TestMarkdown-Analysis

2024-01-21

Software Requirements:

```
- R Installation
- RStudio for IDE
- LaTeX Installation for rendering RMarkdown to PDF (this file)

- GitHub Account for Repository/Data access
```

Setting up the Environment:

Required R Packages:

```
install.packages("rmarkdown")
install.packages("devtools")

install.packages("tidyverse")
install.packages("metafor")
install.packages("meta")

install.packages("tidyr")

install.packages("weightr")
```

Loading first Packages:

```
library(rmarkdown)

library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(metafor)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
```

```
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
## Loading required package: metadat
## Loading required package: numDeriv
##
## Loading the 'metafor' package (version 4.4-0). For an
## introduction to the package please type: help(metafor)
```

```
library(meta)
```

```
## Loading 'meta' package (version 7.0-0).
## Type 'help(meta)' for a brief overview.
## Readers of 'Meta-Analysis with R (Use R!)' should install
## older version of 'meta' package: https://tinyurl.com/dt4y5drs
```

```
library(dmetar)
```

```
## Extensive documentation for the dmetar package can be found at:
##  www.bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/
```

```
library (dplyr)
library(tidyr)

library(stringr)
library(readr)
```

Installing further R Packages:

```
devtools::install_github("MathiasHarrer/dmetar")
```

Loading remaining Packages:

```
library(dmetar)
```

Cloning Repository from Git-Hub (Fetching Data):

```
usethis::create_from_github(
  "https://github.com/VJMeyer/HPVPre_Repo.git",
  destdir = "~/path/to/where/you/want/the/local/repo/"
)
```

Set Working-Directory:

```
setwd("~/path/to/where/you/want/the/local/repo/")
```

Set Private Working-Directory:

Confirm Working-Directory:

Load Data:

```
library(tidyverse)

any_type <- read_csv("Anytype_OLDI.csv")
```

```
## Rows: 87 Columns: 24
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (19): study_id, author_year, title, study_design, world_region, world_su...
## dbl  (5): ending_year, num_older_wom, num_hpv_pos, num_hpv_neg, any_prev
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
hr_type <- read_csv("HR_OLDI.csv")
```

```
## Rows: 127 Columns: 24
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (19): study_id, author_year, title, study_design, world_region, world_su...
## dbl  (5): ending_year, hr_hpv_nr, num_older_wom, num_hr_hpv_pos, hr_prev
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
gtype <- read_csv("gtype.csv")
```

```
## Rows: 239 Columns: 22
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (16): study_id, author_year, title, study_design, world_region, world_su...
## dbl  (6): ending_year, num_older_wom, num_hpv_pos, num_hr_hpv_pos, num_type,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Quantitative Analysis

Data Cleaning

Ensure there are no MISSING values in the important columns of ANYTYPE Data:

```
library(tidyverse)


class(any_type)
```

```
## [1] "spec_tbl_df" "tbl_df"      "tbl"         "data.frame"
```

```
head(any_type)
```

```
## # A tibble: 6 x 24
```

```
##   study_id author_year            title    study_design ending_year world_region
##   <chr>    <chr>                  <chr>    <chr>              <dbl> <chr>
## 1 1O       Tsedenbal et al., 2018 "Human ~ cross-secti~        2017 Asia
## 2 2OG      Çolakoğlu et al., 2017 "Human ~ retrospecti~        2015 Asia
## 3 3O       Demirci et al., 2019   "Human ~ cross-secti~        2017 Asia
## 4 4OG      Dutta et al, 2012      "Preval~ cross-secti~        2010 Asia
## 5 5OH      Herrero et al., 2000   "Popula~ prospective         1994 Americas
## 6 6OHG     Jin et al., 2019       "The pr~ retrospecti~        2018 Asia
## # i 18 more variables: world_subregion <chr>, country <chr>, city_state <chr>,
## #   focus <chr>, overall_cytology <chr>, pap_method <chr>,
## #   recruitment_setting <chr>, hpv_types_reported <chr>,
## #   hpv_types_tested <chr>, hpv_cat <chr>, hpv_test <chr>, test_details <chr>,
## #   num_older_wom <dbl>, num_hpv_pos <dbl>, num_hpv_neg <dbl>, any_prev <dbl>,
## #   risk_of_bias <chr>, Notes <chr>
```

```
str(any_type)
```

```
## spc_tbl_ [87 x 24] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ study_id           : chr [1:87] "1O" "2OG" "3O" "4OG" ...
##  $ author_year        : chr [1:87] "Tsedenbal et al., 2018" "Çolakoğlu et al., 2017" "Demirci et al.
##  $ title              : chr [1:87] "Human papillomavirus genotyping among women with cervical abnorma
##  $ study_design       : chr [1:87] "cross-sectional" "retrospective" "cross-sectional" "cross-sectio
##  $ ending_year        : num [1:87] 2017 2015 2017 2010 1994 ...
##  $ world_region       : chr [1:87] "Asia" "Asia" "Asia" "Asia" ...
##  $ world_subregion    : chr [1:87] "Central Asia" "Western Asia" "Western Asia" "Southern Asia" ...
##  $ country            : chr [1:87] "Mongolia" "Turkiye" "Turkiye" "India" ...
##  $ city_state         : chr [1:87] "Ulaanbaatar" "Adana" "Istanbul" "West Bengal" ...
##  $ focus              : chr [1:87] "Both" "Both" "Both" "Both" ...
##  $ overall_cytology   : chr [1:87] "Predominantly normal" "Predominantly normal" "Predominantly norma
##  $ pap_method         : chr [1:87] "conventional" "LBC" "LBC" "conventional" ...
##  $ recruitment_setting: chr [1:87] "clinical setting" "screening" "clinical setting" "screening" ...
##  $ hpv_types_reported : chr [1:87] "overall" "16, 18, overall" "overall" "16, 18, overall" ...
##  $ hpv_types_tested   : chr [1:87] "6, 11, 16 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55
##  $ hpv_cat            : chr [1:87] "II" "II" "II" "II" ...
##  $ hpv_test           : chr [1:87] "LA" "LA" "LA" "MY09/11" ...
##  $ test_details       : chr [1:87] "L1 consensus PCR with Line-blot assay, Roche Linear Array test" "
##  $ num_older_wom      : num [1:87] 25 52 282 114 417 ...
##  $ num_hpv_pos        : num [1:87] 9 23 63 11 69 745 677 405 6 137 ...
##  $ num_hpv_neg        : num [1:87] 16 29 219 103 348 ...
##  $ any_prev           : num [1:87] 0.36 0.442 0.223 0.096 0.165 ...
##  $ risk_of_bias       : chr [1:87] "low" "high" "low" "high" ...
##  $ Notes              : chr [1:87] "just overall HPV positivity reported by age for 55+; HR-HPV+ and
##  - attr(*, "spec")=
##   .. cols(
##   ..   study_id = col_character(),
##   ..   author_year = col_character(),
##   ..   title = col_character(),
##   ..   study_design = col_character(),
##   ..   ending_year = col_double(),
##   ..   world_region = col_character(),
##   ..   world_subregion = col_character(),
##   ..   country = col_character(),
##   ..   city_state = col_character(),
##   ..   focus = col_character(),
##   ..   overall_cytology = col_character(),
```

```
##   ..   pap_method = col_character(),
##   ..   recruitment_setting = col_character(),
##   ..   hpv_types_reported = col_character(),
##   ..   hpv_types_tested = col_character(),
##   ..   hpv_cat = col_character(),
##   ..   hpv_test = col_character(),
##   ..   test_details = col_character(),
##   ..   num_older_wom = col_double(),
##   ..   num_hpv_pos = col_double(),
##   ..   num_hpv_neg = col_double(),
##   ..   any_prev = col_double(),
##   ..   risk_of_bias = col_character(),
##   ..   Notes = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
summary(any_type)
```

```
##    study_id         author_year           title           study_design
##  Length:87          Length:87          Length:87          Length:87
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1989   Length:87          Length:87          Length:87
##  1st Qu.:2006   Class :character   Class :character   Class :character
##  Median :2010   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2009
##  3rd Qu.:2014
##  Max.   :2019
##  NA's   :4
##   city_state           focus           overall_cytology    pap_method
##  Length:87          Length:87          Length:87          Length:87
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported hpv_types_tested     hpv_cat
##  Length:87           Length:87          Length:87          Length:87
##  Class :character    Class :character   Class :character   Class :character
##  Mode  :character    Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    hpv_test         test_details       num_older_wom      num_hpv_pos
##  Length:87          Length:87          Min.   :    6.0   Min.   :   0.0
##  Class :character   Class :character   1st Qu.:   82.0   1st Qu.:  13.0
##  Mode  :character   Mode  :character   Median :  293.0   Median :  29.0
##                                        Mean   : 1133.6   Mean   : 188.6
```

