



R User's Guide

Draft Version 1.0
December 2022



Expanded
Programme on
Immunization

This page is intentionally blank.

Table of Contents

Acknowledgements.....	1
Chapter 0. Preliminary Material	2
Document Revision History.....	2
R Software is Under Development.....	2
Our R Nomenclature – what to call user-defined input parameters.....	2
Acronym List	3
License Agreement.....	4
Chapter 1. Introduction	6
1.1 VCQI Resources	7
1.2 Differences between the Stata and R versions.....	7
Chapter 2. Overview	9
2.1 Running VCQI	9
2.2 The Indicator Types.....	9
2.3 Files that comprise VCQI	10
2.4 Files Used by VCQI.....	11
2.5 Files Produced by VCQI	11
2.6 Levels of Survey Strata	13
2.7 Program Progress Log	13
2.8 Structure of VCQI Control Programs	13
2.9 Specifying VCQI Input Parameters.....	14
Chapter 3. Analyzing Routine Immunization (RI) Surveys	16
3.1 Dataset File Names (goes in Block D).....	16
3.2 Vaccination Schedule Metadata (goes in Block D).....	17
3.3 Survey Metadata (goes in Block D)	19
3.4 Analysis Metadata and Options (goes in Block D)	22
Chapter 4. Analyzing Tetanus Protection at Birth (TT) Surveys	24
Chapter 5. Analysis of Post-Campaign (SIA) Surveys	25
Chapter 6. Vaccination Coverage Quality Indicator Descriptions.....	26
6.1 Weighted and unweighted analyses	26
6.2 Analysis Counter.....	26
6.3 RI_COVG: RI Survey – Measures Related to Coverage.....	29
6.4 RI_QUAL: RI Survey – Measures Related to quality of Services.....	33

6.5 RI_VCTC: RI Survey – Vaccination Coverage and Timeliness Charts (VCTC)	43
Chapter 7. Examples of Control Programs.....	60
7.1 Block A – Start with clear memory.....	61
7.2 Block B – Specify input/output folders & analysis name	62
7.3 Block C – CD to output folder & open VCQI log	63
7.4 Block D – Specify dataset names & important metadata	64
7.5 Block E – Pre-process survey data.....	77
7.6 Block F – Calculate VCQI indicators requested by the user	80
7.7 Block G – Exit gracefully	85
Chapter 8. Troubleshooting VCQI	86
Annex A. Nested Strata in VCQI	88
Annex B. Controlling How Strata Are Listed in VCQI Output	92
B.0 Demographic stratification using the VCQI_LEVEL4_SET_VARLIST and LAYOUT	93
B.1 Sample listings of stratum name and order datasets for level 1-3.....	97
B.2 Example: Nested output for all Levels: 1, 2, and 3 with additional Level 4 stratification	98
B.3 Example: Nested output for all Levels: 1, 2, and 3	103
B.4 Example: Non-nested output for all Levels: 1, 2, and 3	106
B.5 Example: Output for Level 3 only.....	109
B.6 Example: Output for Level 3 with additional Level 4 stratification	112
B.7 Example: Output for Level 4 Only	116
B.8 A Large Realistic Level 4 LAYOUT Example	119
B.9 Changing Order of Strata in Bar and Unweighted Plots	123
B.10 Additional Options for Customizing VCQI Output	126
Annex C. Customizing VCQI Plot Output.....	127
C.1 Customizing Plot Colors	127
C.2 Changing Confidence Interval Information on Bar Plots	128
C.3 Changing Confidence Interval Information on Double Bar Plots.....	134
C.4 To Customize VCQI Plots in Other Ways.....	139
Annex D. VCQI Output Databases	140
Annex E. VCQI Augmented Dataset	142
Annex F. Additional Tools for Working with VCQI Output.....	143
F.1 Missed Opportunities for Simultaneous Vaccination R-Shiny Application	143
F.2 ICCLOOP Program for Reporting Survey Planning Parameters	148
Annex G. RI Date Data Quality – Method for Detailed Assessment	150

List of Tables

Table 2-1. VCQI control programs consist of seven blocks of code	14
Table 3-1. How RI_RECORDS inputs affect outcome calculation.....	21
Table 3-2. How RI_RECORDS inputs affect the main valid dose outcome (RI_COVG_02).....	21
Table 6-1. VCQI indicators that rely on others being run first.....	28
Table 6-2. Naming convention for RI_COVG_01 databases.....	29
Table 6-3. Naming convention for RI_COVG_02 databases.....	31
Table 6-4. How RI_QUAL_07B uses RI_RECORDS inputs	34
Table 6-5. RI_VCTC_01 inputs that are usually set in the control program.....	46
Table 6-6. RI_VCTC_01 inputs that are often unchanged in globals_for_timeliness_plots.do	47
Table 6-7. RI_VCTC_01 inputs to customize timeliness categories for individual doses.....	51
Table A-1. Overview of three nested levels of administrative hierarchy	88
Table B-1. Level1-3 Names and Order Datasets	92
Table B-2. Nested output for all Levels: 1-3 with Level 4 stratification	100
Table B-3. Nested output for all Levels: 1, 2, and 3	104
Table B-4. Non-nested output for all Levels: 1-3	107
Table B-5. Output for Level 3 only	110
Table B-6. Output for Level 3 with Level 4 stratification	114
Table B-7. Output for Level 4 alone	117

List of Figures

Figure 6-1. Vaccination Coverage & Timeliness Chart (VCTC) for Harmonia using default parameters ...	43
Figure 6-2. VCTC for Harmonia District 03 using customized BCG and HepB timeliness parameters.....	51
Figure A-1. The Fictional Country of Harmonia.....	89
Figure A-2. Harmonia holds a Northern and Southern Province	89
Figure A-3. Each province holds five districts	90
Figure B-1. Bar plot showing nested output in table order for all Levels: 1-3 with Level 4 stratification	102
Figure B-2. Bar plot showing nested output in table order for all Levels: 1, 2, and 3	105
Figure B-3. Bar plot showing non-nested output sorted for all Levels: 1, 2, and 3	108
Figure B-4. Bar plot showing sorted output for Level 3 only	111
Figure B-5. Bar plot showing output in table order for Level 3 with Level 4 stratification	115
Figure B-6. Bar plot showing output in table order for Level 4 only	118
Figure B-7. Bar plot with strata sorted by outcomes, low-to-high, bottom-to-top.....	124
Figure B-8. Bar plot with strata sorted by outcomes, high-to-low, bottom-to-top.....	124
Figure B-9. Bar plot with strata sorted in table order.....	125
Figure C-1. Bar plot demonstrating VCQI_IW PLOT_CITEXT 1.....	129
Figure C-2. Bar plot demonstrating VCQI_IW PLOT_CITEXT 2.....	129
Figure C-3. Bar plot demonstrating VCQI_IW PLOT_CITEXT 3.....	130
Figure C-4. Bar plot demonstrating VCQI_IW PLOT_CITEXT 4.....	131
Figure C-5. Bar plot demonstrating VCQI_IW PLOT_CITEXT 5.....	132
Figure C-6. Double bar plot demonstrating VCQI_DOUBLE_IW PLOT_CITEXT 1.....	135
Figure C-7. Double bar plot demonstrating VCQI_DOUBLE_IW PLOT_CITEXT 2	136
Figure F-1. Missed Opportunities R-Shiny App – Main Bar Chart.....	144
Figure F-2. Missed Opportunities R-Shiny App – Main Results Table.....	145
Figure F-3. Missed Opportunities R-Shiny App – Cumulative Time to Correction Graphs	146
Figure F-4. Missed Opportunities R-Shiny App – Time to Correction Table	147

ACKNOWLEDGEMENTS

In 2022 the Pan American Health Organization funded development of an R version of the Vaccination Coverage Quality Indicators (VCQI) software to complement the Stata version that was developed in 2015 for the World Health Organization. Both versions were developed and are maintained by Biostat Global Consulting.

We express our very fond gratitude for helpful ideas and feedback furnished by Tony Burton, Pierre Clauquin, Felicity Cutts, Carolina Danovaro, Marta Gacic-Dobo, Martha Velandia, Marcela Contreras, John Ndegwa Wagai, David Brown, Kathleen Wannemuehler, David Koffi, Mamadou Diallo, Riswana Soundardjee, Michelle Selim and the dozens of engaging survey analysts who have attended in-person VCQI training events in Europe, Asia, and Africa.

This User's Guide was written by Dale Rhoda, Mary Prier, Mary Kay Trimner, Caitlin Clary, and Mia Yu.

This page is intentionally blank.

CHAPTER 0. PRELIMINARY MATERIAL

Document Revision History

Date	Version	Comment
2022-12-15	1.0	Initial R User's Guide

R Software is Under Development

The R version of VCQI is being developed in stages. There are modules and indicators in the Stata version that do not yet appear in the R version. As time and funding allow, our intention is to implement all of VCQI's capabilities in R, too.

Currently, the R version of VCQI analyzes data from routine immunization (RI) surveys only, and calculates the following indicators:

- RI_COVG_01 – crude coverage
- RI_COVG_02 – valid coverage
- RI_QUAL_07B – what would valid coverage be if every child had received every dose that was due at every one of their documented vaccination visits
- RI_QUAL_08 – percentage of documented vaccination visits that include 1+ missed opportunities for simultaneous vaccination (MOSV)
- RI_QUAL_09 – percentage of children with 1+ missed opportunities for simultaneous vaccination (MOSV)
- RI_VCTC_01 – show coverage and timeliness charts (The R version does not currently show HBR availability or % fully vaccinated or % not vaccinated but includes all other VCTC features.)

Our R Nomenclature – what to call user-defined input parameters

In the Stata version of VCQI the control program uses Stata *scalars* and *global macros* to define VCQI inputs and parameter settings. In the R version, the equivalent constructs should properly be called *R objects* but to be consistent across VCQI, we have adopted the phrase *global value* instead and we sometimes still use the phrase *global macro*. These values are defined in the control program using a new R function named `vcqi_global()`. You will not see the phrase *global value* in other R code documentation, but in our context it means a value that has been assigned in the global environment (using the `<-` operator) and remains defined until its value is changed or until the conclusion of the current VCQI analysis session. Throughout this document the phrase *global value* or simply *global* mean just that. When describing parameters that define the routine immunization schedule, we use the phrase *schedule globals*.

Acronym List

ACC	Access - indicators that measure access to vaccination services
CONT	Continuity - indicators that measure continuity of vaccination services
COVG	Coverage - indicators that estimate vaccination coverage
CCC	Cumulative coverage curve
CIC	Cumulative interval curve
DEFF	Design effect
FVL	Forms and Variable List Document (that accompanies this User's Guide)
HC	Health center (might sometimes be used interchangeably with "health facility")
ICC	Intracluster correlation coefficient
LCB	Lower one-sided confidence bound
MCV	Measles Containing Vaccine
MOV	Missed opportunity for simultaneous vaccination
PCCS	Post-campaign coverage survey
QUAL	Quality - indicators that measure quality of vaccination services
RI	Routine Immunization
SIA	Supplementary Immunization Activity (Vaccination Campaign)
TT	Tetanus toxoid - also used here to mean a survey that measures protection at birth from neonatal tetanus
CI	Two-sided Confidence Interval
UCB	Upper one-sided confidence bound
VCQI	Vaccination Coverage Quality Indicators
VCTC	Vaccination coverage and timeliness chart

License Agreement

(Pending approval by WHO and PAHO)

Copyright (c) 2015-2022, World Health Organization & Pan American Health Organization
All rights reserved.

Redistribution and use of the Vaccination Coverage Quality Indicators (VCQI) programs and documents, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

This page is intentionally blank.

CHAPTER 1. INTRODUCTION

Congratulations! You are going to analyze data from a vaccination coverage survey.

While there is wide agreement on the essentials of what constitutes a *coverage survey*, there are a range of parameters that can reasonably vary from project to project. The Vaccination Coverage Quality Indicators (VCQI¹) is a set of programs to conduct standard analyses and produce standard output tables and figures. VCQI programs are available in Stata and in R.

VCQI accounts for flexibility in the important aspects that may vary from one survey to the next. That flexibility is achieved by allowing the user to modify several dozen software input parameters. You do this by copying a template of a so-called *control program* and customizing the parameter settings that describe 1) the country's vaccination schedule, 2) the design and details of the survey, and 3) the analysis you wish to conduct. Some items that can vary from one VCQI analysis to the next include:

Parameters describing the schedule

- Country vaccination schedule
- List of doses required to be considered "fully vaccinated"
- List of doses considered when defining "zero dose" children

Parameters describing the survey

- Sample design: Stratified? Clustered? Simple random?
- Vaccination evidence sought from card or recall or health facility register
- Evidence sought from BCG scar or not
- SIA evidence sought from finger mark or not
- Ages of eligible respondents

Parameters describing the analysis

- Names of geographic strata (nested up to 3 levels deep)
- Demographic stratifiers (urban/rural, male/female, etc.)
- Order of strata in tables and figures
- List of doses in the analysis & their order in tables and figures
- Should doses be shifted down to fill *holes* in dose series evidence
- Should booster doses be shifted down to count as primary doses
- Which indicators to calculate
- Number of decimal places to include in coverage estimates
- Weighted analysis or no
- Type of confidence interval to report
- Should outcomes from very small samples be suppressed or annotated
- Specific table titles & footnotes

The purpose of this manual is to help you acquire VCQI programs and modify and run them to achieve your goals.

¹ Pronounced "Vicki"

1.1 VCQI Resources

The VCQI programs are freely available, courtesy of the World Health Organization and the Pan-American Health Organization. The R implementation of VCQI is available as an R package, which can be installed from the vcqiR GitHub repository.² Supporting manuals and materials may be downloaded from the vcqiR GitHub repository and or the VCQI Resources website.³ You must have R version 4.2.1 or later to run VCQI. R can be downloaded for free from the Comprehensive R Archive Network (CRAN).⁴ We also recommend downloading and using the RStudio Desktop integrated development environment, which is available for free from Posit.⁵

This guide, for users of the R implementation of VCQI, assumes that you have basic knowledge of how to use R and that you have successfully followed the instructions in the manual entitled: *Getting Started with VCQI – For R Users*. The details included here will help you understand the definitions and run-time options of VCQI’s indicators and understand the flexible features of the software that you may change by editing the code in Blocks B, D, and F of the control program.

There is a VCQI User’s Group hosted on the Technet-21 website⁶. If you have not done so already, join the group to receive updated announcements about VCQI training materials. You may post your questions there and help answer questions posted by other users. Be sure to post celebratory messages there when you accomplish a goal with VCQI.

1.2 Differences between the Stata and R versions

The Stata and R versions of the software produce identical estimates of vaccination coverage when used to analyze the same dataset. The contents of the tables and figures that are produced are meant to be the same, and so far, they are. Currently the R version calculates fewer indicators than the Stata version and although the contents of the tables and figures are the same, there are some small differences in how they are formatted. Over time we hope to make the two sets of output as much alike as is practical.

It is possible to run both the Stata and R versions of VCQI on the same input datasets to compare output, and in the future, we intend to post some demonstrations of how to do that and what we find when we do it both with faux data training datasets and real-world survey datasets.

The early pages of this document list the indicators that have been implemented in R. From the perspective of graphics, R produces organ pipe plots, bar charts for weighted and unweighted outcomes, and vaccination coverage and timeliness charts. It does not yet produce inchworm plots, cumulative coverage curves, or cumulative interval curves.

The control program used to run a VCQI analysis in R is structured in the same manner as one for Stata, but uses different syntax to accomplish the same goals. To analyze a dataset in both R and Stata, you will need to adapt separate control programs.

² <https://github.com/BiostatGlobalConsulting/vcqiR>

³ http://www.biostatglobal.com/VCQI_RESOURCES.html

⁴ <https://cran.r-project.org/>

⁵ <https://posit.co/downloads/>

⁶ <https://www.technet-21.org/en/network/groups/293-vcqi>

Chapter 1. Overview

The R version of VCQI is able to analyze antigen series with up to 9 doses (e.g., penta1-penta9) whereas the Stata version is currently limited to series with up to 3 doses (e.g., penta1-penta3).

Currently the R version of VCQI requires the user to organize the tabular and bar chart output using either a list of Level4 stratification variables or what VCQI calls a *Level4 layout* file. As described in Annex B, the layout can be specified by the user via a small dataset that defines which respondents should be summarized in which rows of output, or via a simple list of stratification variables (e.g., province_name sex maternal_education wealth_quintile). The Stata version of VCQI includes shortcut options for requesting stratification by several levels of geographic stratification. But in the R version, these must be specified via the Level4 variable list or the Level4 layout file.

The R version of VCQI always uses a Level4 layout file to order rows of output in tables and bar charts. If the user provides only a Level4 variable list then a default Level4 layout file is generated and used.

CHAPTER 2. OVERVIEW

This section of the document gives an overview of a) files that comprise VCQI, b) datasets and parameter files that need to be assembled in order to run VCQI, and c) the files that are produced by VCQI (VCQI outputs).

2.1 Running VCQI

The usual practice is to copy a VCQI control program (.R file) from the examples provided, edit the file (to provide the appropriate file locations and analysis parameters), save it and run it in R. Open the resulting spreadsheet and check the log sheet for errors or warnings. If VCQI ran successfully, examine the results to see if they make sense. If yes, you might copy tabulated results or automatically generated figures into a report. Save the control program and output for future reference. To run a second analysis, copy the control program to a new, empty folder; edit the new program to send its output to that new folder where the control program is saved; save the control program, and run it.

Please note that you will need to spend some time customizing the control program for each new project. You need to specify parameters to describe a) the vaccination schedule, b) the coverage survey, and c) the analyses you wish to run. The main purpose of this *User's Guide* is to help you understand the rich set of parameter options available in Blocks B, D, and F of every control program. When you begin a new project the first challenge is to make your dataset compatible with VCQI. The *VCQI Forms and Variable Lists (FVL)* document will help you accomplish that. The second challenge is to edit a control program template and specify parameters that are relevant for YOUR SURVEY. The control program templates are filled with parameters to describe fictional surveys in the fictional country of Harmonia. You must replace the parameters that describe Harmonia's vaccination schedule and survey with parameters to describe the schedule and survey in your country. The chapters that follow will show you how to do it.

VCQI performs a series of checks to be sure the user has defined the necessary inputs and that the input datasets and necessary variables are all present. When something goes wrong, it tries to provide informative error messages both to the R console and in a VCQI log file. If VCQI detects an important error, the log is copied into the output spreadsheet before the program halts. If an unanticipated error occurs, the incomplete log will be a .csv file saved in the VCQI output folder. If you open the spreadsheet and find only placeholder text in the Log worksheet, then close the Excel file and run the command `vcqi_cleanup()` in the R console. In most cases this will cause the log to be closed, processed, and copied to the output spreadsheet file. Re-open the spreadsheet and look at the log tab. Otherwise follow the instructions found in the placeholder Log tab in the spreadsheet.

If you experience problems running VCQI, contact GetVCQIHelp@biostatglobal.com.

2.2 The Indicator Types

Several types of analyses are included with VCQI; they are described in this document using short abbreviations:

- DESC: Descriptive indicators document the composition of the survey sample and summarize responses to multiple-choice questions; these indicators may be calculated for any survey. (Currently in the Stata version only)

Chapter 2. Overview

Within each survey the indicators are organized according to vaccination program attributes that have proven useful in earlier assessments:

- COVG: Indicators related to estimated proportion served, known informally as *coverage* (Some are available in R and Stata and some currently in the Stata version only)
- ACC: Indicators related to whether respondents have *access* to vaccination services (Currently in the Stata version only)
- CONT: Indicators related to whether respondents experience *continuity* of services (Currently in the Stata version only)
- QUAL: Indicators related to the *quality* of vaccination service (Some are available in R and Stata and some currently in the Stata version only)
- CCC: Make *cumulative coverage curves* to summarize vaccination timeliness (Currently in the Stata version only)
- CIC: Make *cumulative interval curves* to summarize vaccination timeliness (Currently in the Stata version only)
- VCTC: Vaccination coverage and timeliness charts

Finally, there are indicators to conduct formal hypothesis tests:

- DIFF: Indicators to estimate *differences* in coverage may be calculated for many outcomes. VCQI calculates differences in coverage a) between strata, and b) between sub-groups within a single stratum. (Currently in the Stata version only)

Additional indicators are added over time.

2.3 Files that comprise VCQI

R Programs

VCQI is a set of R programs that work together to analyze the survey data. No special license key is required to run the VCQI programs – if you have the vcqiR package and you have a copy of R (version 4.2.1 or later), then you can run VCQI.

Control Program

A primary purpose of this *User’s Guide* is to help you understand the VCQI control program and learn to confidently adapt it to meet your needs. A control program is a set of R commands saved in an R script. In most cases a VCQI user will not need to look inside of any of the several hundred VCQI programs except the control program because the control program calls all the other programs that it needs. Every control program alternates between clearly marked blocks of code that the user *should edit* and blocks that they *should not edit*. Portions of the control program that you might edit include those that point to folders and datasets, lines that describe the vaccination schedule and survey, and those that list which indicators you want to calculate. To run VCQI, the user copies a control program and edits the appropriate sections before running the control program in R. Sample control programs are provided in the vcqiR GitHub repository and on the VCQI Resources website⁷ and several are annotated in this *User’s Guide*. See

⁷ See the Control_Programs directory at <https://github.com/BiostatGlobalConsulting/vcqiR> or the demo R control programs at http://www.biostatglobal.com/VCQI_resources.html.

Chapter 7 for detailed examples, but truly the entire document is meant to help you understand how to use the control program.

2.4 Files Used by VCQI

Datasets

You need to assemble a small set of files to run VCQI – precisely which files you need depends on whether you are analyzing data from a routine immunization (RI) survey, a post-supplemental immunization activity (SIA) survey, or a tetanus protection at birth (TT) survey. Details appear in later sections of this document. VCQI assumes that the survey data were collected using the survey questions described in the accompanying document named *Vaccination Coverage Surveys – Forms & Variable Lists (FVL) Structured for Compatibility with VCQI* (described hereafter as ‘the FVL document’). VCQI assumes that variables are named and coded as described there. If the data were collected using other survey instruments, it will be necessary to recode the data to look as though it comes from questions in the FVL document. VCQI can also analyze immunization coverage data collected from other surveys, however, the variables must be renamed and recoded to be consistent with those described in the FVL document.

Parameter Files

The user provides parameter files listing the names of geographic or administrative strata in the survey and the order of them. VCQI requires that these files be in place regardless of whether you are analyzing RI, SIA or TT data. Annex A describes how VCQI structures nested geographic strata and Annex B describes how to specify names and listing orders of each stratum.

VCQI input datasets and parameter files may be saved in several file formats: .rds, .dta, or .csv.

At a minimum, the user must specify:

1. An RI dataset
2. A CM dataset
3. Geographic strata names, ids, and order in:
 - a. level1name dataset
 - b. level2names dataset
 - c. level3names dataset
 - d. level2order dataset
 - e. level3order dataset⁸

2.5 Files Produced by VCQI

VCQI can produce five types of files.

Analysis Datasets

For each indicator, there is usually an intermediate analysis dataset (flat file) produced that includes only the variables required for that indicator. The analysis file usually includes elements from several input datasets (e.g., from the list of households (HH), from the list of household members (HM), from the cluster

⁸ At this time the R code does not use the level2order or level3order datasets to determine the order of output rows in Excel tables, but they are required inputs because they are used by the code to organize datasets and used by the VCTC indicator to name output files.

Chapter 2. Overview

metadata (CM) and from the subject matter dataset: RI, TT, or SIA). The analysis file will also include new so-called *derived variables* that VCQI calculates from the survey data and uses to calculate the indicators.

If the user requests it, VCQI will generate a so-called *augmented dataset* where the original survey responses are merged with the derived variables and indicator outcome variables. The resulting dataset is an excellent resource for a) exporting to other statistical packages to audit VCQI results, b) conducting advanced analyses, like logistic regression to analyze socio-demographic correlates of coverage, or c) generating customized tables or figures. See Annex E for more details.

Output Databases

Most indicators produce one or more flat files that it calls *databases* documenting indicator outcomes. These databases include one row of output per row of the Level4 layout file. Databases are saved as R datasets and are suitable for importing by other programs. They could be used for later calculations or to tabulate or graph results in a way that is not supported by VCQI. The database files have the word *_database* in their filenames. Later sections of this document list the databases saved by each indicator. When VCQI finishes running, it aggregates most of the databases into one large dataset. See Annex D for more details.

Tabulated Output

VCQI saves tabulated output in an Excel file, generating one or more tabs (worksheets) per indicator. The output includes one row per row of the Level4 layout file. It is formatted and ready to be copied and pasted into project reports. The user controls which strata appear in the tables and in what order (see Annex B). Typically, a VCQI control program analyzes only one sort of survey data and produces only one Excel output file. If a survey asked questions about routine immunization, tetanus doses for pregnant women, and campaign coverage, that data would be analyzed using at least three separate control programs and the outputs for each portion of the survey would be saved in a different Excel workbook.

For RI survey analysis, there is an option to summarize the quality of the vaccination data. If the user requests this analysis, its outcomes will be put into a separate Excel spreadsheet with the words “dates_ticks” in the filename.

Graphic Output

Many VCQI indicators generate figures. The control program includes options so the user can stipulate whether the program should make any figures at all, and if so, which types. As a rule, the same strata that appear in the tabulated output also appear in the figures. Each figure is saved as a portable network graphics file (extension .png). The control program has options to save the data underlying the figures as .rds datasets, to facilitate recreating and modifying the figures using R.

At this time VCQI generates four kinds of graphical output, organ pipe plots, bar plots, unweighted sample proportion plots, and vaccination coverage and timeliness charts (VCTCs). Each is described in the *VCQI Results Interpretation Quick-Reference Guide*. Also, Annex B of this *User’s Guide* shows some examples of VCQI figures.

The Stata version of VCQI also makes something called an *inchworm plot*. At this time, the R version does not make inchworm plots; bar plots take the place of inchworm plots.

There is helpful information on organ pipe plots and inchworm plots in the 2018 WHO Vaccination Coverage Cluster Survey Reference Manual⁹ – specifically in Chapter 6 and Annex M. There are helpful conference presentations on organ pipe plots and inchworm plots on the Stata website.^{10,11}

2.6 Levels of Survey Strata

VCQI is flexible and can analyze data from a single geographic region (stratum) or from several strata. If the strata comprise all the pieces of a higher level (e.g., all the provinces in a nation) then VCQI can calculate the aggregated higher-level results as well.

The examples in this user’s guide assume that your survey was conducted in strata at sub-sub-national level (e.g., a separate survey in each health district). It assumes that the districts are nested within provinces, and the survey was conducted in every district in every province in the nation. This document provides examples of estimating results at the district level, at the provincial and national levels.

It is possible to do simpler or even more complex analyses, such as a single survey in a single stratum, or even four nested levels of hierarchy. Common variations are described in Annex B.

2.7 Program Progress Log

Every VCQI session generates a log file with messages to document the user’s inputs and inform the user which programs were used and whether their progress was successful or if they issued errors or warnings. While VCQI is running, the log entries are stored in a .csv file. The many VCQI functions append new comments onto the dataset throughout the run. When VCQI exits, the log entries are copied into the output Excel workbook in a sheet named “Log”.

VCQI users should look at the Log tab in the spreadsheet before focusing on other output. Errors are shaded red and warnings are yellow and all errors and warnings appear at the top of the Log tab. Errors typically must be addressed and VCQI must be re-run. Warnings do not require you to re-run VCQI but they are messages important enough to be brought to your attention before you interpret the VCQI output.

As far as most VCQI users are concerned, the only portion of the log that is of interest is whether there are errors, and if so, how to correct them. The many hundreds of other lines in the log are useful to VCQI developers for debugging problems. You will not need to interpret them, but you may be asked to e-mail your log to the VCQI developers if you have difficulty with a VCQI analysis.

The error messages are meant to be worded in a clear enough manner to help you correct the problem. If the messages are not clear, please send feedback to GetVCQIHelp@biostatglobal.com.

2.8 Structure of VCQI Control Programs

Users should copy and edit the control programs that are provided with VCQI. It is good practice to use a different control program for each analysis and save the control program and resulting output for later reference.

⁹ <https://apps.who.int/iris/handle/10665/272820>

¹⁰ https://www.stata.com/meeting/columbus18/slides/columbus18_Prier.pptx

¹¹ https://www.stata.com/meeting/chicago16/slides/chicago16_rhoda.pptx

Regardless of whether you are analyzing an RI, TT, or SIA survey, the typical VCQI control program consists of seven sections or blocks of code. There are four sections that the user *should not edit* and three that they *should edit*. Chapter 7 lists example control programs, line by line, and describes what they do.

Table 2-1. VCQI control programs consist of seven blocks of code

Block of R Code	User Edits Code in this Block?
A. Initialize VCQI run – clean out old data, programs, and global values	X No
B. User specifies input and output folders and a name for this analysis	✓ Yes
C. Open the log file & document which version of VCQI programs are running	X No
D. User specifies datasets and metadata about survey, schedule and analysis	✓ Yes
E. VCQI checks inputs; pre-processes analysis dataset	X No
F. User specifies which indicators to calculate, and any required inputs	✓ Yes
G. VCQI closes log, deletes temporary files, informs the user of any errors	X No

2.9 Specifying VCQI Input Parameters

In the remainder of this guide there is a lot of specific guidance that describes VCQI input parameters that you specify in the control program. Parameters are specified by creating objects in the R environment.

The syntax to specify a parameter looks like this:

```
PARAMETER_NAME <- parameter_value
```

or

```
vcqi_global(PARAMETER_NAME, parameter_value)
```

R programmers will be familiar with the usual `<-` assignment process to create objects in the R environment. VCQI sometimes uses this familiar method to establish parameter values but it usually uses the function named `vcqi_global` which accomplishes two purposes: 1) it creates an object called `PARAMETER_NAME` to hold the value `parameter_value`, and 2) it writes a message in the VCQI log listing the name of the object and its updated value. Please note that VCQI's user-specified parameters are always listed in upper case letters and the user may not change their names. The names are fixed and hard-coded. In most cases the allowable values are also fixed and described in this User's Guide. If you specify a value that is not allowed, VCQI will issue a clear error message. Although it is possible to edit code in blocks B, D, and F, the user should not change the syntax of `vcqi_global` calls or the `PARAMETER_NAMES`. They should only edit the parameter values. In many cases, to turn on a feature, the user will specify a value of one. For instance, we will see later that to tell VCQI you want VCQI to generate plots in this run, you specify:

```
vcqi_global(MAKE_PLOTS, 1)
```

and to tell VCQI not to make plots , you specify:

```
vcqi_global(MAKE_PLOTS, 0)
```

Only those two values are allowed for MAKE_PLOTS: 1 or 0, which correspond conceptually to YES and NO. When we say that the user may edit Block D, we mean that the user may turn various parameters on or off or select different values for some categorical parameters. But the user should not try to change the names of any VCQI parameters.

The user also specifies information about the routine immunization schedule by setting parameters. For these schedule parameters, the user has some flexibility concerning the names of doses, but the remainder of the object names are fixed. The possibilities include:

<dose>_min_age_days

<dose>_max_age_days

For multi-dose antigens it is valid to specify minimum intervals between doses. For instance, for a three-dose antigen you may specify:

<dose>2_min_interval_days

and

<dose>3_min_interval_days

See Chapters 3 and 7 for examples of how to use vcqi_global parameters to specify the vaccination schedule.

CHAPTER 3. ANALYZING ROUTINE IMMUNIZATION (RI) SURVEYS

To analyze RI survey data, you will use a dedicated control program, copied from an example and modified to fit your survey and dataset. RI control programs differ in several ways from TT and SIA control programs, so you should start with an RI control program template, which you can find in the Control_Programs folder of the vcqiR GitHub repository or download from the VCQI Resources website. The R version of VCQI currently calculates these indicators from RI surveys:

1. Two regarding vaccination coverage (RI_COVG_01 and _02)
2. Three regarding the quality of the vaccination program (RI_QUAL_07B, _08, and _09)
3. One that summarizes coverage and timeliness (RI_VCTC_01)

3.1 Dataset File Names (goes in Block D)

The first lines of code in Block D tell VCQI the names of the (up to 5) survey datasets.

```
# ****
# Code Block: RI-D          (User may change) ----
#
# Specify dataset names and important metadata
# Dataset names should include file extensions
# Accepted file types: .rds, .dta, .csv

# Name of datasets that hold RI data
vcqi_global(VCQI_RI_DATASET, "RI_Harmonia_2015.dta")
vcqi_global(VCQI_RIHC_DATASET, "RIHC_Harmonia_2015.dta")

# Name of dataset that holds cluster metadata
vcqi_global(VCQI_CM_DATASET, "CM_Harmonia_2015.dta")

# If you will describe the dataset using DESC_01 then you need to also
# specify the HH and HM datasets
vcqi_global(VCQI_HH_DATASET, "HH_Harmonia_2015.dta")
vcqi_global(VCQI_HM_DATASET, "HM_Harmonia_2015.dta")
```

Only the surveys that hold vaccination evidence from the household interview (VCQI_RI_DATASET) and that hold the list of cluster metadata (VCQI_CM_DATASET) are strictly required for every RI analysis. If the survey protocol included visits to health centers to collect vaccination evidence from immunization registers, then the RI health center dataset (VCQI_RIHC_DATASET) will be required, too. Both the household listing and household member listing datasets (VCQI_HH_DATASET and VCQI_HM_DATASET) are only required if you wish to run DESC_01 (which has not been implemented in the R version of VCQI yet) to summarize the sample; these are often omitted in the early stages of analysis when there is great enthusiasm to examine output from the coverage indicators. Variables from those datasets may also appear as Level4 stratifiers.

The R version of VCQI does not currently include the DESC_01 indicator, for which the VCQI_HH_DATASET and VCQI_HM_DATASET are required. The user may still wish to specify the HH and HM datasets, as variables from those datasets may be used as Level4 stratifiers.

There is no need to include the data file folder path here. All the input datasets must be present in the VCQI_DATA_FOLDER that was named in Block B. The R implementation of VCQI can use .dta, .rds, and .csv input files; the file extension must be included when defining the dataset names in Block D so that the appropriate package can be used to read those files. Please note that dataset names are case sensitive.

To purposefully omit one of these filenames, simply delete the filename in the template and replace it with NULL. For example, to omit the RIHC dataset, use a line like this:

```
vcqi_global(VCQI_RIHC_DATASET, NULL)
```

(You could also comment out the line or omit it entirely, but that might make it difficult to remember the keywords you need to type if you add an RIHC dataset to the analysis later.)

It is quite common to see code like this at the top of Block D:

```
# ****
# Code Block: RI-D           (User may change) ----
#
# Specify dataset names and important metadata
# Dataset names should include file extensions
# Accepted file types: .rds, .dta, .csv

# Name of datasets that hold RI data
vcqi_global(VCQI_RI_DATASET, "RI_mdy_Harmonia_2015.dta")
vcqi_global(VCQI_RIHC_DATASET, NULL)

# Name of dataset that holds cluster metadata
vcqi_global(VCQI_CM_DATASET, "CM_dataset_Harmonia_2015.dta")

# If you will describe the dataset using DESC_01 then you need to also
# specify the HH and HM datasets
vcqi_global(VCQI_HH_DATASET, NULL)
vcqi_global(VCQI_HM_DATASET, NULL)
```

3.2 Vaccination Schedule Metadata (goes in Block D)

The user must specify the vaccination schedule that was in place over the time being evaluated by the survey. If the target population is children 12-23 months, then the user should specify the vaccination schedule in place over the preceding 24 months. The schedule is defined using R globals, typically in Block D of the control program. The schedule ages vary from country to country.¹²

For single-dose vaccines, specify the minimum age in days at which the dose should be given. E.g.,

```
bcg_min_age_days <- 0    # birth dose
hepb_min_age_days <- 0    # birth dose
opv0_min_age_days <- 0    # birth dose
ipv_min_age_days <- 98    # 14 weeks
mcv1_min_age_days <- 270   # 9 months
yf_min_age_days    <- 270   # 9 months
```

¹² Details available at <https://immunizationdata.who.int/listing.html>.

