

Getting Started with



for R Users

Draft Version 1.0
December 2022



This page is intentionally blank.

Table of Contents

Acknowledgements.....	1
Introduction	3
Goal 1. Download the VCQI Programs	6
Goal 2. Download Demo Datasets.....	7
Goal 3. Download a Demo Control Program	8
Goal 4. Edit the Control Program to use Your File Paths	9
Goal 5. Run VCQI for the first time	10
Goal 6. Look at output	11
Goal 7. Customize the control program & run it again.....	12
Next Steps	14
Annex A. Other VCQI Resources	15

This page is intentionally blank.

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.

The original version of this manual, for Stata users, was written by Dale Rhoda in 2021 with helpful testing and feedback from Shiraz Johnson. The manual was adapted for R users in 2022 by Dale Rhoda, Mia Yu, and Caitlin Clary.

This page is intentionally blank.

INTRODUCTION

The Vaccination Coverage Quality Indicators (VCQI¹) software is a set of programs to make standard output tables and figures using vaccination coverage survey datasets. The VCQI programs were originally developed in Stata; an R implementation of VCQI is in development, with a limited set of indicators currently available.

The VCQI programs are freely available, courtesy of the World Health Organization. Annex A of this manual describes where to find VCQI programs, documentation, demo files, and supporting materials. 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).² We also recommend downloading and using the RStudio Desktop integrated development environment, which is available for free from Posit.³

This document is meant to help you install the VCQI R package on your computer and run it for the first time. The guide assumes that you have basic knowledge of how to use R.

VCQI is organized around three types of surveys:

1. Routine Immunization, abbreviated RI in VCQI documentation
2. Tetanus protection at birth surveys, abbreviated TT, and typically administered to women who had a live birth in the past 12 months
3. Post-campaign coverage surveys – typically administered directly following a vaccination campaign or supplemental immunization activity – abbreviated SIA in VCQI documentation

The R version of VCQI only calculates RI indicators at this time.

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.

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

¹ Pronounced “Vicki”

² <https://cran.r-project.org/>

³ <https://posit.co/downloads/>

Getting Started with VCQI

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, post a message on the VCQI User's Group (See Annex A for details) or contact GetVCQIHelp@biostatglobal.com.

The Indicators

Several types of analyses are included with VCQI; they are described 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.

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*
- ACC: Indicators related to whether respondents have *access* to vaccination services
- CONT: Indicators related to whether respondents experience *continuity* of services
- QUAL: Indicators related to the *quality* of vaccination service

Or by the types of figures that they make:

- CCC: Make *cumulative coverage curves* to summarize vaccination timeliness
- CIC: Make *cumulative interval curves* to summarize vaccination timeliness
- 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.

Additional indicators are added over time.

The indicators currently available in the R implementation of VCQI are: RI_COVG_01, RI_COVG_02, RI_QUAL_07B, RI_QUAL_08, RI_QUAL_09 and RI_VCTC_01. More indicators will be added over time.

Goals of this Manual

VCQI is flexible and all its various options can be overwhelming to a new user, so this document has modest and achievable goals:

1. Download the VCQI R package
2. Download some demo datasets
3. Download a demo control program
4. Edit the control program to point to paths on your computer
5. Run VCQI for the first time
6. Look at output
7. Customize the control program & run it again

For more advanced work, such as learning to make your dataset compatible with VCQI or running an analysis with a more ambitious set of changes to options and indicators, see the section of this manual titled Next Steps and then we refer you to the other more comprehensive documents described in Annex A.

GOAL 1. DOWNLOAD THE VCQI PROGRAMS

The VCQI R package and supporting materials are freely available from the vcqiR GitHub repository.⁴ Supporting materials are also available on the VCQI resources website.⁵

The best way to use the VCQI R package is to install it directly from GitHub. Open RStudio and run the following commands in the console:

```
install.packages("pak")
pak::pkg_install("BiostatGlobalConsulting/vcqiR")
```

Follow the prompts in the console to download and install the vcqiR package. If you are prompted to install and/or update package dependencies, then select the “yes” response in the console – dependencies should be installed.