```
##                                      3rd Qu.:  697.5   3rd Qu.: 126.5
##                                      Max.   :18910.0   Max.   :6199.0
##                                      NA's   :4         NA's   :4
##   num_hpv_neg       any_prev      risk_of_bias          Notes
##  Min.   :    5   Min.   :0.0000   Length:87          Length:87
##  1st Qu.:   63   1st Qu.:0.0815   Class :character   Class :character
##  Median :  225   Median :0.1480   Mode  :character   Mode  :character
##  Mean   :  945   Mean   :0.1787
##  3rd Qu.:  569   3rd Qu.:0.2450
##  Max.   :13973   Max.   :0.6670
##  NA's   :4
```

```
view(any_type)
```

Filter rows where num_older_wom is NA and select study_id

```
na_in_num_older_wom <- any_type %>%
  filter(is.na(num_older_wom)) %>%
  select(study_id)
```

Filter rows where num_hpv_pos is NA and select study_id

```
na_in_num_hpv_pos <-any_type %>%
  filter(is.na(num_hpv_pos)) %>%
  select(study_id)
```

Filter rows where num_hpv_neg is NA and select study_id

```
na_in_num_hpv_neg <- any_type %>%
  filter(is.na(num_hpv_neg)) %>%
  select(study_id)
```

View the results

```
na_in_num_older_wom
```

```
## # A tibble: 4 x 1
##   study_id
##   <chr>
## 1 26OGP
## 2 43OP
## 3 56OHGP
## 4 66OHGP
```

```
na_in_num_hpv_pos
```

```
## # A tibble: 4 x 1
##   study_id
##   <chr>
## 1 26OGP
## 2 43OP
## 3 56OHGP
## 4 66OHGP
```

```
na_in_num_hpv_neg
```

```
## # A tibble: 4 x 1
##   study_id
##   <chr>
## 1 26OGP
## 2 43OP
## 3 56OHGP
## 4 66OHGP
```

there was missing data in one row (63O) of num_hpv_neg which was corrected

Filter out rows where the identifier ends with 'P'

```
any_type_filtered <- any_type %>%
  filter(!str_detect(study_id, "P$"))

summary(any_type_filtered)
```

```
##    study_id         author_year           title           study_design
##  Length:83          Length:83          Length:83          Length:83
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1989   Length:83          Length:83          Length:83
##  1st Qu.:2006   Class :character   Class :character   Class :character
##  Median :2010   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2009
##  3rd Qu.:2014
##  Max.   :2019
##  NA's   :4
##   city_state           focus           overall_cytology    pap_method
##  Length:83          Length:83          Length:83          Length:83
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported hpv_types_tested     hpv_cat
##  Length:83           Length:83          Length:83          Length:83
##  Class :character    Class :character   Class :character   Class :character
##  Mode  :character    Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    hpv_test         test_details       num_older_wom      num_hpv_pos
##  Length:83          Length:83          Min.   :    6.0   Min.   :   0.0
##  Class :character   Class :character   1st Qu.:   82.0   1st Qu.:  13.0
##  Mode  :character   Mode  :character   Median :  293.0   Median :  29.0
```

```
##                                      Mean   : 1133.6   Mean   : 188.6
##                                      3rd Qu.:  697.5   3rd Qu.: 126.5
##                                      Max.   :18910.0   Max.   :6199.0
##
##   num_hpv_neg       any_prev      risk_of_bias          Notes
##  Min.   :    5   Min.   :0.0000   Length:83          Length:83
##  1st Qu.:   63   1st Qu.:0.0790   Class :character   Class :character
##  Median :  225   Median :0.1490   Mode  :character   Mode  :character
##  Mean   :  945   Mean   :0.1812
##  3rd Qu.:  569   3rd Qu.:0.2470
##  Max.   :13973   Max.   :0.6670
##
```

```
nrow(any_type_filtered)
```

```
## [1] 83
```

Extract and view unique study ids in hr_type_filtered to make sure it is correct

```
unique_study_ids_any <- any_type_filtered %>%
  select(study_id) %>%
  distinct()

print(unique_study_ids_any, n = 83)
```

```
## # A tibble: 83 x 1
##    study_id
##    <chr>
##  1 1O
##  2 2OG
##  3 3O
##  4 4OG
##  5 5OH
##  6 6OHG
##  7 7O
##  8 8O
##  9 9O
## 10 10OHG
## 11 11O
## 12 13O
## 13 14O
## 14 15O
## 15 16OHG
## 16 17O
## 17 18OHG
## 18 19OH
## 19 20OH
## 20 21OH
## 21 22O
## 22 23O
## 23 24OH
## 24 25OH
## 25 27OH
## 26 28OH
## 27 29O
```

```
## 28 30OG
## 29 31OG
## 30 32OH
## 31 33OH
## 32 34OHG
## 33 35O
## 34 36O
## 35 37OH
## 36 38OH
## 37 39OG
## 38 40OHG
## 39 41OHG
## 40 44O
## 41 45OHG
## 42 46OH
## 43 47OHG
## 44 48OH
## 45 49OHG
## 46 50OH
## 47 51OHG
## 48 52OHG
## 49 53OH
## 50 54OH
## 51 57OHG
## 52 58OHG
## 53 59OHG
## 54 60OHG
## 55 61OH
## 56 62OH
## 57 63O
## 58 64OH
## 59 65O
## 60 67OHG
## 61 68OHG
## 62 69O
## 63 70OH
## 64 71O
## 65 72OH
## 66 73OH
## 67 74OH
## 68 75OH
## 69 76O
## 70 77OH
## 71 78OG
## 72 79O
## 73 80O
## 74 81OH
## 75 82OH
## 76 83OHG
## 77 84OH
## 78 85OHG
## 79 86OH
## 80 87OH
## 81 88OH
```

```
## 82 89O
## 83 90OH
```

###Ensure there are no MISSING values in the columns of HR TYPE where there should be data

```
class(hr_type)
```

```
## [1] "spec_tbl_df" "tbl_df"      "tbl"         "data.frame"
```

```
head(hr_type)
```

```
## # A tibble: 6 x 24
##   study_id author_year           title     study_design ending_year world_region
##   <chr>    <chr>                 <chr>     <chr>              <dbl> <chr>
## 1 1H       Abulizi et al., 2021  "At what~ cross-secti~        2014 Asia
## 2 2H       Bae et al., 2009      "Natural~ prospective         2004 Asia
## 3 3H       Clarke et al, 2021    "Age-spe~ prospective         2018 Americas
## 4 4HG      Yoshida et al., 2007  "Quantit~ cross-secti~        2004 Asia
## 5 5H       Dalstein., 2009       "HPV Per~ prospective         2002 Europe
## 6 6H       Ernstson et al., 2019 "Detecti~ cross-secti~        2017 Europe
## # i 18 more variables: world_subregion <chr>, country <chr>, city_state <chr>,
## #   focus <chr>, overall_cytology <chr>, pap_method <chr>,
## #   recruitment_setting <chr>, hpv_types_reported <chr>,
## #   hpv_types_tested <chr>, hr_hpv_nr <dbl>, hr_hpv_cat <chr>, hpv_test <chr>,
## #   test_details <chr>, num_older_wom <dbl>, num_hr_hpv_pos <dbl>,
## #   hr_prev <dbl>, risk_of_bias <chr>, Comments <chr>
```