In most countries, most doses do not have a maximum age for valid administration, but if a dose does have a maximum age, specify it thus:

```
opv0_max_age_days <- 14    # valid in 1st 2 weeks  
hepb_max_age_days <- 5     # valid in 1st 5 days
```

For multi-dose vaccines, specify the minimum age in days for the first dose, and the minimum age and interval (also in days) for later doses. E.g.,

```
pental_min_age_days      <- 42    # 6 weeks  
penta2_min_age_days     <- 70    # 10 weeks  
penta2_min_interval_days <- 28    # 4 weeks  
penta3_min_age_days     <- 98    # 14 weeks  
penta3_min_interval_days <- 28    # 4 weeks
```

In many countries the scheduled interval between doses is 28 days, but other countries like to space out the vaccination visits. In the case of doses spaced farther apart, use the min_interval_days schedule global to indicate the minimum interval before the next dose would be considered valid (usually 28 days) and use the min_age_days to indicate when the dose is scheduled to be given. For example, in some of the former Soviet republics, children are scheduled to receive Penta when they are 2, 4 and 6 months old. But the second dose of Penta would be considered a valid dose if at least 28 days had elapsed since a valid first dose, so the values would look like this:

```
pental_min_age_days      <- 60    # 2 months (30.4 days/month)  
penta2_min_age_days     <- 121   # 4 months  
penta2_min_interval_days <- 28    # 4 weeks  
penta3_min_age_days     <- 182   # 6 months  
penta3_min_interval_days <- 28    # 4 weeks
```

Every dose in the analysis must have a_min_age_days. Every second dose in 2- and 3-dose antigens must have a <dose>2_min_interval_days and every third dose in 3-dose antigens must have a <dose>3_min_interval_days. 3+ dose series follow the same pattern.

A Note on Dose Names

The convention in VCQI is for dose names to be expressed using lower case letters in variable names and schedule global names. These are case sensitive. Before running VCQI, you will need to rename the dose-related variables from the FVL document and include the dose names in the variable names. The following lines show how the variables might be renamed for pental. When evaluating dates, VCQI expects to find month, day, and year coded in separate variables named <dose>_date_card_m, <dose>_date_card_d, <dose>_date_card_y, and <dose>_tick_card in the RI dataset,¹³ and corresponding variables with the word ‘register’ substituted for ‘card’ in the RIHC dataset. The user should write a program to either rename the survey variables, or to make new variables to meet that expectation.

In the RI dataset:

¹³ The string <dose> is a placeholder; it might be bcg or mcv1 or pental. The FVL document contains a longer example in the section named “Breaking Dates Into Month, Day and Year Components”.

```
dplyr::rename(
  pental_date_card_d = RI39d,
  pental_date_card_m = RI39m,
  pental_date_card_y = RI39y,
  pental_tick_card = RI40)
```

In the RIHC dataset:

```
dplyr::rename(
  pental_date_register_d = RIHC29d,
  pental_date_register_m = RIHC29m,
  pental_date_register_y = RIHC29y,
  pental_tick_register = RIHC30)
```

The user can name the doses anything they wish (with 6 or fewer characters in the name). For instance, it would be perfectly valid to use the name penta1, dpt1, or dtp1. The doses must be named consistently in the RI and RIHC datasets and in the globals that define the schedule (schedule globals). So if the variables use penta1, then the schedule globals should not say dpt1; they should say penta1.

The only dose name that is hard-coded into VCQI is bcg. If the survey asks interviewers to record whether they saw the BCG scar on the child, then VCQI expects to find a variable named bcg_scar_history. All other doses are free to use any alternate abbreviations.

Dose names should use abbreviations with 6 or fewer characters. If it is a multi-dose sequence, the letter portion of the abbreviation should use 5 or fewer characters. So ‘penta’ is okay and ‘pneum’ is okay, but ‘pneumo’ is too long because when the numbers 1, 2, or 3 are appended for first, second, and third doses, the abbreviation would be 7 characters – too long!

Again, the VCQI convention is for the schedule globals and these date and tick variables to use lower case names.

But in VCQI control programs we sometimes require a single dose name or a list of dose names as inputs and those may be specified in upper or lower case. (VCQI will convert the case to what it needs when it runs.) If you see a dose name being listed using the vcqi_global¹⁴ command, it can be either upper or lower case.

3.3 Survey Metadata (goes in Block D)

There are three categories of information that VCQI requires to describe the survey:

1. What are the earliest and latest allowable dates of vaccination for respondents and doses inquired about in this survey?
2. What are the minimum and maximum age of children eligible for the survey (in days)?
3. Did the survey protocol include seeking vaccination records at health centers, and if yes, for which respondents?

¹⁴ vcqi_global is a program that 1) assigns a R global value the value named in the line of syntax, and 2) writes the new value of the global value in the VCQI log. It has the same consequence as R’s “`<-`” command, with the bonus of documenting the assigned value in the log.

These are specified in Block D of the control program.

Earliest and Latest Allowable Vaccination Dates for this Survey

The user must specify the earliest and latest possible vaccination dates of respondents who are eligible for the RI survey. For surveys that include birth doses, the earliest date will be the same as the earliest possible birth date of survey respondents and the latest date will be the last day of the survey data collection. This information will be used to assess the data quality of dates on cards and registers. If a card or register shows a date that is earlier than the earliest allowable date or later than the latest date, then the date is assumed to contain an error, and VCQI will replace the date with a tick mark.

Specify those dates with the following global values¹⁵ in the control program:

```
vcqi_global(EARLIEST_SVY_VACC_DATE_M, 1)  
vcqi_global(EARLIEST_SVY_VACC_DATE_D, 1)  
vcqi_global(EARLIEST_SVY_VACC_DATE_Y, 2013)  
  
vcqi_global(LATEST_SVY_VACC_DATE_M, 1)  
vcqi_global(LATEST_SVY_VACC_DATE_D, 1)  
vcqi_global(LATEST_SVY_VACC_DATE_Y, 2015)
```

These global values are not dose-specific or child-specific – they apply to all doses and all children. In the (fictional) survey described above, all the field data collection was conducted on January 1, 2015. To be 12-23 months old on January 1, 2015, a child must have been born sometime in the calendar year 2013. The earliest date of vaccination eligibility for that cohort was January 1, 2013. And the survey can consider vaccinations that occurred anytime in the two years from January 1, 2013 through December 31, 2014. Any vaccination date that falls outside the window specified by these scalars will be considered incorrect – vaccination dates that fall outside that window will be treated as tick marks on the card and will not be included in analyses that evaluate date of vaccination.

Eligible Ages for this Survey

The user should specify the age inclusion criteria for the survey using two global macros. If omitted, VCQI assumes that children had to be between 365 and 731 days of age (in case there was a leap year in the past two years). The minimum age of eligibility is used on a dose-by-dose basis to decide which children were age-eligible for which doses. This is particularly relevant in surveys that ask about doses administered in the second year of life.

```
vcqi_global(VCQI_RI_MIN_AGE_OF_ELIGIBILITY, 365)  
vcqi_global(VCQI_RI_MAX_AGE_OF_ELIGIBILITY, 731)
```

Records Sought at Health Centers

The user must specify whether vaccination records were sought at health facilities, by setting one and only one of the following global macros to 1 in the control program with code like the following:

```
vcqi_global(RI_RECORDS_NOT_SOUGHT, 1)  
vcqi_global(RI_RECORDS_SOUGHT_FOR_ALL, 0)
```

¹⁵ By convention, VCQI uses upper-case for global values except for schedule globals which are lower-case.

```
vcqi_global(RI_RECORDS_SOUGHT_IF_NO_CARD, 0)
```

This selection affects calculations for many of the outcomes. All RI indicators interpret this data the same, except for the MOV indicators (RI_QUAL_07B, RI_QUAL_08, RI_QUAL_09). The table below outlines how the data will be used.

Table 3-1. How RI_RECORDS inputs affect outcome calculation

RI_RECORDS_NOT_SOUGHT	RI_RECORDS_SOUGHT_FOR_ALL	RI_RECORDS_SOUGHT_IF_NO_CARD	Outcome is based on:	Notes
1	0	0	Card and History Only	Data from EPI registers is ignored, even if it is present in the RIHC dataset
0	1	0	Card or History or Register	There may be records with data from <u>both</u> card and register. In that case, the indicators set the final outcome to whichever record (card or register) is more favorable to the vaccination program. In other words, it gives the benefit of the doubt to the program and assumes that the source that documents a good outcome is correct.
0	0	1	Card and History if card was seen; Register and History for those without Cards	Only looks at data for the register for respondents who did not furnish a card. If the survey team happens to collect register data for a respondent who also has card data, the register data will be ignored.

For example, for the indicator for valid vaccination coverage (RI_COVG_02), if the card shows that the child received the dose too early to be valid, but the register date indicates that it was valid, then the outcome variables are listed below for each RI_RECORDS_SOUGHT option:

Table 3-2. How RI_RECORDS inputs affect the main valid dose outcome (RI_COVG_02)

RI RECORDS SOUGHT	Card Seen	Valid_dose_by_card	Valid_dose_by_register	Valid_dose_to_analyze (main outcome)
FOR_ALL	Yes	Invalid dose	Valid dose	Valid dose (from register)
NOT_SOUGHT	Yes	Invalid dose	n/a	Invalid dose (from card)
IF_NO_CARD	No	n/a	Valid dose	Valid dose (from register)
IF_NO_CARD	Yes	Invalid dose	Valid dose	Invalid dose (from card)

Changing which of these three inputs is set to 1 will affect the results of the final indicator. The final indicator is sometimes recorded with a variable that uses the suffix “to_analyze” and is often saved in a database with the abbreviation “_a_” in its filename. Chapter 6 contains detailed information about individual indicators.

3.4 Analysis Metadata and Options (goes in Block D)

Lists of Doses

The user must specify the names of the doses in the coverage analysis. This is accomplished in three steps in the control program.

First, specify the names of the single dose vaccines:

```
vcqi_global(RI_SINGLE_DOSE_LIST,  
            c("BCG", "HEPB", "OPV0", "IPV", "MCV1", "YF"))
```

The doses can be listed in any order and in either upper or lower case. The spelling of dose names must correspond exactly to those in the schedule globals and the date and tick variable names in the RI and RIHC datasets. When two or more dose names are in a dose list, they must be inside a concatenate c () call and separated by commas.

Next, specify the name of any two-dose vaccines:

```
vcqi_global(RI_MULTI_2_DOSE_LIST, "ROTA")
```

When defining a multi-dose list you should not specify numbers on the end of the dose names. Do not list ROTA1 and ROTA2; simply list ROTA and VCQI will know that there is a 1 and 2. If there are no two-dose vaccines in the schedule, set the value to be NULL.

Next, specify the name of any three-dose vaccines:

```
vcqi_global(RI_MULTI_3_DOSE_LIST, c("PENTA", "PCV", "OPV"))
```

Again, do not specify numbers on the end of the dose names in a multi-dose list. Do not list c ("PENTA1", "PENTA2", "PENTA3"); simply list PENTA and VCQI will know that there is a 1, 2, and 3 dose.

Note that VCQI in R can handle multi-dose lists with up to nine doses (RI_MULTI_9_DOSE_LIST). The Stata version can handle list with up to three.

It is recommended to list all doses here that appear in the survey questionnaire. In some cases, you might do a limited analysis of a short list of doses and you might be tempted to shorten these lists to include only the doses of interest so VCQI will run faster. But doing so may affect what you are able to calculate in that analysis. If you exclude a dose from this list, no information about its coverage will be available in any of the indicators. And it will be important, in particular, to list all doses from the survey when calculating indicators that summarize missed opportunities for simultaneous vaccination (RI_QUAL_07B, RI_QUAL_08, and RI_QUAL_09).

Defining the rows in output tables and figures

Most of the indicators generate tables and many of them generate bar charts that present results for various user-specified sub-groups of survey respondents. Those groups may be a flexible combination of geographic domains and demographic groups. See Annex B for details on how to use the VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT to control which strata and groups appear in the Excel output and the graphic figures.

Options for Individual Indicators (goes in Block F)

The user specifies a title, subtitle, and as many footnotes as they like for the Excel worksheet that holds the indicator output. These are specified using global values in the control program. For example, the following code specifies the title and two footnotes for the RI_COVG_01 indicator. It specifies an empty subtitle. (The footnotes are long and wrap onto several lines each in this document but they are each specified on a single long line of R code in the sample control program that comes with VCQI.)

```
vcqi_global(RI_COVG_01_TO_TITLE, "Crude Coverage")
vcqi_global(RI_COVG_01_TO_SUBTITLE, NA)

vcqi_global(RI_COVG_01_TO_FOOTNOTE_1, "Abbreviations:
CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper
Confidence Bound; DEFF=Design Effect; ICC=Intracluster Correlation
Coefficient")

vcqi_global(RI_COVG_01_TO_FOOTNOTE_2, "Note: This measure is a
population estimate that incorporates survey weights. The CI, LCB and
UCB are calculated with software that take the complex survey design
into account.")
```

Footnotes are numbered sequentially, and you may specify as many footnotes for a single measure as you wish. It is important not to skip any numbers. Begin with 1 and increase by 1 up to the number that you wish to list.

If you skip a number when specifying footnotes (e.g., 1, 2, 4, and 5) then VCQI will only list the footnotes from before the break (i.e., 1 and 2).

Several indicators include some automatic footnotes, based on user inputs. The logic that produces these is laid out in Chapter 7.

There are no special inputs or metadata required to calculate RI_COVG_01-02. Use the default titles and footnotes in the example control program, or specify new ones if you wish.

The remaining indicators each require the user to specify one or more global values to define precisely what to analyze and how. See the individual descriptions of the indicators in later sections of this document.

CHAPTER 4. ANALYZING TETANUS PROTECTION AT BIRTH (TT) SURVEYS

The R version of VCQI does not yet analyze data from TT surveys. At this time, you would need to use the Stata version of VCQI to accomplish that task.

CHAPTER 5. ANALYSIS OF POST-CAMPAIGN (SIA) SURVEYS

The R version of VCQI does not yet analyze data from SIA surveys. At this time, you would need to use the Stata version of VCQI to accomplish that task.

CHAPTER 6. VACCINATION COVERAGE QUALITY INDICATOR DESCRIPTIONS

The following pages list the individual indicators that are available in VCQI. Each contains an overview, a list and description of required global value inputs (if any), and a short list of outputs that the software generates. The VCQI files that you download include examples of control programs to run each of these indicators.

6.1 Weighted and unweighted analyses

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

Some of the analyses listed below are described as “Weighted: No”. These are usually analyses where only a subset of respondents will be in the denominator, so it could be confusing to draw conclusions about the overall population.

VCQI does not currently provide estimates of sampling error for unweighted analyses. The estimate is a description of a proportion observed in the sample and is reported without an estimate of uncertainty.

6.2 Analysis Counter

Block F of the control program sets a global value named ANALYSIS_COUNTER. It is required, and usually set to 1. In most control programs it will only be set once and never changed.

In the remainder of this chapter, you will note that the analysis counter appears in the names of many VCQI output files and worksheets.

In advanced analyses, the user can conduct sensitivity analyses by running an initial analysis and then changing some of the analysis parameters, changing the analysis counter and re-running the indicator. In the first run, the output files and tabs would list the value 1 for ANALYSIS_COUNTER and in the second run they would be named with the value 2 and would therefore not overwrite the first set of output. This can be accomplished in a single control program.

For example, one could explore how valid coverage changes if we allow a four-day “grace period” whereby we count a dose as valid if the child receives it up to four days before they were scheduled to do so. This can be done with code like the following:

```
# Initial run uses the usual schedule established in Block D

penta1_min_age_days      <- 42 # 6 weeks
penta2_min_age_days      <- 70 # 10 weeks
penta2_min_interval_days <- 28 # 4 weeks
penta3_min_age_days      <- 98 # 14 weeks
penta3_min_interval_days <- 28 # 4 weeks
```

```

# intervening code from Block E goes here
# intervening code from Block E goes here
# intervening code from Block E goes here

# This code in block F accomplishes the original analysis
# Tabular output goes to tab named "RI_COVG_02 1"
# Databases and plots have the ANALYSIS_COUNTER value 1 in filenames

vcqi_global(ANALYSIS_COUNTER, 1)
RI_COVG_02()

# Now re-run using a schedule with a grace period
# Tabular output goes to tab named "RI_COVG_02 2"
# Databases and plots have the ANALYSIS_COUNTER value 2 in filenames

vcqi_global(ANALYSIS_COUNTER, 2)
penta1_min_age_days      <- 38 # 6 weeks minus 4 days
penta2_min_age_days      <- 66 # 10 weeks minus 4 days
penta2_min_interval_days <- 24 # 4 weeks minus 4 days
penta3_min_age_days      <- 94 # 14 weeks minus 4 days
penta3_min_interval_days <- 24 # 4 weeks minus 4 days

RI_COVG_02()

```

This same sensitivity analysis could be accomplished using two CONTROL programs that send output to two different Excel files altogether. In that case, there is no need to change the value of ANALYSIS_COUNTER.

Some indicators use the ANALYSIS_COUNTER to open datasets from indicators that were run earlier, so it is best to experiment carefully with changing the ANALYSIS COUNTER. Note that the indicators in Table 6-1 rely on datasets constructed earlier. The value of ANALYSIS_COUNTER must be the same when the later indicator is run than it was when the earlier indicator was run. In most cases, VCQI will copy the output from the run when ANALYSIS_COUNTER was set to 1 and will put a warning in the VCQI Log. But in some cases, to do the sensitivity analysis, it may be necessary to change ANALYSIS_COUNTER and re-run several indicators or to use a R command to rename copies of earlier datasets. (E.g., to do a sensitivity analysis on RI_QUAL_08 with different inputs, it will be necessary to re-run RI_COVG_02 using the new value of ANALYSIS_COUNTER as well or to copy the dataset named RI_COVG_02_1 to a new dataset named RI_COVG_02_2.)

Chapter 6. VCQI Indicator Descriptions

Table 6-1 lists indicators that rely on output from other indicators.

Table 6-1. VCQI indicators that rely on others being run first

Indicators that use output from RI_COVG_01 (crude coverage)	Indicators that use output from RI_COVG_02 (valid coverage)
RI_COVG_02 RI_VCTC_01	RI_QUAL_07B RI_VCTC_01

6.3 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 according to card, register, history

Vaccines: Calculated for each dose

Time options: By the time of survey

Variations: By card

By history

By register

By card or history (for purpose of comparison with older surveys)

By card or register (i.e., by documented source)

By card or history or register

To analyze (depends on whether RI records were sought at health facilities, and for whom; see Chapter 3.)

User inputs: The dose list (see Chapter 3).

Whether RI records were sought at health facilities. See section 3.3 for a description of the three global macros that describe what was done at health facilities.

Control

Program

Command: RI_COVG_01 ()

Output: This indicator generates databases that summarize crude coverage:

Table 6-2. Naming convention for RI_COVG_01 databases

According to evidence from...	Database Name
Card	RI_COVG_01_<dose>_<analysis counter>_c_database.rds
Caretaker's Verbal History	RI_COVG_01_<dose>_<analysis counter>_h_database.rds
Card or History	RI_COVG_01_<dose>_<analysis counter>_ch_database.rds
Register	RI_COVG_01_<dose>_<analysis counter>_r_database.rds
Card or Register	RI_COVG_01_<dose>_<analysis counter>_cr_database.rds
Card or History or Register	RI_COVG_01_<dose>_<analysis counter>_chr_database.rds
Main Outcome to Analyze	RI_COVG_01_<dose>_<analysis counter>_a_database.rds

How the main outcome for crude coverage is calculated for each respondent depends on whether RI records were sought at health centers, and if so, for whom. This is indicated in the control program by setting one (and only one) of the RECORDS_SOUGHT global macros to 1. See section 3.3 for details on RECORDS_SOUGHT global values in RI Analysis.

Databases for weighted outcomes include the following output fields each outcome and stratum: estimated %, 2-sided 95% CI, 1-sided 95% lower confidence bound (LCB), 1-sided 95% upper confidence bound (UCB), Design Effect (DEFF), Intracluster correlation coefficient (ICC), N (unweighted), N (weighted), the number of clusters in the calculation, and the weighted number of persons with the outcome of interest.

The databases include detailed output fields for every dose and every outcome listed above in every stratum. See Annex D for a description of database contents.

This indicator makes two worksheets. The first worksheet is named RI_COVG_01_BRIEF <*analysis counter*>. It lists the point estimate and two-sided 95% confidence interval for crude coverage for each dose (using all sources of evidence available in the dataset).

The second worksheet is named: RI_COVG_01 <*analysis counter*>. For outcomes by card, history, and register it simply lists estimated % and the two-sided 95% CI. For the main outcome it lists estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted). Note that VCQI always constrains the coverage calculation to use a design effect that is ≥ 1 .

The plots generated by the indicator include one organ pipe plot of the main outcome per dose per stratum and one bar plot per dose summarizing the main crude coverage outcome.

The organ pipe plots are named

RI_COVG_01_<*analysis counter*>_opplot_<*dose*>_<*stratum id*>_<*stratum name*>.png

The bar plots are named RI_COVG_01_<*analysis counter*>_brplot_<*dose*>.png.

In addition to .png files, the user may specify that VCQI should also save the datasets used to create the plots. Users may re-create plots using those .rds dataset files later.

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

Notes: For BCG there is an additional outcome for coverage by scar stored in a database named RI_COVG_01_<*analysis counter*>_BCG_s_database.rds. And for BCG, evidence from the scar is counted in the card or history outcome, the card or history or register outcome and the ‘to analyze’ outcome.

If the survey did not ask for BCG evidence by scar then the RI_COVG_01 table will include scar columns that could be ignored or deleted and will include the words “or scar” in several column labels. In the future, it would be possible to add a user input to tell VCQI that the survey did not include a scar question and then the output could appear without alluding to scars. Alternatively, VCQI could check to see whether BCG scars were observed for any respondents. If not, the scar-related columns could be suppressed and the columns could be relabeled automatically.

RI_COVG_02: Valid coverage

Weighted: Yes

Denominator: Sum of weights for all respondents

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

Vaccines: Calculated for each dose

Time options: By the time of survey, or
By 12 months of ageVariations: By card
By register
By card or register
To analyze

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

User inputs: The dose list (see Chapter 3).

Whether RI records were sought at health facilities. See description above for RI_COVG_01.

The RI dose schedule (see Chapter 3).

Control

Program

Command: RI_COVG_02 ()

Output: This indicator generates databases that summarize valid coverage:

Table 6-3. Naming convention for RI_COVG_02 databases

According to evidence from...	Dataset name
Card	RI_COVG_02_<dose>_<analysis counter>_c_database.rds
Register	RI_COVG_02_<dose>_<analysis counter>_r_database.rds
Card or Register	RI_COVG_02_<dose>_<analysis counter>_cr_database.rds
Main Outcome to Analyze	RI_COVG_02_<dose>_<analysis counter>_a_database.rds
By age 1, according to card	RI_COVG_02_<dose>_<analysis counter>_ca1_database.rds
By age 1, according to register	RI_COVG_02_<dose>_<analysis counter>_ra1_database.rds
By age 1, according to card or register	RI_COVG_02_<dose>_<analysis counter>_cra1_database.rds
Main outcome for valid coverage by age 1	RI_COVG_02_<dose>_<analysis counter>_aa1_database.rds

How the main outcomes for valid coverage and valid coverage by age 1 are calculated for each respondent depends on whether RI records were sought at health centers, and if so, for whom. This is indicated in the control program by setting one (and only one) of the RECORDS_SOUGHT global values to 1. See section 3.3 for details on RECORDS_SOUGHT global values in RI Analysis.

The databases include detailed output fields for every dose and every outcome listed above in every stratum. See Annex D for a description of database contents.

This indicator makes two worksheets. The first is named: RI_COVG_02 <*analysis counter*>. For the outcomes by card, by register, and by card or register it simply lists estimated % and 95% CI. For the main outcomes it lists estimated %, 95% CI, LCB, UCB, DEFF, ICC, N (unweighted), N (weighted).

The second worksheet is named RI_COVG_02_BRIEF <*analysis counter*>. It lists the point estimate and two-sided 95% confidence interval for valid coverage for each dose.

The plots generated by the indicator include one organ pipe plot of the main outcome per dose per stratum and four bar plots per dose: one showing results for the main outcome for valid coverage another for valid coverage by age 1 and two (so-called *double barplots*) showing valid and crude coverage on the same figure for the main outcome and for valid coverage by age 1. Valid coverage appears in color and crude coverage appears in light grey. The indicator does not currently make organ pipe plots of any outcomes by age 1.

The organ pipe plots are named

RI_COVG_02_<*analysis counter*>_opplot_<*dose*>_<*stratum id*>_<*stratum name*>.png

The bar plots are named RI_COVG_02_<*analysis counter*>_brplot_<*dose*>_<*valid or age1 or double or age1_double*>.png.

The bar plots that show both valid and crude coverage on the same plot also have the word “double” in the filenames.

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*> <by 1 year of age>.”

Notes:

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.

6.4 RI_QUAL: RI Survey – Measures Related to quality of Services

RI_QUAL_07B: Valid coverage if there had been no missed opportunities for simultaneous vaccination (MOSV) and no early doses

Weighted: Yes

Denominator: Sum of weights for all respondents

Numerator: Sum of weights for all respondents who would have had a valid dose if every child received every dose they were due at every one of their date-documented vaccination visits

Vaccines: Calculate for each vaccine and dose in the MOV_OUTPUT_DOSE_LIST

User Input: `vcqi_global(MOV_OUTPUT_DOSE_LIST, c("bcg", "hepb", "opv0", "opv1", "opv2", "opv3", "penta1", "penta2", "penta3", "pcv1", "pcv2", "pcv3", "rota1", "rota2", "rota3", "ipv", "mcv1", "yf"))`

The MOV_OUTPUT_DOSE_LIST is the list of doses for which the user wishes to see MOV-related output. If the user does not specify the list, it will default to being equal to the RI_DOSE_LIST. Only doses that appear in the RI_DOSE_LIST may appear in the MOV_OUTPUT_DOSE_LIST but sometimes the analyst requests MOV output for a subset of the RI_DOSE_LIST.

Control

Program

Command: RI_QUAL_07B ()

This indicator uses output from RI_COVG_02, so that must be calculated first. It also uses output from calculate_MOV_flags, so that program must be called in the control program before running this indicator.

Output: This indicator examines the age of the child at each of their date-documented vaccination visits and calculates which doses they would have received if they received all the doses for which they were eligible. (No missing doses (MOSVs) and no early doses...each child vaccinated with all and only the doses they should have received.) It reports the % of children who would have had a valid dose of each if the immunization system had performed perfectly on the days that the child had documented vaccination visits.

This indicator produces a database for each dose in the MOV_OUTPUT_DOSE_LIST. Databases are R datasets named:

`RI_QUAL_07B_<analysis counter>_<dose>_database.rds`

See Annex D for a description of database contents.

The Excel worksheet for this indicator is named: `RI_QUAL_07B <analysis counter>`. Coverage is estimated for each dose in MOV_OUTPUT_DOSE_LIST and for each stratum.

Chapter 6. VCQI Indicator Descriptions

Each dose is summarized in two columns listing: estimated % and two-sided 95% CI. The right-most columns in the worksheet lists N (unweighted) and N (weighted).

How the outcome is calculated for each respondent depends on whether RI records were sought at health centers, and if so, for whom. This is indicated in the control program by setting one (and only one) of the RECORDS_SOUGHT global macros to 1.

Table 6-4. How RI_QUAL_07B uses RI_RECORDS inputs

RI_RECORDS_NOT_SOUGHT	RI_RECORDS_SOUGHT_FOR_ALL	RI_RECORDS_SOUGHT_IF_NO_CARD	Outcome is based on...	Notes
1	0	0	Card Only	Outcome is calculated using the vaccination date on the card
0	1	0	Card; if missing Card or missing Card dose date, then Register	For single-dose vaccines: Dates from register records are used to fill in missing dates on cards; if there is a date on the card, the date from the register is ignored, even if it yields a more favorable outcome than the date on the card. [This could be the topic of a future change.]
				For multi-dose vaccines: Dates from register records are used if the card does not contain any dates, or if the register records more doses for that vaccine series than the card does. [This also could be the topic of a future change to the software.]
0	0	1	Card if card was seen Register for those without cards	Outcome is calculated using the card for respondents who show vaccination cards, and using the register for those without cards, but whose documented vaccination records are collected from health centers

Plots include two bar plots per dose. The first shows what valid coverage would have been if there had been no MOVs and the second is a double bar plot that overlays the valid coverage results from RI_COVG_02 in light grey. The plot files are named RI_QUAL_07B_<analysis counter>_brplot_<dose or dose_double>.png.

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 every child received every dose for which they were eligible on every one of their date-documented vaccination visits.

Notes:

This indicator should be interpreted in light of the % of respondents who showed card or register. VCQI can only generate a positive outcome for this indicator if the respondent has some vaccination dates on their card or register record.

A note regarding RI_QUAL_08 and RI_QUAL_09

RI_QUAL_08 and _09 all summarize missed opportunities for simultaneous vaccination (MOSVs) in the survey dataset.

When interpreting their output 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 MOSV 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 not 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 worse results for the MOV indicators (i.e., more MOSVs). 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

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

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

User inputs:

```
vcqi_global(RI_QUAL_08_VALID_OR_CRUDE,<"CRUDE" or "VALID">)
```

See notes in the section before RI_QUAL_08 regarding CRUDE and VALID.

```
vcqi_global(MOV_OUTPUT_DOSE_LIST, c("bcg", "hepb", "opv0",
"opv1", "opv2", "opv3", "penta1", "penta2", "penta3",
"pcv1", "pcv2", "pcv3", "rotal1", "rota2", "rota3", "ipv",
"mcv1", "yf"))
```

The MOV_OUTPUT_DOSE_LIST is the list of doses for which the user wishes to see MOV-related output. If the user does not specify the list, it will default to being equal to the RI_DOSE_LIST. Only doses that appear in the RI_DOSE_LIST may appear in the MOV_OUTPUT_DOSE_LIST but sometimes the analyst requests MOV output for a subset of the RI_DOSE_LIST.

Control

Program

Command: RI_QUAL_08()

Output: This indicator produces a database for each dose in the MOV_OUTPUT_DOSE_LIST. The database is named: RI_QUAL_08_<analysis counter>_<dose>_database.rds. It lists the number of visits where children were eligible for the dose in question, and the % of those visits where the child had a MOV for every stratum.

The indicator also produces a database that is not dose-specific, named RI_QUAL_08_<analysis counter>_any_database.rds. It lists the total number of visits where a child was eligible for 1+ doses and the percent of those visits where the child had 1+ MOVs.

The indicator also produces a database that is not dose-specific, named RI_QUAL_08_<analysis counter>_rate_database.rds. It lists the total number of visits where a child was eligible for 1+ doses and average number of MOVs per visit.

See Annex D for a description of database contents.

How the outcome is calculated for each respondent depends on whether RI records were sought at health centers, and if so, for whom. This is indicated in the control program by setting one (and only one) of the RECORDS_SOUGHT global values to 1. See

section RI_QUAL_07B for details on how the RECORDS_SOUGHT global values differ on MOV calculations.

The Excel worksheet for this indicator is named: RI_QUAL_08 <analysis counter>. It holds outcomes for all doses in a single very wide table, where each dose has two columns: unweighted number of eligible visits, and % of those visits with MOV. The aggregated data over “all doses” show those two columns plus three additional columns which are the sum of total_movs, the sum of all eligible visits, and the rate of total_movs / total eligible visits.

The indicator generates one plot per dose showing the unweighted % of eligible visits that yielded an MOV from each stratum. The plot files are named RI_QUAL_08_<analysis counter>_uwplot_<dose>.png.

The indicator also generates an overall plot showing the % of visits that had 1+ MOVs for any dose. That plot is named RI_QUAL_08_<analysis counter>_uwplot_any.png.

In addition to .png files, the user may specify that VCQI should also save the datasets used to create the plots. Users may re-create plots using those .rds dataset files later.

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.” (E.g. 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.”)

Notes:

To see the difference between the CRUDE and VALID analysis, simply run the indicator twice. This can be accomplished with the following syntax in the control program:

```
vcqi_global(ANALYSIS_COUNTER, 1)
vcqi_global(RI_QUAL_08_VALID_OR_CRUDE, "CRUDE")
RI_QUAL_08()
vcqi_global(ANALYSIS_COUNTER, 2)
vcqi_global(RI_QUAL_08_VALID_OR_CRUDE, "VALID")
RI_QUAL_08()
```

This will result in two sets of databases and figures, one with the ANALYSIS_COUNTER value of 1 in the filenames and the other with the ANALYSIS_COUNTER value of 2 in the filenames. The tabular output will be summarized in two worksheets named RI_QUAL_08 1 and RI_QUAL_08 2. The crude and valid worksheets will have different footnotes.

RI_QUAL_09: Percent of children with missed opportunity for simultaneous vaccination

Weighted: No

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

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

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

User inputs:

```
vcqi_global(RI_QUAL_09_VALID_OR_CRUDE,<"CRUDE" or "VALID">)
```

See notes section in the section before RI_QUAL_08 regarding CRUDE and VALID.

```
vcqi_global(MOV_OUTPUT_DOSE_LIST, c("bcg", "hepb", "opv0",
"opv1", "opv2", "opv3", "penta1", "penta2", "penta3",
"pcv1", "pcv2", "pcv3", "rota1", "rota2", "rota3", "ipv",
"mcv1", "yf"))
```

The MOV_OUTPUT_DOSE_LIST is the list of doses for which the user wishes to see MOV-related output. If the user does not specify the list, it will default to being equal to the RI_DOSE_LIST. Only doses that appear in the RI_DOSE_LIST may appear in the MOV_OUTPUT_DOSE_LIST but sometimes the analyst requests MOV output for a subset of the RI_DOSE_LIST.

Control

Program

Command: RI_QUAL_09()

Output: This indicator produces a database for each dose in the MOV_OUTPUT_DOSE_LIST.
Each is named: RI_QUAL_09_<analysis counter>_<dose>_database.rds

The database lists output for every level 4 stratum, documenting the number of children who had a recorded date of birth who had 1+ documented vaccinations at an age when they were eligible to receive the dose in question, the number of children who experienced 1+ missed opportunities for the dose, the number of children whose missed opportunities were corrected, and the number of children whose missed opportunity was uncorrected at the time of the survey.

It also produces a database describing the proportion of respondents who experienced 1+ MOSVs for any dose. That database is named:

RI_QUAL_09_<analysis counter>_anydose_database.rds

See Annex D for a description of database contents.

How the outcome is calculated for each respondent depends on whether RI records were sought at health centers, and if so, for whom. This is indicated in the control program by setting one (and only one) of the RECORDS_SOUGHT global values to 1. See section RI_QUAL_07B for details on how the RECORDS_SOUGHT global values differ on MOSV calculations.

The Excel worksheet for this indicator is named: RI_QUAL_09 <analysis counter>.

Reports outcomes for all doses in a single very wide table, where each dose has four columns:

1. The number of children who had at least one visit where they were eligible to receive the dose (this is the number of children for which the indicator is either 0 or 1)
2. The % of those children who had 1+ MOSVs for that dose (this measure)
3. The percent of children with eligible visits who had uncorrected MOSVs
4. The percent of children with eligible visits who had corrected MOSVs.

The latter two figures add up to the percent calculated in this measure.

The data for all doses combined consist of five columns:

1. A total number of children who had dob data and 1+ eligible visits;
2. The percent who had 1+ MOSVs for 1+ doses;
3. The percent for whom all MOSVs were corrected;
4. The percent for whom none of the MOSVs were corrected, and
5. The percent for whom some, but not all of the MOSVs were corrected.

Column 5 is equal to 2 minus 3 minus 4.

The indicator generates two plots for each dose: one that shows the unweighted proportion of respondents who had eligible visits that experienced 1+ MOSVs, and another that shows the proportion of children whose MOSVs were eventually corrected.

The indicator generates two additional plots: One that shows the % of respondents who were eligible for any dose, who experienced 1+ MOSVs, and one that shows the % of respondents who had 1+ MOSVs and later had all of their MOSVs corrected.

The plot files are named RI_QUAL_09_<analysis counter>_uwplot_<dose or anydose>.png and

RI_QUAL_09_<analysis counter>_uwplot_<dose or anydose>_cor.png

To browse output from RI_QUAL_09 visually and to summarize the time-to-MOSV correction, use the Missed Opportunities for Simultaneous Vaccination R-Shiny Application described in Section F.1 of Annex F.

