Introduction to epireview

Sangeeta Bhatia (on behalf of Pathogen Epidemiology Review Group)

20 Feb 2024

Introduction to epireview

- epireview - R package under active development
- one of the key outputs from the Pathogen Epidemiology Review Group
- intended as
    - a one-stop shop for
        curated
        epidemiological parameters for all priority pathogens
    - a 'living repository' that can be easily be updated over time
- ships with a flexible database schema that can be used to `roll your own' database

Features: load curated data

```
r$> marburg <- load_epidata('marburg')
Data loaded for marburg

r$> knitr::kable(head(marburg$articles))


|id                              |first_author_surname | year_publication|article_label |
|:-------------------------------|:--------------------|----------------:|:-------------|
|1aa88a18901c85ccfa6e26f1fc4cc3cb |Martini             |             1968|Martini 1968  |
|bbaf446d6451f09a62498dd2eaeba694 |Mathiot             |             1989|Mathiot 1989  |
|2ad9000587d3800fbdab2d63d98cee72 |Suzuki              |             1997|Suzuki 1997   |
|16ee67946aa737c422fa0cd69d000ad9 |Smith               |             1982|Smith 1982    |
|a1134d5507213435867b532814c839c4 |Johnson             |             1993|Johnson 1993  |
|c2fda585d54e519964f100d5fb86a11a |Gear                |             1975|Gear 1975     |

r$> knitr::kable(head(marburg$params[, 3:7]))


|parameter_type                       | parameter_value|parameter_unit | parameter_lower_bound| parameter_upper_
|:------------------------------------|---------------:|:--------------|---------------------:|----------------
|Human delay - time symptom to careseeking |           4|Days          |                    NA|
|Human delay - incubation period      |              NA|Days          |                     7|
|Human delay - time symptom to outcome |             9|Days          |                    NA|
|Seroprevalence - IFA                 |              NA|NA            |                    NA|
|Attack rate                          |              21|Percentage (%) |                    NA|
|Human delay - incubation period      |              NA|Days          |                     4|

r$> █
```
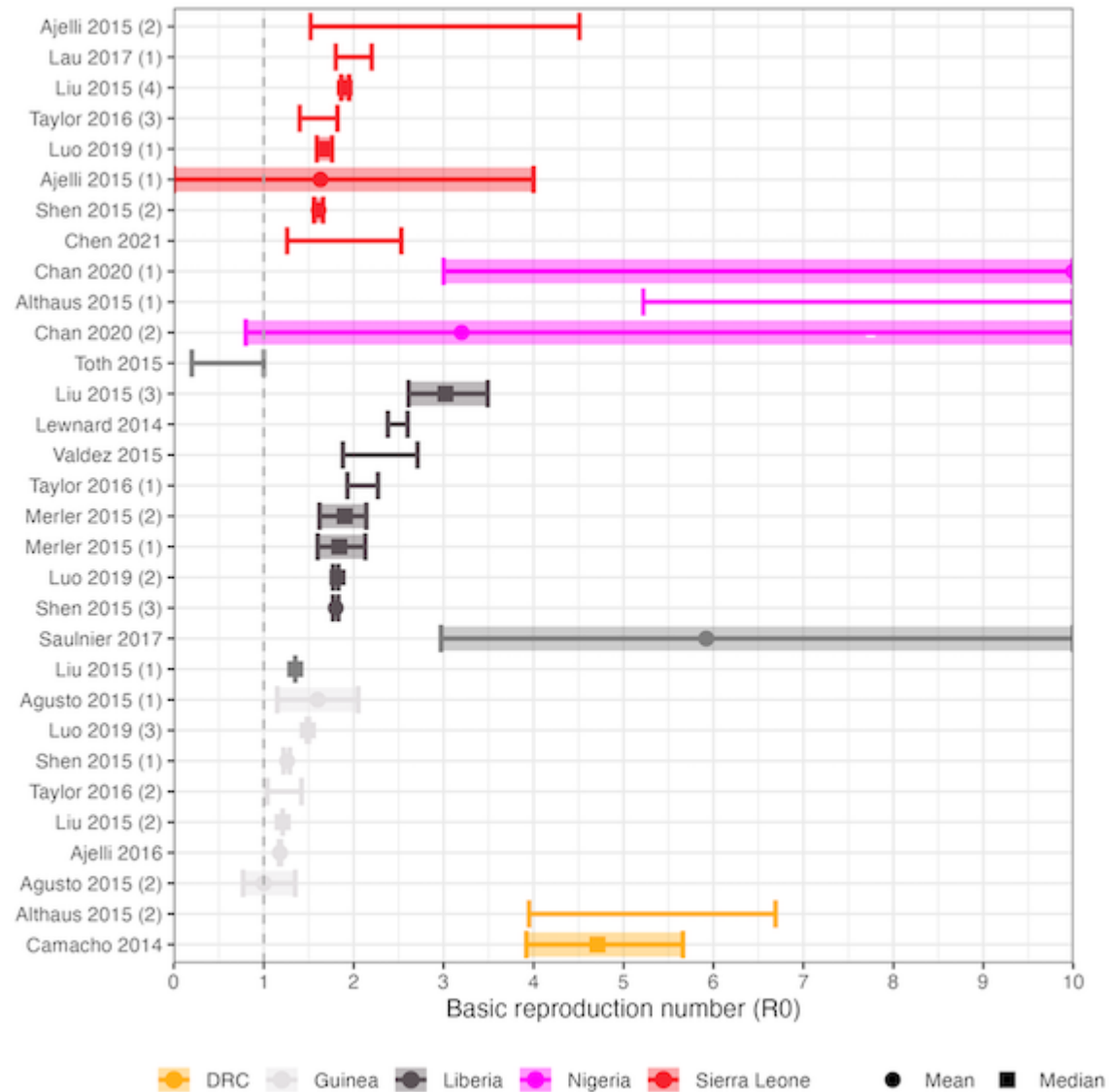
Features: filter data

```
r$>

r$> filter_cols(ebola$params, cols = 'article_qa_score', funs = ">", vals = '50')
# A tibble: 511 × 76
   id      parameter_data_id covidence_id pathogen parameter_type parameter_value exponent parameter_unit
   <chr>   <chr>                    <int> <chr>    <chr>                    <dbl>    <dbl> <chr>