```
str(hr_type)
```

```
## spc_tbl_ [127 x 24] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ study_id           : chr [1:127] "1H" "2H" "3H" "4HG" ...
##  $ author_year        : chr [1:127] "Abulizi et al., 2021" "Bae et al., 2009" "Clarke et al, 2021" "
##  $ title              : chr [1:127] "At what age should the Uyghur minority initiate cervical cancer
##  $ study_design       : chr [1:127] "cross-sectional" "prospective" "prospective" "cross-sectional"
##  $ ending_year        : num [1:127] 2014 2004 2018 2004 2002 ...
##  $ world_region       : chr [1:127] "Asia" "Asia" "Americas" "Asia" ...
##  $ world_subregion    : chr [1:127] "Eastern Asia" "Eastern Asia" "Northern America" "Eastern Asia"
##  $ country            : chr [1:127] "China" "South Korea" "USA" "Japan" ...
##  $ city_state         : chr [1:127] "South Xinjang" "Goyang" "Mississippi" "Gunma" ...
##  $ focus              : chr [1:127] "Both" "Both" "Both" "Both" ...
##  $ overall_cytology   : chr [1:127] "Predominantly normal" "Predominantly normal" "Predominantly norm
##  $ pap_method         : chr [1:127] "LBC" "conventional" "LBC" "LBC" ...
##  $ recruitment_setting: chr [1:127] "clinical setting" "screening" "screening" "clinical setting" ..
##  $ hpv_types_reported : chr [1:127] "HR-HPV" "HR-HPV" "HR-HPV" "HR-HPV" ...
##  $ hpv_types_tested   : chr [1:127] "16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68" "16,18,3
##  $ hr_hpv_nr          : num [1:127] 14 13 14 7 14 14 13 12 13 17 ...
##  $ hr_hpv_cat         : chr [1:127] "III" "III" "III" "II" ...
##  $ hpv_test           : chr [1:127] "other PCR" "HC2" "Multiple\n" "PCR E6/E7" ...
##  $ test_details       : chr [1:127] "careHPVTM\ntest (Qiagen Inc.)" "Full HPV genome detection by HC
##  $ num_older_wom      : num [1:127] 1614 1815 1118 148 77 ...
##  $ num_hr_hpv_pos     : num [1:127] 215 100 185 30 13 27 79 43 8 650 ...
##  $ hr_prev            : num [1:127] 0.133 0.055 0.165 0.203 0.169 0.062 0.142 0.041 0.018 0.351 ...
##  $ risk_of_bias       : chr [1:127] "low" "low" "low" "high" ...
##  $ Comments           : chr [1:127] "Used plot digitizer on Fig. 2" "Age related data only available
##  - attr(*, "spec")=
##   .. cols(
##   ..   study_id = col_character(),
```

```
##   ..   author_year = col_character(),
##   ..   title = col_character(),
##   ..   study_design = col_character(),
##   ..   ending_year = col_double(),
##   ..   world_region = col_character(),
##   ..   world_subregion = col_character(),
##   ..   country = col_character(),
##   ..   city_state = col_character(),
##   ..   focus = col_character(),
##   ..   overall_cytology = col_character(),
##   ..   pap_method = col_character(),
##   ..   recruitment_setting = col_character(),
##   ..   hpv_types_reported = col_character(),
##   ..   hpv_types_tested = col_character(),
##   ..   hr_hpv_nr = col_double(),
##   ..   hr_hpv_cat = col_character(),
##   ..   hpv_test = col_character(),
##   ..   test_details = col_character(),
##   ..   num_older_wom = col_double(),
##   ..   num_hr_hpv_pos = col_double(),
##   ..   hr_prev = col_double(),
##   ..   risk_of_bias = col_character(),
##   ..   Comments = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
summary(hr_type)
```

```
##    study_id         author_year           title           study_design
##  Length:127         Length:127         Length:127         Length:127
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1994   Length:127         Length:127         Length:127
##  1st Qu.:2006   Class :character   Class :character   Class :character
##  Median :2011   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2010
##  3rd Qu.:2015
##  Max.   :2019
##  NA's   :6
##   city_state           focus           overall_cytology    pap_method
##  Length:127         Length:127         Length:127         Length:127
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported hpv_types_tested     hr_hpv_nr
##  Length:127          Length:127         Length:127         Min.   : 2.00
##  Class :character    Class :character   Class :character   1st Qu.:13.00
```

```
##  Mode  :character    Mode  :character   Mode  :character   Median :14.00
##                                                           Mean   :14.16
##                                                           3rd Qu.:15.00
##                                                           Max.   :23.00
##
##   hr_hpv_cat          hpv_test         test_details       num_older_wom
##  Length:127         Length:127         Length:127         Min.   :    12
##  Class :character   Class :character   Class :character   1st Qu.:   187
##  Mode  :character   Mode  :character   Mode  :character   Median :   454
##                                                           Mean   :  5027
##                                                           3rd Qu.:  1138
##                                                           Max.   :299259
##                                                           NA's   :14
##  num_hr_hpv_pos       hr_prev       risk_of_bias         Comments
##  Min.   :    1.0   Min.   :0.0160   Length:127         Length:127
##  1st Qu.:   16.0   1st Qu.:0.0530   Class :character   Class :character
##  Median :   40.0   Median :0.0960   Mode  :character   Mode  :character
##  Mean   :  287.3   Mean   :0.1214
##  3rd Qu.:  117.2   3rd Qu.:0.1530
##  Max.   :12014.0   Max.   :0.6120
##  NA's   :15        NA's   :1
```

```
view(hr_type)
nrow(hr_type)
```

```
## [1] 127
```

Filter rows where num_older_wom is NA and select study_id

```
na_in_num_older_wom_hr <- hr_type %>%
  filter(is.na(num_older_wom)) %>%
  select(study_id)
```

Filter rows where num_hr_hpv_pos is NA and select study_id

```
na_in_num_hr_hpv_pos <- hr_type %>%
  filter(is.na(num_hr_hpv_pos)) %>%
  select(study_id)
```

Filter rows where hr_prev is NA and select study_id

```
na_in_hr_prev <- hr_type %>%
  filter(is.na(hr_prev)) %>%
  select(study_id)
```

View the results

```
na_in_num_older_wom_hr
```

```
## # A tibble: 14 x 1
##    study_id
##    <chr>
##  1 14HP
##  2 16HP
```

```
##  3 28HP
##  4 38HP
##  5 41HGP
##  6 42HGP
##  7 46HGP
##  8 56OHGP
##  9 51HP
## 10 66OHGP
## 11 60HP
## 12 63HGP
## 13 70HGP
## 14 72HP
```

```
na_in_num_hr_hpv_pos
```

```
## # A tibble: 15 x 1
##    study_id
##    <chr>
##  1 14HP
##  2 16HP
##  3 23HG
##  4 28HP
##  5 38HP
##  6 41HGP
##  7 42HGP
##  8 46HGP
##  9 56OHGP
## 10 51HP
## 11 66OHGP
## 12 60HP
## 13 63HGP
## 14 70HGP
## 15 72HP
```

```
na_in_hr_prev
```

```
## # A tibble: 1 x 1
##   study_id
##   <chr>
## 1 23HG
```

we must exclude study id 23HG from the HR analysis because it has crude numbers for all gtypes but not for HR type Filter out rows where the identifier ends with 'P' or is '23HG'

```
hr_type_filtered <- hr_type %>%
  filter(!str_detect(study_id, "P$"), study_id != "23HG")


summary(hr_type_filtered)
```

```
##    study_id         author_year           title           study_design
##  Length:112         Length:112         Length:112         Length:112
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
```

```
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1994   Length:112         Length:112         Length:112
##  1st Qu.:2006   Class :character   Class :character   Class :character
##  Median :2011   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2010
##  3rd Qu.:2015
##  Max.   :2019
##  NA's   :6
##   city_state           focus           overall_cytology    pap_method
##  Length:112         Length:112         Length:112         Length:112
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported hpv_types_tested     hr_hpv_nr
##  Length:112          Length:112         Length:112         Min.   : 2.00
##  Class :character    Class :character   Class :character   1st Qu.:13.00
##  Mode  :character    Mode  :character   Mode  :character   Median :14.00
##                                                            Mean   :14.09
##                                                            3rd Qu.:15.00
##                                                            Max.   :23.00
##
##   hr_hpv_cat          hpv_test         test_details       num_older_wom
##  Length:112         Length:112         Length:112         Min.   :    12
##  Class :character   Class :character   Class :character   1st Qu.:   185
##  Mode  :character   Mode  :character   Mode  :character   Median :   446
##                                                           Mean   :  5066
##                                                           3rd Qu.:  1169
##                                                           Max.   :299259
##
##  num_hr_hpv_pos       hr_prev       risk_of_bias         Comments
##  Min.   :    1.0   Min.   :0.0160   Length:112         Length:112
##  1st Qu.:   16.0   1st Qu.:0.0530   Class :character   Class :character
##  Median :   40.0   Median :0.0960   Mode  :character   Mode  :character
##  Mean   :  287.3   Mean   :0.1226
##  3rd Qu.:  117.2   3rd Qu.:0.1510
##  Max.   :12014.0   Max.   :0.6120
##
```

```
nrow(hr_type_filtered)
```

```
## [1] 112
```

Extract and view unique study ids in hr_type_filtered to make sure it is correct