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

Chapter 6. VCQI Indicator Descriptions

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+ MOSVs for any doses, X% had all of their MOSVs 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+ MOSVs for any doses, X% had all of their MOSVs 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+ MOSVs for any doses, X% had some but not all of their MOVs corrected by the time of the survey.”

Notes:

To see the difference between the CRUDE and VALID analysis, simply run the indicator twice. This can be accomplished with the following syntax in the control program:

```
vcqi_global(ANALYSIS_COUNTER, 1)
vcqi_global(RI_QUAL_09_VALID_OR_CRUDE, "CRUDE")
RI_QUAL_09()
vcqi_global(ANALYSIS_COUNTER, 2)
vcqi_global(RI_QUAL_09_VALID_OR_CRUDE, "VALID")
RI_QUAL_09()
```

This will result in two sets of databases and figures, one with the ANALYSIS_COUNTER value of 1 in the filenames and the other with value of 2 in the filenames. The tabular output will be summarized in two worksheets named RI_QUAL_09 1 and RI_QUAL_09 2. The crude and valid worksheets will have different footnotes.

6.5 RI_VCTC: RI Survey – Vaccination Coverage and Timeliness Charts (VCTC)

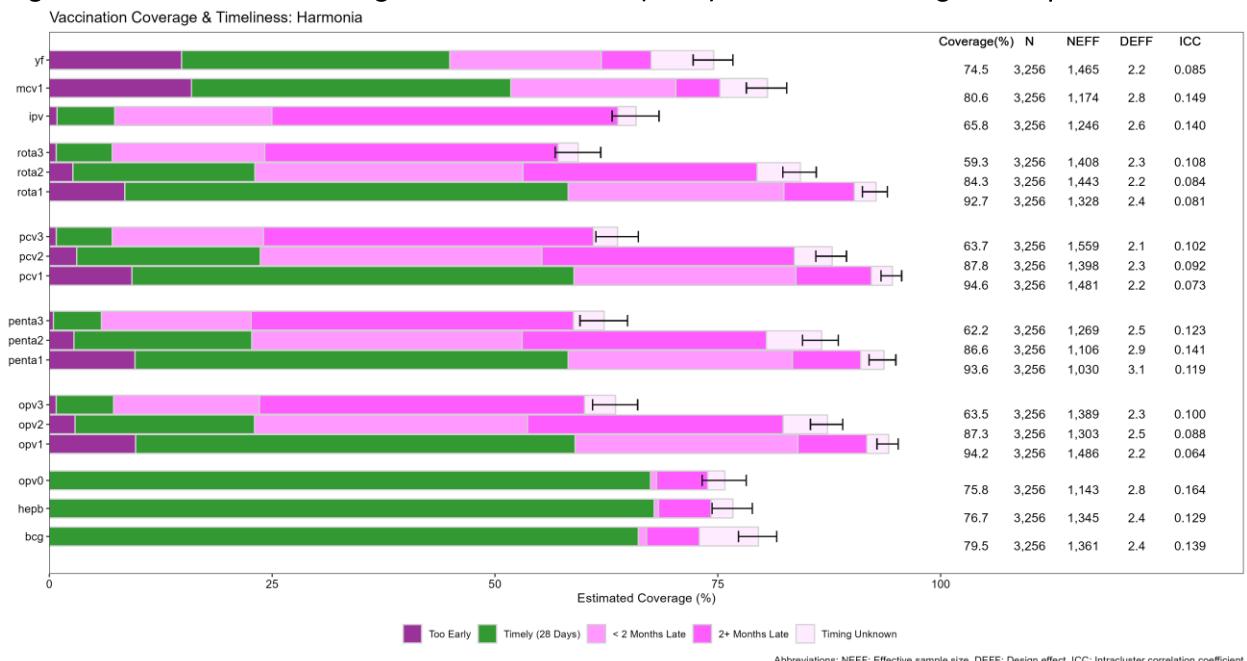
RI_VCTC_01: Vaccination Coverage and Timeliness Stacked Bar Charts

The figure below is an example of a stacked bar chart to summarize vaccination coverage and timeliness results from:

- RI_COVG_01 (crude coverage – or % with any evidence of vaccination) using age at vaccination calculated in
- RI_COVG_02 (timeliness) to categorize the continuous timeliness information from vaccination dates

Each figure (chart) represents data for a single stratum. Each dose is represented in a single row. The coverage estimates are weighted. The length of the full bar agrees precisely with coverage estimates from RI_COVG_01, but the bar is divided into categories of timeliness. This figure shows the default categories. The VCQI user may customize the categories. Each dose is annotated with optional tabular summaries of coverage (%), unweighted sample size (N), effective unweighted sample size (NEFF), design effect (DEFF) and intracluster correlation coefficient (ICC).

Figure 6-1. Vaccination Coverage & Timeliness Chart (VCTC) for Harmonia using default parameters



Chapter 6. VCQI Indicator Descriptions

Weighted: Yes

Denominator: Sum of weights for all eligible respondents in the stratum

Numerator: Sum of weights for respondents whose vaccination evidence indicates that they received the dose in one of the timeliness categories. Timeliness is measured with respect to the minimum age for a valid dose as specified in the vaccination schedule. The software assigns each child to the appropriate category based the age at which they received the dose.

Vaccines: The user specifies the list (and order) of doses to include in the charts. In the figure above the doses are grouped by antigen. It would also be possible to order them by schedule age (put OPV1, PENTA1, PCV1, and ROTA1 beside each other, followed by dose 2 of all 4 antigens, followed by dose 3).

User inputs:

Conceptually, each dose is represented by a bar graph that stacks several tiles from left-to-right. Each tile represents a category of timeliness. In the example above, the left-most tile represents children who were vaccinated too early (if any). The second tile (in green) represents those who were vaccinated within 28 days of the age when they were scheduled to receive that dose. The third tile represents those who were vaccinated between 28 and 56 days late, so < 2 months late. The fourth tile represents those who were vaccinated between more than 56 days late, and the fifth tile quantifies those who have some evidence of vaccination, but that evidence does not allow VCQI to calculate the child's age at the time of vaccination. In other words, the evidence is from a tick mark, or an incomplete date or from caregiver recall.

Nearly every aspect of the figure can be modified by the user, including the timeliness categories, tile colors, dose order, labels, line colors, etc. These input parameters are described below in three tables.

These charts summarize a lot of information and VCQI provides a lot of flexibility to customize the charts, so this indicator has more input parameters than any other. If the user is happy with defaults, they do not need to specify any input parameters at all. VCQI will use default parameters to make a sensible chart.

If the user wishes to change some of the default characteristics, they may copy the parameter file that holds the settings listed in the second table below, make some changes, and save that parameter file in the VCQI_OUTPUT_FOLDER with the filename as described below.

Finally, if the user wishes to define new categories of timeliness for some or all doses, they will need to use the parameters listed in the third table below. These parameters are demonstrated in a block of code that is commented out in the parameter file that is included with the VCQI software.

Feel free to correspond with the VCQI developers with questions or with suggestions for additional features of these very informative plots.

Most inputs in the R version of VCQI act just like their corresponding ones in the Stata version. The one input with noticeable difference is `TIMELY_LEGEND_ORDER`. This single input in the R version takes the role of two sets of inputs in the Stata version: `TIMELY_DT_LEGEND_ORDER` and `TIMELY_CD_<DOSE>_LEGEND_ORDER`, thus it shows up in both Table 6-18 and Table 6-19. See the examples provided below for how to use this new input that is unique to the R version of VCQI.

This first table lists the input parameters that the user usually specifies explicitly, in the control program via global values, before calling the `RI_VCTC_01` program.

Chapter 6. VCQI Indicator Descriptions

Table 6-5. RI_VCTC_01 inputs that are usually set in the control program

Global Macro	Acceptable Values	Description	Notes
RI_VCTC_01_LEVELS	Any combination of the integers 1 or 2 or 3	VCQI will produce a separate chart for every stratum in the levels specified.	Defaults to 3. VCQI does not currently have the capability to make VCTCs for each row in the Level4 layout dataset; it only makes them for strata defined by the variables level1id, level2id, or level3id.
TIMELY_DOSE_ORDER	Dose names, lower-case, formatted like dose list global values	Dose names, usually listed with lower-case letters, listed in the order you want the bars to appear from bottom-to-top in the chart	Defaults to the RI_DOSE_LIST set by VCQI, which will usually not be in a helpful order. The user may use this parameter to override the default and specify a subset or a different order.
TIMELY_Y_COORDS	Positive numbers	Y-coordinates of the dose bars	Defaults to evenly-spaced bars at Y= 1, 2, 3, ... up to the number of doses in the chart. Users sometimes specify unequally spaced values to visually <i>group</i> doses in a series together. See the example in the control program template.

The next table lists the default global value parameters that are stored with the VCQI programs in a function called `VCTC_default_globals()`.

If the user wishes to change any of these default parameters, before running VCQI, make edits to a file named “`globals_for_timeliness_plots.R`” which is provided with the demo control program; save the file and edit the control program to define the global value `VCTC_globals_path` with the path to the .R file.

Note that the VCQI RI control program template over-rides one default by stipulating:

```
vcqi_global(TIMELY_BARWIDTH, 6.7)
```

The user may over-ride other defaults in a similar fashion.

Table 6-6. RI_VCTC_01 inputs that are often unchanged in globals_for_timeliness_plots.do

Global Values	Acceptable Values	Description	Notes
TIMELY_N_DTS	Positive integer that is greater than 1	Number of default tiles (or timeliness categories)	Current default is 5 for these five categories: 1 – Too early 2 – Timely (28 days) 3 - < 2 months late 4 – 2+ months late 5 – Timing unknown
TIMELY_DT_UB_1 TIMELY_DT_UB_2 ... Up to The number of TIMELY_N_DTS minus 1	0 or positive integer	DT_UB stands for “default tile upper bound” – these bounds are the upper bound of the timeliness categories specified in days relative to when the dose is scheduled. You specify this parameter for integers 1 up to TIMELY_N_DTS – 1	The category will capture doses received <u>before</u> the scheduled age of vaccination plus this upper bound. So for default category 1, the value is 0 days: the category captures doses received before the dose was scheduled. The value for tile _2 is 28 so it captures doses given before 28 days had passed. For _3 it is 56 so it captures doses given before the scheduled age plus 56 days. For _4 it is 100000 days, so it captures all doses given at all ages from 12-23m. We do not need to specify a bound for tile _5 because that final tile represents respondents whose timing is not known.
TIMELY_DT_COLOR_1 TIMELY_DT_COLOR_2 ... Up to TIMELY_N_DTS	Valid R color	Color of the default tiles	Either a valid color from colors() or a valid hex color code.
TIMELY_DT_LCOLOR_1 TIMELY_DT_LCOLOR_2 ... Up to TIMELY_N_DTS	Valid R color	Color of the line that forms the tile border	Either a valid color from colors() or a valid hex color code.

Chapter 6. VCQI Indicator Descriptions

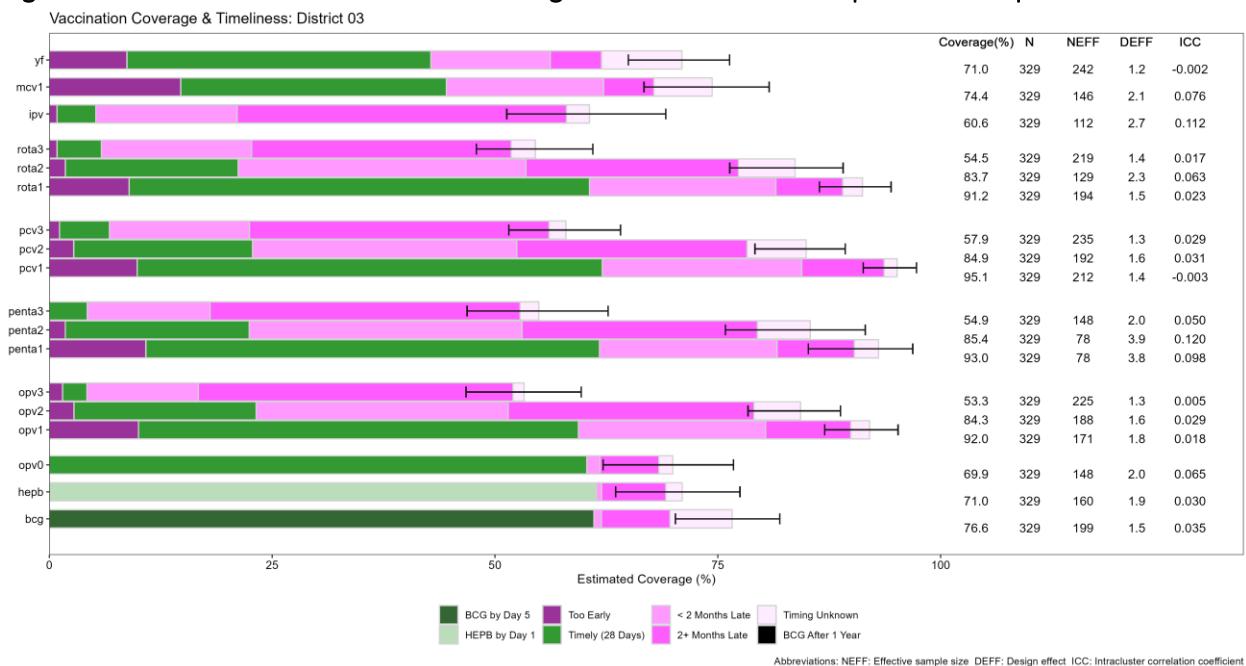
Global Values	Acceptable Values	Description	Notes
TIMELY_DT_LABEL_1 TIMELY_DT_LABEL_2 ... Up to TIMELY_N_DTS	String	String that appears in Excel table to label the category	
TIMELY_DT_LEGEND_LABEL_1 TIMELY_DT_LEGEND_LABEL_2 ... Up to TIMELY_N_DTS	String	String that appears in plot legend for this category	May differ (sometimes is more brief) than the corresponding TIMELY_DT_LABEL_n
TIMELY_LEGEND_ORDER	String	Which legend should be listed in below the graph, and in what order	The default setting is: c("DT_1", "DT_2", "DT_3", "DT_4", "DT_5")
TIMELY_XLABEL_SIZE	Positive number	Size of font listing numbers along the x axis	10 is the default
TIMELY_XLABEL_COLOR	Valid R text color	X label text color	Either a valid color from colors() or a valid hex color code. "black" is the default
TIMELY_YLABEL_SIZE	Positive number	Size of the numbers along the x axis	10 is the default
TIMELY_YLABEL_COLOR	Valid R text color	Y label text color	Either a valid color from colors() or a valid hex color code. "black" is the default
TIMELY_BARWIDTH	Positive number	Bar width	0.67 is the default, so the bar will be about twice as wide as the space between bars, if the bars fall on the integer y = 1, 2, 3, etc. values.
TIMELY_CI_LCOLOR	Valid R color	Confidence interval (CI) line color	Either a valid color from colors() or a valid hex color code. "grey8" is the default
TIMELY_CI_LWIDTH	Positive number	Line cap marker size	4 is the default

Global Values	Acceptable Values	Description	Notes
TIMELY_TEXTBAR_ORDER	Set of some or all of the following strings, in any order, separated by comma: c("COVG", "N", "NEFF", "DEFF", "ICC")	Which elements should be listed in tabular form, and in what order	Default is: c("COVG", "N", "NEFF", "DEFF", "ICC")
TIMELY_TEXTBAR_X_COVG	Positive number	X coordinate for COVG string, if it is listed in the TIMELY_TEXTBAR_ORDER	104 is the default
TIMELY_TEXTBAR_X_N	Positive number	X coordinate for N string, if it is listed in the TIMELY_TEXTBAR_ORDER	108 is the default
TIMELY_TEXTBAR_X_NEFF	Positive number	X coordinate for NEFF string, if it is listed in the TIMELY_TEXTBAR_ORDER	116 is the default
TIMELY_TEXTBAR_X_DEFF	Positive number	X coordinate for DEFF string, if it is listed in the TIMELY_TEXTBAR_ORDER	122 is the default
TIMELY_TEXTBAR_X_ICC	Positive number	X coordinate for ICC string, if it is listed in the TIMELY_TEXTBAR_ORDER	128 is the default
TIMELY_TEXTBAR_LABEL_COVG _N _NEFF _DEFF _ICC	Strings	Short string to appear at the top of the tabular columns	Defaults are: Coverage (%) N NEFF DEFF ICC
TIMELY_XSCALE_MAX	Positive number	x-coordinate of the far right end of the tabular output	134 is the default

Chapter 6. VCQI Indicator Descriptions

Global Values	Acceptable Values	Description	Notes
TIMELY_TEXTBAR_COLOR_COVG_N_NEFF_DEFF_ICC	Valid R text colors	Color of the font for tabular output	Either a valid color from colors() or a valid hex color code. "black" is the default
TIMELY_TEXTBAR_COVG_DEC_DIGITS	Non-negative integer, usually 1 or 0	Number of digits to show after the decimal point for tabular coverage estimates	1 is the default
TIMELY_TEXTBAR_COVG_ICC_DIGITS	Non-negative integer, usually 3	Number of digits to show after the decimal point for tabular ICC estimates	3 is the default
TIMELY_PLOT_NOTE	Strings	The footnote of the plot.	Default is: "Abbreviations: NEFF: Effective sample size DEFF: Design effect ICC: Intraclass correlation coefficient".
TIMELY_PLOT_WIDTH	Positive number	The width of the saved .png file; unit is inch.	15 is the default
TIMELY_PLOT_HEIGHT	Positive number	The height of the saved .png file; unit is inch.	8 is the default

Figure 6.2 below shows a VCTC that specifies three new default timeliness categories: Two for the dose BCG and one for HEPB. For OPV doses, it uses the default set of categories. Note that the legend lists the new categories.

Figure 6-2. VCTC for Harmonia District 03 using customized BCG and HepB timeliness parameters

An additional set of parameters must be defined to specify customized timeliness categories and colors and legend entries and integrate them with the default categories. The parameters associated with *default tiles* had the letters _DT_ in their names. Those associated with *customized doses* have the letters _CD_ in their names. The parameter file named “globals_for_timeliness_plots.R” delivered with the VCQI demo control program contains a set of customized dose definitions for BCG and HEPBO in a block of code that is commented out. They are listed below the table should serve as illustrative examples.

Table 6-7. RI_VCTC_01 inputs to customize timeliness categories for individual doses

Global Value	Acceptable Values	Description	Notes
TIMELY_CD_LIST	Dose names list formatted like TIMELY_DOSE_ORDER	List of doses that will not use the default tile definitions	
TIMELY_CD_<DOSE>_NTILES	Positive integer that is greater than 1	Number of tiles in the stacked bar for <DOSE> (specify for each dose in TIMELY_CD_LIST)	

Global Value	Acceptable Values	Description	Notes
TIMELY_CD_<DOSE>_UB_1	0 or positive number	Set of upper-bound definitions to specify the timeliness categories, specified in the same manner as TIMELY_DT_UB_1, etc.	Specify values for UB_1, UB_2, etc. all the way up thru TIMELY_CD_<DOSE>_N TILES – 1. The final (right-most) tile in <u>every</u> stacked bar will represent children whose timing is unknown.
TIMELY_CD_<DOSE>_COLOR_1	Valid R color	Tile color specified in the same manner as TIMELY_DT_COLOR_1, etc.	Either a valid color from colors() or a valid hex color code.
TIMELY_CD_<DOSE>_LCOLOR_1	Valid R color	Tile border line color specified in the same manner as TIMELY_DT_LCOLOR_1, etc.	Either a valid color from colors() or a valid hex color code.
TIMELY_CD_<DOSE>_LABEL_1	String	Specified in the same manner as TIMELY_DT_LABEL_1, etc.	
TIMELY_CD_<DOSE>_LEGEND_LABEL_1	String	Specified in the same manner as TIMELY_DT_LEGEND_LABEL_1, etc.	
TIMELY_LEGEND_ORDER	String	Which legend should be listed in below the graph, and in what order	An Example is: c("CD_BCG_1","CD_HEPB_1","DT_1","DT_2","DT_3","DT_4","DT_5","CD_BCG_4")

And note that if you want any of the customized dose tiles to appear in the legend then you need to specify updated value for `TIMELY_LEGEND_ORDER` as shown in the example below (**Examples of customized dose tile parameters**). In the “globals_for_timeliness_plots.R” file, there is also a line of commented code below the default setting code demonstrating how to set the value of `TIMELY_LEGEND_ORDER` with customized doses.

Examples of default tile parameters

The default parameters make a stack of up to five *tiles* per dose, as specified by `TIMELY_N_DTS` below. The text below shows how they are defined and describes the logic of the parameters.

```
assign("TIMELY_N_DTS", 5, envir = .GlobalEnv)
```

```
assign("TIMELY_DT_UB_1", 0, envir = .GlobalEnv)
assign("TIMELY_DT_COLOR_1", "#993399", envir = .GlobalEnv)
assign("TIMELY_DT_LABEL_1", "Too Early", envir = .GlobalEnv)
```

The upper bound on default tile 1 is 0 days. That means the first tile represents children who were vaccinated before the (scheduled age plus 0 days) which means they were vaccinated too early. They will appear in a tile that is colored "#993399" and with a label that says "Too Early". That label will appear first in the legend.

```
assign("TIMELY_DT_UB_2", 28, envir = .GlobalEnv)
assign("TIMELY_DT_COLOR_2", "#339933", envir = .GlobalEnv)
assign("TIMELY_DT_LABEL_2", "Timely (28 days)", envir = .GlobalEnv)
```

The upper bound on default tile 2 is 28 days, so the second tile represents children who were vaccinated before the (scheduled age plus 28 days). They will appear in a tile that is colored "#339933" with a label that says 'Timely (28 Days)'. That label will appear second in the legend.

```
assign("TIMELY_DT_UB_3", 56, envir = .GlobalEnv)
assign("TIMELY_DT_COLOR_3", "#FF99FF", envir = .GlobalEnv)
assign("TIMELY_DT_LABEL_3", "< 2 Months Late", envir = .GlobalEnv)
```

The upper bound on default tile 3 is 56 days, so it represents children who were vaccinated before the (scheduled age plus 56 days). They will appear in a tile that is colored "#FF99FF" with a label that says '< 2 Months Late'. That label will appear third in the legend.

```
assign("TIMELY_DT_UB_4", 100000, envir = .GlobalEnv)
assign("TIMELY_DT_COLOR_4", "#FF5CFF", envir = .GlobalEnv)
assign("TIMELY_DT_LABEL_4", "2+ Months Late", envir = .GlobalEnv)
```

The upper bound on the fourth tile is 100000 days so it represents children who were vaccinated before the (scheduled age plus 100000 days). The oldest child in this survey is two years (~730 days) so this tile will represent all the children for whom we know the age at vaccination, who were vaccinated 56+ days after the scheduled age. They will appear in a tile that is colored "#FF5CFF" with a label that says '2+ Months Late'. That label will appear fourth in the legend.

```
assign("TIMELY_DT_COLOR_5", "#FFEBFF", envir = .GlobalEnv)
assign("TIMELY_DT_LEGEND_LABEL_5", "Timing Unknown", envir =
.GlobalEnv)
```

Note that there is no upper bound for tile 5. Many features of these charts are customizable, but one feature is always the same. The last tile in the stack represents children whose vaccination timing is unknown. The user specifies the aesthetic properties of this tile (color, border line properties) and its label and order in the legend, but the definition of what it represents is always the difference between the next-to-last tile (all children whose vaccination age is known) and all the children with evidence of vaccination.

According to the parameters listed above, this tile will appear with color "#FFEBFF" and a label that says 'Timing Unknown'. The label will appear fifth in the legend.

The default setting of TIMELY_LEGEND_ORDER looks like:

Chapter 6. VCQI Indicator Descriptions

```
assign("TIMELY_LEGEND_ORDER", c("DT_1","DT_2","DT_3","DT_4","DT_5"),  
envir = .GlobalEnv)
```

Examples of customized dose tile parameters

Figure 6.2 shows customized tiles for BCG and HEPB. Those were defined using the parameters listed below.

Modifications for BCG are as follows:

1. There is no 'Too Early' tile because it is due at birth – so cannot be early.
2. The definition of timeliness is that it must be received on day 0-4 of life, so before day 5.
3. Because this is a new category of timely, assign the tile a new shade of green.
4. Assign a clear label and list it in position 1 in the legend. (This means the default tile legend entries will need to shift.)
5. And for BCG we are especially interested to see if any doses are given after the age of 1 year, so assign a new tile to show those vaccinated between the ages of 365 days and 1000 days.
6. The tile that shows those children, if any, will be black. (There are no such children in Figure 6.2 above.)

Modifications for HEPB include:

7. No 'Too Early' tile is needed because the dose is due at birth.
8. The dose is timely if received within 24 hours of birth, so let's define this as day 0 or 1 of life (i.e., before age 2 days).
9. Assign this unique category of timely a new green color.

If you have difficulty specifying customized dose definitions, contact GetVCQIHelp@biostatglobal.com.

```
# Define customized tiles for BCG & HEPB

assign("TIMELY_CD_LIST", c("bcg", "hepb"), envir = .GlobalEnv)
# customized definitions for BCG & HEPB

# Note that the dose names appear in all of the customized parameters below

assign("TIMELY_CD_BCG_NTILES", 5, envir = .GlobalEnv) # BCG still has 5 tiles

# First tile is for given <= target age (0 days) plus 5 days
assign("TIMELY_CD_BCG_UB_1", 5, envir = .GlobalEnv)
# Second is for given < 2 months late
assign("TIMELY_CD_BCG_UB_2", 56, envir = .GlobalEnv)
# Third is for given 2+ months late but within a year
assign("TIMELY_CD_BCG_UB_3", 365, envir = .GlobalEnv)
# Fourth is for doses given after age 1 year
assign("TIMELY_CD_BCG_UB_4", 100000, envir = .GlobalEnv)

#use a dark green for this special BCG timely category
assign("TIMELY_CD_BCG_COLOR_1", "#336633", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_COLOR_2", "#FF99FF", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_COLOR_3", "#FF5CFF", envir = .GlobalEnv)
#very late BCG shows in a BLACK bar
assign("TIMELY_CD_BCG_COLOR_4", "black", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_COLOR_5", "#FFEBFF", envir = .GlobalEnv)
```

Chapter 6. VCQI Indicator Descriptions

```

assign("TIMELY_CD_BCG_LCOLOR_1", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LCOLOR_2", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LCOLOR_3", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LCOLOR_4", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LCOLOR_5", "lightgrey", envir = .GlobalEnv)

assign("TIMELY_CD_BCG_LABEL_1", "BCG by day 5", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LABEL_2", "< 2 Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LABEL_3", "2+ Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LABEL_4", "After 1 Year (BCG only)", envir =
.GlobalEnv)
assign("TIMELY_CD_BCG_LABEL_5", "Timing Unknown", envir = .GlobalEnv)

assign("TIMELY_CD_BCG_LEGEND_LABEL_1", "BCG by Day 5", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LEGEND_LABEL_2", "< 2 Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LEGEND_LABEL_3", "2+ Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_BCG_LEGEND_LABEL_4", "BCG After 1 Year", envir =
.GlobalEnv)
assign("TIMELY_CD_BCG_LEGEND_LABEL_5", "Timing Unknown", envir = .GlobalEnv)

*****
# Parameters for HEPB tiles

assign("TIMELY_CD_HEPB_NTILES", 4, envir = .GlobalEnv)

# HEPB is timely if given on day 0 or 1
assign("TIMELY_CD_HEPB_UB_1", 2, envir = .GlobalEnv)
# < 2 months late
assign("TIMELY_CD_HEPB_UB_2", 56, envir = .GlobalEnv)
# 2+ months late
assign("TIMELY_CD_HEPB_UB_3", 1000, envir = .GlobalEnv)

#use yet another green for HEPB timely
assign("TIMELY_CD_HEPB_COLOR_1", "#BCDDBC", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_COLOR_2", "#FF99FF", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_COLOR_3", "#FF5CFF", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_COLOR_4", "#FFEBFF", envir = .GlobalEnv)

assign("TIMELY_CD_HEPB_LCOLOR_1", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LCOLOR_2", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LCOLOR_3", "lightgrey", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LCOLOR_4", "lightgrey", envir = .GlobalEnv)

assign("TIMELY_CD_HEPB_LABEL_1", "Timely (within 1 day)", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LABEL_2", "< 2 Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LABEL_3", "2+ Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LABEL_4", "Timing Unknown", envir = .GlobalEnv)

assign("TIMELY_CD_HEPB_LEGEND_LABEL_1", "HEPB by Day 1", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LEGEND_LABEL_2", "< 2 Months Late", envir =
.GlobalEnv)
assign("TIMELY_CD_HEPB_LEGEND_LABEL_3", "2+ Months Late", envir = .GlobalEnv)
assign("TIMELY_CD_HEPB_LEGEND_LABEL_4", "Timing Unknown", envir = .GlobalEnv)

```

```
*****  
#Legend order with customized doses:  
assign("TIMELY_LEGEND_ORDER",  
c("CD_BCG_1","CD_HEPB_1","DT_1","DT_2","DT_3","DT_4","DT_5","CD_BCG_4"),  
envir = .GlobalEnv)
```

Chapter 6. VCQI Indicator Descriptions

Control

Program

Command: RI_VCTC_01()

Output: This indicator produces a coverage and timeliness chart (or plot) for each stratum in RI_VCTC_01_LEVELS. It plots doses in TIMELY_DOSE_ORDER (from bottom to top).

Plots are saved in a folder named PLOTS_VCTC with the following naming convention:
VCTC_01_<analysis counter>_level_<level id>_id_<id number within the
level>_<nation/zone/stratum name>.png

The indicator also produces a worksheet named RI_VCTC_01_<analysis counter> that lists the quantitative aspects of each tile in the chart. In addition to some relevant labels, it lists the percent of children in the stratum represented in each tile (in columns labeled: 'Pct width of tile for <dose>') and what percent are represented cumulatively by that tile and all those that appear to the left of it (in columns labeled 'Cum pct for <dose>').

Interpretation: The overall length of each bar and its 2-sided confidence interval agree exactly with those in RI_COVG_01, so they show the estimated % of the target population with any evidence of vaccination, stratified by timeliness category. The precise percentage figures for each timeliness category is not listed on the chart, but may be read from the corresponding Excel worksheet. The relative lengths of bars for early and later doses give a sense of drop-out and the timeliness tiles give a sense of what portion received the dose at an age that was earlier than recommended (Too Early) and what percent received it more than 28 days late (< 2 Months Late) and very very late (2+ Months Late).

Notes:

The user must run RI_COVG_01 and RI_COVG_02 before running RI_VCTC_01.

Although RI_VCTC_01 uses some timeliness variables that are calculated by RI_COVG_02 (valid coverage), the coverage and timeliness charts do not directly summarize the % of children who received a valid dose.

The timeliness categories are defined with respect to the age at which the dose is scheduled to be given. Some children are delayed in starting their vaccination sequence and we expect those children to also be delayed in receiving later doses. If the 3-dose Penta series is scheduled to be given at ages 42, 70, and 98 days, but for some reason a child receives her first dose at age 100 days (2+ Months Late) then of course that child should not receive her second dose until age 128 days and her third dose until age 156 days. In the vaccination coverage and timeliness chart, she would appear in the darkest magenta tile (2+ Months Late) for all three doses, because she received all three doses very late with respect to the national schedule even though there is a sense in which she received doses 2 and 3 precisely when she was expected to, based on the timing of dose 1. So, bear in mind that the interpretation is of timeliness with respect to the national immunization programme schedule, and not with respect to the age at which the child received his or her early doses of vaccine.

At this time the Stata VCTCs include three features that will not be available in the R version until sometime in 2023:

- 1) % of respondents who showed home-based records (HBRs)
- 2) % of respondents who are fully vaccinated (listed in a footnote)
- 3) % of respondents who are not vaccinated (listed in a footnote)

CHAPTER 7. EXAMPLES OF CONTROL PROGRAMS

As described at the end of Chapter 2, VCQI control programs usually consist of seven blocks of code. Three blocks are edited and customized by the user and four blocks are usually not edited but are necessary for the program to run correctly.

This chapter shows examples of each of the seven blocks for an RI control program.

7.1 Block A – Start with clear memory

The first block of a VCQI control program clears out old global values and ensures that the output that goes to the screen will not pause during the run. The code below shows the top of a RI program.

```
# User's Guide RI Control Program R version 1.00 - Biostat Global Consulting - 2022-12-01
#
# Vaccination Coverage Quality Indicators (VCQI) control program to analyze data
# from a routine immunization survey
#
# Change log
#
# Date          Version Number      Name           What Changed
# 2022-12-01    1.00              BGC            Original R Version
#
# ****
# Code Block: RI-A             (Do not change) ----
#
# Load the VCQI package
library(vcqiR, attach.required = TRUE)

# Start with clear memory
cleanup_VCQI_globals()
```

7.2 Block B – Specify input/output folders & analysis name

Block B is the first of three sections that the user edits. It consists of three lines of code, two name folders where VCQI will find the survey datasets, where she will put the output files, and the third line gives a name to the analysis. The analysis name will appear in the name of the output spreadsheet. This page shows Block B from a RI control program.

```
# ****
# Code Block: RI-B          (User may change) ----

# Specify input/output folders and analysis name

# Where should the programs look for datasets?
VCQI_DATA_FOLDER <- "Q:/- Folders shared outside BGC/BGC Team - WHO Software/Test datasets/2020-10-16"
# Where should the programs put output?
VCQI_OUTPUT_FOLDER <- "Q:/- Folders shared outside BGC/BGC Team - WHO Software/Working folder - Dale/VCQI test output/RI test"

# Establish analysis name (used in log file name and Excel file name)
VCQI_ANALYSIS_NAME <- "RI_Test"

# Set VCQI_CHECK_INSTEAD_OF_RUN value to 1 to test all metadata
# and code that makes datasets and calculates derived variables,
# without running the indicators or generating output
# Note: checks are not fully implemented and tested in the R version
# of VCQI
VCQI_CHECK_INSTEAD_OF_RUN <- 0
```

Enter the full paths to the VCQI_DATA_FOLDER and VCQI_OUTPUT_FOLDER here.
Note that these two folders cannot be the same .

List a short text string here for the VCQI_ANALYSIS_NAME. Your tabulated output will be stored in the VCQI_OUTPUT_FOLDER in a file named <VCQI_ANALYSIS_NAME>_TO.xlsx

Important note!! Every time you do a new VCQI analysis:

1. Make an empty new folder.
2. Copy your project-specific control program there.
3. Edit the program and point the VCQI_OUTPUT_FOLDER to the new output folder.
4. Make other changes in Blocks B or D or F.
5. Save the program and run VCQI.

7.3 Block C – CD to output folder & open VCQI log

Block C has R change the working directory to be the output folder specified in Block B. Then it deletes any old copies of the Excel output file so this new run will be putting output into a new file. Next it opens the VCQI log, putting some initial messages in there to document the user-inputs that have been specified up to this point.

```
# ****
# Code Block: RI-C          (Do not change) ----

# ** CD to output folder and open VCQI log
setwd(VCQI_OUTPUT_FOLDER)

# Start with a clean, empty Excel file for tabulated output (TO)
unlink(paste0(VCQI_OUTPUT_FOLDER, "/", VCQI_ANALYSIS_NAME, "_TO.xlsx"), force = TRUE)

# Give the current program a name, for logging purposes
VCP <- "RI_Control_Program"

# Open the VCQI log and put a comment in it
vcqi_log_comment(VCP, 3, "Comment", "Run begins...log opened...")

# Document the global macros that were defined before the log opened
vcqi_log_global(VCQI_DATA_FOLDER)
vcqi_log_global(VCQI_OUTPUT_FOLDER)
vcqi_log_global(VCQI_ANALYSIS_NAME)

# Write an entry in the log file documenting the vcqiR package version
vcqi_log_comment(VCP, 3, "Package",
                  paste0("vcqiR package version ", utils::packageVersion("vcqiR")))
```

The global value VCP stands for VCQI Current Program; it holds the name of the program that is currently running. When a message is posted to the log file, the message lists the name of the program that made the message. This is accomplished by passing VCP to the program that writes the log.