 1 3fcec…  ad84262af2e94475…           30 Ebola v… Human delay –…              NA        0 Days
 2 3fcec…  f6a75a765034c39f…           30 Ebola v… Reproduction …             NA        0 No units
 3 3fcec…  cc4b1927b74f5083…           30 Ebola v… Human delay –…              NA        0 Days
 4 303ce…  07919213f2fdbf6d…           65 Ebola v… Reproduction …           2.01        0 No units
 5 4fc94…  877f0a1a8a355def…          104 Ebola v… Human delay –…            3.91        0 Days
 6 4fc94…  50d1d9811e75ffbe…          104 Ebola v… Human delay –…             4.2        0 Days
 7 4fc94…  516579f27e9805dc…          104 Ebola v… Severity – ca…            74.2        0 Percentage (%)
 8 4fc94…  973cb2ce99eec1cb…          104 Ebola v… Risk factors                NA        0 NA
 9 4fc94…  a4565f450352666d…          104 Ebola v… Risk factors                NA        0 NA
10 54207…  ea2a4be346093b2d…          163 Ebola v… Human delay –…               9        0 Days
# i 501 more rows
# i 67 more variables: parameter_upper_bound <dbl>, parameter_value_type <chr>, parameter_uncertainty_sin
#   parameter_uncertainty_singe_type <chr>, parameter_uncertainty_lower_value <dbl>, parameter_uncertaint
#   parameter_uncertainty_type <chr>, cfr_ifr_numerator <int>, cfr_ifr_denominator <int>, distribution_ty
#   distribution_par1_value <dbl>, distribution_par1_type <chr>, distribution_par1_uncertainty <lgl>,
#   distribution_par2_value <dbl>, distribution_par2_type <chr>, distribution_par2_uncertainty <lgl>,
#   method_from_supplement <lgl>, method_moment_value <chr>, cfr_ifr_method <chr>, method_r <chr>, …
# i Use `print(n = ...)` to see more rows

r$> █
```

Features: visualise data

```
:$>

:$> params <- filter_cols(params, "article_qa_score", funs = ">", vals = "50")

:$> forest_plot_r0(params, col_by = "population_country", shape_by = "parameter_value_type")
Warning messages:
1: Removed 11 rows containing missing values (`geom_point()`).
2: Removed 11 rows containing missing values (`geom_segment()`).

:$>
```

Features: visualise data

Development Roadmap

- currently 'experinemtal' but a stable release with each new pathogen as our group works down the list of pathogens

    - Ebola and Lassa 'nearly done' - epireview stable release planned end of March

    - SARS-CoV-1 `in progess'

    - followed by Zika, Henipa, RFV, CCHF, and MERS-CoV (order may change)

- utilities to summarise and visualise data

- user-friendly interface to update data or add new pathogen

- smooth interoperability with other tools (epiverse)

- extensive documentation

- what else can we do to make this a useful tool for the community?