Earliest Possible Date	Dates Past Survey Date	Dates Out of Order Within Series	Same Dates Within Series	To be consistent with evidence of a later dose
BCG	0	0	41	0			
HEPB	0	0	47	0			
IPV	0	0	0	7			
MCV1	0	0	0	0			
OPV0	0	0	44	0			
OPV1	0	0	0	0	16	0	0
OPV2	0	0	0	4	19	0	0
OPV3	0	0	0	9	13	0	
PCV1	0	0	0	0	4	0	1
PCV2	0	0	0	3	4	0	0
PCV3	0	0	0	1	0	0	
PENTA1	0	0	0	0	22	0	1
PENTA2	0	0	0	3	32	0	0
PENTA3	0	0	0	6	23	0	
ROTA1	0	0	0	0	2	0	0
ROTA2	0	0	0	3	12	0	0
ROTA3	0	0	0	4	10	0	
YF	0	0	0	0			

Note: These are the details around why VCQI changed dates to tick marks for each dose.
Note: For multi dose series the second dose is also changed to a tick if dose 1 and 3 are out of order and dose 2 is within range.

Number of Dates VCQI Changed to Tick by Dose and Reason	Register						
	Partial Dates	Nonsensical Dates	Dates Before Earliest Possible Date	Dates Past Survey Date	Dates Out of Order Within Series	Same Dates Within Series	To be consistent with evidence of a later dose
BCG	0	0	16	0			
HEPB	0	0	22	0			
IPV	0	0	0	11			
MCV1	0	0	0	0			
OPV0	0	0	22	0			
OPV1	0	0	0	0	26	0	0
OPV2	0	0	0	5	42	0	0
OPV3	0	0	0	11	29	0	
PCV1	0	0	0	0	8	0	0
PCV2	0	0	0	4	7	0	0
PCV3	0	0	0	1	5	0	
PENTA1	0	0	0	0	29	0	0
PENTA2	0	0	0	4	50	1	0
PENTA3	0	0	0	8	37	1	
ROTA1	0	0	0	0	2	0	0
ROTA2	0	0	0	4	17	0	0
ROTA3	0	0	0	9	15	0	
YF	0	0	0	0			

Note: These are the details around why VCQI changed dates to tick marks for each dose.
Note: For multi dose series the second dose is also changed to a tick if dose 1 and 3 are out of order and dose 2 is within range.