7.4 Block D – Specify dataset names & important metadata

Block D holds the second set of lines that a user typically edits. The user specifies the names of the R/Stata/csv datasets that hold the coverage survey data. Variable names and coding conventions for those datasets are in VCQI's Forms and Variable List (FVL) document. The user also specifies some parameters or metadata to describe the vaccination schedule, the coverage survey, and some parameters to control what VCQI generates and how it looks.

Block D – Code common to RI, TT, and SIA analyses

```
# ****
# Code Block: RI-D          (User may change) ----
#
# Specify dataset names and important metadata
# Dataset names should include file extensions
# Accepted file types: .rds, .dta, .csv

# Name of dataset that holds cluster metadata
vcqi_global(VCQI_CM_DATASET, "CM_faux_dataset.dta")

# The R version of VCQI does not use the HH or HM
# datasets except sometimes to provide a variable
# that is used by the level4 layout dataset.
vcqi_global(VCQI_HH_DATASET, "HH_faux_dataset.dta")
vcqi_global(VCQI_HM_DATASET, "HM_faux_dataset.dta")

#
# Parameters to describe the analysis being requested
# .....

vcqi_global(LEVEL2_ORDER_DATASET, paste0(VCQI_DATA_FOLDER, "/level2order.dta"))
vcqi_global(LEVEL3_ORDER_DATASET, paste0(VCQI_DATA_FOLDER, "/level3order.dta"))

vcqi_global(LEVEL1_NAME_DATASET, paste0(VCQI_DATA_FOLDER, "/level1name.dta"))
vcqi_global(LEVEL2_NAME_DATASET, paste0(VCQI_DATA_FOLDER, "/level2names.dta"))
vcqi_global(LEVEL3_NAME_DATASET, paste0(VCQI_DATA_FOLDER, "/level3names.dta"))
```

See Annex B for a description and of the _ORDER_ and _NAME_ datasets. They are required inputs.

```

# LEVEL4 parameters
# The LEVEL4 parameters determine the geographic and/or demographic strata for
# which results are displayed in tabular output and plots. To use the R version
# of VCQI, there must be at least one variable listed in
# VCQI_LEVEL4_SET_VARLIST. The user may specify a single stratifier (like
# urban/rural) or a set of several stratifiers (like urban/rural and sex and
# household wealth).
#
# For example, setting vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("level1name",
# "level3name")) will produce output for the level 1 stratum (overall/national)
# and each level 3 stratum (e.g. each state). If VCQI_LEVEL4_SET_VARLIST is
# populated and VCQI_LEVEL4_SET_LAYOUT is not defined, then VCQI will generate a
# default layout for tables and figures. That layout file will be saved in the
# VCQI_OUTPUT_FOLDER.
#
# The user may create their own VCQI_LEVEL4_SET_LAYOUT file defining the
# conditions, preferred order, and row labels for the LEVEL4 strata and point to
# that layout file in the control program, e.g.
# vcqi_global(VCQI_LEVEL4_SET_LAYOUT, "Q:/My_VCQI_Output/my_level4_layout.rds").
# See the VCQI User's Guide for more details on creating a layout file.

```

```

vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("urban_cluster", "RI20", "RI01"))
vcqi_global(VCQI_LEVEL4_SET_LAYOUT, paste0(VCQI_DATA_FOLDER, "/Level4_Layout_Test3.dta"))

```

SET_VARLIST is always required while SET_LAYOUT is optional in the R version of VCQI. See Annex B for examples where the user requested output for different demographic strata.

```
# User specifies survey::svydesign syntax to describe the complex sample
# The data argument in survey::svydesign should *not* be specified here
vcqi_global(VCQI_SVYDESIGN_SYNTAX, list(ids = ~clusterid, weights = ~psweight, strata = ~stratumid))

# User specifies the method for calculating confidence intervals
# Valid choices are "Logit", "Wilson", "Jeffreys" or "Clopper"; our default recommendation is "Wilson"
vcqi_global(VCQI_CI_METHOD, "Wilson")

# Specify whether the code should export to excel, or not (usually 1)
vcqi_global(EXPORT_TO_EXCEL, 1) ← Set this parameter to 1 to generate tabular output in an Excel file. Set it to 0 if you only wish to make figures or database output.
```

```
# User specifies the number of digits after the decimal place in coverage outcomes
```

```
vcqi_global(VCQI_NUM_DECIMAL_DIGITS, 1) ← VCQI usually provides one digit after the decimal point in tabular and graphic output, but sometimes we like to change this to zero (0) to produce compact tables.
```

```

# Specify whether the code should make plots, or not (usually 1)
# MAKE_PLOTS must be 1 for any plots to be made
vcqi_global(MAKE_PLOTS, 1)

# Set PLOT_OUTCOMES_IN_TABLE_ORDER to 1 if you want inchworm and
# unweighted plots to list strata in the same order as the tables;
# otherwise the strata will be sorted by the outcome and shown in
# bottom-to-top order of increasing indicator performance
vcqi_global(PLOT_OUTCOMES_IN_TABLE_ORDER, 1)

```

If this is set to 0, VCQI will not make any plots or figures or charts.

If this parameter is 0, strata will be sorted in figures in order of their outcomes (those with best outcomes at the top and worst at the bottom). If it is set to 1, they will be sorted in the same order in plots and tables. See Annex C.

The next two pages show options for controlling bar plots, unweighted proportion plots, and organ pipe plots.

You may suppress individual types of plots.

Turn on or off:

- bar plots (VCQI_MAKE_IW_PLOTS)
- unweighted proportion plots (VCQI_MAKE_UW_PLOTS)
- organ pipe plots (VCQI_MAKE_OP_PLOTS)

Note: The letters IW stand for inchworm; that abbreviation is used here because VCQI's bar plots are an alternative to the inchworm plots. See the Stata user's guide or the WHO 2018 Vaccination Coverage Survey Reference Manual for more details on inchworm plots.

<http://www.biostatglobal.com/downloads/VCQI%20User's%20Guide.pdf>

<https://apps.who.int/iris/handle/10665/272820>

```

# Make inchworm/bar plots? Set to 1 for yes.
vcqi_global(VCQI_MAKE_IW_PLOTS, 1)

# Text at right side of inchworm/bar plots
# 1 1-sided 95% LCB | Point Estimate | 1-sided 95% UCB
# 2 Point Estimate (2-sided 95% Confidence Interval) [THIS IS THE DEFAULT]
# 3 Point Estimate (2-sided 95% Confidence Interval) (0, 1-sided 95% UCB)
# 4 Point Estimate (2-sided 95% Confidence Interval) [1-sided 95% UCB, 100)
# 5 Point Estimate (2-sided 95% CI) (0, 1-sided 95% UCB] [1-sided 95% LCB, 100)
vcqi_global(VCQI_IWPLT_CITEXT, 2) ←

# Text at right side of double inchworm/bar plots
# 1 (default) means show both point estimates
# 2 means show both point estimates and both 2-sided 95% CIs
# 3 means do not show any text
vcqi_global(VCQI_DOUBLE_IWPLT_CITEXT, 1) ←

# IWPLT_SHOWBARS = 0 means show inchworm distributions
# IWPLT_SHOWBARS = 1 means show horizontal bars instead of inchworms
# For the current version of R VCQI, please always set this to be 1
vcqi_global(IWPLT_SHOWBARS, 1) ←

```

Control what text is listed at the far right of each bar plot and each double-bar plot.

Always set this option to 1 for now because VCQI R does not make inchworm plots.

Unweighted proportion plots have four input parameters:

```

# Make unweighted sample proportion plots? Set to 1 for yes.
vcqi_global(VCQI_MAKE_UW_PLOTS, 1)

```

```
#Annotate text in the plot for small sample sizes? Set 1 for yes.
vcqi_global(UWPLT_ANNOTATE_LOW_MED, 0)
```

```
#Add square brackets around N < UWPLT_ANNOTATE_LOW_N
vcqi_global(UWPLT_ANNOTATE_LOW_N, NA) ←
```

```
#Add parentheses around N < UWPLT_ANNOTATE_MED_N
vcqi_global(UWPLT_ANNOTATE_MED_N, NA) ←
```

UWPLT_ANNOTATE_MED_N needs to be \geq UWPLT_ANNOTATE_LOW_N

Note: For some survey reports from DHS or MICS we see tables and figures suppress output based on very small sample sizes (i.e., smaller than 25 persons) and if the results are based on a moderately small number (i.e., from 25 to 49) the point estimates are surrounded in parentheses. (e.g., (78.3%)). To turn on these suppression/annotation options, set UWPLT_ANNOTATE_LOW_MED to 1 and set appropriate values of _LOW_N and _MED_N (e.g., 25 and 50, respectively). We plan to elaborate on these options with an entire appendix and examples of inputs and outputs in a forthcoming update to this guide.

```
# Make organ pipe plots? Set to 1 for yes.  
vcqi_global(VCQI_MAKE_OP_PLOTS, 1)  
  
# Save the data underlying each organ pipe plot? Set to 1 for yes.  
#  
# Recall that organ pipe plots do not include many quantitative details  
# and do not list the cluster id for any of the bars.  
#  
# If this option is turned on, (set to 1) then the organ pipe plot program  
# will save a dataset in the Plots_OP folder for each plot. The dataset will  
# list the cluster id for each bar in the plot along with its height and width.  
# This makes it possible to identify which cluster id goes with which bar in  
# the plot and to understand the quantitative details of each bar.
```

```
vcqi_global(VCQI_SAVE_OP_PLOT_DATA, 1)
```

This option tells VCQI to save an individual dataset holding detailed data for each organ pipe plot.

69

```
# Save the data underlying inchworm plots/bar plots? Set to 1 for yes.  
# If this option is turned on, inchworm and barplot programs will save a dataset  
# in the Plots_IW_UW folder that makes it possible to understand the quantitative  
# details of each plot component and can be used to recreate the plot.  
vcqi_global(VCQI_SAVE_IW_PLOT_DATA, 1)
```

This option tells VCQI to save an individual dataset holding detailed data for each bar plot and each double bar plot.

```
# Save the data underlying unweighted plots? Set to 1 for yes.  
# If this option is turned on, unweighted programs will save a dataset  
# in the Plots_IW_UW folder that makes it possible to understand the quantitative  
# details of each plot component and can be used to recreate the plot.  
vcqi_global(VCQI_SAVE_UW_PLOT_DATA, 1)
```

This option tells VCQI to save an individual dataset holding detailed data for each unweighted plot.

```

# Specify whether the code should save VCQI output databases
#
# WARNING!! If this macro is set to 1, VCQI will delete ALL files that
# end in _database.rds in the VCQI_OUTPUT_FOLDER at the end of the run
# If you want to save the databases, change the value to 0.
# (Usually 1)
vcqi_global(DELETE_VCQI_DATABASES_AT_END, 1) ←

# If you wish to aggregate files that end
# in _database.rds into a single dataset, set the
# DELETE_VCQI_DATABASES_AT_END option to 0 and set the
# AGGREGATE_VCQI_DATABASES option to 1.
vcqi_global(AGGREGATE_VCQI_DATABASES, 1)

```

VCQI generates datasets (flat files) of analysis results that it calls *databases*. See Annex D for details. When this parameter is set to 1, VCQI deletes them because we assume that most users want to look at graphical or tabular output rather than unformatted flat files. If you wish to keep the database files, set this parameter to 0.

```

# Specify whether the code should delete intermediate datasets
# at the end of the analysis (Usually 1)
# If you wish to keep them for additional analysis or debugging,
# set the option to 0.
vcqi_global(DELETE_TEMP_VCQI_DATASETS, 1) ←

```

This parameter is usually set to 1 so VCQI will delete its temporary datasets when it is finished running. A user might set it to 0 to keep those datasets for the purpose of debugging a program or following along to understand some of VCQI's intermediate work products.

```

# Set this global to 1 if you would like to create an augmented dataset
# that merges survey dataset with derived variables calculated by VCQI.
# Default value is 0 (no)
vcqi_global(VCQI_MAKE_AUGMENTED_DATASET, 0) ←

```

Set this to 1 if you want to generate an augmented dataset. See Annex E for details.

A full RI analysis can result in many dozens of _database.rds files in the VCQI_OUTPUT_FOLDER, each documenting a single outcome, often for a single dose. If the user will make use of the databases, it will usually be efficient to combine or aggregate the contents of all those datasets into a single large dataset and then to tidy up by deleting the individual constituent files. To do this, set `DELETE_VCQI_DATABASES_AT_END` to 0 and set `AGGREGATE_VCQI_DATABASES` to 1.

Block D for an RI survey analysis, continued

```
# Name of datasets that hold RI data
vcqi_global(VCQI_RI_DATASET, "RI_mdy.dta")
vcqi_global(VCQI_RIHC_DATASET, "RIHC_mdy.dta")

# .....  
# Parameters to describe RI schedule  
# .....  
# These parameters may change from survey to survey

# See:  
# http://www.who.int/immunization/policy/Immunization_routine_table2.pdf?ua=1  
# http://apps.who.int/immunization_monitoring/globalsummary/schedules  

# Single-dose antigens will use a parameter named <dose>_min_age_days (required)  
# Single-dose antigens may use a parameter named <dose>_max_age_days (optional)  
# Note: If a dose is not considered valid AFTER a certain age, then specify  
#       that maximum valid age using the _max_age_days parameter.  
#       If the dose is considered late, but still valid, then do not specify  
#       a maximum age.

bcg_min_age_days      <- 0    # birth dose
hepb_min_age_days     <- 0    # birth dose
opv0_min_age_days     <- 0    # birth dose

# Note: In this country, opv0 and hepb0 are only considered valid
#       if given in the first two weeks of life
opv0_max_age_days     <- 14   # birth dose
hepb_max_age_days     <- 14   # birth dose

penta1_min_age_days   <- 42   # 6 weeks
penta2_min_age_days   <- 70   # 10 weeks
penta2_min_interval_days <- 28   # 4 weeks
pcv1_min_age_days     <- 42   # 6 weeks
opv1_min_age_days     <- 42   # 6 weeks
rota1_min_age_days    <- 42   # 6 week
```

The person who makes the VCQI-compatible datasets will decide what to name these datasets. They may have any name that is valid for a R dataset file. Note that the user must specify the file extension when you list the filename(s) here.

See Chapter 3 for more detail on how to parameterize the routine immunization schedule.

```

pcv2_min_age_days      <- 70  # 10 weeks
pcv2_min_interval_days <- 28  # 4 weeks
opv2_min_age_days      <- 70  # 10 weeks
opv2_min_interval_days <- 28  # 4 weeks
rota2_min_age_days     <- 70  # 10 weeks
rota2_min_interval_days <- 28  # 4 weeks

penta3_min_age_days    <- 98  # 14 weeks
penta3_min_interval_days <- 28  # 4 weeks
pcv3_min_age_days      <- 98  # 14 weeks
pcv3_min_interval_days <- 28  # 4 weeks
opv3_min_age_days      <- 98  # 14 weeks
opv3_min_interval_days <- 28  # 4 weeks
rota3_min_age_days     <- 98  # 14 weeks
rota3_min_interval_days <- 28  # 4 weeks

ipv_min_age_days        <- 98  # 14 weeks; may be co-administered w/ OPV

mcv1_min_age_days      <- 270  # 9 months
yf_min_age_days         <- 270  # 9 months

# ..... .
# Parameters to describe survey
# ..... .
# Specify the earliest and latest possible vaccination date for this survey.

# The software assumes this survey includes birth doses, so the earliest date
# is the first possible birthdate for RI survey respondents and the latest
# date is the last possible vaccination date for this dataset - the latest
# date might be the date of the final survey interview.

vcqi_global(EARLIEST_SVY_VACC_DATE_M, 1)
vcqi_global(EARLIEST_SVY_VACC_DATE_D, 1)
vcqi_global(EARLIEST_SVY_VACC_DATE_Y, 2013)

vcqi_global(LATEST_SVY_VACC_DATE_M, 1)
vcqi_global(LATEST_SVY_VACC_DATE_D, 1)
vcqi_global(LATEST_SVY_VACC_DATE_Y, 2015)

```

```

# These parameters indicate the eligible age range for survey respondents
# (age expressed in days)

vcqi_global(VCQI_RI_MIN_AGE_OF_ELIGIBILITY, 365)
vcqi_global(VCQI_RI_MAX_AGE_OF_ELIGIBILITY, 731)

# These following parameters help describe the survey protocol with regard to whether they:
# a) skipped going to health centers to find RI records (RI_RECORDS_NOT_SOUGHT 1)
# b) looked for records for all respondents (RI_RECORDS_SOUGHT_FOR_ALL 1)
# c) looked for records for respondents who didn't present vaccination cards
#      during the household interview (RI_RECORDS_SOUGHT_IF_NO_CARD 1)

# These are mutually exclusive, so only one of them should be set to 1.


```

These are the VCQI default values: children are eligible for the survey if they are at least 12 months old and not yet 24 months old.

```

vcqi_global(RI_RECORDS_NOT_SOUGHT, 0)
vcqi_global(RI_RECORDS_SOUGHT_FOR_ALL, 1)
vcqi_global(RI_RECORDS_SOUGHT_IF_NO_CARD, 0)

# .....
# Which doses should be included in the analysis?
# .....

# Note that these abbreviations must correspond to those used in the names of
# the dose date and dose tick variables *AND* the names used above in the
# schedule globals (<dose>_min_age_days and <dose>_min_interval_days and
# <dose>_max_days. The variables are named using lower-case acronyms. The
# globals here may be upper or mixed case; they will be converted to lower case
# in the software.

vcqi_global(RI_SINGLE_DOSE_LIST, c("BCG", "HEPB", "OPV0", "IPV", "MCV1", "YF"))
vcqi_global(RI_MULTI_2_DOSE_LIST, c())
vcqi_global(RI_MULTI_3_DOSE_LIST, c("PENTA", "PCV", "OPV", "ROTA"))

# In this example we do not have any two-dose vaccine series to analyze. The
# RI_MULTI_2_DOSE_LIST is defined as an empty list above - it could also be
# NULL: vcqi_global(RI_MULTI_2_DOSE_LIST, NULL), or the line defining
# RI_MULTI_2_DOSE_LIST could be omitted entirely.

# The R VCQI software can handle dose lists with up to 9 doses (RI_MULTI_9_DOSE_LIST)

```

Populate the dose shifting parameters to indicate whether you would like to shift evidence from later doses *down* to fill holes in evidence for earlier doses. For example, if a child has evidence of penta1 and penta3, but not of penta2, shifting the evidence will move the dose 3 evidence into the variables for dose 2. In some countries there are booster doses. To move evidence from variables for boosters, down into the variables that hold evidence of the primary dose series, populate the parameters described here. The SHIFTTO parameter holds the names of the primary dose series and the SHIFTFROM parameter holds the names of the booster doses.

```
# .....  
# Do you want to shift doses?  
# .....  
  
# This can be done with multi-dose vaccines and/or boosters  
  
# Number of dose series you would like to shift  
# Set to 0 if you do not wish to implement any shifts  
vcqi_global(NUM_DOSE_SHIFTS, 0)  
  
vcqi_global(SHIFTTO_1, c("penta1", "penta2", "penta3")) # List of doses where evidence will be shifted *to*  
vcqi_global(SHIFTFROM_1, c("penta4", "penta5")) # List of doses where evidence will be shifted *from*  
vcqi_global(SHIFTWITHIN_1, 0) # Set to 1 to shift dates in a series down to fill holes in evidence  
vcqi_global(DROPDUP_1, 0) # Set to 1 to convert duplicate dates in a series to missing  
  
vcqi_global(SHIFTTO_2, c("polio1", "polio2", "polio3")) # List of doses where evidence will be shifted *to*  
vcqi_global(SHIFTFROM_2, c("polio4", "polio5")) # List of doses where evidence will be shifted *from*  
vcqi_global(SHIFTWITHIN_2, 0) # Set to 1 to shift dates in a series down to fill holes in evidence  
vcqi_global(DROPDUP_2, 0) # Set to 1 to convert duplicate dates in a series to missing
```

Because NUM_DOSE_SHIFTS is 0 here, no shifting will occur, but we show possible values for the SHIFTTO and SHIFTFROM parameters to aid in future use. If NUM_DOSE_SHIFTS were changed from 0 to 2, then shifting would occur for both the penta and polio dose series. Here we will use the penta dose series as an example to demonstrate how changing SHIFTWITHIN_1 and DROPDUP_1 settings will affect the results (note that 0 is the default setting for SHIFTWITHIN and DROPDUP):

If SHIFTWITHIN_1 is set to 1 and DROPDUP_1 is set to 0:

* Raw data:	Shifted data:
* penta1 – missing	penta1 – 1/1/2021
* penta2 -- 1/1/2021	penta2 – 3/1/2021
* penta3 – missing	penta3 – 3/1/2021
* penta4 – 3/1/2021	penta4 – missing
* penta5 – 3/1/2021	penta5 – missing

If SHIFTWITHIN_1 is set to 0 and DROPDUP_1 is set to 1:

* Raw data:	Shifted data:
* penta1 – missing	penta1 – missing
* penta2 -- 1/1/2021	penta2 – 1/1/2021
* penta3 – missing	penta3 – 3/1/2021
* penta4 – 3/1/2021	penta4 – missing
* penta5 – 3/1/2021	penta5 – 3/1/2021

If SHIFTWITHIN_1 is set to 0 and DROPDUP_1 is set to 0, which is the default setting for both:

* Raw data:

* penta1 – missing	----->	Shifted data:
* penta2 -- 1/1/2021	----->	penta1 – missing
* penta3 – missing	----->	penta2 – 1/1/2021
* penta4 – 3/1/2021	----->	penta3 – 3/1/2021
* penta5 – 3/1/2021	----->	penta4 – missing
		penta5 – 3/1/2021

If SHIFTWITHIN_1 is set to 1 and DROPDUP_1 is set to 1:

* Raw data:

* penta1 – missing	----->	Shifted data:
* penta2 -- 1/1/2021	----->	penta1 – 1/1/2021
* penta3 – missing	----->	penta2 – 3/1/2021
* penta4 – 3/1/2021	----->	penta3 – missing
* penta5 – 3/1/2021	----->	penta4 – missing
		penta5 – missing

7.5 Block E – Pre-process survey data

Block E should not be changed by the user. The code varies across RI, TT and SIA surveys. The R version of VCQI does not yet analyze data from SIA/TT surveys. At this time, we have pasted Block E from RI survey only here.

The global value named RI_DOSE_LIST holds a (lower-case) list of all the doses in the analysis. The order that doses will appear in tabulated output is determined by the order they appear in the list. By default, single-doses are listed before multi-dose series and multi-dose series are listed sequentially. The R version of VCQI accepts up to 9-dose series.

The user may prefer a different dose order...perhaps doses should appear in tables in the order that they are scheduled: (birth doses followed by 6-week doses, then 10-week, then 14-week, then 9-month doses, then later boosters).

If the user wishes to set a preferred left-to-right dose order for tabular output, they should put a line of code in the top of Block F to fix (over-ride) RI_DOSE_LIST. See the example below in the section on Block F for RI control programs.

```
# ****
# Code Block: RI-E          (Do not change) ----
# ** Format the VCQI dose list and pre-process survey data

# Construct the global RI_DOSE_LIST from what the user specified above
# VCQI currently handles single-dose vaccines and multi-dose vaccines
# with up to nine doses in the series

# First, list single dose vaccines
if (vcqi_object_exists("RI_SINGLE_DOSE_LIST")){
  RI_DOSE_LIST <- stringr::str_to_lower(RI_SINGLE_DOSE_LIST)
} else{
  RI_DOSE_LIST <- NULL
}

# Second, list doses in multi-dose lists
for(i in 2:9){
  if(vcqi_object_exists(paste0("RI_MULTI_", i, "_DOSE_LIST"))){
    dl <- get(paste0("RI_MULTI_", i, "_DOSE_LIST"))
    if(!is.null(dl) & length(dl) > 0){
      RI_DOSE_LIST <- c(RI_DOSE_LIST, paste0(rep(stringr::str_to_lower(dl), each = i), 1:i)))
    }
  }
}
```

```
# Put a copy of the dose list in the log
vcqi_log_global(RI_DOSE_LIST)

# .....
# Check the user's metadata for completeness and correctness
# .....

# Starting RI_TEMP_DATASETS as an empty object before starting to record temp dataset list
RI_TEMP_DATASETS <- NULL

check_RI_schedule_metadata()
check_RI_survey_metadata()
check_RI_analysis_metadata()

# Run the program to look at date of birth (from history, card, and register)
# and look at dates of vaccination from cards and register. This program
# evaluates each date and checks to see that it occurred in the period
# allowed for respondents eligible for this survey. It also checks to see
# that doses in a sequence were given in order. If any vaccination date
# seems to be outside the right range or recorded out of sequence, the date
# is stripped off and replaced with a simple yes/no tick mark. This step
# means less date-checking is necessary in subsequent programs.

cleanup_RI_dates_and_ticks()

# The name of the datasets coming out of these cleanup steps are:
#   "{VCQI_OUTPUT_FOLDER}/{VCQI_DATASET}_clean" &
#   "{VCQI_OUTPUT_FOLDER}/{VCQI_RIHC_DATASET}_clean"
```

```
# .....  
# Establish unique IDs  
# .....  
  
# The name of the dataset coming out of the ID step is RI_with_ids  
establish_unique_RI_ids()  
  
# If the user requests a check instead of a run, then turn off  
# flags that result in databases, excel output, and plots  
  
if(VCQI_CHECK_INSTEAD_OF_RUN == 1){  
  vcqi_log_comment(VCP, 3, "Comment",  
    "The user has requested a check instead of a run.")  
  VCQI_PREPROCESS_DATA <- 0  
  VCQI_GENERATE_DVS <- 0  
  VCQI_GENERATE_DATABASES <- 0  
  EXPORT_TO_EXCEL <- 0  
  MAKE_PLOTS <- 0  
}  
79
```

7.6 Block F – Calculate VCQI indicators requested by the user

Block F is the third and final section that the user edits.

Broadly speaking, there are four steps to run an indicator:

1. Specify required (and optional) inputs via vcqi_global function.
2. Specify the title, subtitle, and footnotes for the Excel worksheet that will hold tabular results.
3. Call the function that calculates the indicator and generates output.
4. If you will calculate this indicator again later in the same control program, clear out the input global values so old values are not mistakenly used again.

These steps are quite similar across indicators, but the details of the code in Block F differs substantially across TT, RI and SIA surveys. We have pasted some example code for RI analysis here. See the control programs that accompany this guide for full examples of how to run the VCQI indicators.

Note: Shaded lines below are wrapped onto multiple lines in this document but appear on a single line of code (each) in R.

80

```
# Over-ride the default RI_DOSE_LIST so we can be in charge of the left-to-right order
# in which dose results will appear in tabulated output.
# (Remember that RI_DOSE_LIST must be lower case.)
# This new order starts with birth doses, proceeds to dose series, and ends with IPV, MCV1, and YF

vcqi_global(RI_DOSE_LIST, c("bcg", "hepb", "opv0", "opv1", "opv2", "opv3", "pental", "penta2", "penta3",
"pcv1", "pcv2", "pcv3", "rotal", "rota2", "rota3", "ipv", "mcv1", "yf"))

# Alternatively, some users might prefer to have the doses listed strictly in order of the age
# at which they are due. If you prefer that, then this code might be appropriate:

vcqi_global(RI_DOSE_LIST, c("bcg", "hepb", "opv0", "opv1", "pental", "pcv1", "rotal", "opv2", "penta2",
"pcv2", "rota2", "opv3", "penta3", "pcv3", "rota3", "ipv", "mcv1", "yf"))
```

Note: Shaded lines below are wrapped onto multiple lines in this document but appear on a single line of code (each) in R.

```
# Estimate crude dose coverage for all the doses in the RI_DOSE_LIST  
  
vcqi_global(RI_COVG_01_TO_TITLE, "Crude Coverage")  
vcqi_global(RI_COVG_01_TO_SUBTITLE, NA)  
  
vcqi_global(RI_COVG_01_TO_FOOTNOTE_1, "Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound;  
UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intracluster Correlation Coefficient")  
  
vcqi_global(RI_COVG_01_TO_FOOTNOTE_2, "Note: This measure is a population estimate that incorporates survey  
weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.")  
  
vcqi_global(SORT_PLOT_LOW_TO_HIGH, 1) # 1 means show strata w/ low outcomes at bottom and high at top; 0 is  
the opposite
```

Note: Shaded lines below are wrapped onto multiple line in this document but appear on a single line of code (each) in R. Note also that Block F uses some logic to change the wording of FOOTNOTE_7 depending on whether the user asks for a CRUDE or VALID analysis.

```
# Estimate the proportion of children who experienced 1+ MOVs
vcqi_global(RI_QUAL_09_VALID_OR_CRUDE, "CRUDE") # Set to CRUDE or VALID

vcqi_global(RI_QUAL_09_TO_TITLE, "Percent of Respondents with MOVs")
vcqi_global(RI_QUAL_09_TO_SUBTITLE, NA)
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_1, "Percent of respondents who had date of birth and visit date data who failed to receive a vaccination for which they were eligible on an occasion when they received another vaccination.")
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_2, "An uncorrected MOV means that the respondent had still not received a valid dose at the time of the survey.")
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_3, "A corrected MOV means that the respondent had received a valid dose by the time of the survey.")
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_4, "The denominator for Had MOV (%) is the number of respondents who had visits eligible.")
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_5, "The denominator for MOV uncorrected and corrected (%) is the number of MOVs.")
vcqi_global(RI_QUAL_09_TO_FOOTNOTE_6, "Note that for individual doses, the % MOV uncorrected + % MOV corrected adds up to 100%.")

if(stringr::str_to_upper(RI_QUAL_09_VALID_OR_CRUDE) == "VALID"){
  vcqi_global(RI_QUAL_09_TO_FOOTNOTE_7, "Note: Early doses are ignored in this analysis; the respondent is considered to have not received them.")
}

if(stringr::str_to_upper(RI_QUAL_09_VALID_OR_CRUDE) == "CRUDE"){
  vcqi_global(RI_QUAL_09_TO_FOOTNOTE_7, "Note: Early doses are accepted in this analysis; all doses are considered valid doses.")
}

# This indicator makes plots (1) if any MOV and (2) if corrected. These are sorted in opposite directions, so global SORT_PLOT_LOW_TO_HIGH is set inside RI_QUAL_09_06PO.R rather than here by the user.

RI_QUAL_09()
```

This page shows an example of an indicator that uses several user-specified inputs.

Note: Shaded lines below are wrapped onto multiple line in this document but appear on a single line of code (each) in R.

```
# .....  
# Make Coverage and Timeliness Charts  
# .....  
  
# Specify 1 or 2 or 3 here to make charts for every level 1, 2 or 3 stratum.  
RI_VCTC_01_LEVELS <- 3  
# You may also specify a combination like c(1,3)  
# RI_VCTC_01_LEVELS <- c(1,3)  
  
#Specify which doses to show in the chart and the order, from bottom to top  
  
TIMELY_DOSE_ORDER <-  
c("bcg", "hepb", "opv0", "opv1", "opv2", "opv3", "penta1", "penta2", "penta3", "pcv1", "pcv2", "pcv3", "rota1", "rota2",  
"rota3", "ipv", "mcv1", "yf")  
  
# Specify the y-coordinates for the bars. If you want them to be spaced evenly, you may omit this global  
(leave it empty)  
# In this example, we use irregular spacing to group the different dose series.  
TIMELY_Y_COORDS <- c(10, 20, 30, 43, 50, 57, 73, 80, 87, 103, 110, 117, 133, 140, 147, 160, 170,  
180)  
  
# You may customize the parameters in the .R file and define the full  
# path to the .R file below as VCTC_globals_path. Or you may re-specify them  
# in code after VCTC_default_global() was called  
  
# Include the user-specified parameters, if want to customize any settings  
vcqi_global(VCTC_globals_path, NA)  
  
if (vcqi_object_exists("VCTC_globals_path")) {  
  source(file = VCTC_globals_path)  
} else {  
  VCTC_default_global()  
}  
  
# Over-ride one default parameters:
```

```
# Because we are spacing the bars about every y=10 units instead of the
# default Y=1, specify a bar width that is 10X the default.
vcqi_global(TIMELY_BARWIDTH, 6.7)

# Do the calculations and make the charts
RI_VCTC_01()
```

7.7 Block G – Exit gracefully

Block G is the same across all three kinds of control programs. It makes the augmented dataset, if the user asked for one, then it calls a program that cleans up after VCQI, moving the log file into Excel and, if the user wishes, deleting temporary files. The VCQI log is moved into a worksheet of the Excel output file; errors are shaded red and warnings are shaded yellow; the log is sorted so errors and warnings appear at the top of the log worksheet.

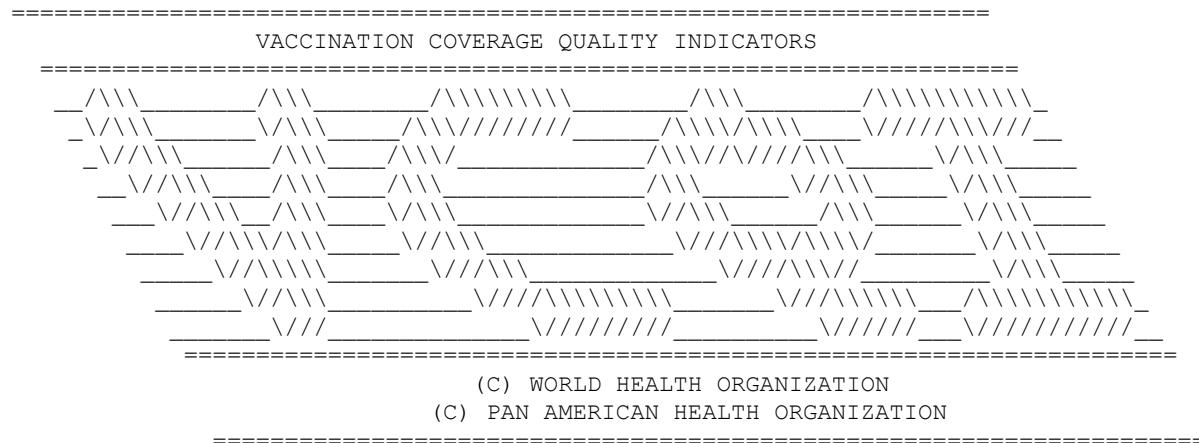
```
# ****
# Code Block: RI-G          (Do not change) ----
#
# Exit gracefully

# Make augmented dataset for additional analysis purposes if user requests it.
if(vcqi_object_value("VCQI_MAKE_AUGMENTED_DATASET", 1) &
   !vcqi_object_value("VCQI_CHECK_INSTEAD_OF_RUN", 1)){
  make_RI_augmented_dataset(outpath = NA)
}

# Close the datasets that hold the results of hypothesis tests and put them into
# the output spreadsheet
# Close the log file and put it into the output spreadsheet
# Clean up extra files
# Send a message to the screen if there are warnings or errors in the log
vcqi_cleanup()
```

CHAPTER 8. TROUBLESHOOTING VCQI

As mentioned in section 2.1, VCQI trap obvious errors and omissions concerning input parameter values and files and tries to provide helpful error messages. If you see red error messages in the R console window, followed by the VCQI billboard like this:



then either scroll back through the console window to read the error messages, or open VCQI's tabulated output spreadsheet (named <VCQI_ANALYSIS_NAME>_TO.xlsx) and read the error messages and warnings in the worksheet named "Log". If the messages are confusing, provide some feedback or ask for assistance from the VCQI software developers or other members of the Technet-21 VCQI User's Group. (See Annex A of *Getting Started with VCQI* for details on how to join the User's Group.)

If R encounters a hard error and halts without showing the VCQI billboard, you will need to type the command "vcqi_cleanup()" at the R command prompt to cue VCQI to put the log and its informative messages into the output spreadsheet file. Try to follow the instructions in those error messages, if any. If you are puzzled, send a note to the User's Group or to GetVCQIHelp@biostatglobal.com.

The remainder of this chapter describes several challenges that can occur with VCQI and recommends steps to address them.

1. R halts with an error message like:

Error in gzfile(file, mode) : cannot open the connection

In addition: Warning message:

In gzfile(file, mode) :

cannot open compressed file <file path>, probable reason 'Invalid argument'

This problem is usually caused by a momentary conflict between R and your cloud-based syncing or backup service, like Dropbox. Here at Biostat Global Consulting, when we run VCQI, we usually pause Dropbox syncing first, which prevents this error. When VCQI is finished we resume Dropbox syncing.

