

Nov 20, 2020

## Quality of Care Index (QCI)

Esmaeil Mohammadi<sup>1</sup>, Erfan Ghasemi<sup>1</sup>, Sahar Saeedi Moghaddam<sup>1</sup>, Moein Yoosefi<sup>1</sup>, Shahin Roshani<sup>1</sup>, Naser Ahmadi<sup>1</sup>, Masoud Masinaei<sup>1</sup>, Maryam Nasserinejad<sup>1</sup>, Narges Ebrahimi<sup>1</sup>, Sina Azadnajafabad<sup>1</sup>, Mahtab Rouhifard Khalilabad<sup>1</sup>, Negar Rezaei<sup>1</sup>, Ali H Mokdad<sup>2</sup>, Bagher Larijani<sup>3</sup>, Farshad Farzadfar<sup>1</sup>

<sup>1</sup>Non-Communicable Diseases Research Center (NCDRC), Endocrinology and Metabolism Research Institute (EMRI), Tehran University of Medical Sciences (TUMS), Tehran, Iran;

<sup>2</sup>Institute for Health Metrics and Evaluation (IHME), University of Washington (UW), Seattle, Washington, USA;

<sup>3</sup>Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute (EMRI), Tehran University of Medical Sciences (TUMS), Tehran, Iran

Farshad Farzadfar: \* Corresponding author;

*In Development* dx.doi.org/10.17504/protocols.io.bprjmm4n

Esmaeil



Esmaeil Mohammadi

Non-communicable Diseases Research Center (NCDRC), Endocrino...

### ABSTRACT

#### Abstract

This protocol and set of codes bring forward a newly introduced index and system that can assess the quality of care given to health-seekers on a large-scale, named the quality of care index (QCI). Foremost, QCI is designated for data miners using a large amount of data entry. Dimension reduction approaches are utilized to reduce and ease the complexity of such environments. QCI codes have been trained and built based on the Global Burden of Disease (GBD) database structure [<https://vizhub.healthdata.org/gbd-compare/>]. Other data sources can be fed to the code upon re-structuring to the same format as GBD's. Only aggregates are compatible for analysis and individual data sets are not appropriate commodities. Codes are considerably embedded in R language, easily transformable to Python. Classically, we use STATA file extensions (.dta) and efforts should be carried out to reformat files into other extensions if desired. Stages of QCI computation are 1. data acquisition 2. data curation 3. data analysis and/or 4. visualization.

### DOI

[dx.doi.org/10.17504/protocols.io.bprjmm4n](https://dx.doi.org/10.17504/protocols.io.bprjmm4n)

### PROTOCOL CITATION

Esmaeil Mohammadi, Erfan Ghasemi, Sahar Saeedi Moghaddam, Moein Yoosefi, Shahin Roshani, Naser Ahmadi, Masoud Masinaei, Maryam Nasserinejad, Narges Ebrahimi, Sina Azadnajafabad, Mahtab Rouhifard Khalilabad, Negar Rezaei, Ali H Mokdad, Bagher Larijani, Farshad Farzadfar 2020. Quality of Care Index (QCI). **protocols.io**  
<https://dx.doi.org/10.17504/protocols.io.bprjmm4n>

### KEYWORDS

Quality of care index (QCI), Epidemiology, cancer, quality of care, health care

### LICENSE

— This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

### CREATED

Nov 16, 2020

LAST MODIFIED

Nov 20, 2020

PROTOCOL INTEGER ID

44555

#### MATERIALS TEXT

Data available from the [Global Burden of Disease \(GBD\) data repository](#) [linked to the latest estimates of GBD; for older 2017 estimates please visit [here](#)]

[Sample Brain and other CNS Cancers \(CNS Cancers\) data in 2017](#)

[Sample compressed CNS Cancers data in 2017](#)

ISO codes of nations for [195 countries](#) or [full list](#)

R language available from [CRAN](#) based on your machine interface

R studio [optional] is available from [here](#)

Packages: code has been prepared to be a pipeline and requires to install packages. It needs an active internet connection.