```
unique_study_ids_hr <- hr_type_filtered %>%
  select(study_id) %>%
  distinct()

print(unique_study_ids_hr, n = 112)
```

```
## # A tibble: 112 x 1
```

```
##     study_id
##     <chr>
##   1 1H
##   2 2H
##   3 3H
##   4 4HG
##   5 5H
##   6 6H
##   7 7H
##   8 8HG
##   9 5OH
##  10 6OHG
##  11 9H
##  12 10H
##  13 11H
##  14 12H
##  15 13H
##  16 15H
##  17 10OHG
##  18 17HG
##  19 18H
##  20 19H
##  21 16OHG
##  22 20H
##  23 21HG
##  24 22H
##  25 24HG
##  26 25HG
##  27 26H
##  28 27H
##  29 18OHG
##  30 19OH
##  31 20OH
##  32 21OH
##  33 24OH
##  34 25OH
##  35 29HG
##  36 30HG
##  37 31H
##  38 27OH
##  39 28OH
##  40 32H
##  41 33H
##  42 32OH
##  43 34H
##  44 33OH
##  45 34OHG
##  46 35H
##  47 36H
##  48 37H
##  49 37OH
##  50 38OH
##  51 40OHG
##  52 41OHG
```

```
##  53 39H
##  54 40H
##  55 45OHG
##  56 46OH
##  57 47OHG
##  58 43HG
##  59 48OH
##  60 49OHG
##  61 44H
##  62 50OH
##  63 51OHG
##  64 45H
##  65 52OHG
##  66 47H
##  67 48HG
##  68 53OH
##  69 54OH
##  70 49H
##  71 57OHG
##  72 50HG
##  73 58OHG
##  74 52HG
##  75 59OHG
##  76 60OHG
##  77 61OH
##  78 53HG
##  79 62OH
##  80 54HG
##  81 55H
##  82 56H
##  83 64OH
##  84 57H
##  85 58H
##  86 67OHG
##  87 68OHG
##  88 59H
##  89 70OH
##  90 72OH
##  91 73OH
##  92 74OH
##  93 75OH
##  94 61H
##  95 62H
##  96 77OH
##  97 64H
##  98 65HG
##  99 81OH
## 100 82OH
## 101 66H
## 102 83OHG
## 103 67H
## 104 84OH
## 105 85OHG
## 106 86OH
```

```
## 107 68H
## 108 87OH
## 109 88OH
## 110 69H
## 111 90OH
## 112 71HG
```

Filter out rows where the identifier ends with 'P' for prevalance only for gtype_filtered

###trim whitespaces

```
gtype <- gtype %>%
  mutate(study_id = trimws(study_id))

gtype_filtered <- gtype %>%
  filter(!str_detect(study_id, "P$"))

class(gtype_filtered)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

```
head(gtype_filtered)
```

```
## # A tibble: 6 x 22
##   study_id author_year             title   study_design ending_year world_region
##   <chr>    <chr>                   <chr>   <chr>              <dbl> <chr>
## 1 4HG      Yoshida et al., 2007    "Quant~ cross-secti~        2004 Asia
## 2 2OG      Çolakoğlu et al., 2017  "Human~ cross-secti~        2015 Asia
## 3 4OG      Dutta et al, 2012       "Preva~ cross-secti~        2010 Asia
## 4 4OG      Dutta et al, 2012       "Preva~ cross-secti~        2010 Asia
## 5 8HG      Hermansson et al., 2018 "HPV p~ retrospecti~        2015 Europe
## 6 8HG      Hermansson et al., 2018 "HPV p~ retrospecti~        2015 Europe
## # i 16 more variables: world_subregion <chr>, country <chr>, city_state <chr>,
## #   focus <chr>, overall_cytology <chr>, pap_method <chr>,
## #   recruitment_setting <chr>, hpv_types_reported <chr>, hpv_test <chr>,
## #   num_older_wom <dbl>, num_hpv_pos <dbl>, num_hr_hpv_pos <dbl>,
## #   hpv_type <chr>, num_type <dbl>, type_prev <dbl>, notes <chr>
```

```
str(gtype_filtered)
```

```
## tibble [229 x 22] (S3: tbl_df/tbl/data.frame)
##  $ study_id           : chr [1:229] "4HG" "2OG" "4OG" "4OG" ...
##  $ author_year        : chr [1:229] "Yoshida et al., 2007" "Çolakoğlu et al., 2017" "Dutta et al, 20
##  $ title              : chr [1:229] "Quantitative real-time polymerase chain reaction analysis of th
##  $ study_design       : chr [1:229] "cross-sectional" "cross-sectional" "cross-sectional" "cross-sec
##  $ ending_year        : num [1:229] 2004 2015 2010 2010 2015 ...
##  $ world_region       : chr [1:229] "Asia" "Asia" "Asia" "Asia" ...
##  $ world_subregion    : chr [1:229] "Eastern Asia" "Western Asia" "Southern Asia" "Southern Asia" ..
##  $ country            : chr [1:229] "Japan" "Turkey" "India" "India" ...
##  $ city_state         : chr [1:229] "Gunma" "Adana" "West Bengal" "West Bengal" ...
##  $ focus              : chr [1:229] "Both" "Both" "Both" "Both" ...
##  $ overall_cytology   : chr [1:229] "Predominantly normal" "Predominantly normal" "Predominantly nor
##  $ pap_method         : chr [1:229] "LBC" "LBC" "Conventional" "Conventional" ...
##  $ recruitment_setting: chr [1:229] "clinical setting" "screening" "screening" "screening" ...
##  $ hpv_types_reported : chr [1:229] "HR-HPV, 16" "16, 18 overall" "16, 18, overall" "16, 18, overall
##  $ hpv_test           : chr [1:229] "PCR E6/E7" "LA" "MY09/11" "MY09/11" ...
```

```
##  $ num_older_wom      : num [1:229] 148 52 114 114 1051 ...
##  $ num_hpv_pos        : num [1:229] NA 23 11 11 NA NA NA NA NA NA ...
##  $ num_hr_hpv_pos     : num [1:229] 30 NA NA NA 43 43 43 43 43 43 ...
##  $ hpv_type           : chr [1:229] "16" "16,18" "16" "18" ...
##  $ num_type           : num [1:229] 3.6 7 1 1 12 5 6 9 3 4 ...
##  $ type_prev          : num [1:229] 0.02432 0.13462 0.00877 0.00877 0.01142 ...
##  $ notes              : chr [1:229] "numbers are not reported in table format but as figures therefo
```

```
summary(gtype_filtered)
```

```
##    study_id         author_year           title           study_design
##  Length:229         Length:229         Length:229         Length:229
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1993   Length:229         Length:229         Length:229
##  1st Qu.:2005   Class :character   Class :character   Class :character
##  Median :2011   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2010
##  3rd Qu.:2017
##  Max.   :2019
##  NA's   :4
##   city_state           focus           overall_cytology    pap_method
##  Length:229         Length:229         Length:229         Length:229
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported   hpv_test         num_older_wom
##  Length:229          Length:229         Length:229         Min.   :    7
##  Class :character    Class :character   Class :character   1st Qu.:  276
##  Mode  :character    Mode  :character   Mode  :character   Median : 1458
##                                                            Mean   : 3406
##                                                            3rd Qu.: 3859
##                                                            Max.   :18910
##
##   num_hpv_pos     num_hr_hpv_pos     hpv_type           num_type
##  Min.   :   2.0   Min.   :   4.0   Length:229         Min.   :   0.0
##  1st Qu.:  40.0   1st Qu.:  33.0   Class :character   1st Qu.:   4.0
##  Median : 133.5   Median : 104.0   Mode  :character   Median :  16.0
##  Mean   : 850.6   Mean   : 363.2                      Mean   : 107.6
##  3rd Qu.: 713.0   3rd Qu.: 405.0                      3rd Qu.:  41.0
##  Max.   :6199.0   Max.   :2836.0                      Max.   :5054.5
##  NA's   :97       NA's   :24
##    type_prev           notes
##  Min.   :0.000000   Length:229
##  1st Qu.:0.004905   Class :character
##  Median :0.012000   Mode  :character
##  Mean   :0.032086
```

```
##  3rd Qu.:0.034980
##  Max.   :0.413793
##
```

```
view(gtype_filtered)

summary(gtype_filtered)
```

```
##    study_id         author_year           title           study_design
##  Length:229         Length:229         Length:229         Length:229
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1993   Length:229         Length:229         Length:229
##  1st Qu.:2005   Class :character   Class :character   Class :character
##  Median :2011   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2010
##  3rd Qu.:2017
##  Max.   :2019
##  NA's   :4
##   city_state           focus           overall_cytology    pap_method
##  Length:229         Length:229         Length:229         Length:229
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported   hpv_test         num_older_wom
##  Length:229          Length:229         Length:229         Min.   :    7
##  Class :character    Class :character   Class :character   1st Qu.:  276
##  Mode  :character    Mode  :character   Mode  :character   Median : 1458
##                                                            Mean   : 3406
##                                                            3rd Qu.: 3859
##                                                            Max.   :18910
##
##   num_hpv_pos     num_hr_hpv_pos     hpv_type           num_type
##  Min.   :   2.0   Min.   :   4.0   Length:229         Min.   :   0.0
##  1st Qu.:  40.0   1st Qu.:  33.0   Class :character   1st Qu.:   4.0
##  Median : 133.5   Median : 104.0   Mode  :character   Median :  16.0
##  Mean   : 850.6   Mean   : 363.2                      Mean   : 107.6
##  3rd Qu.: 713.0   3rd Qu.: 405.0                      3rd Qu.:  41.0
##  Max.   :6199.0   Max.   :2836.0                      Max.   :5054.5
##  NA's   :97       NA's   :24
##    type_prev           notes
##  Min.   :0.000000   Length:229
##  1st Qu.:0.004905   Class :character
##  Median :0.012000   Mode  :character
##  Mean   :0.032086
##  3rd Qu.:0.034980
##  Max.   :0.413793
```

```
##
```

```
nrow(gtype_filtered)
```

```
## [1] 229
```