2. R shows a warning message like:

Warning message:

In file.create(to[okay]) :

cannot create file <path to the Excel file>, reason 'Permission denied'

This warning is usually caused by having an Excel file that VCQI writes to opened while trying to run VCQI again. Please make sure that output Excel files are closed before running VCQI again.

This list of possible problems may grow over time, so check this chapter in future versions of the *User's Guide*, too. If you encounter challenges, send a message to the VCQI User's Group or to the VCQI developers at GetVCQIHelp@biostatglobal.com.

ANNEX A. NESTED STRATA IN VCQI

The notion of geographic strata is embedded deeply in the VCQI code and the strata are described as residing at one of three levels. With the R version of VCQI, we have required the user to provide a list of Level4 stratification variables and they will probably also usually provide a Level4 layout file. In the Stata version, there are options for ordering output by geographic strata, and those options may be included in the R version in the future. For both the Stata and R versions, it is required to define the Level3, Level2, and Level1 strata in the so-called _name(s) and _order datasets described below.

Table A-1 lists the vocabulary associated with the three nested levels of strata in our example. VCQI always requires the Level1, Level2 and Level3 datasets. And for R users, Level4 variables are required, also.

In the R version of VCQI, levels 1 and 2 and 3 are currently only used by the RI module that generates vaccination coverage and timeliness charts (RI_VCTC_01). But the parameter files are required even if the user is not running RI_VCTC_01.

Table A-1. Overview of three nested levels of administrative hierarchy

Name	Description	Note 1	Note 2
Level 1	Entire country	There is only one level 1 stratum per VCQI analysis	Calculating results for level 1 makes sense if the level 2 strata are exhaustive (comprise the entire country).
Level 2	Sub-national strata (e.g., provinces)	All level 2 strata are contained within level 1; level 2 strata are mutually exclusive, (meaning that each level 3 stratum is part of only one level 2 stratum) but level 2 strata do not have to be exhaustive (you do not have to do the survey in every province in the nation).	Calculating results for level 2 makes sense if the level 3 strata are exhaustive (comprise the entire level 2 stratum). If you do a survey only of high-risk districts at level 3, then it may not make sense to calculate results at levels 2 or 1.
Level 3	Sub-sub-national (i.e., health districts nested within provinces)	Each level 3 stratum is contained within a level 2 stratum; level 3 strata are mutually exclusive, (each cluster appears in only one level 3 stratum) but they do not have to be exhaustive (you do not need to do a survey in every district in the province).	Level 3 is typically the lowest administrative level at which the survey was conducted. Level 2 is constructed by aggregating data from a set of level 3 strata, and level 1 is constructed, if appropriate, by aggregating all the data from all level 2 strata.

Name	Description	Note 1	Note 2
Level 4	Demographic variable that defines sub-groups within Levels 1-3	The user specifies one or more categorical variables to define Level 4 strata. This variable might code the sex of the respondent, or whether they live in an urban or rural cluster	The Level 4 stratification variables are required for the R version.

Figure A-1. The Fictional Country of Harmonia



Figure A-2. Harmonia holds a Northern and Southern Province

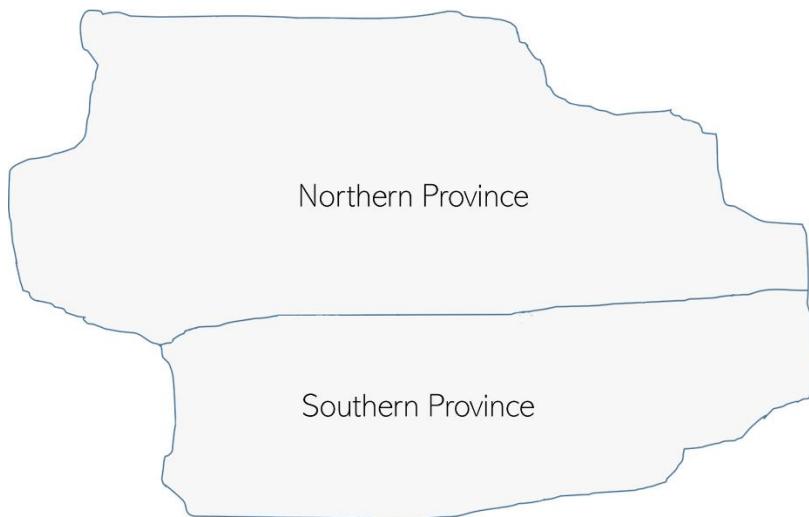
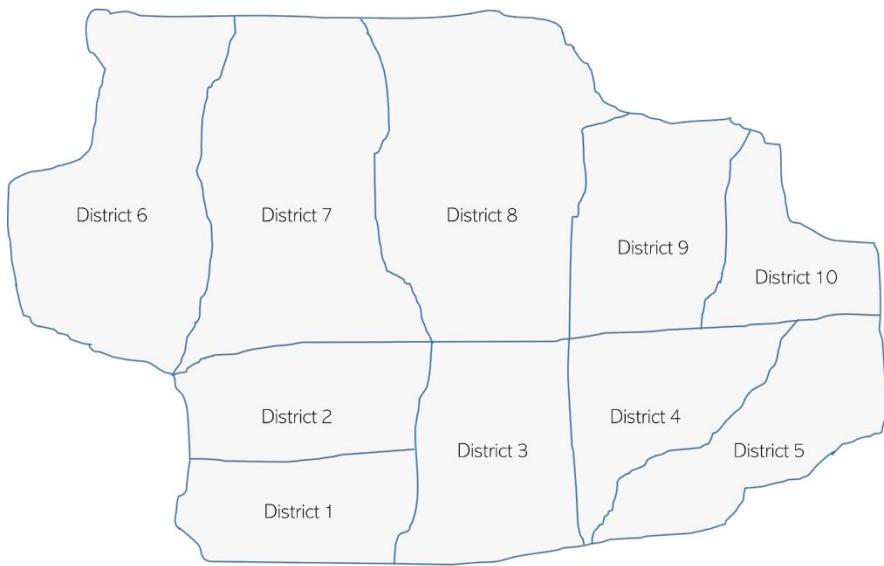


Figure A-3. Each province holds five districts



VCQI is designed to be able to analyze vaccination coverage surveys conducted at any of these three levels of administrative hierarchy. The terminology can be a little confusing because the terms Level 1, Level 2, and Level 3 can refer to different things depending on how the survey was done. Table A-1 indicates what level of hierarchy the three terms would refer to under several common survey scenarios.

Important Note!

Level 3 always refers to the lowest level of administrative hierarchy where survey reports will be reported in separate strata. So if the survey is designed to yield district-level results, then Level 3 is the district level. If it is designed to yield provincial results (but not district level) then Level 3 is the provincial level. If it is only a national survey with no sub-national strata, then Level 3 is the national level.

If the Level 3 strata are nested within a higher level, and if the survey is conducted in every Level 3 stratum, then it may make sense to report aggregated results at Level 2 and possibly at Level 1.

Table A-2. Levels 1-3 vary depending on where the survey was conducted

Survey conducted in :	Level 1	Level 2	Level 3
Each health district in each province in the nation	Nation	Province	District
Each province in the nation (alternative #1)		Nation	Province
Each province in the nation (alternative #2)	Nation		Province
A subset of health districts (not all in any province)			District
All health districts in one province		Province	District
National survey only (no sub-national strata)			Nation
Single district survey			District

VCQI also alludes to something called Level 4, which gives users the flexibility to stratify results by geographic and/or demographic sub-groups. Common demographic sub-groups might include urban/rural, boy/girl, literate caregiver/illiterate caregiver, wealthy household/poorer household, etc.

To calculate results using a demographic stratification variable, first be sure to establish a variable that codes the demographic sub-groups. Once the stratification variable is defined for every respondent, then tell VCQI to use it by including the following line in the control program:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "stratification_variable_name")
```

e.g.,

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "urban_cluster")
```

or

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "gender")
```

or

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "literacy_status")
```

See Annex B for examples of tabulated output calculated using Level4 layout files that list strata from geographic Levels 1-3 plus urban/rural sub-groups.

ANNEX B. CONTROLLING HOW STRATA ARE LISTED IN VCQI OUTPUT

The user has control over the names of strata at all levels. The user has control over which strata will be listed, and in what order, in Excel tables. The order for tabulated output is controlled by VCQI_LEVEL4_SET_LAYOUT. If user wants the output to be in a specific order, then the user needs to build and define a VCQI_LEVEL4_SET_LAYOUT file. See Section B.0 for more information on how to build a VCQI_LEVEL4_SET_LAYOUT dataset. In bar plots and unweighted proportion plots, the default setting is for strata to be sorted by estimated coverage with the lowest levels at the bottom of the figure and higher levels near the top. The user can reverse that sort order or opt to plot strata in the same order as Excel tables; see Section B.9 for more details.

Even though the order and name datasets for level 1-3 are not used to control the output order anymore in the R version of VCQI, they are still required. Table B-1 lists how stratum names, and their order datasets are defined:

Table B-1. Level1-3 Names and Order Datasets

Level	Names	Order	Notes
1	Dataset: level1name	Not applicable	There is NO 's' on the end of the dataset name as there can only be one Level 1 stratum.
2	Dataset: level2names	Dataset: level2order	There is an 's' on the end of the dataset name.
3	Dataset: level3names	Dataset: level3order	VCQI obtains stratum names from variables like RI02, TT02, or SIA02, which hold the survey stratum name
4	If the user does <u>not</u> specify the name of a VCQI_LEVEL4_SET_LAYOUT dataset then VCQI obtains the LEVEL4 names from value label of the variable that defines the sub-groups or if the variable is a string variable, then the names are the strings themselves. If the user names a VCQI_LEVEL4_SET_LAYOUT dataset, then VCQI takes the LEVEL4 names from that file.	If the user does <u>not</u> specify the name of a VCQI_LEVEL4_SET_LAYOUT dataset then VCQI obtains the LEVEL4 order from the order in which stratifiers are listed in the VCQI_LEVEL4_SET_VARLIST and the order of values of those variables. If the user names a VCQI_LEVEL4_SET_LAYOUT dataset, then VCQI takes the LEVEL4 order from that file.	Examples: urban_cluster, gender If the string variable <i>gender</i> took on string values of <i>Male</i> and <i>Female</i> then the Level 4 names would be <i>Male</i> and <i>Female</i> . If the variable <i>gender</i> was an integer with a value label where 1 is labeled <i>Boys</i> and 2 is labeled <i>Girls</i> then once again, the Level 4 names would be 1: <i>Boys</i> and 2: <i>Girls</i> .

B.0 Demographic stratification using the VCQI_LEVEL4_SET_VARLIST and LAYOUT

There are two global values that control stratifiers. The first is VCQI_LEVEL4_SET_VARLIST. This is a required global value for the R version of VCQI and it must be populated with one or more variable names. If the user wants to have output at level 1-3, user need to include the corresponding level identifier in VCQI_LEVEL4_SET_VARLIST.

If the demographic stratifier variables have clear and succinct variable labels and value labels then it may not be necessary to also define a VCQI_LEVEL4_SET_LAYOUT dataset. When the user does not define such a dataset, VCQI builds one and saves it in the VCQI_OUTPUT_FOLDER. The file will be named VCQI_LEVEL4_LAYOUT_automatic.rds. This dataset controls the layout of demographic strata in the tables. If the user wants to modify how the strata are listed, they may edit this dataset and rename it and then tell VCQI to use their modified layout dataset. One thing to note is that, for the R version of VCQI, if user wants to have one lower level nested in within a higher level, then the user must define a layout file.

Structure of a LEVEL4 LAYOUT dataset

The LAYOUT dataset holds four variables.

1. order is a numeric variable that takes integer values starting with 1 and increasing by 1. It indicates the order in which the rows should appear in VCQI tables.
2. label is a string variable that holds the label, if any, that should appear in this row in the table
3. rowtype is a string variable that takes three possible values:
 - a. LABEL_ONLY means the row contains a label (i.e., Sex)
 - b. DATA_ROW means the row contains a condition (i.e., sex == 1)
 - c. BLANK_ROW means the user wants tables to include an extra blank row
4. condition is a string variable that holds R syntax to identify the demographic sub-group. For urban respondents, the condition might read “urban_cluster == 1” and for rural respondents it might read “urban_cluster == 0”.

So a simple LAYOUT dataset might look like this:

order	label	condition	rowtype
1	Is the cluster urban?		LABEL_ONLY
2	0: Rural	urban_cluster == 0	DATA_ROW
3	1: Urban	urban_cluster == 1	DATA_ROW

And if the user did not want the initial label to appear, s/he could edit the dataset to look like this:

order	label	condition	rowtype
1	0: Rural	urban_cluster == 0	DATA_ROW
2	1: Urban	urban_cluster == 1	DATA_ROW

Annex C. Customizing VCQI Plot Output

And if the user wanted the urban row to appear first and the rural row to appear second, the dataset might look like this:

order	label	condition	rowtype
1	1: Urban	urban_cluster == 1	DATA_ROW
2	0: Rural	urban_cluster == 0	DATA_ROW

If the user specified three demographic stratifiers, like this:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("urban_cluster", "RI20", "RI136"))
vcqi_global(VCQI_LEVEL4_SET_LAYOUT, NA)
```

If the variables in the Level4 varlist have variable labels and value labels (as they might if the input datasets are Stata .dta files), then VCQI would write a file named VCQI_LEVEL4_LAYOUT_automatic.rds that incorporates those labels and looks like this:

order	label	condition	rowtype
1	Is the cluster urban?		LABEL_ONLY
2	0: Rural	urban_cluster == 0	DATA_ROW
3	1: Urban	urban_cluster == 1	DATA_ROW
4	Sex		LABEL_ONLY
5	1: Male	RI20 == 1	DATA_ROW
6	2: Female	RI20 == 2	DATA_ROW
7	Did anyone from this household travel for 1+ months of the last 12 months?		LABEL_ONLY
8	1: Yes	RI136 == 1	DATA_ROW
9	2: No	RI136 == 2	DATA_ROW
10	99: Do not know	RI136 == 99	DATA_ROW

If the urban_cluster, RI20, and RI136 variables did *not* have variable and value labels, then the VCQI_LEVEL4_LAYOUT_automatic.rds file would look like this:

order	label	condition	rowtype
1	urban_cluster		LABEL_ONLY
2	0	urban_cluster == 0	DATA_ROW
3	1	urban_cluster == 1	DATA_ROW
4	RI20		LABEL_ONLY
5	1	RI20 == 1	DATA_ROW
6	2	RI20 == 2	DATA_ROW
7	RI136		LABEL_ONLY
8	1	RI136 == 1	DATA_ROW
9	2	RI136 == 2	DATA_ROW
10	99	RI136 == 99	DATA_ROW

The appearance of the VCQI_LEVEL4_LAYOUT files automatically created by VCQI will differ depending on whether the stratification variables have variable and/or value labels in the input datasets or whether they do not.

The user might edit the dataset to remove and revise some labels, like this:

order	label	condition	rowtype
1	0: Rural	urban_cluster == 0	DATA_ROW
2	1: Urban	urban_cluster == 1	DATA_ROW
3			BLANK_ROW
4	1: Male	RI20 == 1	DATA_ROW
5	2: Female	RI20 == 2	DATA_ROW
6	Anyone travel for 1+ months?		LABEL_ONLY
7	1: Yes	RI136 == 1	DATA_ROW
8	2: No	RI136 == 2	DATA_ROW
9	99: Do not know	RI136 == 99	DATA_ROW

The user could rename the dataset “layout_edited.rds” and re-run VCQI, specifying:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("urban_cluster", "RI20", "RI136"))
vcqi_global(VCQI_LEVEL4_SET_LAYOUT,
            paste0(VCQI_OUTPUT_FOLDER, "/layout_edited.rds"))
```

And the resulting Excel table would look like this:

Crude Coverage

	BCG crude coverage (%)	95% CI (%)
0: Rural	73.1	(69.3, 76.7)
1: Urban	74.3	(71.1, 77.3)
1: Male	73.2	(70.2, 76.0)
2:		
Female	74.5	(71.8, 77.0)
Anyone travel for 1+ months?		
1: Yes	75.6	(69.6, 80.7)
2: No	73.7	(71.4, 75.8)

The ability to edit the layout file gives the user substantial flexibility in specifying which stratifiers will appear in the tables and in what order. The conditions in the example above are simple, but the user could specify more complex strata, like this:

order	label	condition	rowtype
1	Rural Males (Families who Travel)	urban_cluster == 0 & RI20 == 1 & RI136 == 1	DATA_ROW
2	Rural Females (Families who Travel)	urban_cluster == 0 & RI20 == 2 & RI136 == 1	DATA_ROW

Annex C. Customizing VCQI Plot Output

Which results in a table like this:

Crude Coverage

	BCG crude coverage (%)	95% CI (%)
Rural Males (Families who Travel)	79.0	(65.2, 88.3)
Rural Females (Families who Travel)	73.3	(59.7, 83.6)

B.1 Sample listings of stratum name and order datasets for level 1-3

The following section shows the datasets for level 1-3 listed above for the sample dataset analyzed in this user's guide.

Dataset: level1name

level1id	level1name
1	Harmonia

Dataset: level2names

level2id	level2name
1	Southern Province
2	Northern Province

Dataset: level2order

level2id	level2order
1	1
2	2

Dataset: level3names

level3id	level3name
1	District 01
2	District 02
3	District 03
4	District 04
5	District 05
6	District 06
7	District 07
8	District 08
9	District 09
10	District 10

Dataset: level3order

level3id	level3order
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Note that it is not possible to examine these datasets to learn which Level 3 strata are nested inside which Level 2 strata. That nesting information is defined by the relationship between variables HH01 and province_id in the Cluster Metadata (CM) dataset. See the *VCQI Forms and Variable Lists (FVL)* document for CM details.

The purposes of the datasets described here are to a) specify how the stratum names should appear in VCQI output, and b) specify the relative order in which the strata should appear in tabulated Stata output. For both R and Stata, these files are also used to define output filenames and labels for RI_VCTC_01.

In this simple example, the level2order is the same as level2id, but that is not required. If the user wanted tables to show outcomes from the Northern Province first, s/he could reverse the order of level2order in the level2order dataset.

Similarly, level3order is not required to fall in the same sort order as level3id. The user may rearrange level3order in any manner they wish, as long as every stratum in the survey is represented in this dataset and each row has a unique value of level3order.

B.2 Example: Nested output for all Levels: 1, 2, and 3 with additional Level 4 stratification

This is an example which will show output for every level, 1-3, and for each of the Level 4 sub-groups in each, using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT. In this example, the Level 4 stratifier is the variable that codes whether the cluster is urban or rural.

Set VCQI_LEVEL4_SET_VARLIST like following:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("level1id", "level2id",
"level3id", "urban_cluster"))
```

To have level 3 nested under level 2 and level 4 nested under each of level 1-3, user must define a layout dataset. An example of VCQI_LEVEL4_SET_LAYOUT looks like this:

order	label	condition	rowtype
1	Harmonia	level1id == 1	DATA_ROW
2	Harmonia Rural	level1id == 1 & urban_cluster == 0	DATA_ROW
3	Harmonia Urban	level1id == 1 & urban_cluster == 1	DATA_ROW
4			BLANK_ROW
5	Southern Province	level2id == 1	DATA_ROW
6	Southern Province Rural	level2id == 1 & urban_cluster == 0	DATA_ROW
7	Southern Province Urban	level2id == 1 & urban_cluster == 1	DATA_ROW
8			BLANK_ROW
9	District 01	level3id == 1	DATA_ROW
10	District 01 Rural	level3id == 1 & urban_cluster == 0	DATA_ROW
11	District 01 Urban	level3id == 1 & urban_cluster == 1	DATA_ROW
12			BLANK_ROW
13	District 02	level3id == 2	DATA_ROW
14	District 02 Rural	level3id == 2 & urban_cluster == 0	DATA_ROW
15	District 02 Urban	level3id == 2 & urban_cluster == 1	DATA_ROW
16			BLANK_ROW
17	District 03	level3id == 3	DATA_ROW
18	District 03 Rural	level3id == 3 & urban_cluster == 0	DATA_ROW
19	District 03 Urban	level3id == 3 & urban_cluster == 1	DATA_ROW
20			BLANK_ROW
21	District 04	level3id == 4	DATA_ROW
22	District 04 Rural	level3id == 4 & urban_cluster == 0	DATA_ROW
23	District 04 Urban	level3id == 4 & urban_cluster == 1	DATA_ROW
24			BLANK_ROW
25	District 05	level3id == 5	DATA_ROW
18	District 05 Rural	level3id == 5 & urban_cluster == 0	DATA_ROW
19	District 05 Urban	level3id == 5 & urban_cluster == 1	DATA_ROW
20			BLANK_ROW
21	Northern Province	level2id == 2	DATA_ROW
22	Northern Province Rural	level2id == 2 & urban_cluster == 0	DATA_ROW
23	Northern Province Urban	level2id == 2 & urban_cluster == 1	DATA_ROW
24			BLANK_ROW
25	District 06	level3id == 6	DATA_ROW
26	District 06 Rural	level3id == 6 & urban_cluster == 0	DATA_ROW
27	District 06 Urban	level3id == 6 & urban_cluster == 1	DATA_ROW
28			BLANK_ROW
29	District 07	level3id == 7	DATA_ROW
30	District 07 Rural	level3id == 7 & urban_cluster == 0	DATA_ROW
31	District 07 Urban	level3id == 7 & urban_cluster == 1	DATA_ROW

order	label	condition	rowtype
32			BLANK_ROW
33	District 08	level3id == 8	DATA_ROW
34	District 08 Rural	level3id == 8 & urban_cluster == 0	DATA_ROW
35	District 08 Urban	level3id == 8 & urban_cluster == 1	DATA_ROW
36			BLANK_ROW
37	District 09	level3id == 9	DATA_ROW
38	District 09 Rural	level3id == 9 & urban_cluster == 0	DATA_ROW
39	District 09 Urban	level3id == 9 & urban_cluster == 1	DATA_ROW
40			BLANK_ROW
41	District 10	level3id == 10	DATA_ROW
42	District 10 Rural	level3id == 10 & urban_cluster == 0	DATA_ROW
43	District 10 Urban	level3id == 10 & urban_cluster == 1	DATA_ROW

The following pages show national, provincial, and district level results, each broken out by urban and rural sub-groups.¹⁶

Figure B-1 shows coverage estimates in table order. Thus, the national coverage appears first, followed by Southern Province and districts in Southern Province. Northern Province and all the districts for the Northern Province appear at the bottom of the page. Within each district, or province, or national level, the rural sub-group appeared first then the urban sub-group.

¹⁶ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%).”

Annex C. Customizing VCQI Plot Output

Table B-2. Nested output for all Levels: 1-3 with Level 4 stratification

Crude Coverage

	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
Harmonia	73.9	(71.7, 75.9)	1.1	72.0	75.6	2.0	0.0855	3,256	6,857,718
Harmonia Rural	73.1	(69.3, 76.7)	1.9	69.9	76.1	2.2	0.0733	1,236	2,733,272
Harmonia Urban	74.3	(71.1, 77.3)	1.6	71.6	76.9	2.7	0.0938	2,020	4,124,446
Southern Province	63.7	(60.6, 66.6)	1.5	61.1	66.1	1.7	0.0497	1,683	3,405,594
Southern Province Rural	63.6	(59.0, 68.0)	2.3	59.8	67.4	1.7	0.0393	727	1,513,999
Southern Province Urban	63.7	(59.3, 67.8)	2.2	60.0	67.1	1.9	0.0593	956	1,891,595
District 01	57.5	(51.2, 63.7)	3.1	52.3	62.7	1.4	0.0290	353	1,174,632
District 01 Rural	57.8	(48.2, 66.8)	4.6	49.8	65.3	1.4	0.0238	156	514,308
District 01 Urban	57.4	(48.8, 65.5)	4.1	50.3	64.2	1.4	0.0403	197	660,324
District 02	69.5	(62.8, 75.5)	3.1	64.0	74.6	1.7	0.0546	375	344,454
District 02 Rural	72.8	(60.8, 82.2)	5.3	63.0	80.9	1.8	0.0738	127	117,252
District 02 Urban	67.8	(59.6, 75.1)	3.8	61.0	74.0	1.7	0.0492	248	227,201
District 03	73.9	(67.9, 79.1)	2.8	68.9	78.2	1.3	0.0257	329	699,989
District 03 Rural	73.0	(65.3, 79.5)	3.5	66.7	78.5	1.2	0.0099	188	398,457
District 03 Urban	75.0	(65.4, 82.6)	4.3	67.1	81.5	1.4	0.0582	141	301,532
District 04	57.1	(51.4, 62.6)	2.8	52.3	61.7	1.0	-0.0102	296	579,974
District 04 Rural	59.6	(50.8, 67.8)	4.2	52.3	66.5	1.1	-0.0024	151	292,888
District 04 Urban	54.4	(46.3, 62.3)	3.5	47.6	61.1	1.0	-0.0180	145	287,086
District 05	66.7	(59.1, 73.4)	3.5	60.4	72.4	1.8	0.0678	330	606,545
District 05 Rural	60.4	(49.2, 70.7)	5.4	51.1	69.1	1.3	0.0390	105	191,094
District 05 Urban	69.5	(59.9, 77.7)	4.4	61.6	76.4	2.1	0.0763	225	415,452
Northern Province	83.9	(81.2, 86.3)	1.3	81.7	86.0	1.9	0.0384	1,573	3,452,123
Northern Province Rural	84.9	(80.9, 88.3)	1.9	81.6	87.8	1.4	0.0059	509	1,219,273
Northern Province Urban	83.4	(79.7, 86.5)	1.7	80.3	86.1	2.3	0.0529	1,064	2,232,850
District 06	80.7	(75.5, 85.1)	2.4	76.4	84.4	1.1	0.0081	303	226,281
District 06 Rural	80.6	(72.5, 86.7)	3.4	73.9	85.8	1.0	0.0102	118	86,671
District 06 Urban	80.8	(73.5, 86.5)	3.2	74.8	85.7	1.2	0.0131	185	139,610
District 07	80.1	(73.9, 85.0)	2.7	75.1	84.3	1.6	0.0561	336	1,005,880

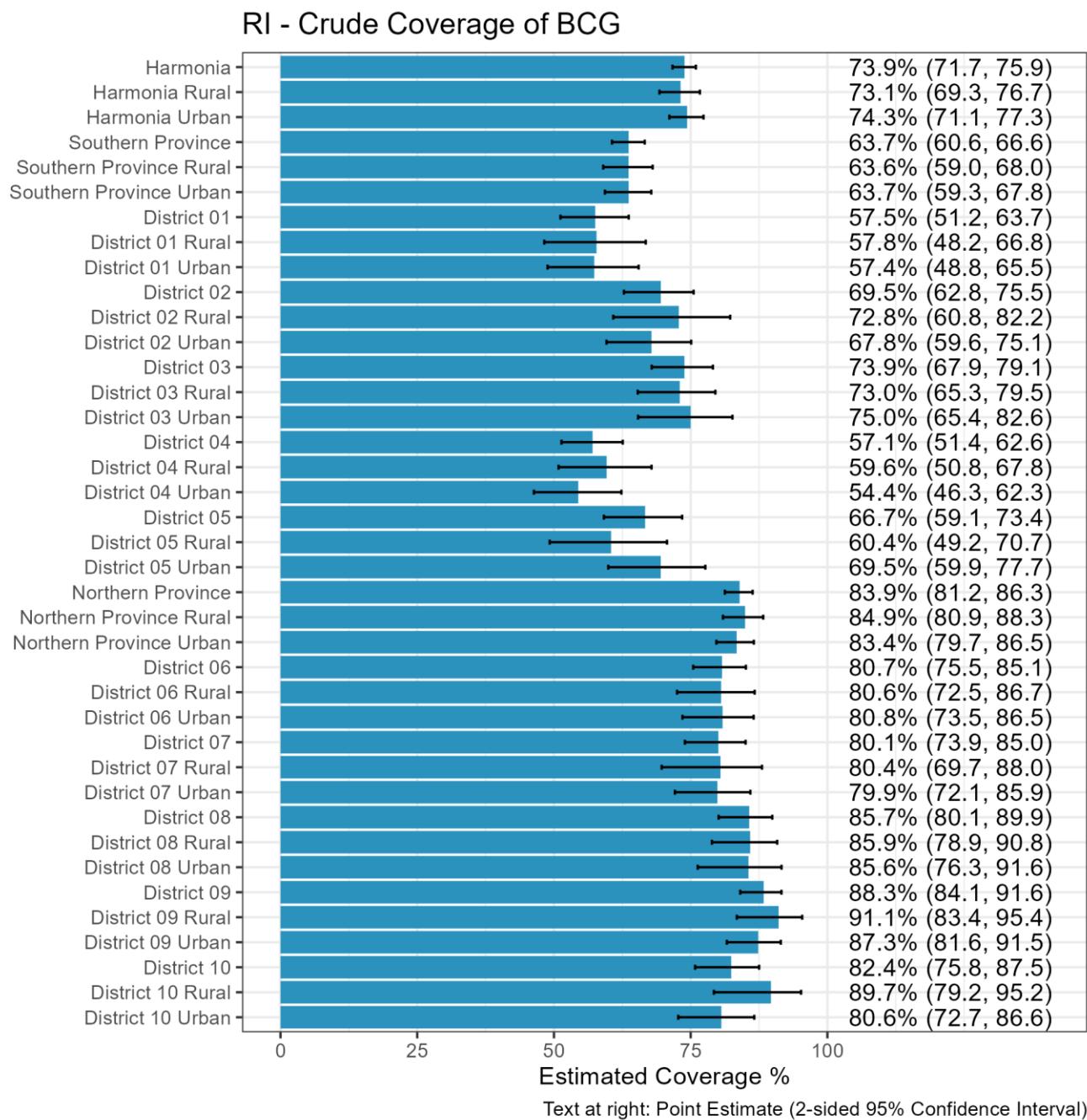
Annex B. Controlling How Strata Are Listed in VCQI Output

District 07 Rural	80.4	(69.7, 88.0)	4.5	71.7	87.0	1.4	0.0494	111	325,490
District 07 Urban	79.9	(72.1, 85.9)	3.4	73.5	85.0	1.6	0.0657	225	680,390
District 08	85.7	(80.1, 89.9)	2.4	81.1	89.3	1.4	0.0414	299	1,303,848
District 08 Rural	85.9	(78.9, 90.8)	2.7	80.1	90.2	1.0	-0.0009	131	580,696
District 08 Urban	85.6	(76.3, 91.6)	3.7	78.1	90.8	1.9	0.0829	168	723,152
District 09	88.3	(84.1, 91.6)	1.8	84.9	91.1	1.1	0.0146	343	632,694
District 09 Rural	91.1	(83.4, 95.4)	2.0	84.9	94.9	1.0	-0.0607	91	169,850
District 09 Urban	87.3	(81.6, 91.5)	2.4	82.7	90.9	1.3	0.0353	252	462,844
District 10	82.4	(75.8, 87.5)	2.9	77.0	86.8	1.7	0.0458	292	283,420
District 10 Rural	89.7	(79.2, 95.2)	2.8	81.2	94.6	1.0	-0.0912	58	56,565
District 10 Urban	80.6	(72.7, 86.6)	3.4	74.2	85.7	1.7	0.0611	234	226,855

Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intraclass Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Figure B-1. Bar plot showing nested output in table order for all Levels: 1-3 with Level 4 stratification



B.3 Example: Nested output for all Levels: 1, 2, and 3

This is an example to show output for every level, 1-3, with level 3 output nested under level 2, but no additional stratification by sub-group, using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT.

Set VCQI_LEVEL4_SET_VARLIST like following:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST,
c("level1id", "level2id", "level3id"))
```

To have level 3 nested under level 2, the user must define a layout dataset. An example of VCQI_LEVEL4_SET_LAYOUT looks like this:

order	label	condition	rowtype
1	Harmonia	level1id == 1	DATA_ROW
2			BLANK_ROW
3	Southern Province	level2id == 1	DATA_ROW
4	District 01	level3id == 1	DATA_ROW
5	District 02	level3id == 2	DATA_ROW
6	District 03	level3id == 3	DATA_ROW
7	District 04	level3id == 4	DATA_ROW
8	District 05	level3id == 5	DATA_ROW
9			BLANK_ROW
10	Northern Province	level2id == 2	DATA_ROW
11	District 06	level3id == 6	DATA_ROW
12	District 07	level3id == 7	DATA_ROW
13	District 08	level3id == 8	DATA_ROW
14	District 09	level3id == 9	DATA_ROW
15	District 10	level3id == 10	DATA_ROW

The following page shows national, provincial, and district level results.¹⁷

Figure B-2 shows coverage estimates sorted in table order. Thus, the national coverage appears first, followed by Southern Province and districts in Southern Province. Northern Province and all the districts for the Northern Province appear at the bottom of the page.

¹⁷ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%”).

Annex C. Customizing VCQI Plot Output

Table B-3. Nested output for all Levels: 1, 2, and 3

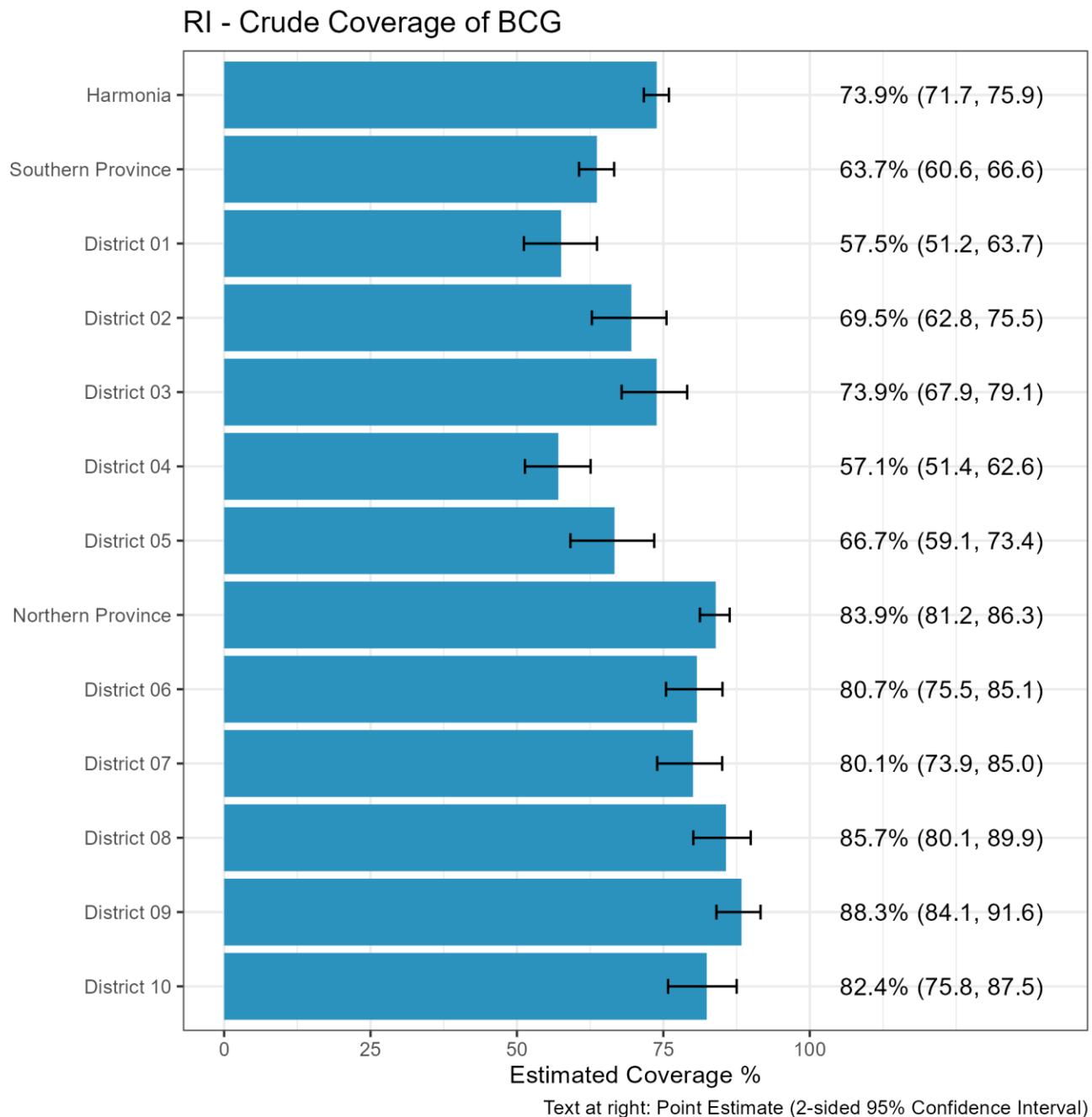
Crude Coverage

	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
Harmonia	73.9	(71.7, 75.9)	1.1	72.0	75.6	2.0	0.0855	3,256	6,857,718
Southern Province	63.7	(60.6, 66.6)	1.5	61.1	66.1	1.7	0.0497	1,683	3,405,594
District 01	57.5	(51.2, 63.7)	3.1	52.3	62.7	1.4	0.0290	353	1,174,632
District 02	69.5	(62.8, 75.5)	3.1	64.0	74.6	1.7	0.0546	375	344,454
District 03	73.9	(67.9, 79.1)	2.8	68.9	78.2	1.3	0.0257	329	699,989
District 04	57.1	(51.4, 62.6)	2.8	52.3	61.7	1.0	-0.0102	296	579,974
District 05	66.7	(59.1, 73.4)	3.5	60.4	72.4	1.8	0.0678	330	606,545
Northern Province	83.9	(81.2, 86.3)	1.3	81.7	86.0	1.9	0.0384	1,573	3,452,123
District 06	80.7	(75.5, 85.1)	2.4	76.4	84.4	1.1	0.0081	303	226,281
District 07	80.1	(73.9, 85.0)	2.7	75.1	84.3	1.6	0.0561	336	1,005,880
District 08	85.7	(80.1, 89.9)	2.4	81.1	89.3	1.4	0.0414	299	1,303,848
District 09	88.3	(84.1, 91.6)	1.8	84.9	91.1	1.1	0.0146	343	632,694
District 10	82.4	(75.8, 87.5)	2.9	77.0	86.8	1.7	0.0458	292	283,420

Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intraclass Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Figure B-2. Bar plot showing nested output in table order for all Levels: 1, 2, and 3



B.4 Example: Non-nested output for all Levels: 1, 2, and 3

This is an example to show output for every level, 1-3, with level 3 output listed underneath level 2, but not nested, and without stratification by sub-group, using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT.