If the package installation was successful, then lines similar to the following will print to the R console:

```
✓ Built vcqiR 0.1.0
✓ Installed vcqiR 0.1.0
✓ 1 pkg + 71 deps
```

⁴ <https://github.com/BiostatGlobalConsulting/vcqiR>

⁵ http://www.biostatglobal.com/VCQI_RESOURCES.html

GOAL 2. DOWNLOAD DEMO DATASETS

The VCQI resources website⁶ has several sets of files for VCQI demonstration purposes. The datasets for fictional District 20 are simple surveys with a single stratum each. The datasets for the fictional country of Harmonia are larger, with ten strata nested within two provinces, nested within the country. Demo RI survey datasets for both District 20 and Harmonia can be found on the VCQI Resources website.⁷ Start by downloading and unzipping the file with District 20 RI files. Again, make a note to remember where you saved the files.

The current R package version of VCQI does not have functions to analyze SIA or TT surveys.

Path to folder where I saved the District 20 demo data files:

e.g., C:/Users/Dale/VCQI Work/VCQI Demo Data/District 20 RI Data Files

⁶ http://www.biostatglobal.com/VCQI_RESOURCES.html

⁷ The RI demo datasets are also available on the vcqiR GitHub repository, in the Demo_Datasets folder.

GOAL 3. DOWNLOAD A DEMO CONTROL PROGRAM

Each set of demo datasets is accompanied by one or more demo control programs. Download the zip file holding “District 20 RI Demo Control Program - vcqiR.R”. Make a new empty folder on your computer and unzip the files into that folder.⁸

Path to folder where I saved the control program:

e.g., C:/Users/Dale/VCQI Work/VCQI Demo Output/District 20 RI Run 01

Note: On the pages that follow, some lines in this document represent lines of R code in a VCQI control program. Those are formatted using the Courier New font and use a light gray background, like this:

```
# This is a line of R code in the VCQI control program
```

⁸ There will be two zipped files: a control program and a program named “globals_for_timeliness_plots.R”.

GOAL 4. EDIT THE CONTROL PROGRAM TO USE YOUR FILE PATHS

Navigate to the folder whose path you wrote down in Goal 3. Open the “District 20 RI Demo Control Program - vcqiR.R” file in R.

Take a moment to scroll up and down. The control program is several hundred lines long. Don’t worry: you will only need to modify two lines of code. Notice that the program is divided into sections named Blocks. Each alternating block begins with some comments that tells you whether the user is allowed to change the code in that section or not. The user may change code in Blocks B, D, and F. You should not change code in Blocks A, C, E, or G.

Scroll to Block B and make these changes:

1. Find the lines of code (around line 39) that say:

```
# Where should the programs look for datasets?
VCQI_DATA_FOLDER <- "your path"
```

Replace the string "your path" with the full path you wrote under Goal 2: the path to the VCQI datasets. On my computer this would say:

```
VCQI_DATA_FOLDER <- "C:/Users/Dale/VCQI Work/VCQI Demo Data/District 20 RI
Data Files"
```

2. Scroll down farther (around line 42); find the lines of code that say:

```
# Where should the programs put output?
VCQI_OUTPUT_FOLDER <- "your path"
```

Replace the string "your path" with the full path you wrote under Goal 3: the path to the folder where you saved the control program. Note that the output folder cannot be the same as the data folder. On my computer, this would say:

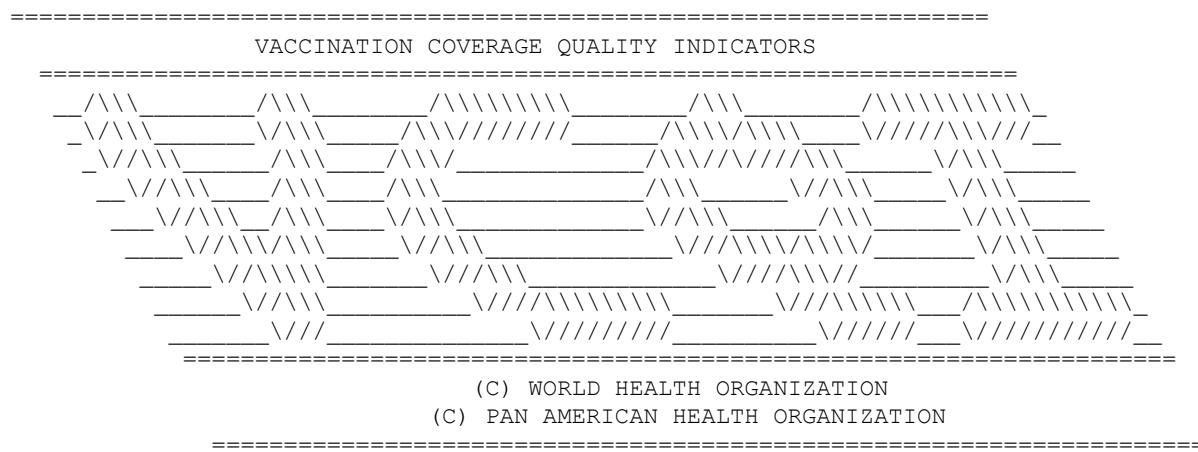
```
VCQI_OUTPUT_FOLDER <- "C:/Users/Dale/VCQI Work/VCQI Demo Output/District 20
RI Run 01"
```