Extract and view unique study ids in gtype_filtered to make sure it is correct

```
unique_study_ids_gtype <- gtype_filtered %>%
  select(study_id) %>%
  distinct()

nrow(unique_study_ids_gtype)
```

```
## [1] 43
```

```
print(unique_study_ids_gtype, n = 43)
```

```
## # A tibble: 43 x 1
##    study_id
##    <chr>
##  1 4HG
##  2 2OG
##  3 4OG
##  4 8HG
##  5 6OHG
##  6 10OHG
##  7 71HG
##  8 17HG
##  9 16OHG
## 10 24HG
## 11 25HG
## 12 18OHG
## 13 29HG
## 14 30HG
## 15 30OG
## 16 31OG
## 17 34OHG
## 18 39OG
## 19 40OHG
## 20 41OHG
## 21 45OHG
## 22 47OHG
## 23 43HG
## 24 49OHG
## 25 51OHG
## 26 48HG
## 27 52OHG
## 28 57OHG
## 29 50HG
## 30 58OHG
## 31 52HG
## 32 59OHG
## 33 60OHG
## 34 53HG
```

```
## 35 54HG
## 36 67OHG
## 37 68OHG
## 38 78OG
## 39 65HG
## 40 83OHG
## 41 85OHG
## 42 21HG
## 43 23HG
```

```
#Data frames for quantitative analysis ----
# any hpv type analysis ---> any_type_filtered
#hr type analysis ---> hr_type_filtered
#type-specific analysis --->  gtype_filtered
```

Meta-analysis model for any_type_filtered —-

```
summary(any_type_filtered)
```

```
##    study_id         author_year          title           study_design
##  Length:83          Length:83          Length:83          Length:83
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ending_year   world_region       world_subregion      country
##  Min.   :1989   Length:83          Length:83          Length:83
##  1st Qu.:2006   Class :character   Class :character   Class :character
##  Median :2010   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2009
##  3rd Qu.:2014
##  Max.   :2019
##  NA's   :4
##   city_state           focus           overall_cytology    pap_method
##  Length:83          Length:83          Length:83          Length:83
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  recruitment_setting hpv_types_reported hpv_types_tested     hpv_cat
##  Length:83           Length:83          Length:83          Length:83
##  Class :character    Class :character   Class :character   Class :character
##  Mode  :character    Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    hpv_test         test_details       num_older_wom      num_hpv_pos
##  Length:83          Length:83          Min.   :    6.0   Min.   :   0.0
##  Class :character   Class :character   1st Qu.:   82.0   1st Qu.:  13.0
```

```
##  Mode  :character   Mode  :character   Median :  293.0   Median :  29.0
##                                        Mean   : 1133.6   Mean   : 188.6
##                                        3rd Qu.:  697.5   3rd Qu.: 126.5
##                                        Max.   :18910.0   Max.   :6199.0
##
##   num_hpv_neg       any_prev      risk_of_bias          Notes
##  Min.   :    5   Min.   :0.0000   Length:83          Length:83
##  1st Qu.:   63   1st Qu.:0.0790   Class :character   Class :character
##  Median :  225   Median :0.1490   Mode  :character   Mode  :character
##  Mean   :  945   Mean   :0.1812
##  3rd Qu.:  569   3rd Qu.:0.2470
##  Max.   :13973   Max.   :0.6670
##
```

```
# there are zero event values in the num_hpv_pos?
```

Identifying zero-event cases

```
zero_event_cases <- any_type_filtered[any_type_filtered$num_hpv_pos == 0, ]
```

Counting the number of zero-event cases

```
num_zero_event_cases <- nrow(zero_event_cases)
```

Printing the number of zero-event cases

```
print(paste("Number of zero-event cases:", num_zero_event_cases))
```

```
## [1] "Number of zero-event cases: 3"
```

Printing the study IDs of zero-event cases

```
if(num_zero_event_cases > 0) {
  print("Study IDs with zero-event cases:")
  print(zero_event_cases$author_year)
  print(zero_event_cases$study_id)
} else {
  print("There are no zero-event cases.")
}
```

```
## [1] "Study IDs with zero-event cases:"
## [1] "Becker et al., 1991" "Eren et al., 2010"   "Rahmat et al., 2021"
## [1] "230" "440" "710"
```