To set VCQI_LEVEL4_SET_VARLIST for every level, non-nested:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST,
            c("level1id", "level2id", "level3id"))
```

The corresponding VCQI_LEVEL4_SET_LAYOUT dataset looks like this:

order	label	condition	rowtype
1	level1		LABEL_ONLY
2	Harmonia	level1id == 1	DATA_ROW
3	level2		LABEL_ONLY
4	Southern Province	level2id == 1	DATA_ROW
5	Northern Province	level2id == 2	DATA_ROW
6	level3		LABEL_ONLY
7	District 01	level3id == 1	DATA_ROW
8	District 02	level3id == 2	DATA_ROW
9	District 03	level3id == 3	DATA_ROW
10	District 04	level3id == 4	DATA_ROW
11	District 05	level3id == 5	DATA_ROW
12	District 06	level3id == 6	DATA_ROW
13	District 07	level3id == 7	DATA_ROW
14	District 08	level3id == 8	DATA_ROW
15	District 09	level3id == 9	DATA_ROW
16	District 10	level3id == 10	DATA_ROW

However, if level 3 is not nested within level 2, then a VCQI_LEVEL4_SET_LAYOUT dataset is not required as VCQI will generate one for the user. The VCQI_LEVEL4_SET_LAYOUT generated by VCQI will have a slightly different “label” column than the example above.

The following page shows national, provincial, and district level results.¹⁸ The order in which results are listed is controlled by VCQI_LEVEL4_SET_LAYOUT dataset.

¹⁸ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%”).

Table B-4. Non-nested output for all Levels: 1-3

Crude Coverage

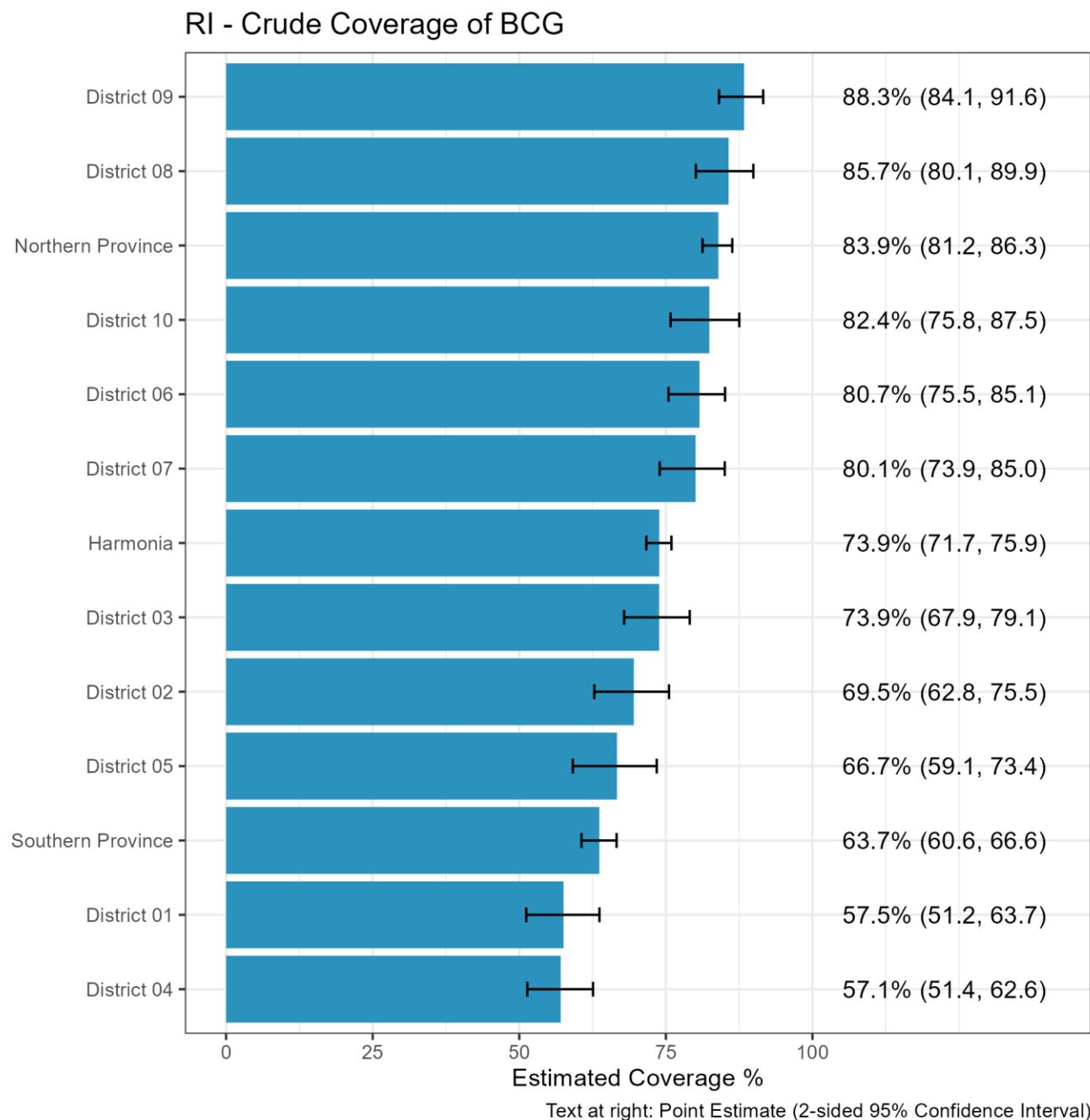
	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
level1									
Harmonia	73.9	(71.7, 75.9)	1.1	72.0	75.6	2.0	0.0855	3,256	6,857,718
level2									
Southern Province	63.7	(60.6, 66.6)	1.5	61.1	66.1	1.7	0.0497	1,683	3,405,594
Northern Province	83.9	(81.2, 86.3)	1.3	81.7	86.0	1.9	0.0384	1,573	3,452,123
level3									
District 01	57.5	(51.2, 63.7)	3.1	52.3	62.7	1.4	0.0290	353	1,174,632
District 02	69.5	(62.8, 75.5)	3.1	64.0	74.6	1.7	0.0546	375	344,454
District 03	73.9	(67.9, 79.1)	2.8	68.9	78.2	1.3	0.0257	329	699,989
District 04	57.1	(51.4, 62.6)	2.8	52.3	61.7	1.0	-0.0102	296	579,974
District 05	66.7	(59.1, 73.4)	3.5	60.4	72.4	1.8	0.0678	330	606,545
District 06	80.7	(75.5, 85.1)	2.4	76.4	84.4	1.1	0.0081	303	226,281
District 07	80.1	(73.9, 85.0)	2.7	75.1	84.3	1.6	0.0561	336	1,005,880
District 08	85.7	(80.1, 89.9)	2.4	81.1	89.3	1.4	0.0414	299	1,303,848
District 09	88.3	(84.1, 91.6)	1.8	84.9	91.1	1.1	0.0146	343	632,694
District 10	82.4	(75.8, 87.5)	2.9	77.0	86.8	1.7	0.0458	292	283,420

Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intracluster Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Figure B-3 is the same as Figure B-2 but with the Block D parameter PLOT_OUTCOMES_IN_TABLE_ORDER set to 0. The plot is sorted in order of estimated coverage. If the parameter is set to 1, then bar plots will list strata in the same order as tables. See section B.8 for an example.

Figure B-3. Bar plot showing non-nested output sorted for all Levels: 1, 2, and 3



B.5 Example: Output for Level 3 only

This is an example to have level 3 only result using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT.

To set VCQI_LEVEL4_SET_VARLIST for level 3 only:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "level3id")
```

An example of the VCQI_LEVEL4_SET_LAYOUT dataset looks like this:

order	label	condition	rowtype
1	District		LABEL_ONLY
2	District 01	level3id == 1	DATA_ROW
3	District 02	level3id == 2	DATA_ROW
4	District 03	level3id == 3	DATA_ROW
5	District 04	level3id == 4	DATA_ROW
6	District 05	level3id == 5	DATA_ROW
7	District 06	level3id == 6	DATA_ROW
8	District 07	level3id == 7	DATA_ROW
9	District 08	level3id == 8	DATA_ROW
10	District 09	level3id == 9	DATA_ROW
11	District 10	level3id == 10	DATA_ROW

The following page shows district level results.¹⁹ The order in which results are listed is controlled by VCQI_LEVEL4_SET_LAYOUT dataset.

If the user chooses to let VCQI auto-generate a VCQI_LEVEL4_SET_LAYOUT dataset, the labels will look slightly different.

¹⁹ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%”).

Annex C. Customizing VCQI Plot Output

Table B-5. Output for Level 3 only

Crude Coverage

District	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
District 01	57.5	(51.2, 63.7)	3.1	52.3	62.7	1.4	0.0290	353	1,174,632
District 02	69.5	(62.8, 75.5)	3.1	64.0	74.6	1.7	0.0546	375	344,454
District 03	73.9	(67.9, 79.1)	2.8	68.9	78.2	1.3	0.0257	329	699,989
District 04	57.1	(51.4, 62.6)	2.8	52.3	61.7	1.0	-0.0102	296	579,974
District 05	66.7	(59.1, 73.4)	3.5	60.4	72.4	1.8	0.0678	330	606,545
District 06	80.7	(75.5, 85.1)	2.4	76.4	84.4	1.1	0.0081	303	226,281
District 07	80.1	(73.9, 85.0)	2.7	75.1	84.3	1.6	0.0561	336	1,005,880
District 08	85.7	(80.1, 89.9)	2.4	81.1	89.3	1.4	0.0414	299	1,303,848
District 09	88.3	(84.1, 91.6)	1.8	84.9	91.1	1.1	0.0146	343	632,694
District 10	82.4	(75.8, 87.5)	2.9	77.0	86.8	1.7	0.0458	292	283,420

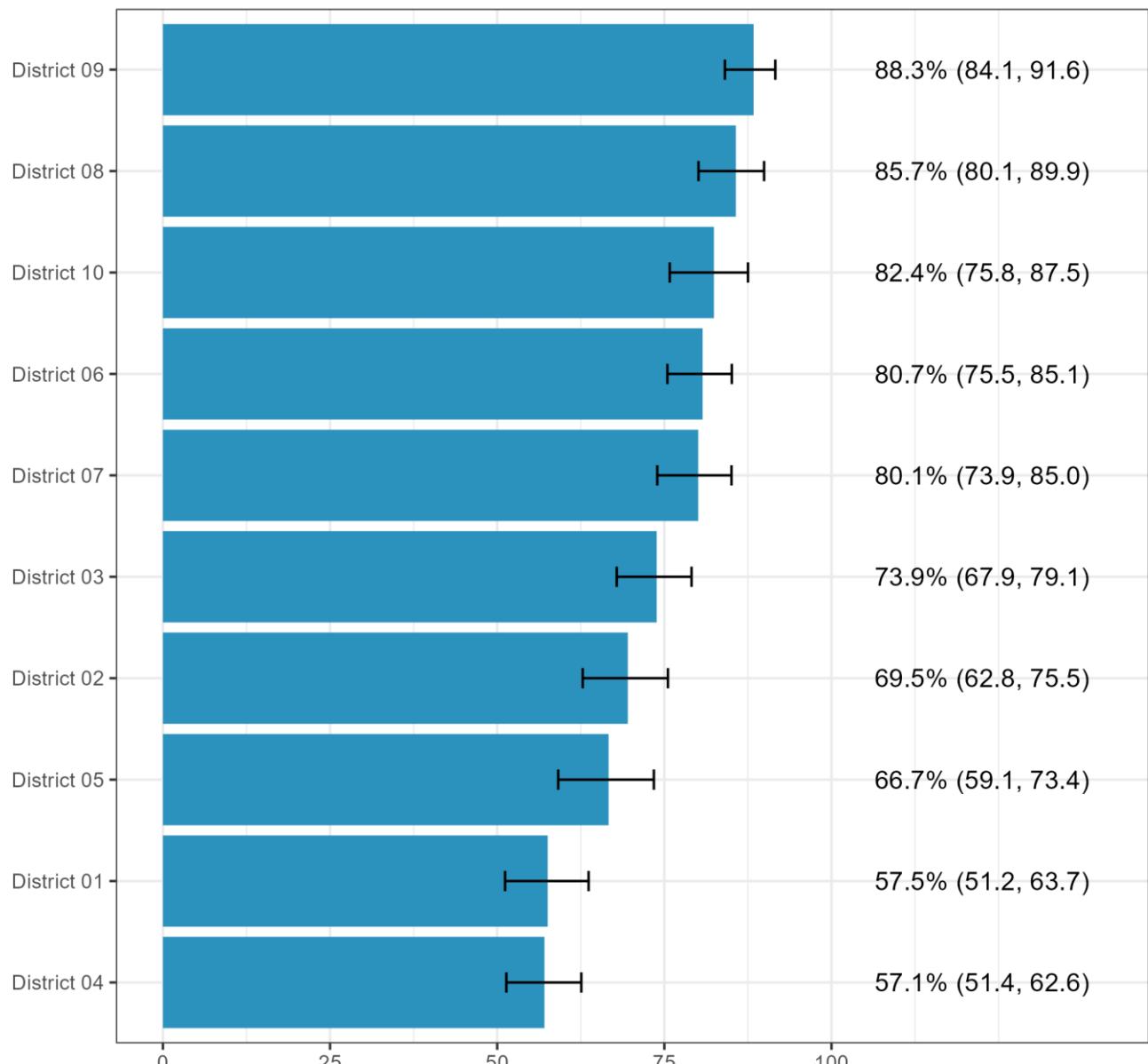
Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intracluster Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Figure B-4 page shows Level 3 strata sorted in order of estimated coverage. Neither the table above nor Figure B-4 make any reference whatsoever to Level 2 strata.

Figure B-4. Bar plot showing sorted output for Level 3 only

RI - Crude Coverage of BCG



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

B.6 Example: Output for Level 3 with additional Level 4 stratification

This is an example to show output for level 3 and for the urban and rural sub-groups in each district, using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT.

Set VCQI_LEVEL4_SET_VARLIST like following:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("level3id", "urban_cluster"))
```

To have level 4 nested within level3, the user must have a layout file. An example of the VCQI_LEVEL4_SET_LAYOUT dataset looks like this:

order	label	condition	rowtype
1	District 1		LABEL_ONLY
2	District 01	level3id == 1	DATA_ROW
3	District 01 Rural	level3id == 1 & urban_cluster == 0	DATA_ROW
4	District 01 Urban	level3id == 1 & urban_cluster == 1	DATA_ROW
5	District 02		LABEL_ONLY
6	District 02	level3id == 2	DATA_ROW
7	District 02 Rural	level3id == 2 & urban_cluster == 0	DATA_ROW
8	District 02 Urban	level3id == 2 & urban_cluster == 1	DATA_ROW
9	District 03		LABEL_ONLY
10	District 03	level3id == 3	DATA_ROW
11	District 03 Rural	level3id == 3 & urban_cluster == 0	DATA_ROW
12	District 03 Urban	level3id == 3 & urban_cluster == 1	DATA_ROW
13	District 04		LABEL_ONLY
14	District 04	level3id == 4	DATA_ROW
15	District 04 Rural	level3id == 4 & urban_cluster == 0	DATA_ROW
16	District 04 Urban	level3id == 4 & urban_cluster == 1	DATA_ROW
17	District 05		LABEL_ONLY
18	District 05	level3id == 5	DATA_ROW
19	District 05 Rural	level3id == 5 & urban_cluster == 0	DATA_ROW
20	District 05 Urban	level3id == 5 & urban_cluster == 1	DATA_ROW
21	District 06		LABEL_ONLY
22	District 06	level3id == 6	DATA_ROW
23	District 06 Rural	level3id == 6 & urban_cluster == 0	DATA_ROW
24	District 06 Urban	level3id == 6 & urban_cluster == 1	DATA_ROW
25	District 07		LABEL_ONLY
26	District 07	level3id == 7	DATA_ROW
27	District 07 Rural	level3id == 7 & urban_cluster == 0	DATA_ROW
28	District 07 Urban	level3id == 7 & urban_cluster == 1	DATA_ROW
29	District 08		LABEL_ONLY
30	District 08	level3id == 8	DATA_ROW
31	District 08 Rural	level3id == 8 & urban_cluster == 0	DATA_ROW
32	District 08 Urban	level3id == 8 & urban_cluster == 1	DATA_ROW
33	District 09		LABEL_ONLY
34	District 09	level3id == 9	DATA_ROW
35	District 09 Rural	level3id == 9 & urban_cluster == 0	DATA_ROW
36	District 09 Urban	level3id == 9 & urban_cluster == 1	DATA_ROW
37	District 10		LABEL_ONLY
38	District 10	level3id == 10	DATA_ROW
39	District 10 Rural	level3id == 10 & urban_cluster == 0	DATA_ROW
40	District 10 Urban	level3id == 10 & urban_cluster == 1	DATA_ROW

The following page shows district level results.²⁰ The order in which results are listed is controlled by VCQI_LEVEL4_SET_LAYOUT dataset.

Figure B-5 shows estimated coverage in the table order. Neither the table above nor Figure B-5 make any reference to Level 2 strata.

²⁰ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%).”

Annex C. Customizing VCQI Plot Output

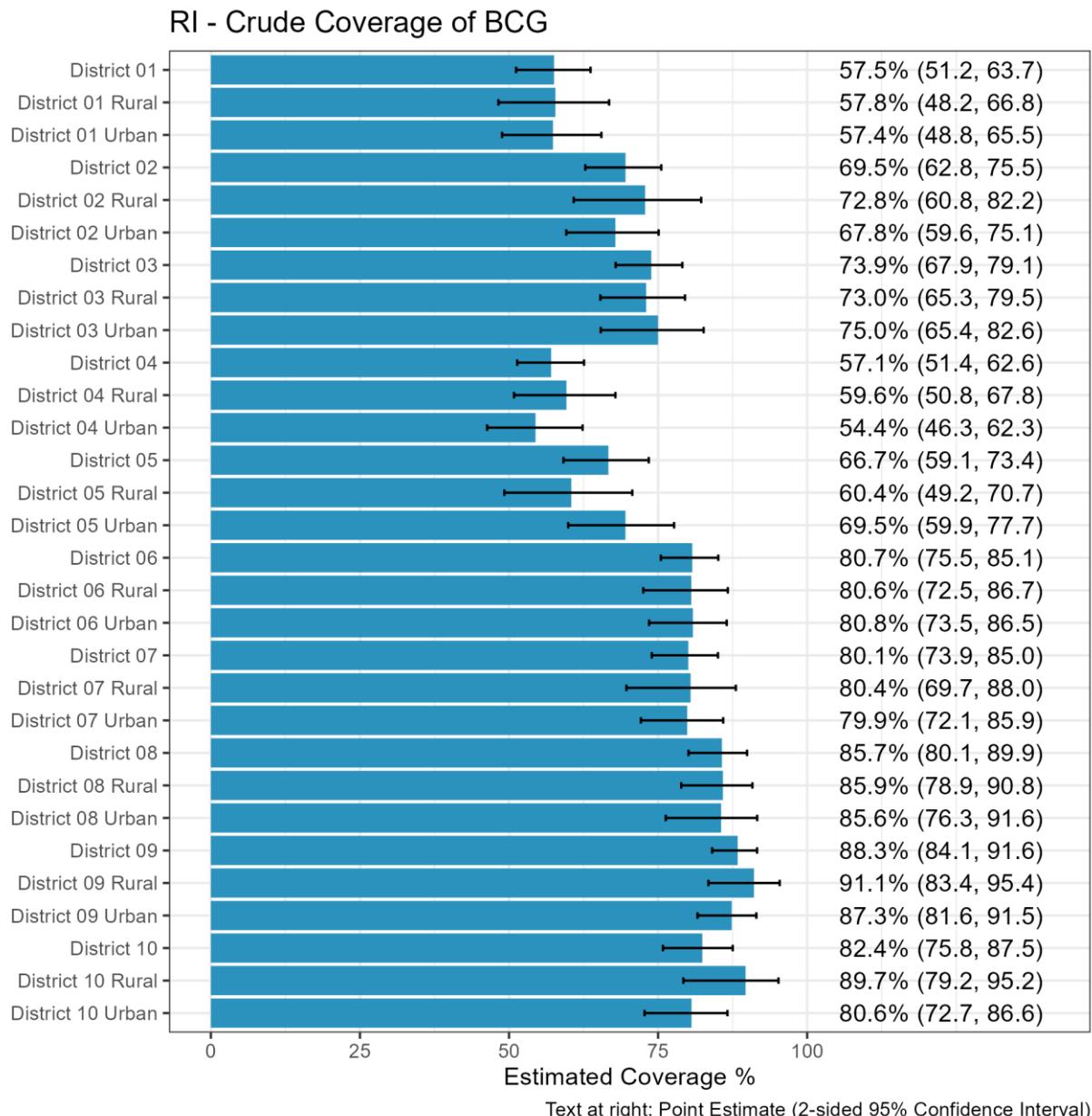
Table B-6. Output for Level 3 with Level 4 stratification

Crude Coverage	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
District 1									
District 01	57.5	(51.2, 63.7)	3.1	52.3	62.7	1.4	0.0290	353	1,174,632
District 01 Rural	57.8	(48.2, 66.8)	4.6	49.8	65.3	1.4	0.0238	156	514,308
District 01 Urban	57.4	(48.8, 65.5)	4.1	50.3	64.2	1.4	0.0403	197	660,324
District 02									
District 02	69.5	(62.8, 75.5)	3.1	64.0	74.6	1.7	0.0546	375	344,454
District 02 Rural	72.8	(60.8, 82.2)	5.3	63.0	80.9	1.8	0.0738	127	117,252
District 02 Urban	67.8	(59.6, 75.1)	3.8	61.0	74.0	1.7	0.0492	248	227,201
District 03									
District 03	73.9	(67.9, 79.1)	2.8	68.9	78.2	1.3	0.0257	329	699,989
District 03 Rural	73.0	(65.3, 79.5)	3.5	66.7	78.5	1.2	0.0099	188	398,457
District 03 Urban	75.0	(65.4, 82.6)	4.3	67.1	81.5	1.4	0.0582	141	301,532
District 04									
District 04	57.1	(51.4, 62.6)	2.8	52.3	61.7	1.0	-0.0102	296	579,974
District 04 Rural	59.6	(50.8, 67.8)	4.2	52.3	66.5	1.1	-0.0024	151	292,888
District 04 Urban	54.4	(46.3, 62.3)	3.5	47.6	61.1	1.0	-0.0180	145	287,086
District 05									
District 05	66.7	(59.1, 73.4)	3.5	60.4	72.4	1.8	0.0678	330	606,545
District 05 Rural	60.4	(49.2, 70.7)	5.4	51.1	69.1	1.3	0.0390	105	191,094
District 05 Urban	69.5	(59.9, 77.7)	4.4	61.6	76.4	2.1	0.0763	225	415,452
District 06									
District 06	80.7	(75.5, 85.1)	2.4	76.4	84.4	1.1	0.0081	303	226,281
District 06 Rural	80.6	(72.5, 86.7)	3.4	73.9	85.8	1.0	0.0102	118	86,671
District 06 Urban	80.8	(73.5, 86.5)	3.2	74.8	85.7	1.2	0.0131	185	139,610
District 07									
District 07	80.1	(73.9, 85.0)	2.7	75.1	84.3	1.6	0.0561	336	1,005,880
District 07 Rural	80.4	(69.7, 88.0)	4.5	71.7	87.0	1.4	0.0494	111	325,490
District 07 Urban	79.9	(72.1, 85.9)	3.4	73.5	85.0	1.6	0.0657	225	680,390
District 08									
District 08	85.7	(80.1, 89.9)	2.4	81.1	89.3	1.4	0.0414	299	1,303,848
District 08 Rural	85.9	(78.9, 90.8)	2.7	80.1	90.2	1.0	-0.0009	131	580,696
District 08 Urban	85.6	(76.3, 91.6)	3.7	78.1	90.8	1.9	0.0829	168	723,152
District 09									
District 09	88.3	(84.1, 91.6)	1.8	84.9	91.1	1.1	0.0146	343	632,694
District 09 Rural	91.1	(83.4, 95.4)	2.0	84.9	94.9	1.0	-0.0607	91	169,850
District 09 Urban	87.3	(81.6, 91.5)	2.4	82.7	90.9	1.3	0.0353	252	462,844
District 10									
District 10	82.4	(75.8, 87.5)	2.9	77.0	86.8	1.7	0.0458	292	283,420
District 10 Rural	89.7	(79.2, 95.2)	2.8	81.2	94.6	1.0	-0.0912	58	56,565
District 10 Urban	80.6	(72.7, 86.6)	3.4	74.2	85.7	1.7	0.0611	234	226,855

Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intraclass Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Figure B-5. Bar plot showing output in table order for Level 3 with Level 4 stratification



B.7 Example: Output for Level 4 Only

This is an example to have level 4 only result using VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT.

To set VCQI_LEVEL4_SET_VARLIST with the desired level 4 variables:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("urban_cluster", "RI20"))
```

An example of the VCQI_LEVEL4_SET_LAYOUT dataset looks like this:

order	label	condition	rowtype
1	Is the cluster urban?		LABEL_ONLY
2	0: Rural	urban_cluster == 0	DATA_ROW
3	1: Urban	urban_cluster == 1	DATA_ROW
4	Child's Sex		LABEL_ONLY
5	Male	RI20 == 1	DATA_ROW
6	Female	RI20 == 2	DATA_ROW

This will produce output like Table B-7.²¹ The order in which results are listed is controlled by VCQI_LEVEL4_SET_LAYOUT dataset.

If the user chooses to let VCQI auto-generate a VCQI_LEVEL4_SET_LAYOUT dataset, the labels may look slightly different. As discussed in Section B.0, the appearance of the automatic Level4 layout dataset depends on whether the stratification variables have variable and/or value labels.

²¹ The table shows columns for the main result of BCG crude coverage only that considers evidence from card and history. It omits columns for intermediate variables like “BCG crude coverage, by card (%”).

Table B-7. Output for Level 4 alone

Crude Coverage

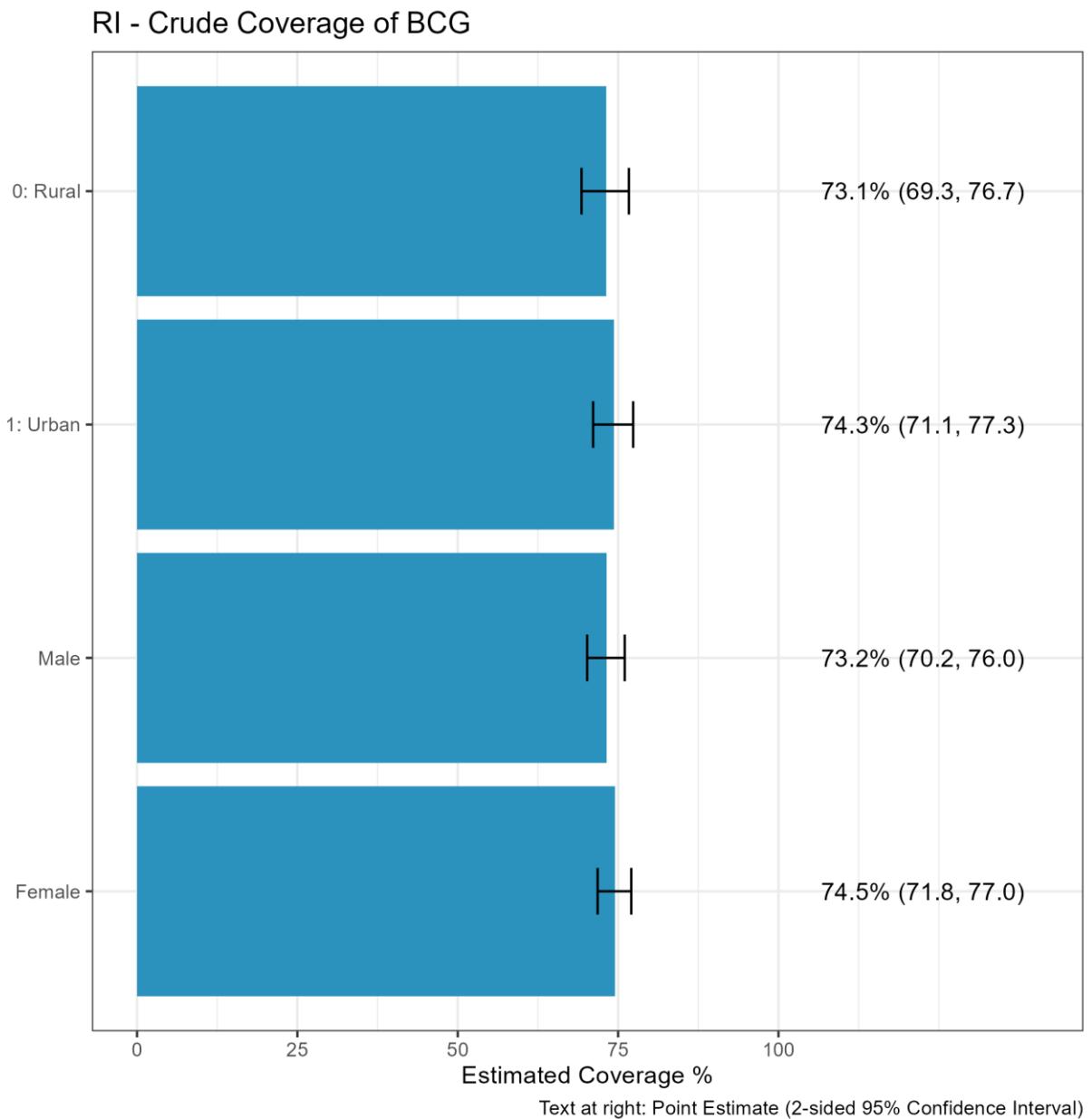
	BCG crude coverage (%)	95% CI (%)	StdErr (%)	95% LCB (%)	95% UCB (%)	DEFF	ICC	N	Weighted N
Is the cluster urban?									
0: Rural	73.1	(69.3, 76.7)	1.9	69.9	76.1	2.2	0.0733	1,236	2,733,272
1: Urban	74.3	(71.1, 77.3)	1.6	71.6	76.9	2.7	0.0938	2,020	4,124,446
Child's Sex									
Male	73.2	(70.2, 76.0)	1.5	70.7	75.6	1.8	0.0743	1,608	3,389,901
Female	74.5	(71.8, 77.0)	1.3	72.3	76.7	1.5	0.1041	1,648	3,467,817

Abbreviations: CI=Confidence Interval; LCB=Lower Confidence Bound; UCB=Upper Confidence Bound; DEFF=Design Effect; ICC=Intraclass Correlation Coefficient

Note: This measure is a population estimate that incorporates survey weights. The CI, LCB and UCB are calculated with software that take the complex survey design into account.

Annex C. Customizing VCQI Plot Output

Figure B-6. Bar plot showing output in table order for Level 4 only



B.8 A Large Realistic Level 4 LAYOUT Example

The LAYOUT dataset examples in this *User's Guide* are mostly brief because of space considerations, but the full power of the Level 4 LAYOUT dataset is best demonstrated with a large realistic example from Nigeria. In the 2016-17 Multi-Indicator Cluster Survey/National Immunization Coverage Survey (MICS/NICS),²² the table shells followed a common UNICEF MICS layout with tables split across two pages for each outcome. The first table lists seven stratifiers: sex, geopolitical zone, urban/rural, caretaker's education, caretaker's age, household wealth and ethnicity. The second table shows outcomes for the entire country and for each of Nigeria's 36 states plus the federal capital territory, arranged by geopolitical zone. The following pages specify the table layout using customized VCQI Level 4 LAYOUT files. You might follow these examples in your own work.

²² Survey report: <https://www.nigerianstat.gov.ng/nada/index.php/catalog/59/download/573>

Rhoda, D.A., Wagai, J.N., Beshanski-Pedersen, B.R., Yusafari, Y., Sequeira, J., Hayford, K., Brown, D.W., Danovaro-Holliday, M.C., Braka, F., Ali, D. and Shuaib, F., 2020. Combining cluster surveys to estimate vaccination coverage: Experiences from Nigeria's multiple indicator cluster survey/national immunization coverage survey (MICS/NICS), 2016–17. *Vaccine*, 38(39), pp.6174-6183. <https://doi.org/10.1016/j.vaccine.2020.05.058>

Wagai, J.N.; Rhoda, D.A.; Prier, M.L.; Trimner, M.K.; Clary, C.B.; Oteri, J.; Okposen, B.; Adeniran, A.; Danovaro-Holliday, M.C.; Cutts, F.T. Implementing WHO Guidance on Conducting and Analysing Vaccination Coverage Cluster Surveys: Two Examples from Nigeria. Preprints 2020, 2020090645 (doi: 10.20944/preprints202009.0645.v2).