3. Save the modified program.

GOAL 5. RUN VCQI FOR THE FIRST TIME

Run the control program that you just edited. If you see a red error message in the R console, you probably made a typographical error. Check your two file paths and be sure you typed them correctly and put the paths in quotes. If you still receive an error message, e-mail GetVCQIHelp@biostatglobal.com for some assistance.

If you have followed the directions above, VCQI should run successfully. The entire run will take several minutes, and the R console window will update with messages to announce VCQI progress. When the program finishes, it will show a VCQI billboard like this:



Congratulations! You have successfully run VCQI!

Take a moment to join the VCQI User's Group on the TechNet-21 website. (See Annex A to learn how to join.) Post a message there introducing yourself and letting others know that you have conducted your first successful run.

GOAL 6. LOOK AT OUTPUT

Look in the folder that you wrote down in Goal 3: the VCQI output folder. You will see not only the control program .R file that you edited and ran, but also a set of output files and folders.

To learn how to interpret the VCQI tabular and graphical output, open the document named the *VCQI Results Interpretation Quick-Reference Guide*. You can download this document from the vcqiR GitHub repository⁹ or from the VCQI Resources website.¹⁰

To explore the graphic output, navigate into one of the folders with the word *Plots_* in its name and double-click on the files with .png extension.

Spend a few minutes looking at output in:

1. The folder that holds the vaccination coverage and timeliness chart: Plots_VCTC. Double-click on the file(s) with extension .png.
2. The folder that holds organ pipe plots: Plots_OP.
3. The folder that holds unweighted proportion plots: Plots_IW_UW. Double click on the files with .png extensions.

To explore VCQI's tabular output, double-click on:

4. The spreadsheet that holds tabular output from VCQI indicators: District_20_RI_TO.xlsx

Make a list of questions that the output raises in your mind. Try to answer them using the Results Interpretation Guide. If some questions persist, post them on the VCQI User's Group on TechNet-21.

⁹ <https://github.com/BiostatGlobalConsulting/vcqiR>

¹⁰ http://www.biostatglobal.com/VCQI_RESOURCES.html

GOAL 7. CUSTOMIZE THE CONTROL PROGRAM & RUN IT AGAIN.

Now that you have successfully run VCQI, spend some time reading through the VCQI Users' Guide and select one or more user inputs to modify in Block D or Block F. Do not edit the control program that you already ran. Instead generate a new empty output folder and copy the control program from the output folder you wrote in Goal 3 to the new folder. Write the name of that folder here:

Path to new empty output folder where I copied the control program.

e.g., C:/Users/Dale/VCQI Work/VCQI Demo Output/District 20 RI Run 02

(Note: Text shaded yellow in this description of Goal 7 represents characters that differ from those used in Run 01.)

Open the file you just copied into the new folder and make some changes there. As an illustrative example, let's say that we wish to view outcomes stratified by the sex of the child.