Random effects MA, Freeman-Tukey double arcsine transformation, restricted maximum likelihood estimator, Knapp-Hartung adjustment

```
any_type_ma <- metaprop(event = num_hpv_pos,
                        n = num_older_wom,
                        studlab = author_year,
                        data = any_type_filtered,
                        sm = "PFT",
                        method.tau = "REML",
```

```
                        method.ci = "NAsm",
                        add = 0,
                        fixed = FALSE,
                        random = TRUE,
                        hakn = TRUE,
                        title = "Anytype HPV Prevalence in Women (50+) with Predominantly Normal Cytolo

summary(any_type_ma)
```

```
## Review:     Anytype HPV Prevalence in Women (50+) with Predominantly Normal  ...
##
##                               proportion           95%-CI %W(random)
## Tsedenbal et al., 2018            0.3600 [0.1812; 0.5598]        1.0
## Çolakoğlu et al., 2017            0.4423 [0.3091; 0.5797]        1.1
## Demirci et al., 2019              0.2234 [0.1766; 0.2740]        1.3
## Dutta et al, 2012                 0.0965 [0.0482; 0.1583]        1.2
## Herrero et al., 2000              0.1655 [0.1313; 0.2027]        1.3
## Jin et al., 2019                  0.4025 [0.3802; 0.4249]        1.3
## Li, P. et al., 2021               0.1880 [0.1754; 0.2009]        1.3
## Li, XF. et al., 2021              0.1235 [0.1125; 0.1350]        1.3
## Nuñez-Troconis et al., 2009       0.2400 [0.0895; 0.4295]        1.0
## Richter et al., 2013              0.4215 [0.3683; 0.4757]        1.3
## Souho et al., 2016                0.5167 [0.4533; 0.5798]        1.3
## Tezcan et al., 2014               0.2453 [0.1376; 0.3711]        1.1
## Donkoh et al., 2022               0.3662 [0.2574; 0.4821]        1.2
## Maehama et al., 2002              0.1050 [0.0933; 0.1173]        1.3
## Brotherton et al., 2015           0.1704 [0.1112; 0.2389]        1.2
## Ahmadi et al., 2020               0.1500 [0.0534; 0.2798]        1.1
## Andujar et al., 2020              0.0806 [0.0677; 0.0944]        1.3
## Ardhaoui et al., 2016             0.1702 [0.0743; 0.2927]        1.1
## Balanda et al., 2016              0.0774 [0.0515; 0.1079]        1.3
## Baloch et al., 2017               0.1667 [0.1017; 0.2434]        1.2
## Bansal et al., 2014               0.0718 [0.0473; 0.1009]        1.3
## Becker et al., 1991               0.0000 [0.0000; 0.0515]        1.0
## Bell et al., 2007                 0.1667 [0.0396; 0.3464]        0.9
## Bi et al., 2015                   0.2866 [0.2620; 0.3119]        1.3
## Castellsague et al., 2012         0.0705 [0.0486; 0.0960]        1.3
## Castellsague et al., 2001         0.2195 [0.1042; 0.3607]        1.1
## Castle et al., 2006               0.1222 [0.0968; 0.1501]        1.3
## Cathro et al., 2009               0.0682 [0.0089; 0.1656]        1.1
## Centurioni et al., 2005           0.1553 [0.1139; 0.2017]        1.3
## Chan et al., 2002                 0.0298 [0.0112; 0.0560]        1.3
## Chansaeroj et al., 2010           0.0678 [0.0470; 0.0919]        1.3
## Chen et al., 2015                 0.1473 [0.1227; 0.1736]        1.3
## Chong et al., 2010                0.3913 [0.1996; 0.6006]        0.9
## Coser et al., 2013                0.2449 [0.1643; 0.3354]        1.2
## Dai et al., 2006                  0.1429 [0.0776; 0.2231]        1.2
## DeVuyst et al., 2003              0.2500 [0.0393; 0.5392]        0.7
## Debrah et al., 2021               0.2857 [0.0098; 0.6822]        0.6
## Demers et al., 2012               0.0939 [0.0580; 0.1371]        1.3
## Dufit et al., 2016                0.4396 [0.3387; 0.5429]        1.2
## Eren et al., 2010                 0.0000 [0.0000; 0.2680]        0.5
## Foliaki et al., 2014              0.1857 [0.1022; 0.2863]        1.2
## Gravitt et al., 2013              0.1533 [0.1146; 0.1965]        1.3
```

```
## Hamlin-Douglas et al., 2008      0.1983 [0.1317; 0.2745]        1.2
## Hernandez-Rosas et al., 2021     0.0607 [0.0367; 0.0901]        1.3
## Herrero et al., 2005             0.2845 [0.2626; 0.3068]        1.3
## Hong et al., 2015                0.1415 [0.0976; 0.1919]        1.3
## Hooi et al., 2018                0.1705 [0.1302; 0.2149]        1.3
## Jiang et al., 2011               0.1479 [0.1256; 0.1717]        1.3
## Shen et al., 2021                0.1373 [0.0868; 0.1967]        1.2
## Xiao et al, 2016                 0.3458 [0.3244; 0.3675]        1.3
## Klug et al., 2007                0.0254 [0.0184; 0.0333]        1.3
## Kobetz et al., 2012              0.1852 [0.0913; 0.3012]        1.1
## Lee et al., 2012                 0.3278 [0.3211; 0.3345]        1.3
## Leinonen et al., 2013            0.0486 [0.0452; 0.0522]        1.3
## Levert et al., 2000              0.1485 [0.1223; 0.1769]        1.3
## Li et al., 2019                  0.2857 [0.2537; 0.3188]        1.3
## Li et al., 2011                  0.0625 [0.0331; 0.0999]        1.3
## Liu et al., 2014                 0.2724 [0.2382; 0.3081]        1.3
## López Rivera et al., 2012        0.0927 [0.0595; 0.1323]        1.3
## Moore et al., 2009               0.1160 [0.0980; 0.1353]        1.3
## Mudderis et al., 2019            0.1781 [0.0978; 0.2751]        1.2
## Anh et al., 2003                 0.0422 [0.0273; 0.0600]        1.3
## Pista et al., 2011               0.0566 [0.0364; 0.0808]        1.3
## Rahmat et al., 2021              0.0000 [0.0000; 0.1507]        0.7
## Schmitt et al., 2013             0.3367 [0.2721; 0.4046]        1.3
## Shakya et al., 2017              0.1293 [0.0890; 0.1758]        1.3
## Sukvirach et al., 2003           0.0447 [0.0300; 0.0620]        1.3
## Sun et al., 2014                 0.4094 [0.3652; 0.4543]        1.3
## Tang et al., 2017                0.3341 [0.3166; 0.3518]        1.3
## Thomas et al., 2004              0.2491 [0.2012; 0.3004]        1.3
## Vu et al., 2013                  0.0857 [0.0687; 0.1043]        1.3
## Wang et al., 2018                0.1885 [0.1758; 0.2014]        1.3
## Wei et al., 2014                 0.0230 [0.0206; 0.0255]        1.3
## Wu et al., 2013                  0.1156 [0.0915; 0.1422]        1.3
## Xue et al., 2015                 0.1152 [0.0972; 0.1346]        1.3
## Yip et al., 2010                 0.1275 [0.0965; 0.1621]        1.3
## Zhao et al., 2009                0.0489 [0.0326; 0.0682]        1.3
## Zhu et al., 2021                 0.2349 [0.2051; 0.2660]        1.3
## Zoa Assoumou et al., 2016        0.6667 [0.5159; 0.8024]        1.1
## Giorgi Rossi et al., 2010        0.0465 [0.0299; 0.0665]        1.3
## Giuliano et al., 2005            0.0604 [0.0270; 0.1051]        1.2
## Giuliano et al., 2001            0.0441 [0.0056; 0.1088]        1.2
## Jin et al., 2010                 0.0905 [0.0574; 0.1301]        1.3
##
## Number of studies: k = 83
## Number of observations: o = 94086
## Number of events: e = 15654
##
##                       proportion           95%-CI
## Random effects model     0.1603 [0.1338; 0.1887]
##
## Quantifying heterogeneity:
##  tau^2 = 0.0255 [0.0187; 0.0369]; tau = 0.1596 [0.1369; 0.1922]
##  I^2 = 99.4% [99.3%; 99.4%]; H = 12.41 [11.99; 12.83]
##
## Test of heterogeneity:
```

```
##         Q d.f. p-value
##  12620.36   82       0
##
## Details on meta-analytical method:
## - Inverse variance method
## - Restricted maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 82)
## - Freeman-Tukey double arcsine transformation
## - Normal approximation confidence interval for individual studies
```

ANYTYPE FOREST PLOT —-

Create a forest plot based on the meta-analysis results

```
pdf("forestplot.pdf", width=8, height=18)

forest(any_type_ma,
       common = TRUE,
       print.tau2 = TRUE,
       print.Q = TRUE,
       print.pval.Q = TRUE,
       print.I2 = TRUE,
       rightcols = FALSE,
       pooled.totals = TRUE,
       weight.study = "random",
       leftcols = c("studlab", "num_hpv_pos", "num_older_wom", "effect", "ci"),
       leftlabs = c("Study", "HPV+ Cases", "Total", "Prevalence", "95% C.I."),
       xlab = "Pooled Prevalence Rate",
       smlab = "",
       xlim = c(0,1),
       pscale = 1,
       squaresize = 0.5,
       fs.hetstat = 10,
       digits = 2,
       col.square = "navy",
       col.square.lines = "navy",
       col.diamond = "maroon",
       col.diamond.lines = "maroon",
       sortvar = TE
       )
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
```

```
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>
```

```
dev.off()
```

```
## pdf
##   2
```

```
pdf("forestplot2.pdf", width=8, height=18)

forest(any_type_ma,
       common = TRUE,
       print.tau2 = TRUE,
       print.Q = TRUE,
       print.pval.Q = TRUE,
       print.I2 = TRUE,
       rightcols = FALSE,
       pooled.totals = TRUE,
       weight.study = "random",
       xlab = "Pooled Prevalence Rate",
       leftcols = c("studlab", "num_hpv_pos", "n", "effect", "ci"),
       leftlabs = c("Study", "HPV+ Cases", "Total", "Prevalence", "95% C.I."),
       smlab = "",
       xlim = c(0,1),
       pscale = 1,
       squaresize = 0.5,
       fs.hetstat = 10,
       digits = 2,
       sortvar = n,
       col.square = "navy",
       col.square.lines = "navy",
       col.diamond = "maroon",
       col.diamond.lines = "maroon"
)
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <c4>
```

```
## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Çolakoğlu et al., 2017' in 'mbcsToSbcs': dot substituted
## for <9f>
```

```
dev.off()
```

```
## pdf
##   2
```

Show Forest Plot

```
system2('open', args = 'forestplot2.pdf', wait = FALSE)
```

Explanation & Interpretation of Forest Plot

Heterogeneity: Higgins & Thompson's (https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/references.html#ref-higgins2002quantifying) Statistic I^2 = 99%, Tau^2 = 0.0255, Chi^2 82 = 12620.36 (p = 0)

In the meta-analysis you provided, the between-study heterogeneity is quantified using several statistical measures, which are crucial for interpreting the results:

- **I^2 Statistic**: Higgins & Thompson's I^2 statistic is 99%, which suggests that 99% of the variability in effect estimates is due to true differences in effect sizes across studies rather than chance. An I^2 value of this magnitude indicates substantial heterogeneity, necessitating careful consideration in the analysis and interpretation of results.
- **Tau² (tau²)**: The tau^2 value is 0.0255, indicating the estimated variance of underlying effect sizes across studies. A tau^2 greater than zero suggests that there is variation in effect sizes that the model attributes to heterogeneity rather than sampling error.
- **Chi-Square (Chi²)**: The chi^2 statistic, with a value of 12620.36 and 82 degrees of freedom, is highly significant (p = 0). This reinforces the presence of heterogeneity, as a large chi^2 relative to its degrees of freedom typically indicates that the variability in observed effects is greater than what would be expected by sampling error alone.