Annex C. Customizing VCQI Plot Output

This Level 4 LAYOUT dataset describes the table row order for seven stratification variables in Nigeria's 2016-17 MICS/NICS Survey

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST,
c("MICS_5_hl4", "MICS_5_zone", "urban_cluster", "caretaker_education",
"caretaker_age_category", "MICS_5_windex", "MICS_5_ethnicity"))
```

order	label	condition	rowtype
1	Sex		LABEL_ONLY
2	Male	MICS_5_hl4 == 1	DATA_ROW
3	Female	MICS_5_hl4 == 2	DATA_ROW
4	Geopolitical Zone		LABEL_ONLY
5	North Central	MICS_5_zone == 1	DATA_ROW
6	North East	MICS_5_zone == 2	DATA_ROW
7	North West	MICS_5_zone == 3	DATA_ROW
8	South East	MICS_5_zone == 4	DATA_ROW
9	South South	MICS_5_zone == 5	DATA_ROW
10	South West	MICS_5_zone == 6	DATA_ROW
11	Area		LABEL_ONLY
12	Urban	urban_cluster == 1	DATA_ROW
13	Rural	urban_cluster == 2	DATA_ROW
14	Caretaker's Education		LABEL_ONLY
15	Primary	caretaker_education == 1	DATA_ROW
16	Secondary/technical	caretaker_education == 2	DATA_ROW
17	Higher	caretaker_education == 3	DATA_ROW
18	Non-formal	caretaker_education == 4	DATA_ROW
19	Missing	caretaker_education == 9	DATA_ROW
20	Caretaker's Age		LABEL_ONLY
21	15-19	caretaker_age_category == 1	DATA_ROW
22	20-29	caretaker_age_category == 2	DATA_ROW
23	30-39	caretaker_age_category == 3	DATA_ROW
24	40-49	caretaker_age_category == 4	DATA_ROW
25	50+	caretaker_age_category == 5	DATA_ROW
26	DNK	caretaker_age_category == 6	DATA_ROW
27	Wealth index quintile		LABEL_ONLY
28	Poorest	MICS_5_windex5 == 1	DATA_ROW
29	Second	MICS_5_windex5 == 2	DATA_ROW
30	Middle	MICS_5_windex5 == 3	DATA_ROW
31	Fourth	MICS_5_windex5 == 4	DATA_ROW
32	richest	MICS_5_windex5 == 5	DATA_ROW
33	Ethnicity		LABEL_ONLY
34	Hausa	MICS_5_ethnicity == 1	DATA_ROW
35	Igbo	MICS_5_ethnicity == 2	DATA_ROW
36	Yoruba	MICS_5_ethnicity == 3	DATA_ROW
37	Other	!(MICS_5_ethnicity %in% c(1:3))	DATA_ROW

This Level 4 LAYOUT dataset describes a table row order where national results are followed by results for 36 states and the federal capital territory, sorted by geopolitical zone and then alphabetically, for Nigeria's 2016-17 MICS/NICS Survey

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "MICS_5_hh7")
```

order	label	condition	rowtype
1	Nigeria	MICS_5_hh7 %in% c(1:37)	DATA_ROW
2	State		LABEL_ONLY
3	North Central		LABEL_ONLY
4	FCT-Abuja	MICS_5_hh7 == 37	DATA_ROW
5	Benue	MICS_5_hh7 == 7	DATA_ROW
6	Kogi	MICS_5_hh7 == 22	DATA_ROW
7	Kwara	MICS_5_hh7 == 23	DATA_ROW
8	Nasarawa	MICS_5_hh7 == 25	DATA_ROW
9	Niger	MICS_5_hh7 == 26	DATA_ROW
10	Plateau	MICS_5_hh7 == 31	DATA_ROW
11	North East		LABEL_ONLY
12	Adamawa	MICS_5_hh7 == 2	DATA_ROW
13	Bauchi	MICS_5_hh7 == 5	DATA_ROW
14	Borno	MICS_5_hh7 == 8	DATA_ROW
15	Gombe	MICS_5_hh7 == 15	DATA_ROW
16	Taraba	MICS_5_hh7 == 34	DATA_ROW
17	Yobe	MICS_5_hh7 == 35	DATA_ROW
18	North West		LABEL_ONLY
19	Jigawa	MICS_5_hh7 == 17	DATA_ROW
20	Kaduna	MICS_5_hh7 == 18	DATA_ROW
21	Kano	MICS_5_hh7 == 19	DATA_ROW
22	Katsina	MICS_5_hh7 == 20	DATA_ROW
23	Kebbi	MICS_5_hh7 == 21	DATA_ROW
24	Sokoto	MICS_5_hh7 == 33	DATA_ROW
25	Zamfara	MICS_5_hh7 == 36	DATA_ROW
26	South East		LABEL_ONLY
27	Abia	MICS_5_hh7 == 1	DATA_ROW
28	Anambra	MICS_5_hh7 == 4	DATA_ROW
29	Ebonyi	MICS_5_hh7 == 11	DATA_ROW
30	Enugu	MICS_5_hh7 == 14	DATA_ROW
31	Imo	MICS_5_hh7 == 16	DATA_ROW
32	South South		LABEL_ONLY
33	Akwa Ibom	MICS_5_hh7 == 3	DATA_ROW
34	Bayelsa	MICS_5_hh7 == 6	DATA_ROW
35	Cross River	MICS_5_hh7 == 9	DATA_ROW
36	Delta	MICS_5_hh7 == 10	DATA_ROW
37	Edo	MICS_5_hh7 == 12	DATA_ROW
38	Rivers	MICS_5_hh7 == 32	DATA_ROW
39	South West		LABEL_ONLY
40	Ekiti	MICS_5_hh7 == 13	DATA_ROW
41	Lagos	MICS_5_hh7 == 24	DATA_ROW
42	Ogun	MICS_5_hh7 == 27	DATA_ROW
43	Ondo	MICS_5_hh7 == 28	DATA_ROW
44	Osun	MICS_5_hh7 == 29	DATA_ROW
45	Oyo	MICS_5_hh7 == 30	DATA_ROW

Annex C. Customizing VCQI Plot Output

The flexibility afforded by Level 4 LAYOUT should enable you to construct tables with whatever table rows are suitable for your survey report. If one of the Level 1-4 combinations defined above in sections B.2 through B.6 is what you need, specify the input for VCQI_LEVEL4_SET_VARLIST and VCQI_LEVEL4_SET_LAYOUT you see there. But if you wish to customize the structure of table rows, specify the list of stratification variables in the VCQI_LEVEL4_SET_VARLIST and then customize the default LAYOUT dataset that VCQI produces in your first run.

Workflow to create the default LAYOUT and then customize it:

1. Specify a list of stratification variables in the VCQI_LEVEL4_SET_VARLIST and define the VCQI_LEVEL4_SET_LAYOUT as NA.
2. Run VCQI. It will generate a default LAYOUT dataset in your VCQI_OUTPUT_FOLDER named VCQI_LEVEL4_SET_LAYOUT_automatic.rds.
3. Copy that dataset; name the copy something like VCQI_LEVEL4_LAYOUT_customized.rds. Import the dataset in R to modify the layout and labels to suit you using options described in section B.0 and what you see in the Nigeria example above. Save the edited dataset in your VCQI_OUTPUT_FOLDER.
4. Edit the control program and update the LAYOUT line to point to your customized dataset:
`vcqi_global(VCQI_LEVEL4_SET_LAYOUT, paste0(VCQI_OUTPUT_FOLDER,
"/VCQI_LAYOUT_customized.rds"))`
5. Re-run VCQI. The output table will appear as defined by the customized LAYOUT dataset.

B.9 Changing Order of Strata in Bar and Unweighted Plots

By default, strata in bar plots, double bar plots, and unweighted plots are sorted by the estimated outcome with the poorest performing strata at the bottom of the figure and highest performers near the top. Users can change this behavior in two ways: by using the SORT_PLOT_LOW_TO_HIGH global to reverse the default plot order, or by using the PLOT_OUTCOMES_IN_TABLE_ORDER global to plot strata in the same order they appear in tabular output.

Note that when PLOT_OUTCOMES_IN_TABLE_ORDER is set to 1, then strata will be plotted in table order regardless of the value of the SORT_PLOT_LOW_TO_HIGH global.

In VCQI's template control programs, the global SORT_PLOT_LOW_TO_HIGH is reset for each outcome in Block F to 1 or 0 depending on what the outcome represents. For indicators that summarize a desirable outcome, like vaccination coverage, the value is 1, meaning that strata with lower outcomes are plotted at the bottom of the figure. But for outcomes that represent an undesirable outcome, like drop-out, low values represent the desired condition, so the global is set to 0 so strata with high outcomes (i.e., poor performance) are plotted at the bottom of the figure and those with low outcomes (which is what we hope to see) are plotted at the top.

Two RI_COVG_01 plots below illustrate the SORT_PLOT_LOW_TO_HIGH options. (In both cases, the global PLOT_OUTCOMES_IN_TABLE_ORDER is set to 0.)

Annex C. Customizing VCQI Plot Output

Figure B-7. Bar plot with strata sorted by outcomes, low-to-high, bottom-to-top

```
vcqi_global(SORT_PLOT_LOW_TO_HIGH, 1)
```

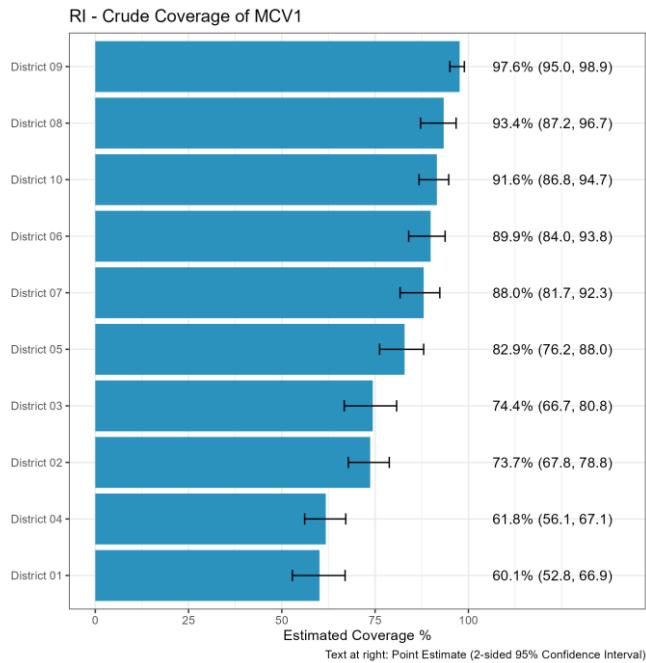
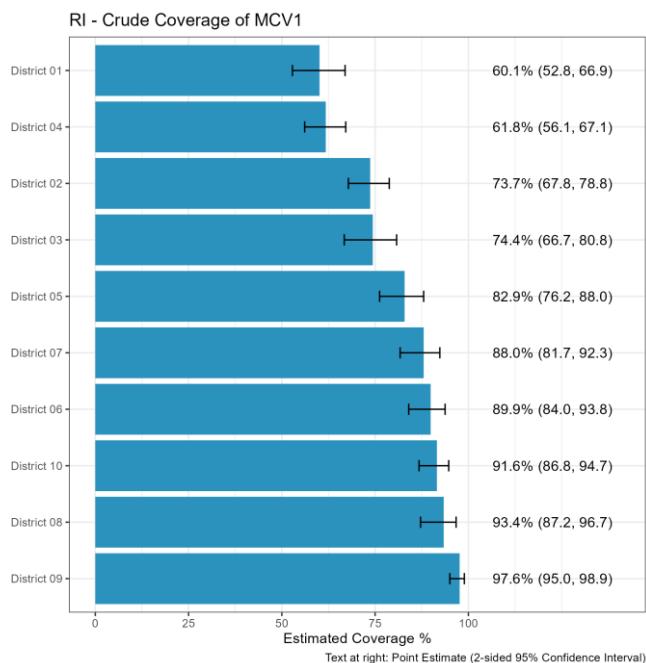


Figure B-8. Bar plot with strata sorted by outcomes, high-to-low, bottom-to-top

```
vcqi_global(SORT_PLOT_LOW_TO_HIGH, 0)
```



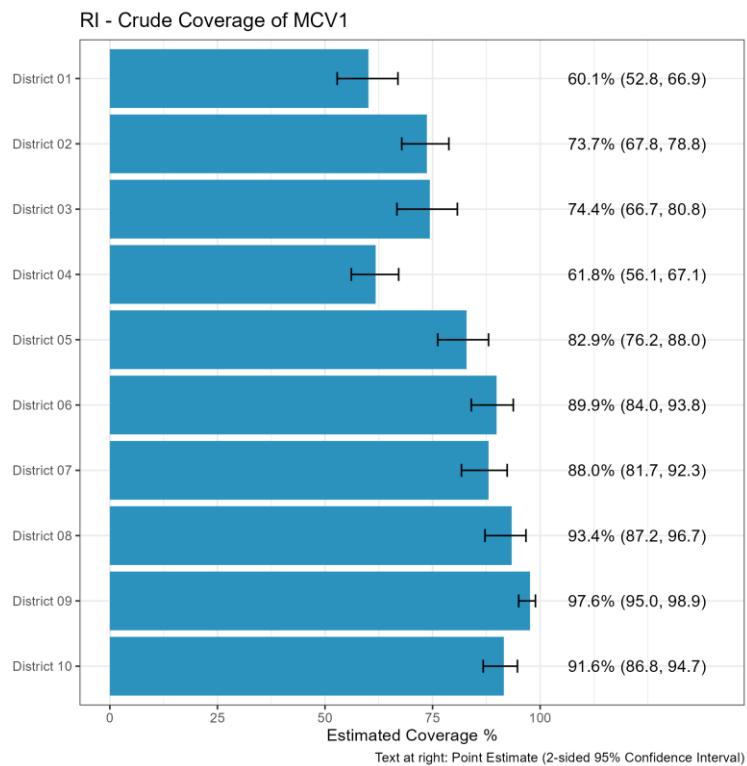
Sorting strata from worst-to-best performance yields visually pleasing figures, but means that the position of each stratum varies from figure to figure. For example, the district listed at the top of the MCV1 crude coverage bar plot, because it had the highest coverage of measles vaccine, might appear second or third or fourth in the figure showing OPV1 crude coverage. Indeed, that district might appear at the very bottom of the figure for some other outcome if its performance is worse than all the other districts. Some viewers find it confusing to have to search for their favorite stratum – to have to scan up and down the list of district names at the left side of the figure – time and time again rather than have each fall in the same predictable location on each figure.

If the user wishes to plot strata in a predictable order, then set PLOT_OUTCOMES_IN_TABLE_ORDER to 1 and the order of strata in every bar plot, double-bar plot, and unweighted proportion plot will be precisely the same order that the strata appear in VCQI's Excel format tabulated outcome tables.

This figure demonstrates the result, listing the districts in order 1 thru 10.

Figure B-9. Bar plot with strata sorted in table order

```
vcqi_global(PLOT_OUTCOMES_IN_TABLE_ORDER, 1)
```



(Again, as a reminder, if the user sets PLOT_OUTCOMES_IN_TABLE_ORDER to 1 then VCQI will ignore the value of SORT_PLOT_LOW_TO_HIGH.)

B.10 Additional Options for Customizing VCQI Output

If VCQI's options to define table layouts are not to your liking, then you have the option of writing a program to access the results in the database files and construct customized tables of your own.

Similarly, you may wish to construct bar plots that use a different order or different set of colors than the VCQI default. In that case you might wish to re-run VCQI and export datasets that save plot inputs (setting the `VCQI_SAVE_IW_PLOT_DATA` and/or `VCQI_SAVE_UW_PLOT_DATA` globals to 1 in the control program), load a plot dataset into R, and write code to produce a customized plot. Annex C sections C.2, C.3, and C.4 contain example ggplot2 code to reproduce VCQI's bar charts using the plot datasets.

VCQI outcome databases may be imported into Excel or R or Tableau or any visualization and table-making software that the user has mastered in order to customize every aspect of how results are presented to vaccination stakeholders. Do not hesitate to contact the VCQI developers if you have questions or suggestions about how to summarize and visualize survey results.

ANNEX C. CUSTOMIZING VCQI PLOT OUTPUT

C.1 Customizing Plot Colors

VCQI's default plot color for the bar plots is:

VCQI Blue 1 (RGB: 49, 130, 189; HEX: #2b92be).



VCQI's default plot colors for the double bar plots are:

VCQI Blue 1 (RGB: 49, 130, 189; HEX: #2b92be)



"lightgrey"(RGB: 211,211,211; HEX: #d3d3d3)



To customize plot colors for bar plots, double bar plots or unweighted plots, set the global VCQI_SAVE_IW_PLOT_DATA or VCQI_SAVE_UW_PLOT_DATA to 1 and then load the saved .rds dataset file into R. User can modify the ggplot codes provided in Section C.3, Section C.4 or Section C.5 and recreate the plots with the color the user prefers.

For now, the user cannot change the colors for organ pipe plots. The default colors for opplot are:

#9ecae1 (RGB: 158, 202, 225)



##f0f0f0 (RGB: 240, 240, 240)



C.2 Changing Confidence Interval Information on Bar Plots

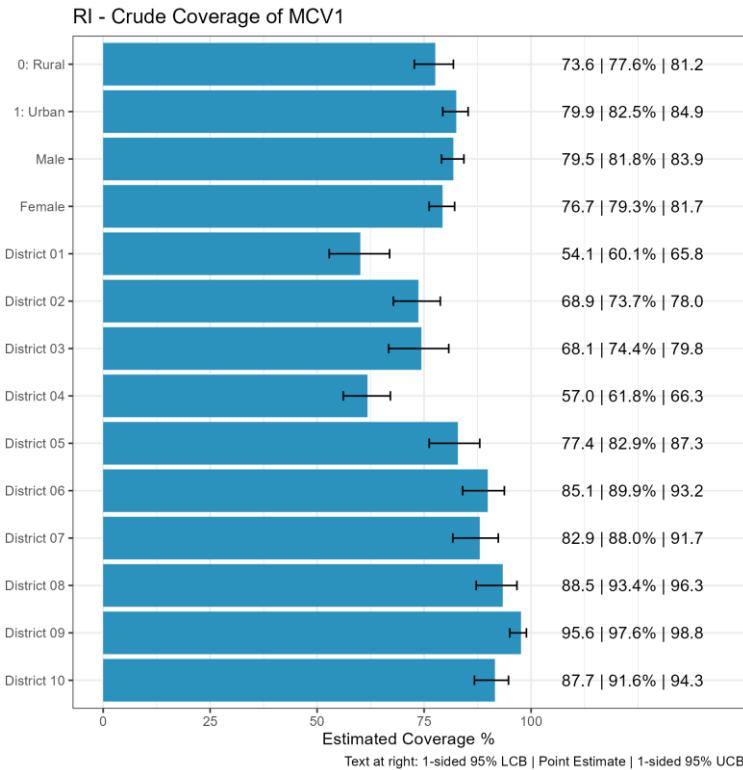
The VCQI_IWPLLOT_CITEXT global allows users to customize the confidence interval information included on single bar plots. This global can take the integer values 1, 2, 3, 4, or 5, which produce the following results:

1. 1-sided 95% LCB | Point Estimate | 1-sided 95% UCB
Example: 74.5 | 75.9% | 77.3
2. Point Estimate (2-sided 95% Confidence Interval)
Example: 75.9% (74.2, 77.6)
Option 2 is the VCQI default.
3. Point Estimate (2-sided 95% Confidence Interval) (0, 1-sided 95% UCB]
Example: 75.9% (74.2, 77.6) (0, 77.3]
4. Point Estimate (2-sided 95% Confidence Interval) [1-sided 95% UCB, 100)
Example: 75.9% (74.2, 77.6) [74.5, 100)
5. Point Estimate (2-sided 95% CI) (0, 1-sided 95% UCB] [1-sided 95% LCB, 100)
Example: 75.9% (74.2, 77.6) (0, 77.3] [74.5, 100)

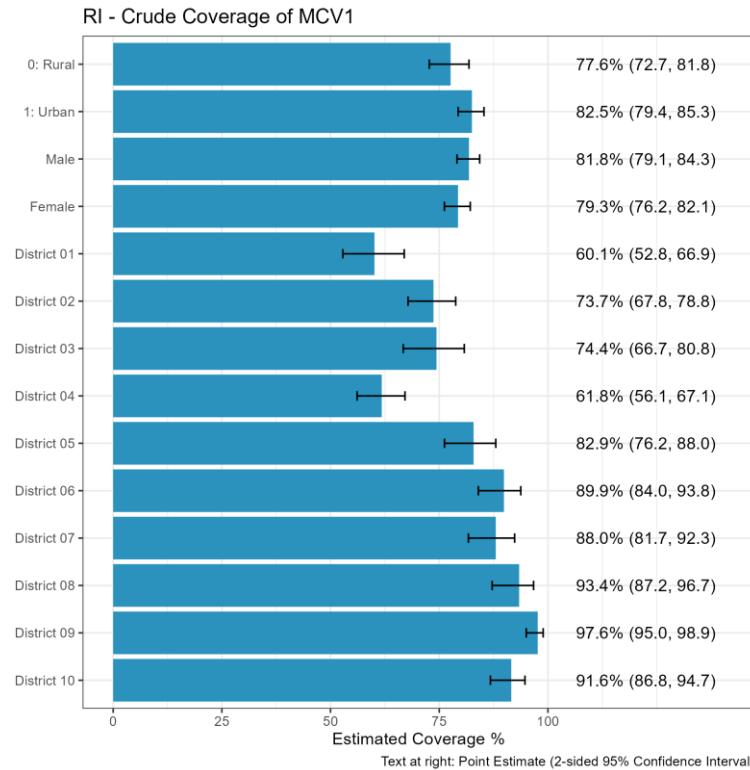
Illustrations of each of these options follow.

Figure C-1. Bar plot demonstrating VCQI_IWPLLOT_CITEXT 1

```
vcqi_global(VCQI_IWPLLOT_CITEXT, 1)
```

**Figure C-2. Bar plot demonstrating VCQI_IWPLLOT_CITEXT 2**

```
vcqi_global(VCQI_IWPLLOT_CITEXT, 2)
```



Annex C. Customizing VCQI Plot Output

Figure C-3. Bar plot demonstrating VCQI_IWPLLOT_CITEXT 3

```
vcqi_global(VCQI_IWPLLOT_CITEXT, 3)
```

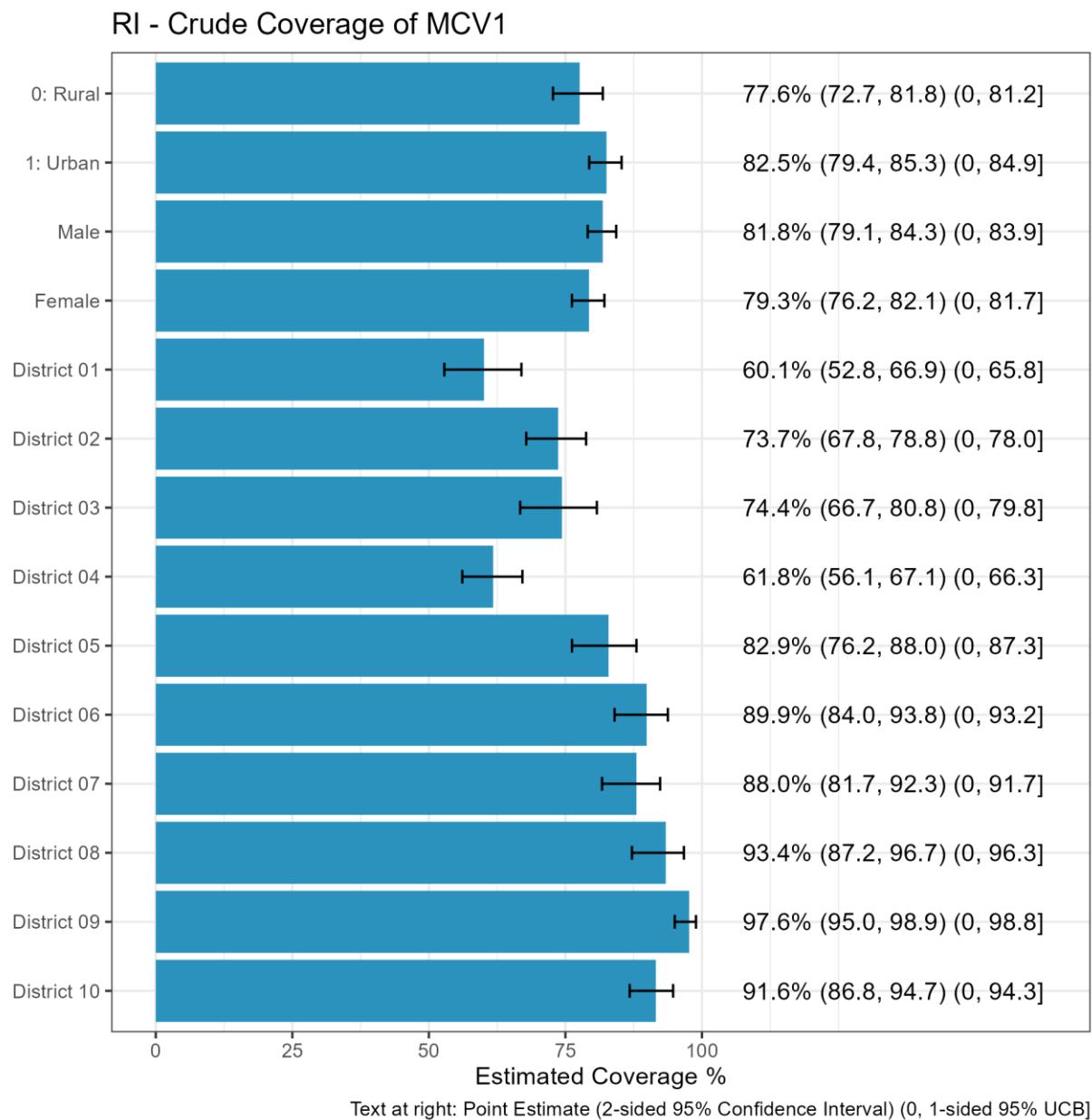
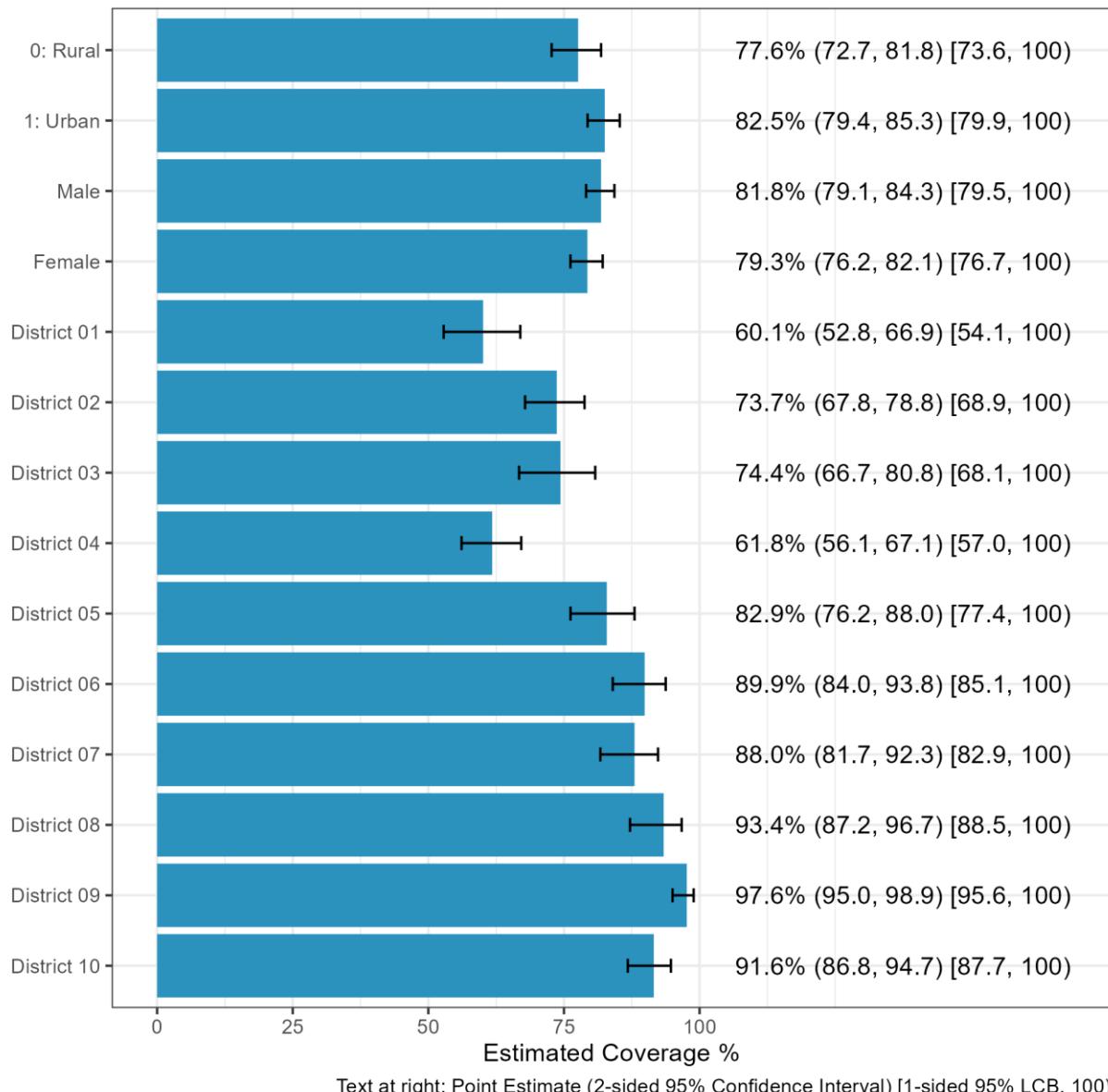


Figure C-4. Bar plot demonstrating VCQI_IWPLLOT_CITEXT 4

```
vcqi_global(VCQI_IWPLLOT_CITEXT, 4)
```

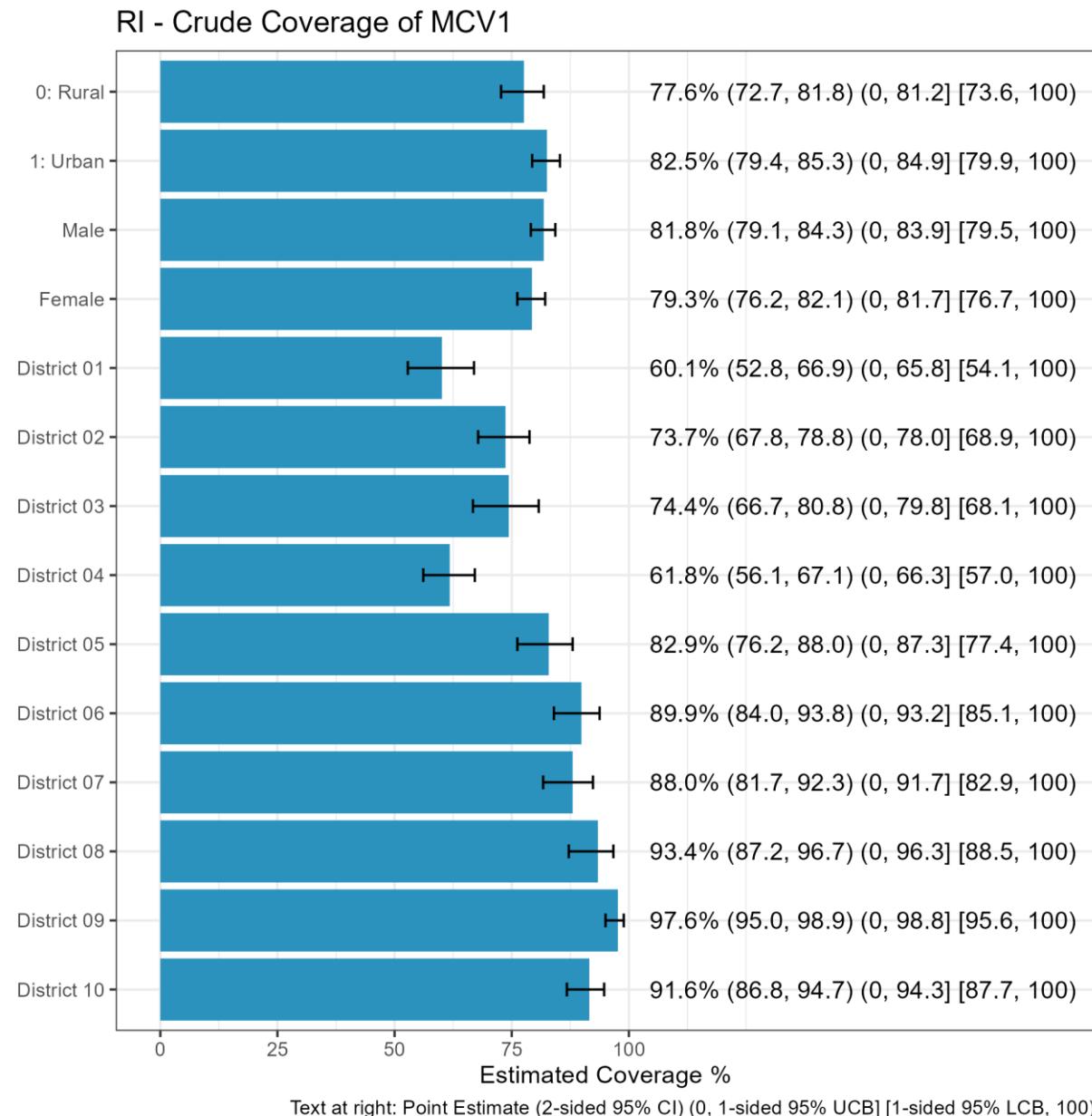
RI - Crude Coverage of MCV1



Annex C. Customizing VCQI Plot Output

Figure C-5. Bar plot demonstrating VCQI_IW PLOT_CITEXT 5

```
vcqi_global(VCQI_IW PLOT_CITEXT, 5)
```



To recover text clipped at the right side of a figure:

Note that in some cases, the longer annotations used in options 3, 4, and 5 are clipped in the .png output when the plot is saved. To correct this problem, set the global VCQI_SAVE_IW_PLOT_DATA to 1 in the control program and then load the saved .rds dataset file into R and modify the following ggplot code to recreate and adjust the plot:

```

library(ggplot2)
dat <- readRDS(<path to plot data>)

extraspace <- max(nchar(dat$text))

if (is.na(dat$graphtitle[1])){
  title <- NULL
} else {
  title <- dat$graphtitle[1]
}

if (is.na(dat$graphsubtitle[1])){
  subtitle <- NULL
} else {
  subtitle <- dat$graphsubtitle[1]
}

if (is.na(dat$graphcaption[1])){
  note <- NULL
} else {
  note <- dat$graphcaption[1]
}

ggplot(dat, aes(x = as.factor(rowid), y = estimate * 100)) +
  geom_col(fill = "#2b92be") +
  geom_errorbar(aes(ymin = cill * 100, ymax = ciul * 100),
                 width = .2,
                 position = position_dodge(.9)) +
  geom_text(aes(
    x = as.factor(rowid),
    y = 100 + 1.25*extraspace,
    label = text
  )) +
  coord_flip() +
  labs(
    y = "Estimated Coverage %",
    x = "",
    title = title,
    subtitle = subtitle,
    caption = note
  ) +
  scale_x_discrete(labels = dat$name) +
  scale_y_continuous(limits = c(0, 100 + 2.25*extraspace),
                     breaks = c(0, 25, 50, 75, 100)) +
  theme(plot.caption = element_text(hjust = 0)) +
  theme_bw()

ggsave(<filename.png>, width = <customize width>, height = <customize height>,
       units = "in")

```

C.3 Changing Confidence Interval Information on Double Bar Plots

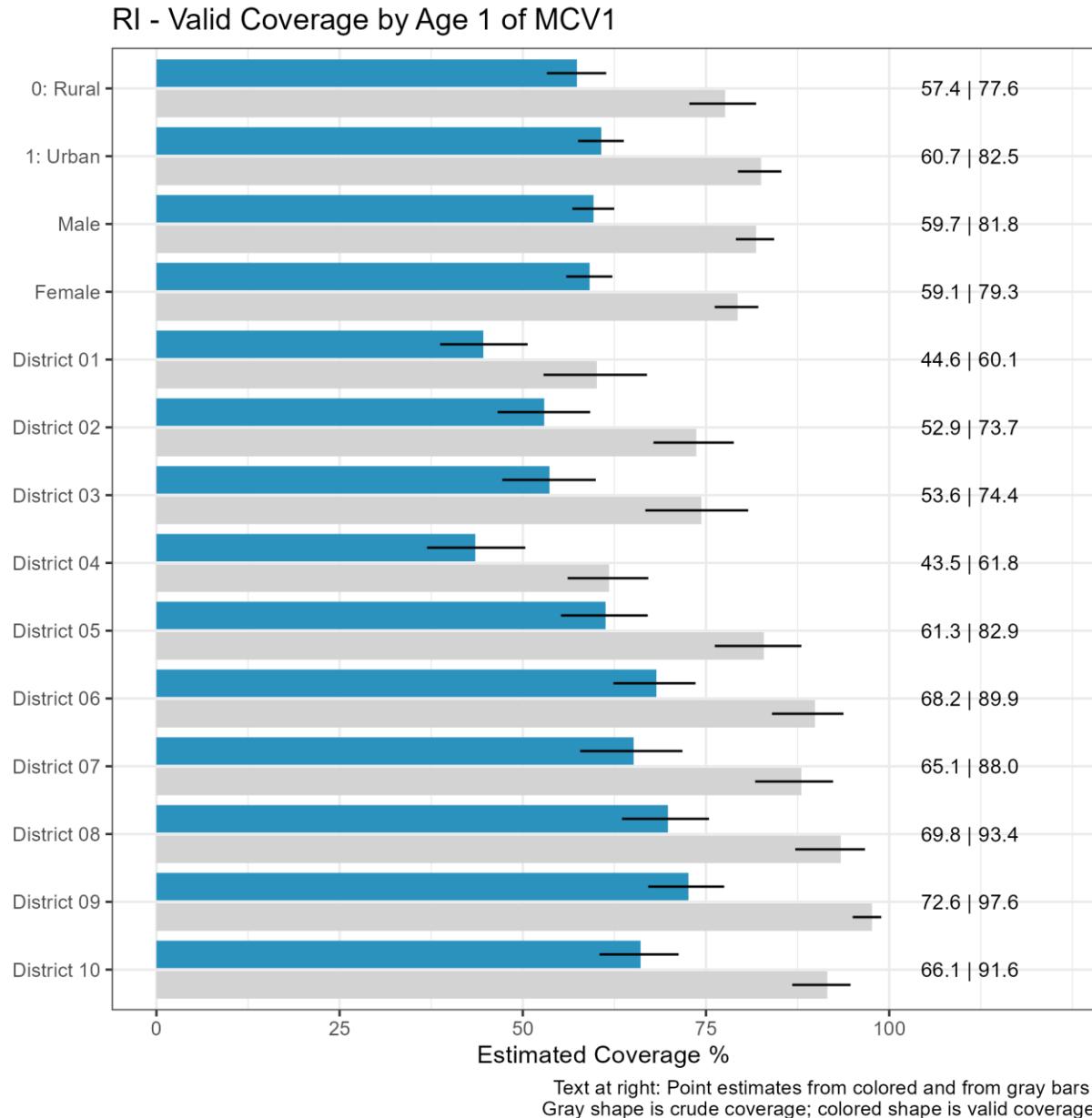
The VCQI_DOUBLE_IWPLLOT_CITEXT global allows users to customize the confidence interval information printed at the right side of double bar plots. This global can take the integer values 1, 2, or 3 which produce the following results:

1. Point Estimate for shaded distribution | Point estimate for unshaded hollow distribution
Example: 75.9% | 77.3%
Option 1 is the VCQI default.
2. Shaded distribution Point Estimate & (2-sided 95% Confidence Interval) and |
Unshaded hollow distribution Point Estimate & (2-sided 95% Confidence Interval)
Example: 75.9% (74.2, 77.6) | 88.2% (86.1, 89.9)
3. No text at the right side of the plot.