1. First, change the path to the output folder in Block B. Change the name of the folder:

```
VCQI_OUTPUT_FOLDER <- "C:/Users/Dale/VCQI Work/VCQI Demo Output/District  
20 RI Run 02"
```

2. Scroll down to around line #243 and find the line that says:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, "level3name")
```

Edit this line so it says:

```
vcqi_global(VCQI_LEVEL4_SET_VARLIST, c("level3name", "RI20"))
```

By adding "RI20" to SET_VARLIST, rather than simply showing output from the survey stratum District 20, which is defined through variable "level3name", we want to show output stratified by child sex too. Read Annexes A and B of the VCQI User's Guide for more information on levels of strata and how to ask for output for the levels that interest you.

3. Save the updated control program. Re-run it and look at the output again. All the tables will list outcomes for all strata and list additional rows to show outcomes for boys and for girls. Many (but not all) of the figures will also show the additional stratified output.

Congratulations! You have just modified a VCQI control program and re-run.

You may repeat Goal 7 many many times, experimenting with different sets of inputs in Blocks D and F.

Remember that for each new run, it is a best practice to make a new empty output folder, copy the control program there, edit it to point the VCQI_OUTPUT_FOLDER in Block B to the new folder, and then make other changes in Blocks B or D or F.

A list of possible changes you might like to explore in addition to adding the demographic stratifier, RI20:

- a) Ask VCQI to annotate text in the unweighted plot for small sample sizes. Change the Block D line that reads:

```
vcqi_global(UWPLOT_ANNOTATE_LOW_MED, 0)
```

to read:

```
vcqi_global(UWPLOT_ANNOTATE_LOW_MED, 1)
```

- b) Turn off organ pipe plots. Change the Block D line that reads:

```
vcqi_global(VCQI_MAKE_OP_PLOTS, 1)
```

to read:

```
vcqi_global(VCQI_MAKE_OP_PLOTS, 0)
```

- c) Change the number of digits after the decimal place from 1 to 0. Change the Block D line that reads:

```
vcqi_global(VCQI_NUM_DECIMAL_DIGITS, 1)
```

to read:

```
vcqi_global(VCQI_NUM_DECIMAL_DIGITS, 0)
```

NEXT STEPS

1. To build more familiarity and skills with the VCQI control program, download a dataset from Harmonia, where there are ten survey strata. This will yield longer tables and figures. Download the demo control program for Harmonia and work through Annex B of the User's Guide to master producing output for the strata that interest you, in the order that you wish to see them. Alternate between showing strata in the figures sorted by outcomes:

```
vcqi_global(PLOT_OUTCOMES_IN_TABLE_ORDER, 0)
```

to listing the strata in the same order as they appear in the tables:

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

2. Next, learn to make datasets compatible with VCQI by studying the *Forms and Variable List* document and by using Stata to explore the datasets from District 20 and Harmonia.

Write your notes & questions on the final page of this document and then post them (along with announcements of your successes) in the *Discussion* portion of the VCQI User's Group.

ANNEX A. OTHER VCQI RESOURCES

Five documents describe VCQI:

1. This *Getting Started with VCQI* guide (there are separate Getting Started guides for R users and for Stata users).
2. The *VCQI User's Guide* is a 200+ page reference manual that helps you understand how to modify a VCQI control program template to achieve your analysis goals (there are separate User's Guides for R users and for Stata users).
3. The *VCQI Results Interpretation Quick-Reference Guide* tells the reader how to interpret output in VCQI's tables and figures. It also describes the RI date data quality report.
4. The document named *Vaccination Coverage Surveys – Forms & Variable Lists (FVL) Structured for Compatibility with VCQI* specifies variable names and attributes of coverage survey datasets that are compatible with VCQI.
5. And finally, the *VCQI Working List of Vaccination Survey Analyses and Software Specifications* lists coverage indicators and describes how software can calculate them.

You may download the latest versions of VCQI software, documentation, and demo datasets at the VCQI resources website.¹¹

You may also join a VCQI User's Group hosted on the TechNet-21 website.¹² The User's Group is useful for its discussion and announcement features.

You are welcome to contact the VCQI developers with any questions. E-mail the VCQI development team at Biostat Global Consulting (Dale Rhoda, Mary Kay Trimner, Caitlin Clary, and Mia Yu) (GetVCQIHelp@biostatglobal.com) or Carolina Danovaro at the World Health Organization (danovaroc@who.int).

¹¹ http://www.biostatglobal.com/VCQI_RESOURCES.html

¹² <https://www.technet-21.org/en/network/groups/293-vcqi>

Notes & Questions