[Sample PCA output](#)

Code files: Are attached to the protocol and also available from [here](#)

#### SAFETY WARNINGS

- A large amount of hardware is required for the best yield. Do not try on old systems.
- This approach is exploratory and should be worked with caution
- Public data repositories like GBD are readily and abundantly prepared for well-organized data structures that come from models. The accuracy of their estimates are upon themselves and are prone to errors.
- This has been validated on GBD causes: CNS, thyroid, pancreas, gastric, and other cancers of the human body. For 'less burdensome' causes with lower mortality rates it might give not sensible outputs.
- Several packages have been used in the body of this code that are open-access yet prone to copyright.
- First-time experience needs to install packages, so an active internet connection is required.

#### ABSTRACT

##### Abstract

This protocol and set of codes bring forward a newly introduced index and system that can assess the quality of care given to health-seekers on a large-scale, named the quality of care index (QCI). Foremost, QCI is designated for data miners using a large amount of data entry. Dimension reduction approaches are utilized to reduce and ease the complexity of such environments. QCI codes have been trained and built based on the Global Burden of Disease (GBD) database structure [<https://vizhub.healthdata.org/gbd-compare/>]. Other data sources can be fed to the code upon re-structuring to the same format as GBD's. Only aggregates are compatible for analysis and individual data sets are not appropriate commodities. Codes are considerably embedded in R language, easily transformable to Python. Classically, we use STATA file extensions (.dta) and efforts should be carried out to reformat files into other extensions if desired. Stages of QCI computation are 1. data acquisition 2. data curation 3. data analysis and/or 4. visualization.

#### BEFORE STARTING

The sample data provided in this protocol comes from Brain and other CNS Cancers in 2017.

The sample ISO-3 file provided is also for 2017. In that year, GBD corresponded to 195 nations while in their latest update [2019] it has been extended to 204 countries. To access the latest ISO-3 attributes please visit [this](#) or [this](#).

#### Data acquisition

1. A. Consider retrieving data from the GBD data repository [<https://vizhub.healthdata.org/gbd-compare/>].  
B. Use the left sidebars and menus to navigate for causes.  
C. Primary parameters are required 1. age-standardized mortality rate 2. age-standardized prevalence rate 3. age-standardized incidence rate 4. age-standardized Disability-adjusted life years (DALYs) rate 5. age-standardized years of life lost (YLLs) rate and 6. age-standardized years lived with disability (YLDs) rate.  
D. The use of age-standardized measures is not mandatory and confounding effects of age and un-even distribution of population pyramid are weighted out.  
E. Manually inspect each parameter for zero estimates. Zeros interfere with the code and yield odd outputs.  
F. If nulls were encountered you might proceed with the elimination of the definite stratum with zero or very low estimates. For example, you need to limit your analysis on adults when you are working on gastric cancers as children

are much less likely to have it and estimates of GBD (or any other source) are redundant  
G. Download from the right upper corner (in CSV format)  
H. You can download a clean template file from this [link](#) for 'CNS cancers'.

#### First time run

- 2 A. First time running the code requires installing some packages. You can do it either manually or let the code take care of it.  
B. Run the attached code to let R get ready [ [@ Installing Packages.R](#) or [this link](#) ]

#### Data curation

- 3 A. GBD data export is large and hard to digest as it encompasses many regional aggregates and sub-national reports [\[intact CNS cancers sample in 2017\]](#). We recommend this step to lighten the size of exported data and prune unnecessary details but it is not mandatory.  
B. We used the iso-3 naming of nations for merging and sub-setting. Any other way is just fine unless changing the cell names or column orders.  
C. The iso-3 names are attached and can be easily integrated into the main file with *merge* command [ *by location\_name\*or location\_id* ] [\[CNS cancers sample for reduced version in 2017\]](#)  
D. Attached file is in STATA extension and can be imported to R with *read\_dta13* command [ [@ ISO3.dta](#) or [this link](#) ]  
E. Use of iso-3 names is recommended especially for merging two data or visualization on maps as the original full names may differ.