Illustrations of each of these options follow.

Figure C-6. Double bar plot demonstrating VCQI_DOUBLE_IW PLOT_CITEXT 1

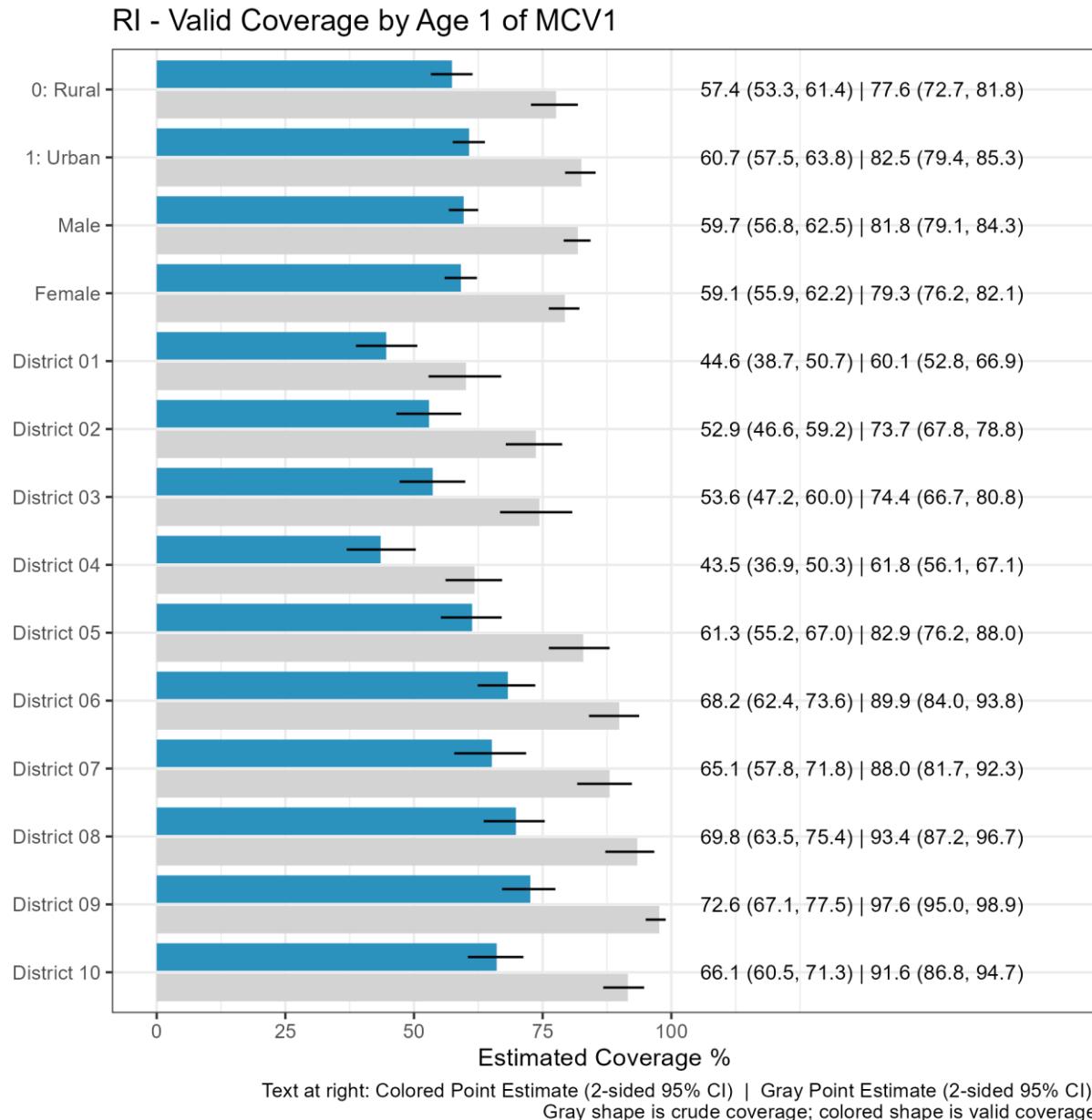
```
vcqi_global(VCQI_DOUBLE_IW PLOT_CITEXT, 1)
```



Annex C. Customizing VCQI Plot Output

Figure C-7. Double bar plot demonstrating VCQI_DOUBLE_IW PLOT_CITEXT 2

```
vcqi_global(VCQI_DOUBLE_IW PLOT_CITEXT, 2)
```



To recover text clipped at the right side of a figure:

If the right-side text is clipped in the .png files, set the global VCQI_SAVE_IW_PLOT_DATA to 1 in the control program and then load the saved .rds dataset file into R and modify the following ggplot code to recreate and adjust the plot:

```

library(ggplot2)
dat <- readRDS(<path to plot data>)

extraspace <- max(nchar(dat$text))
group.colors <- c(dat = "#2b92be", dat2 = "lightgrey")

if (is.na(dat$graphtitle[1])){
  title <- NULL
} else {
  title <- dat$graphtitle[1]
}

if (is.na(dat$graphsubtitle[1])){
  subtitle <- NULL
} else {
  subtitle <- dat$graphsubtitle[1]
}

if (is.na(dat$graphcaption[1])){
  note <- NULL
} else {
  note <- dat$graphcaption[1]
}

ggplot(dat, mapping = aes(x = as.factor(rowid), y = estimate * 100, fill =
source)) +
  scale_fill_manual(name = "", values = group.colors, guide = "none") +
  geom_col(position = position_dodge2(width = 0.5, preserve = "single")) +
  geom_linerange(aes(ymin = cill * 100, ymax = ciul * 100),
                 position = position_dodge(.9)) +
  geom_text(aes(x = as.factor(rowid),
                y = 100 + extraspace,
                label = text),
            size = 3.25) +
  coord_flip() +
  labs(y = "Estimated Coverage %",
       x = "",
       title = title,
       subtitle = subtitle,
       caption = note) +
  scale_x_discrete(labels = dat$name[!is.na(dat$name)]) +
  scale_y_continuous(limits = c(0, 100 + 2*extraspace),
                     breaks = c(0, 25, 50, 75, 100)) +
  theme(plot.caption = element_text(hjust = 0)) +
  theme_bw()

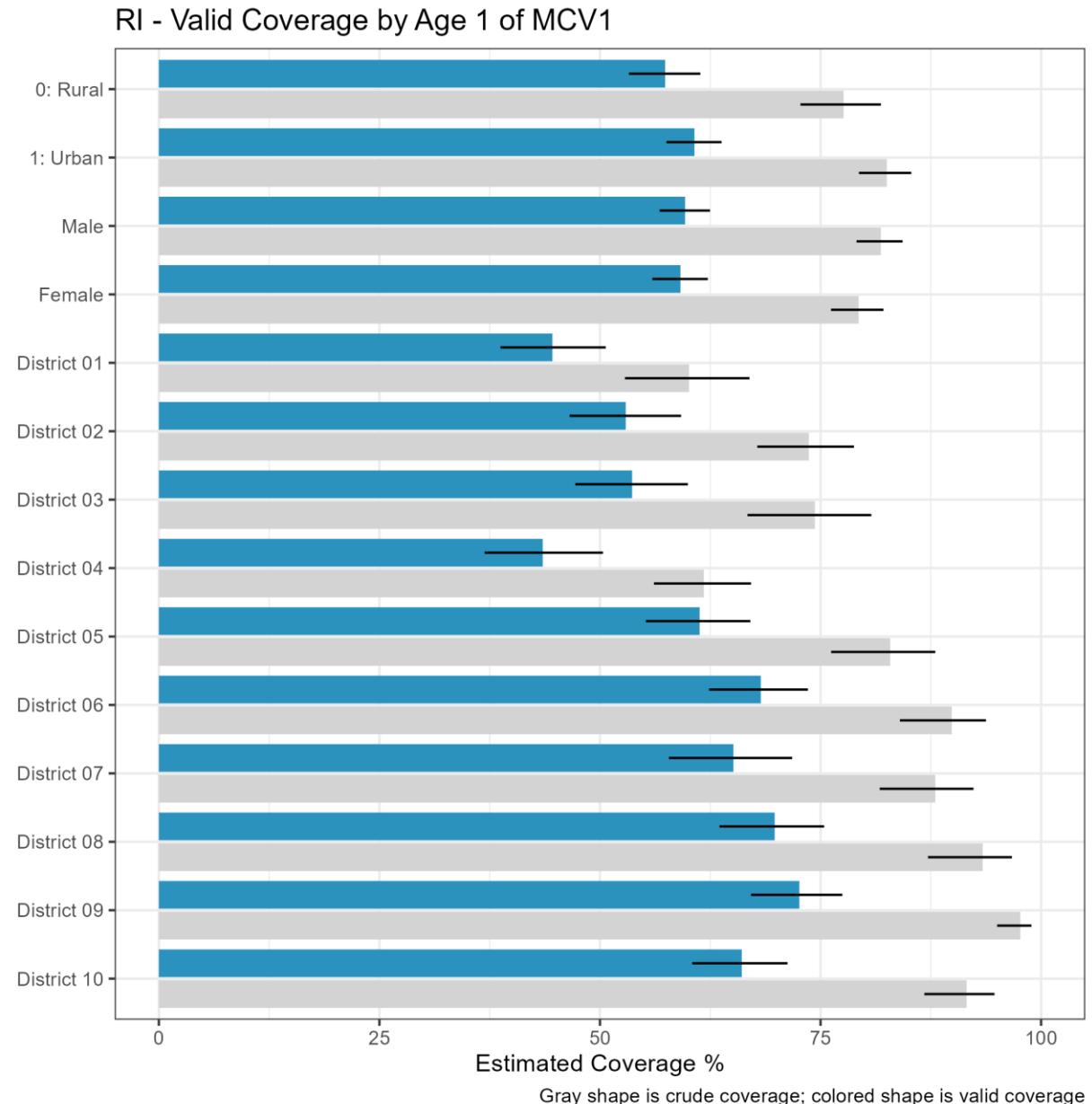
ggsave(<filename.png>, width = <customize width>, height = <customize height>,
units = "in")

```

Annex C. Customizing VCQI Plot Output

Figure C-8. Double bar plot demonstrating VCQI_DOUBLE_IW PLOT_CITEXT 3

```
vcqi_global(VCQI_DOUBLE_IW PLOT_CITEXT, 3)
```



C.4 To Customize VCQI Plots in Other Ways

You may wish to make other changes to VCQI figures, modifying text in titles or labels or visual attributes of the plots. For weighted bar plots and double bar plots, set the global VCQI_SAVE_IW_PLOT_DATA to 1 in the control program and then load the saved .rds dataset file into R; use the ggplot code in section C.3 (**To recover text clipped at the right side of a figure**) or in section C.4 (**To recover text clipped at the right side of a figure**) above with your own settings. To modify unweighted plots, set the global VCQI_SAVE_UW_PLOT_DATA to 1 in the control program and then load the saved .rds dataset file into R; use the following ggplot code with your own settings:

```
library(ggplot2)
dat <- readRDS(<path to plot data>)

extraspace <- max(nchar(dat$text))

if (is.na(dat$graphtitle[1])){
  title <- NULL
} else {
  title <- dat$graphtitle[1]
}

if (is.na(dat$graphcaption[1])){
  note <- NULL
} else {
  note <- dat$graphcaption[1]
}

ggplot(dat, aes(x = as.factor(rowid), y = estimate * 100)) +
  geom_col(fill = "#2b92be") +
  geom_text(aes(x = as.factor(rowid),
                y = 100 + extraspace,
                label = text)) +
  coord_flip() +
  labs(y = "Sample Proportion %",
       x = "",
       title = title,
       caption = note) +
  scale_x_discrete(labels = dat$name) +
  scale_y_continuous(limits = c(0, 100 + 1.75*extraspace),
                     breaks = c(0, 25, 50, 75, 100)) +
  theme(plot.caption = element_text(hjust = 0))+
  theme_bw()

ggsave(<filename.png>, width = <customize width>, height = <customize height>,
units = "in")
```

You may also contact the VCQI developers for assistance.

ANNEX D. VCQI OUTPUT DATABASES

In addition to standardized tables and plots, most VCQI indicators produce one or more datasets to store indicator outcomes for each stratum. VCQI's term for these outcome datasets is *database* but they are simply R datasets with one row per stratum per outcome. Broadly speaking, nearly all of the numbers that appear in VCQI tables and plots are also stored in a database. Users may import these results into whatever software they like to use to make customized tables and figures. Note that each indicator section in chapter 6 of this document lists the names of the database files produced by that indicator.

In a VCQI control program, the `DELETE_VCQI_DATABASES_AT_END` parameter in Block D controls whether these database files are deleted, or whether they are consolidated and saved. If `DELETE_VCQI_DATABASES_AT_END` is set to 1, then the databases are discarded at the end of the VCQI run. If `DELETE_VCQI_DATABASES_AT_END` is set to 0 and `AGGREGATE_VCQI_DATABASES` is set to 1, then VCQI combines and saves most of the database files in a single dataset named `VCQI_aggregated_databases_all.rds`.

The variable named `db_name` in `VCQI_aggregated_databases_all.rds` lists the original database name. Users can use this variable to extract individual databases from the aggregated databases file. For instance, to extract the database that lists crude coverage for dose PCV2, type:

```
dat <-  
  readRDS("C:/VCQI Test Output/VCQI_aggregated_databases_all.rds")  
dat <- subset(dat, db_name == "RI_COVG_01_1_pcv2_a_database")
```

Database contents differ for weighted and unweighted indicators: VCQI currently provides estimates of uncertainty for weighted indicators, but not for unweighted indicators. The table below summarizes which indicators produce weighted, unweighted, or custom databases.

VCQI's Weighted, Unweighted, and Custom Databases

Indicator*	Weighted or Unweighted	Database Type
RI_COVG_01	Weighted	Weighted
RI_COVG_02	Weighted	Weighted
RI_QUAL_07B	Weighted	Weighted
RI_QUAL_08	Unweighted	Unweighted**
RI_QUAL_09	Unweighted	Custom

* Not listed in this table is indicator RI_VCTC_01 which does not produce databases.

** The estimate in the RI_QUAL_08 database is an unweighted *count*, rather than a proportion

The standard weighted and unweighted databases that VCQI produces have the following variables:

Variable Name	Variable Label	In Unweighted Database	In Weighted Database
level	Stratum geographic label	✓	✓
name	Sub-stratum name	✓	✓
level4id	Sub-stratum ID	✓	✓
outcome	Outcome variable name	✓	✓
estimate	Estimated proportion	✓	✓
stderr	Standard error		✓
cilevel	Confidence level		✓
cill	2-sided CI lower bound		✓
ciul	2-sided CI upper bound		✓
lcb	1-sided lower confidence bound		✓
ucb	1-sided upper confidence bound		✓
deff	Design effect		✓
icc	Intracluster correlation coefficient		✓
n	Sample size (unweighted)	✓	✓
nwtd	Sample size (weighted)		✓
nclusters	Number of clusters		✓
nwtd_est	Persons with outcome (weighted)		✓

Most but not all VCQI databases are appended together into the aggregated database file.

At the time of this writing (December 2022) few users outside of Biostat Global Consulting have used VCQI databases to make customized tables or figures. We would be happy to collaborate with you on ideas and tips for how to use them.

ANNEX E. VCQI AUGMENTED DATASET

If desired, VCQI can merge the input survey dataset with derived variables calculated by various VCQI indicators to create an *augmented dataset*. This dataset can be used to make custom figures and tables and for follow-up analyses. When the user calls for an augmented dataset to be made, then after all indicators have been run, VCQI will collect and merge the indicator datasets with the analysis dataset(s). Only indicators that produce output at the respondent level will be included.

To generate the augmented dataset, set the global VCQI_MAKE_AUGMENTED_DATASET to 1 in Block D of the control program.

For most purposes, the names of derived variables are unique for each VCQI indicator, but if a VCQI dataset produced by one indicator does include variables with the same name but *different* values as those produced by another indicator, then both copies of the variable are kept in the augmented dataset. Both variables will be renamed, with a name ending with _ac<ANALYSIS_COUNTER> (e.g., _ac1 and _ac2) to indicate which indicator produced it

With the appropriate R syntax, it should be possible for an analyst to load the augmented dataset and reproduce any of the analysis output numbers that appear in VCQI tables or figures. When a survey steering committee produces table shells for a report that follow a different layout or format than VCQI's default tables, the staff at Biostat Global Consulting often start by producing VCQI tables and a VCQI augmented dataset, and then write a custom program to populate the customized table shells. It is possible to cross-check the customized tables with VCQI's standardized tables to confirm that the outcomes are the same. In many cases it would also be possible to populate customized table shells using a program that manipulates VCQI's aggregated databases, which are described in Annex D.

Finally, augmented datasets can serve as the starting point for logistic regression analyses or time-to-event analyses to explore factors associated with whether a child received a particular dose (or combination of doses) or associated with the timeliness of vaccination.

ANNEX F. ADDITIONAL TOOLS FOR WORKING WITH VCQI OUTPUT

F.1 Missed Opportunities for Simultaneous Vaccination R-Shiny Application

The RI_QUAL_09 indicator summarizes missed opportunities for simultaneous vaccination (MOVs) for every child in the survey dataset whose vaccination evidence includes vaccination dates from a card or register. For each dose in the dose list²³, VCQI calculates whether the child had a missed opportunity to be vaccinated with that dose, and if so, whether the MOV was corrected or uncorrected by the time of the survey. For children who experienced a corrected MOV, VCQI also calculates the time to correction: the number of days between the initial MOV and when the child eventually received the dose.

It can be useful to compare the prevalence of MOVs across different doses and different strata. There is an interactive R Shiny tool for exploring output from RI_QUAL_09 to enable those comparisons; the application is available online at https://biostat-global-consulting.shinyapps.io/MOV_Tool_Public/.

A series of short YouTube videos provides information on this tool, demonstrating the features of the MOV and Time to Correction tabs, and introducing the detailed data requirements for the application:

- Introduction: <https://youtu.be/GQ7Hcmh2czs>
- Exploring MOVs: <https://youtu.be/oyyBJ-NzNug>
- Time to Correction: <https://youtu.be/TS0ePSsbZRk>
- Data Requirements: <https://youtu.be/gppzvsmKoVU>

To produce a dataset compatible with the MOV application, users can do a VCQI run including RI_QUAL_09²⁴ and either:

- a) Set the DELETE_TEMP_VCQI_DATASETS global to 0 so that the RI_QUAL_09 dataset is saved, or
- b) Set the VCQI_MAKE_AUGMENTED_DATASET global to 1 to save a dataset that includes the RI_QUAL_09 variables

Use the interface on the application's *Getting Started* tab to upload either your output dataset named RI_QUAL_09_1.dta²⁵ or the one named RI_augmented_dataset.dta. The *MOV Occurrence* tab summarizes MOVs by stratum and by dose, and the *Time to Correction* tab visualizes the time it takes for MOVs to be corrected.

²³ If the user wishes to see MOV output for a subset of doses, define the MOV_OUTPUT_DOSE_LIST in Block D or F before the control program calls the program named calculate_mov_flags. Otherwise VCQI will produce MOV output for every dose that is included in the global named RI_DOSE_LIST.

²⁴ Recall from Chapter 6 that RI_QUAL_09 can calculate MOVs using either CRUDE or VALID definitions, as specified by the VCQI global named RI_QUAL_09_VALID_OR_CRUDE. CRUDE calculations give credit for doses that are administered too early and VALID calculations do not. The R Shiny application can summarize output from CRUDE and VALID calculations, so the user may wish to run RI_QUAL_09 twice: once for CRUDE calculations and once for VALID calculations. To do so, remember to change the value of ANALYSIS_COUNTER as described in Section 6.2 of this User's Guide.

²⁵ The format of the RI_QUAL_09 filenames is RI_QUAL_09_<ANALYSIS_COUNTER>. In this RI_QUAL_09_1.dta filename, the “_1” component indicates the ANALYSIS_COUNTER value.

Annex F. Additional Tools for Working with VCQI Output

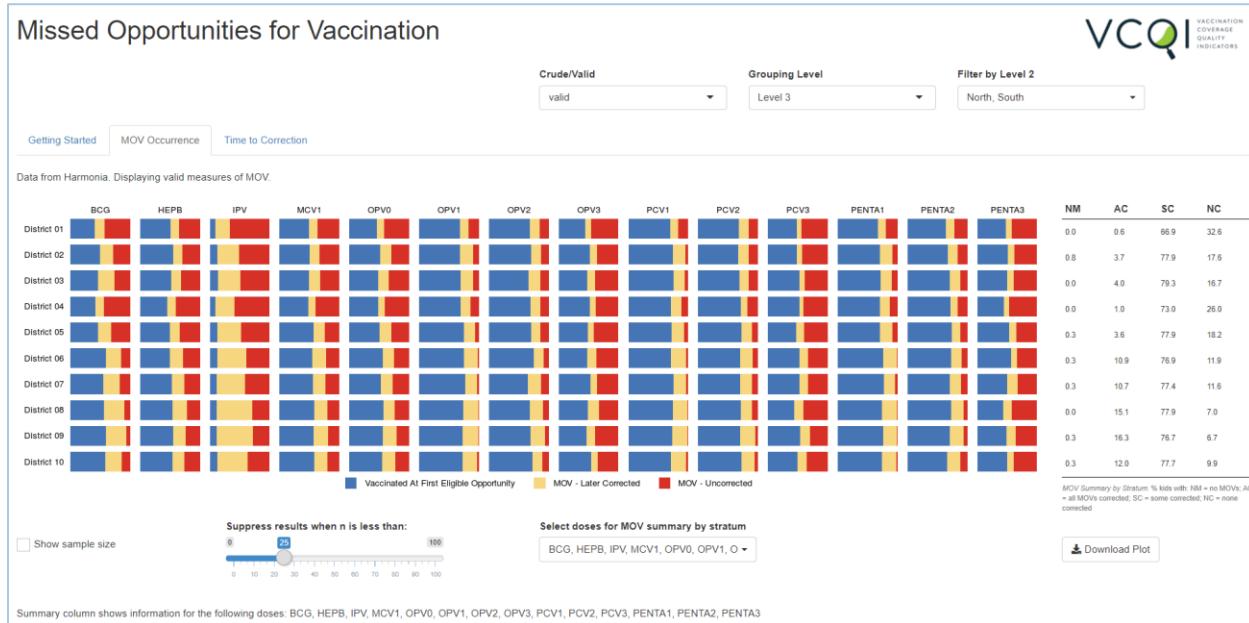
At this time the R Shiny MOV tool only accepts uploads of Stata (.dta) datasets. In the future, this tool will be updated to also accept R (.rds) datasets produced by the R version of VCQI. In the meantime, the RI_QUAL_09_1.rds and RI_augmented_dataset.rds datasets that the R version of VCQI produces can be converted to Stata (.dta) datasets (e.g. by using the write_dta command from the haven package) and used in the MOV application.

MOV Occurrence Tab

The centerpiece of this tab is a figure showing, for each stratum and dose combination, the proportions of children who were vaccinated at the first eligible opportunity (blue), who had a corrected MOV (yellow), and who had an uncorrected MOV (red). The stratum summary table to the right of the figure considers all doses and shows the percentage of respondents who had no MOV for any dose (NM), who had MOVs that were all corrected (AC), who had some MOVs corrected and some uncorrected (SC), and who had none of their MOVs corrected (NC).

By default, level 3 strata are displayed in the figure. The *Grouping Level* input allows users to view results for the level 1 stratum or for level 2 strata instead. Users can also view only the level 3 strata within a particular level 2 stratum by using the *Filter by Level 2* input.

Figure F-1. Missed Opportunities R-Shiny App – Main Bar Chart



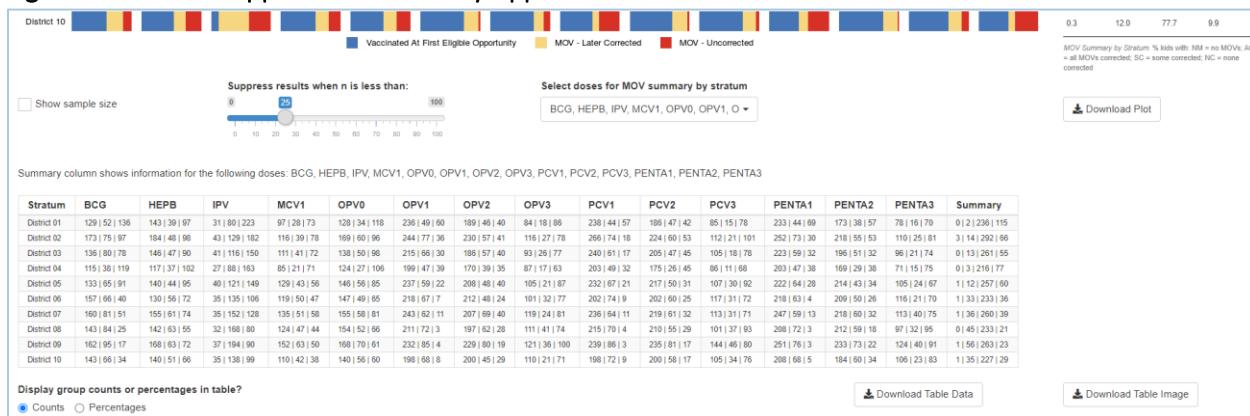
Using the controls underneath the figure, users may:

- display sample sizes for each panel in the figure,
- suppress results in panels with a sample size smaller than a user-specified threshold, and
- select which doses to include in the calculations for the stratum summary table on the right

Tabular counts and percentages appear below the figure. Radio buttons under the table allow users to toggle whether the table shows counts or percentages.

Annex F. Additional Tools for Working with VCQI Output

Figure F-2. Missed Opportunities R-Shiny App – Main Results Table



The user may download the figure and the table.

Annex F. Additional Tools for Working with VCQI Output

Time to Correction Tab

For MOVs which were corrected by the time of the survey, VCQI calculates the time elapsed between the initial MOV and the correction. The figure in the Time to Correction tab shows the cumulative percentage of children whose MOVs were corrected over time.

Users have the option to show the sample size in each panel (the total number of corrected MOVs for that stratum and dose combination), and to suppress outcomes in panels whose sample size falls below a user-specified threshold. Users may also change the limits of the X axis to view different parts of the cumulative coverage curves more clearly. Lines showing the 25th, 50th, 75th, or 90th percentile for days until correction may be overlaid on the data, allowing, for example, a comparison of the median time to correction across strata and doses.

Figure F-3. Missed Opportunities R-Shiny App – Cumulative Time to Correction Graphs

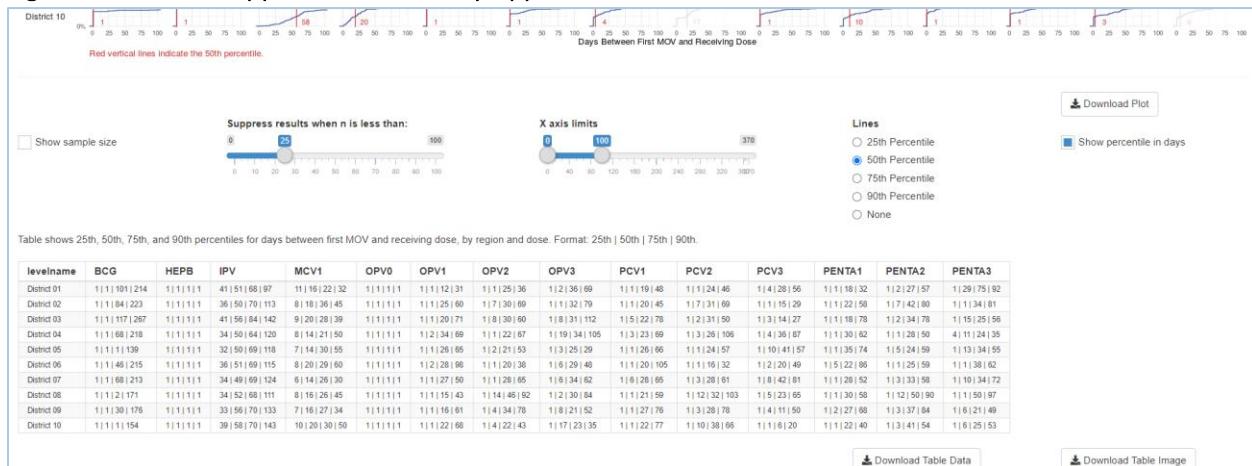


Annex F. Additional Tools for Working with VCQI Output

Underneath the figure, a table shows the 25th, 50th, 75th, and 90th percentiles for days to correction for each stratum and dose.

The user may download the figure and table from this tab, too.

Figure F-4. Missed Opportunities R-Shiny App – Time to Correction Table



See the peer-reviewed manuscript²⁶ for examples of output from this R Shiny MOV application.

²⁶ <https://www.preprints.org/manuscript/202009.0645>

F.2 ICCLOOP Program for Reporting Survey Planning Parameters

When planning a cluster survey, several parameters must be assembled to calculate an appropriate sample size. Annex B1 of the 2018 WHO Vaccination Coverage Cluster Survey Reference Manual²⁷ provides guidance on estimating those parameters, which include:

- The intracluster correlation coefficient (ICC), a measure of spatial heterogeneity of the outcome
- The coefficient of variation of weights (CVw), a measure of the variability of survey weights
- The design effect (DEFF), which quantifies how much larger a complex survey sample must be compared to a simple random sample to achieve a desired level of precision

After conducting the survey, it may be useful to calculate and report the observed values of these planning parameters, both to check the accuracy of the planning process and, perhaps more importantly, to furnish helpful inputs for the team who will plan the next coverage survey. VCQI provides estimates of ICC and DEFF in databases generated by weighted indicators like RI_COVG_01, so it is possible to use VCQI output to construct a table of those values for the survey report. At this time, VCQI does not summarize CVw in its output.

In other work with the World Health Organization, Biostat Global Consulting has written a program named **iccloop** to summarize these helpful parameters from a vaccination coverage survey dataset. The input dataset does not have to come from VCQI, but VCQI output datasets are compatible with iccloop.

As you design your survey report, consider including an annex that lists CVw, and the average number of respondents per cluster, the ICC, and DEFF for the main outcomes.

Datasets that are compatible with iccloop have one row per respondent, and include variables capturing the following concepts:

- Unique ID for each survey sampling stratum (level 3 strata, in VCQI terminology)
- Unique ID for each cluster within each stratum
- Binary outcome variable(s) of interest, coded as
 - 1 if the respondent is in the eligible group and experienced the outcome of interest,
 - 0 if the respondent is in the eligible group and did not experience the outcome, and
 - Missing (NA) if the respondent is not in the eligible group
- Survey weight variable indicating the relative number of respondents in the population represented by each respondent (if the survey was self-weighted or unweighted, this variable does not need to be present)

It may also be useful for the input dataset to have other stratification variables of interest, e.g. a level 2 variable or a level 4 stratifier such as a variable indicating whether a cluster is urban or rural.

The ICCLOOP program loops through each level of a specified stratification variable and each outcome variable that the user specifies and calculates survey parameters. ICCLOOP's output dataset contains one row for each combination of stratum and outcome:

- For a single stratum with one outcome, the program makes one row of output
- For a single stratum survey with five outcomes, the program makes five rows of output
- For a survey with five strata and five outcomes, the program makes 25 rows of output

²⁷ <https://apps.who.int/iris/handle/10665/272820>

In addition to the output dataset, ICCLOOP produces organ pipe plots for each outcome and stratum.

The dataset produced by iccloop contains:

- Identifying information for each row of output (which stratum level, which outcome variable, the file path to the dataset that was analyzed, and so on)
- Stratum-specific information: number of clusters, number of respondents, respondents per cluster (mean, minimum, maximum, standard deviation), how many clusters in the stratum had 0% coverage and how many had 100% coverage
- Coverage calculations for each stratum level: point estimate, standard deviation, and confidence interval bounds (each of the following confidence intervals is calculated: Wilson, Logit, Agresti-Coull, Clopper-Pearson, Fleiss-Cuzick, Jeffreys, and Wald)
- Design effect, effective sample size, coefficient of variation on weights, ICC, ICC confidence interval bounds, and ICC CI method
- ANOVA mean square between and mean square within
- Indicator of homogenous coverage across clusters

You may download the iccloop source code and description with supporting materials in Stata and R from the Biostat Global Consulting GitHub site.²⁸

²⁸²⁸ <https://github.com/BiostatGlobalConsulting/iccloop>

ANNEX G. RI DATE DATA QUALITY – METHOD FOR DETAILED ASSESSMENT

Chapter 7 indicates that the user can request a date data quality report by including the following code in Block D of the control program:

```
vcqi_global(VCQI_REPORT_DATA_QUALITY, 1)
```

The VCQI Results Quick Interpretation Guide has instructions for reviewing that report for evidence of incomplete or faulty or discordant dates of birth or vaccination. The date data quality report programs are not currently available in the R version of VCQI; users should use the Stata version of VCQI to produce this report.

If the date data quality report indicates that there are a notable number of dates with problems, the user may produce yet another report that lists line-by-line details of the questionable data values and provides an Excel-based opportunity to make corrections to the dates. This more detailed report is not routinely run by VCQI, but is possible using a companion program that is distributed with the Stata version of VCQI. The program is named: vcqi_ri_dose_assertlist.ado. Currently, the R version of VCQI has not yet implemented the assertlist program. At this time, you would need to use the Stata version of VCQI to accomplish this task.