The choice of an inverse variance method in this analysis is justified as it weights each study by the inverse of its variance, giving more weight to studies with greater precision. The restricted maximum-likelihood estimator for tau^2 is used for its property of being unbiased in large samples, and the Q-Profile method provides confidence intervals for tau^2 and tau, considered more accurate for small numbers of studies.

The Hartung-Knapp adjustment for the random effects model is justified by the need to correct the tendency of the conventional random-effects model to under-estimate the variance of the pooled effect size when there is significant heterogeneity. This adjustment provides a more conservative estimate that accounts for the random-effects distribution.

The Freeman-Tukey double arcsine transformation was appropriately chosen to stabilize the variances of studies that report proportions near 0 or 1, ensuring a more accurate estimation of the pooled effect size. This is particularly relevant when dealing with proportions, as in prevalence studies.

The significant heterogeneity suggested by these statistics implies that the pooled prevalence rate of 16.03% for HPV among women aged 50 years and older with predominantly normal cytology might not reflect the true rate for any specific study or population. Subgroup analyses or meta-regression could be explored to understand the sources of heterogeneity, like differences in population characteristics, study designs, or methodologies. The clinical and public health implications should be interpreted with caution, considering

the substantial variation in study outcomes. This heterogeneity must be acknowledged when applying the findings to specific populations or settings.

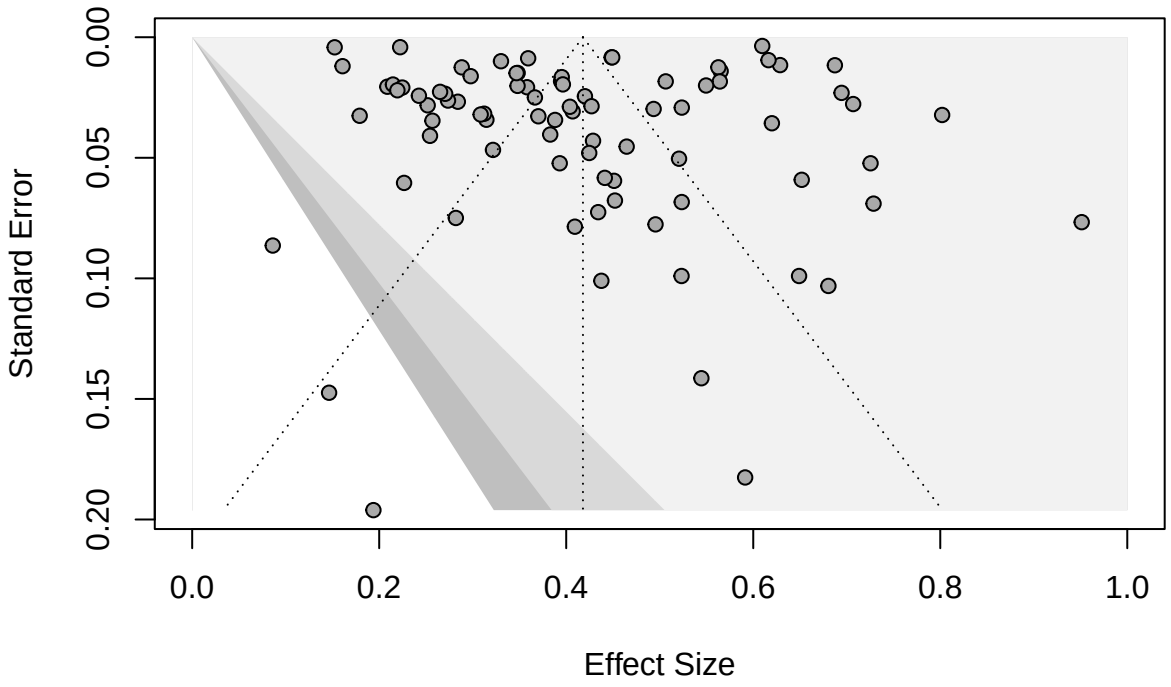
Addressing Risks of Bias from Publication Bias

Funnel Plot

```
library(meta)

# Define fill colors for contour
col.contour = c("gray75", "gray85", "gray95")

funnel(any_type_ma, xlim = c(0, 1),
       main = "Funnel Plot of HPV Prevalence Studies",
       xlab = "Effect Size",
       slab = paste(any_type_ma$author_year, any_type_ma$study_id),
       ci.lim = TRUE,
       shade = TRUE,
       contour = c(0.9, 0.95, 0.99),
       col.contour = col.contour
)
```



Egger's test for publication bias

```
# Egger's test for publication bias using meta
library(meta)
egger_test <- metabias(any_type_ma, method.bias = "Egger")
print(egger_test)
```

```
## Review:     Anytype HPV Prevalence in Women (50+) with Predominantly Normal  ...
##
```

```
## Linear regression test of funnel plot asymmetry
##
## Test result: t = 0.67, df = 81, p-value = 0.5042
## Bias estimate: 1.2805 (SE = 1.9088)
##
## Details:
## - multiplicative residual heterogeneity variance (tau^2 = 154.9461)
## - predictor: standard error
## - weight:    inverse variance
## - reference: Egger et al. (1997), BMJ
```

Description:

The test statistic (t-value) is 0.67, with 81 degrees of freedom, resulting in a p-value of 0.5042. The bias estimate is 1.2805, with a standard error (SE) of 1.9088. The Egger's test for publication bias yielded a non-significant result (p-value = 0.5042), with a bias estimate of 1.2805.

Explanation:

Egger's test assesses the symmetry of the funnel plot, which can be an indicator of publication bias. The test statistic and p-value indicate whether there is statistically significant evidence of asymmetry. A p-value greater than 0.05 suggests no significant evidence of publication bias. The bias estimate represents the extent of bias detected, with its standard error indicating the variability of this estimate. Egger's test assesses funnel plot asymmetry to detect publication bias. The non-significant p-value suggests that there is no strong evidence of bias in the studies included in the meta-analysis. The bias estimate quantifies the potential bias detected, but its non-significant nature implies that it may not be a reliable indicator of actual bias.

Interpretation:

In your meta-analysis, the high p-value (0.5042) suggests that there is no significant evidence of publication bias. This implies that the studies included in your meta-analysis are not overly skewed towards reporting only significant or positive results, and the overall results can be considered more reliable. The results indicate no substantial publication bias in your meta-analysis. This suggests a balanced representation of studies, including those with both significant and non-significant findings. It supports the reliability of your meta-analysis results, assuming that unpublished studies or those with negative results are not disproportionately missing.