\*Script 'Georgia' corresponds to both a nation and a state (in the United States). Use of *location\_name* merges these two and should be recoded before

#### Data analysis


- 4 - Data analysis is made of two steps that are simultaneously performed 1. secondary parameters calculation 2. principal component analysis (PCA) on the secondary parameters. [ [@ QCI.R](#) or [this link](#) ]  
- You only need to define age-groups related to your desired cause on line 44 of the code. By default, it is >1-year-old\*. You can choose the options from code line 51. Choose them wisely and based on the GBD Compare tool for non-null estimates. [mind the quotation marks!]  
- Secondary parameters include four 1. mortality-to-incidence ratio (MIR)\*\* 2. prevalence-to-incidence ratio 3. YLLs-to-YLDs ratio and 4. DALYs-to-prevalence ratio. These four indirectly show the quality of care. We aim to composite these four into one QCI.  
- The first component of PCA on these four measures are then saved and re-scaled into 0-100 and named QCI  
- This is a loop, takes some time to calculate based on the dimension of arrays you've chosen [age groups × sex groups].

\* Age-standardized values are used for the calculation of QCI in a nation as a whole, aimed for comparison purposes. For age-group strata, data are not age-standardized and comparisons should be made with caution.



\*\* For some portion of GBD's cause list, estimates of Death are close to zero, or the condition is not lethal. In such a situation, the MIR calculation yields lots of zeros and interfere with PCA. In this condition, code must be modified and MIR needs to be removed.

§§§



- A. Recruit the libraries [Line 1-27]  
B. Define age groups on line 44  
C. Line 64, if you plan to use age-standardized estimates, by default it uses '*Rate*'; otherwise, for all-age estimates, use '*Number*'  
D. Define your function. The *create\_data* function on line 28th  
E. Run the *create\_data* function  
F. It would ask you for your main data file [in .dta format] that you have downloaded from GBD, paste the location of file on your machine

- G. Then it would ask for iso-3, paste the address [  **ISO3.dta** or [this link](#)]
- H. Then it would ask for location of folder you plan to save your outputs to.
- I. If it loops smoothly, it is working fine. Otherwise, a null has encountered and check your data
- J. Wait until it exports the outputs into your working directory. The output file is named 'main\_data.dta' [[Sample output of CNS cancers](#)].
- K. The other exported file is 'variance\_first\_component.dta' that resmbles the proportion of information [variance] of all data points of your original data enclosed in QCI.

#### Visualization (QCI map)

- 5
  - A. Visualization and maps creation is carried out with the 'ggplot2' and 'sf' packages
  - B. As this step is obligatory, commands for installing codes have not been implemented in Step 1
  - C. Use the attached file for the set of codes [  **Map.R** or [this link](#)]
  - D. First time usage requires running the first line for installation of packages, otherwise start from line 4
  - E. Recruit libraries from lines 4-13
  - F. Run line 14 and choose the output of Step 4 (file named as 'main\_data.dta' [For a sample output of CNS cancers use [this link](#)])
  - G. Run line 15 and choose iso-3 file [  **ISO3.dta** or [this link](#)]
  - H. Choose your desired year and sex (either 'Male', 'Female', or 'Both') from line 20
  - I. Run the code line by line to the end
  - J. A map is created and you can save it [[Sample output](#)]
  - K. You can save the output either manually or using the *ggsave* command

#### Gender Disparity map

- 6
  - A. Gender disparity ratio (GDR) is calculated from the QCI of males and females. [  **GDR.R** or [this link](#)]
  - B. First time usage requires running the first line for installation of packages, otherwise start from line 4
  - C. Recruit libraries from lines 4-13
  - D. Run line 14 and choose the output of Step 4 (file named as 'main\_data.dta' [For a sample output of CNS cancers use [this link](#)])
  - E. Run line 15 and choose iso-3 file [  **ISO3.dta** or [this link](#)]
  - F. Choose your desired year from line 20
  - G. Run the code line by line to the end
  - H. A map is created and you can save it [[Sample output](#)]
  - I. You can save the output either manually or using the *ggsave* command