Meta-analysis provides the best and most reliable scientific evidence

Strongest

Meta-analyses & systematic reviews

Randomized controlled trials

Cohort studies

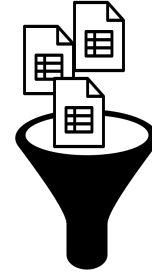Case-control studies

Cross sectional studies

Animal trials & *in vitro* studies

opinion papers, and letters
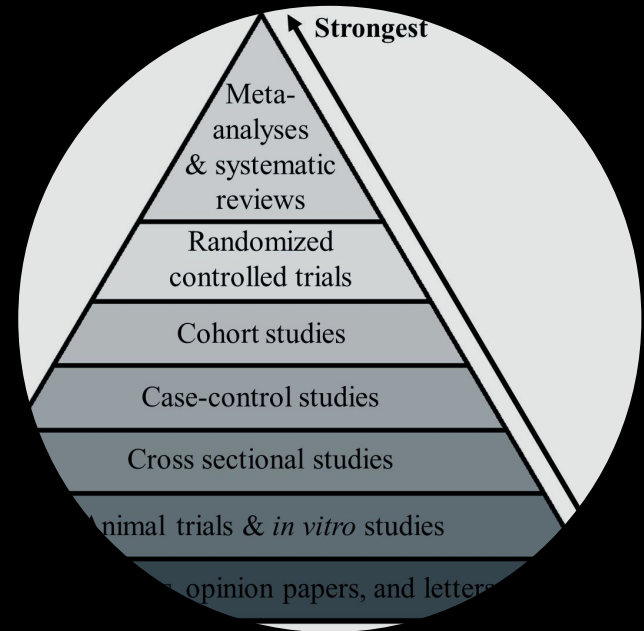
Meta-analysis aggregates evidence across several studies
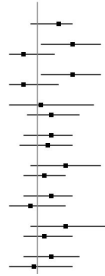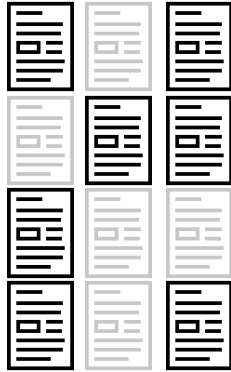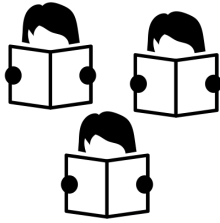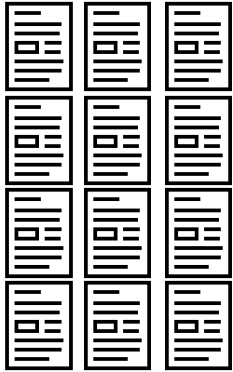
1. Search literature
2. Screen articles for eligibility

3. Select the studies of interest
4. Extract data from published articles

5. Calculate effect sizes and aggregate across studies
6. Statistical inference

There are standards that regulate how to do a meta-analysis properly

- A protocol defines the aims and methods
- Decisions about article screening and data extraction should be made at least by two independent researchers
- But usually the amount of work requires more collaborators

Doing a meta-analysis is not easy

Domain specific
knowledge

Methodological
knowledge

Teamwork

**Data management
and organization**

There should be a package for this!

- Reduce project management overhead
- File management should be automated and made reproducible
- Not straightforward to coordinate human and machine generated outputs
- Integrate with a GUI that is easy to use
- Currently, there is only one other (great) package {metagear} for meta-analysis workflow management, but it is not using the tidyverse principles, and it has a different philosophy and focus

The {metamanager} is here to help!

- Search scientific databases and get article meta-data
- Automate file creation for manual coding (via Google drive)
- Aggregate, evaluate, validate human coded data
- Clear and reproducible outputs
- Embracing the tidyverse principles (piping, readability, simplicity of functions)

Setting up a meta-analysis project

`init_gdrive()`: Creates a folder structure on google drive. Can be self specified or a default structure.

```
literature_search/
screening/
fulltext/
screening_consensus/
extraction/
extraction_consensus/
```

	A	B	C	D
1	name	role	screening_effort	extraction_effort
2	Ali	trainee	0.25	0
3	Andreas	trainee	0.25	0
4	Lydia	trainee	0.25	0
5	Tamas	trainer	0.25	0.1
6	Kristoffer	trainee	0	0.4
7	Marcell	trainee	0	0.3
8	Zoli	trainer	0	0.2

`init_team()`: Creates a google sheet with team member data. Accepts a tibble or opens a google sheet with header

