**Purpose:** Check the master doc about why these are in two separate folders

# Key Scripts:

* [data\_prep.R](#_data_prep.R)
* [gbd\_data\_sub.R](#_gbd_data_sub.R)
* gbd\_outputs.R
* plot\_output.R
* read\_spec\_object.R

## data\_prep.R

**Created by:** Maggie Walters ([mwalte10@uw.edu](mailto:mwalte10@uw.edu))

**Documentation updated: 5/5/2020**

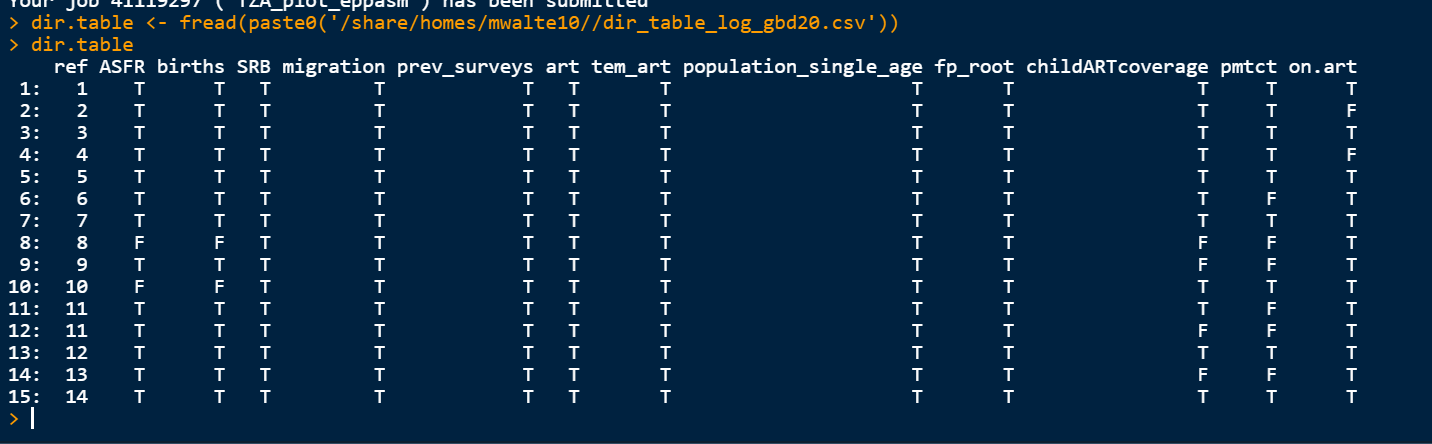
**Script updated: 5/5/2020**

**Purpose:** This script uses toggles to indicate what file paths should be used by read\_spec\_object(). Currently, T indicates the most up-to-date filepath while F pulls those used in gbd19. This file reads in the highest “ref” number to determine which toggles are true or false.

**Called by:** main.R

**Inputs:**

* /share/hiv/epp\_input/gbd20/eppasm\_run\_table.csv
* /share/homes/mwalte10/dir\_table\_log\_gbd20.csv
  + This is currently in my home directory to allow for different people to run different tests.



**Outputs:**

Outputs are the filepaths for ASFR, births, SRB, migration, prev\_surveys, art.dt (adult art, corresponds to column “art”), tem\_art (Tembisa ART estimates for ZAF), population\_single\_age, art (child ART coverage), and pmtct.

The following outputs are filepaths that are based on the fp\_root argument, which indicates where parameters for children are stored. These include: artdist (child ART distribution), artelig (child ART eligibility), percbf (percent breastfed), mort.art (child mortality on ART), prog (child progress parameters through CD4 stages), mort.offart (mortality not on ART), and dropout (dropout rates for PMTCT).

The following parameters change each year, and so will need their filepath updated:

**Future innovation:**

The structure of dir\_table\_log\_gbd20.csv could be improved so that it can be stored centrally and used by multiple people at the same time to conduct simultaneous tests. To do this I think the following columns should be added: user, run\_name, location.

## gbd\_data\_sub.R

**Created by:**

**Documentation updated: 5/14/2020**

**Script updated: 5/5/2020**

**Purpose:** This script contains the functions that are called in read\_spec\_object() to pull in the correct parameters for a given location.

**Called by:** read\_spec\_object() in main.R

**Key functions:**

* extend\_years(): copies over the most recent value to proj.end
  + Arguments: dt, years
* extrapolate\_years(): extrapolates forward using a moving average to proj.end.
  + Arguments: dt, end\_year, years\_to\_average, trans\_vars, id\_vars
    - Trans\_vars: variable you are extrapolating
    - Id\_vars: variables that the data is stratified by (commonly: age, sex, location. Cannot be year)
* append.ciba.incrr(): used for group 2 countries
  + Arguments: dt, loc, run.name
* append.diag(): includes case notification, used for group 2 countries
  + Arguments: dt, loc, run.name
* append.vr(): includes VR data, used for group 2 countries
  + Arguments: dt, loc, run.name
* append.deaths()
  + Arguments: dt, loc, run.name, start,year, stop.year
* convert\_paed\_cd4(): converts CD4 counts to categories
  + Arguments: dt, agegr
* sub.paeds(): substitutes in pediatric specific arguments. Calls a lot of the other functions in gbd\_data\_sub.R
  + Arguments: dt, loc, k, start.year, stop.year
* sub.pop.params.specfp(): substitutes in demographic parameters including: population\_single\_age, surv, ASFR, births, SRB, migration
  + Arguments: fp, loc, k
* sub.prev(): old version of sub.prev.granular, no longer used
  + Arguments: loc, dt
* add\_frr\_noage\_fp(): appends region-specific fertility rate ratio parameters which were provided by Jeff Eaton in June of 2019. These replace those that are provided in the UNAIDS files which are generally more of an estimated value. More information about this process can be found at: I:\RTs\_and\_Projects\GBD\Teams\HIV-TB-SelectID\HIV\Presentations and RW\spectrum-frr-estimates\_Eaton\_2018-05-30.pptx
  + Arguments: obj
* sub.prev.granular(): substitutes in the prevalence data using the prev\_surveys fp. This keeps age, sex-specific surveys when available while also retaining aggregate surveys if they are the only ones there.
  + Arguments: dt, loc
* sub.off.art(): mortality rates for adults that aren’t on ART
  + Arguments: dt, loc, k
* sub.on.art(): pulls in MRBRT/BRADMOD mortality rates for those that are on ART. This fp must be updated whenever this mortality model is updated, as in new data becomes available for mortality when on ART. Currently we use a MRBRT model that consists of 90 models that are all different across regions (high-income, SSA, and other), sex, 3 categories of time since ART initiation, and 5 age categories. Joy handled this model in GBD20.
  + Arguments: dt, loc, k
* sub.cd4.prog()
  + Arguments: dt, loc, k
* geo\_adj()
  + Arguments: loc, dt, i, uncertainty
* geo\_adj\_old()
  + Arguments: loc, dt, I, uncertainty
* sub.art()
  + Arguments: dt, loc, use.recent.unaids
* sub.sexincrr()
  + Arguments: dt, loc, i
* sub.anc.prior()
  + Arguments: dt, loc

**Inputs:**

**Outputs:**

**Future innovation:**