`init_meta()`: Does both, using defaults

Create article screening files

```
merged_records %>%
    assign_articles(team_df, "screening_effort", seed = 42) %>%
    add_columns(c("decision", "reason")) %>%
    save_locally(local_path = "temp/screening/",
                 nesting = "reviewer",
                 postfix = "screening",
                  overwrite = TRUE)
# A tibble: 4,390 x 15
   decision reason identifier id      source title journal authors  year abstract    type   level position
   <chr>    <chr>  <chr>      <chr>   <chr>  <chr> <chr>   <chr>   <dbl> <chr>       <chr> <chr> <chr>
 1 NA       NA     pmid       24357~  pubmed 'Res~ Journa~ Silk, ~  2013 "Whilst n~  NA     NA    reviewe~
 2 NA       NA     pq_id      33710~  proqu~ 'U G~ NA      Young,~  2007 Backgroun~  THES   NA    reviewe~
 3 NA       NA     doi        10.10~  Scopus "\"B~ Accide~ Mammar~  2013 Two studi~  Arti~  NA    reviewe~
 4 NA       NA     doi        10.11~  Scopus "\"B~ Entrep~ Wolfe ~  2015 In this s~  Arti~  NA    reviewe~
 5 NA       NA     doi        10.10~  Scopus "\"C~ Femini~ De E.N.  2012 "This art~  Arti~  NA    reviewe~
 6 NA       NA     doi        10.10~  Scopus "\"G~ Comput~ Charsk~  2011 Does usin~  Arti~  NA    reviewe~
 7 NA       NA     doi        10.10~  Scopus "\"I~ Journa~ Balzar~  2014 "Human-hu~  Arti~  NA    reviewe~
 8 NA       NA     doi        10.11~  Scopus "\"I~ Artifi~ Davies~  2011 Simple di~  Arti~  NA    reviewe~
 9 NA       NA     doi        10.10~  Scopus "\"J~ Games ~ Lin J.~  2015 "Objectiv~ Arti~  NA    reviewe~
10 NA       NA     pmid       11850~  Scopus "\"S~ Americ~ Howard~  2002 "We recen~ Arti~  NA    reviewe~
# ... with 4,380 more rows, and 2 more variables: reviewer <chr>, name <chr>
```

Uploading and downloading gdrive folders

```
#This also converts files to google sheets that can be edited online
upload_to_gdrive(local_path = "temp/screening/",
                 gdrive_path = "research/meta-analysis/screening",
                 overwrite = TRUE)

# Downloads all files from a google folder as a nested dataframe
screening <- get_from_gdrive(gdrive_path = "research/meta-analysis/screening",
                             all_char = FALSE)
```

Why google drive?

- There are built in functions to convert data files to google sheets
- Google sheets provide collaborative editing with version control
- It is user friendly
- As all files are also saved locally, other cloud solutions can be used too, like github or dropbox

Google Drive

Evaluate screening performance

```
screening %>%
    unnest() %>%
    calculate_agreement()

# A tibble: 6 x 6
  name_pair       irr include_both exclude_both no_agreement invalid_decision
  <chr>         <dbl>        <dbl>        <dbl>        <dbl>            <dbl>
1 Ali_Lydia      0.52           23          122           50               11
2 Ali_Tamas      0.79           32          111           43                5
3 Ali_Andreas    0.82           15          203           19               12
4 Lydia_Tamas    0.66           26           63           53               10
5 Lydia_Andreas  0.56           33           95           32                0
6 Tamas_Andreas  0.89           22           99           11                0
```

Correcting and highlighting human errors

```
# Correct typos in the reasons for excluded studies
correct_terms <- read_lines("valid_exclude_reasons.txt")
screening %>%
    unnest() %>%
    mutate(exclude_reason = correct_categories(exclude_reason, correct_terms)) %>%
    count(exclude_reason)

# Safely convert col_types with human readable errors

articles %>%
    conversion_errors(c("year", "measurement_n"), as.integer)

# A tibble: 2 x 3
  file               fun          conversion_errors
  <chr>              <chr>              <dbl>
1 Ali_extract.csv    as.integer          25
2 Tamas_extract.csv  as.integer           3
```

Create data extraction files

```
variables_to_add <- c("subsample_id", "subsample_age", "subsample_female%",
"measurement_point", "measurement_n", "measurement_mean", "measurement_sd",
"study_design", "study_quality")

get_from_gdrive(screening_path) %>%
    unnest() %>%
    filter(is.na(exclude_reason)) %>%
    assign_articles(team_df, "extraction_effort", seed = 42) %>%
    add_columns(variables_to_add) %>%
    save_locally(local_path = "temp/extraction/", nesting = "reviewer", postfix =
"screening", overwrite = TRUE)

upload_to_gdrive(local_path = "temp/extraction/",
                 gdrive_path = "research/meta-analysis/extraction",
                 overwrite = TRUE)
```

Future plans

- Create functions for ML augmented screening process
- Outsource all literature search functions to a separate package
- ggplot2 based figures (flow diagrams, funnel plots, forest plots) because most meta-analysis packages use base plotting
- Handle article pdfs
- Hex sticker?

How to get / contribute to {metamanager}?

Please star, fork, open issues, PR on github!
https://github.com/nthun/metamanager

It is a work in progress, so don't expect perfection just yet

Installation:
devtools::install_github("nthun/metamanager")

Twitter: @nagyt