Sensitivity Analysis for Meta-Analysis

Excluding Studies with Zero Events

```r
# Create a vector of study identifiers with zero events
studies_with_zero_events <- c("Becker et al., 1991", "Eren et al., 2010", "Rahmat et al., 2021")

# Filter out these studies
any_type_filtered_sensitivity <- any_type_filtered[!any_type_filtered$author_year %in% studies_with_zero

# Re-run the meta-analysis
any_type_ma_sens <- metaprop(event = num_hpv_pos,
                             n = num_older_wom,
                             studlab = author_year,
                             data = any_type_filtered_sensitivity,
                             sm = "PFT",
                             method.tau = "REML",
```

```
                            add = 0,
                            fixed = FALSE,
                            random = TRUE,
                            hakn = TRUE,
                            title = "Sensitivity Analysis - Excluding Zero Events")
summary(any_type_ma_sens)
```

```
## Review:     Sensitivity Analysis - Excluding Zero Events
##
##                               proportion           95%-CI %W(random)
## Tsedenbal et al., 2018            0.3600 [0.1797; 0.5748]        1.0
## Çolakoğlu et al., 2017            0.4423 [0.3047; 0.5867]        1.1
## Demirci et al., 2019              0.2234 [0.1762; 0.2766]        1.3
## Dutta et al, 2012                 0.0965 [0.0492; 0.1661]        1.2
## Herrero et al., 2000              0.1655 [0.1311; 0.2047]        1.3
## Jin et al., 2019                  0.4025 [0.3800; 0.4252]        1.3
## Li, P. et al., 2021               0.1880 [0.1754; 0.2012]        1.3
## Li, XF. et al., 2021              0.1235 [0.1124; 0.1353]        1.3
## Nuñez-Troconis et al., 2009       0.2400 [0.0936; 0.4513]        1.0
## Richter et al., 2013              0.4215 [0.3672; 0.4773]        1.3
## Souho et al., 2016                0.5167 [0.4515; 0.5814]        1.3
## Tezcan et al., 2014               0.2453 [0.1376; 0.3828]        1.1
## Donkoh et al., 2022               0.3662 [0.2550; 0.4890]        1.2
## Maehama et al., 2002              0.1050 [0.0933; 0.1176]        1.3
## Brotherton et al., 2015           0.1704 [0.1112; 0.2446]        1.3
## Ahmadi et al., 2020               0.1500 [0.0571; 0.2984]        1.1
## Andujar et al., 2020              0.0806 [0.0677; 0.0950]        1.3
## Ardhaoui et al., 2016             0.1702 [0.0765; 0.3081]        1.1
## Balanda et al., 2016              0.0774 [0.0516; 0.1106]        1.3
## Baloch et al., 2017               0.1667 [0.1019; 0.2506]        1.2
## Bansal et al., 2014               0.0718 [0.0475; 0.1035]        1.3
## Bell et al., 2007                 0.1667 [0.0474; 0.3738]        1.0
## Bi et al., 2015                   0.2866 [0.2618; 0.3124]        1.3
## Castellsague et al., 2012         0.0705 [0.0487; 0.0980]        1.3
## Castellsague et al., 2001         0.2195 [0.1056; 0.3761]        1.1
## Castle et al., 2006               0.1222 [0.0967; 0.1516]        1.3
## Cathro et al., 2009               0.0682 [0.0143; 0.1866]        1.1
## Centurioni et al., 2005           0.1553 [0.1138; 0.2047]        1.3
## Chan et al., 2002                 0.0298 [0.0121; 0.0604]        1.3
## Chansaeroj et al., 2010           0.0678 [0.0471; 0.0938]        1.3
## Chen et al., 2015                 0.1473 [0.1226; 0.1747]        1.3
## Chong et al., 2010                0.3913 [0.1971; 0.6146]        0.9
## Coser et al., 2013                0.2449 [0.1636; 0.3421]        1.2
## Dai et al., 2006                  0.1429 [0.0783; 0.2319]        1.2
## DeVuyst et al., 2003              0.2500 [0.0549; 0.5719]        0.7
## Debrah et al., 2021               0.2857 [0.0367; 0.7096]        0.6
## Demers et al., 2012               0.0939 [0.0583; 0.1413]        1.3
## Dufit et al., 2016                0.4396 [0.3356; 0.5475]        1.2
## Foliaki et al., 2014              0.1857 [0.1028; 0.2966]        1.2
## Gravitt et al., 2013              0.1533 [0.1145; 0.1992]        1.3
## Hamlin-Douglas et al., 2008       0.1983 [0.1314; 0.2806]        1.2
## Hernandez-Rosas et al., 2021      0.0607 [0.0369; 0.0932]        1.3
## Herrero et al., 2005              0.2845 [0.2625; 0.3072]        1.3
## Hong et al., 2015                 0.1415 [0.0976; 0.1958]        1.3
```

```
## Hooi et al., 2018                0.1705 [0.1300; 0.2175]        1.3
## Jiang et al., 2011               0.1479 [0.1255; 0.1726]        1.3
## Shen et al., 2021                0.1373 [0.0870; 0.2021]        1.3
## Xiao et al, 2016                 0.3458 [0.3242; 0.3679]        1.3
## Klug et al., 2007                0.0254 [0.0185; 0.0339]        1.3
## Kobetz et al., 2012              0.1852 [0.0925; 0.3143]        1.1
## Lee et al., 2012                 0.3278 [0.3211; 0.3346]        1.3
## Leinonen et al., 2013            0.0486 [0.0452; 0.0522]        1.3
## Levert et al., 2000              0.1485 [0.1221; 0.1782]        1.3
## Li et al., 2019                  0.2857 [0.2534; 0.3197]        1.3
## Li et al., 2011                  0.0625 [0.0337; 0.1045]        1.3
## Liu et al., 2014                 0.2724 [0.2378; 0.3092]        1.3
## López Rivera et al., 2012        0.0927 [0.0597; 0.1359]        1.3
## Moore et al., 2009               0.1160 [0.0980; 0.1360]        1.3
## Mudderis et al., 2019            0.1781 [0.0984; 0.2853]        1.2
## Anh et al., 2003                 0.0422 [0.0275; 0.0616]        1.3
## Pista et al., 2011               0.0566 [0.0366; 0.0831]        1.3
## Schmitt et al., 2013             0.3367 [0.2710; 0.4075]        1.3
## Shakya et al., 2017              0.1293 [0.0890; 0.1794]        1.3
## Sukvirach et al., 2003           0.0447 [0.0301; 0.0635]        1.3
## Sun et al., 2014                 0.4094 [0.3645; 0.4554]        1.3
## Tang et al., 2017                0.3341 [0.3165; 0.3520]        1.3
## Thomas et al., 2004              0.2491 [0.2007; 0.3028]        1.3
## Vu et al., 2013                  0.0857 [0.0687; 0.1052]        1.3
## Wang et al., 2018                0.1885 [0.1758; 0.2017]        1.3
## Wei et al., 2014                 0.0230 [0.0206; 0.0256]        1.3
## Wu et al., 2013                  0.1156 [0.0914; 0.1436]        1.3
## Xue et al., 2015                 0.1152 [0.0972; 0.1353]        1.3
## Yip et al., 2010                 0.1275 [0.0964; 0.1642]        1.3
## Zhao et al., 2009                0.0489 [0.0327; 0.0699]        1.3
## Zhu et al., 2021                 0.2349 [0.2049; 0.2670]        1.3
## Zoa Assoumou et al., 2016        0.6667 [0.5045; 0.8043]        1.1
## Giorgi Rossi et al., 2010        0.0465 [0.0300; 0.0684]        1.3
## Giuliano et al., 2005            0.0604 [0.0280; 0.1116]        1.3
## Giuliano et al., 2001            0.0441 [0.0092; 0.1236]        1.2
## Jin et al., 2010                 0.0905 [0.0576; 0.1339]        1.3
##
## Number of studies: k = 80
## Number of observations: o = 94036
## Number of events: e = 15654
##
##                       proportion           95%-CI
## Random effects model     0.1662 [0.1396; 0.1947]
##
## Quantifying heterogeneity:
##  tau^2 = 0.0247 [0.0181; 0.0355]; tau = 0.1573 [0.1344; 0.1885]
##  I^2 = 99.4% [99.3%; 99.4%]; H = 12.63 [12.21; 13.07]
##
## Test of heterogeneity:
##         Q d.f. p-value
##  12604.40   79       0
##
## Details on meta-analytical method:
## - Inverse variance method
```

```
## - Restricted maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 79)
## - Freeman-Tukey double arcsine transformation
## - Clopper-Pearson confidence interval for individual studies
```

Results:

This analysis included 80 studies with a total of 94,036 observations and 15,654 events. The summary of the meta-analysis using a random effects model provided the following results: Estimated HPV prevalence (proportion) is 16.62% with a 95% confidence interval (CI) of [13.96%; 19.47%]. High heterogeneity is indicated by an I^2 statistic of 99.4% and a tau^2 value of 0.0247. The Q test for heterogeneity was significant ($Q = 12604.40$, $df = 79$, $p < 0.001$).

Explanation:

The proportion of 16.62% indicates that the estimated prevalence of HPV in the specified population, after excluding zero-event studies, is about 16.62%. The high I^2 value suggests that there is substantial variability in the effect sizes across the included studies, which is not attributable to chance. This indicates that factors other than sampling error contribute to the observed differences. The tau^2 value provides an estimate of the between-study variance in true effect sizes. The Q test's significance indicates that the observed variability in effect sizes is more than what would be expected by chance, reinforcing the presence of heterogeneity.

Interpretation:

The estimated prevalence of HPV in the target population is significant, suggesting a considerable public health concern. The high level of heterogeneity suggests that individual study characteristics or other contextual factors might be influencing the HPV prevalence rates. It implies the need for further exploration into potential sources of this heterogeneity, such as differences in study design, population demographics, or HPV testing methods. The results underscore the importance of careful interpretation of meta-analysis findings, especially in the presence of high heterogeneity.

Excluding Studies with Small Sample Sizes

```r
# Define threshold for small studies
threshold <- 100

# Filter out small studies
filtered_data <- any_type_filtered[any_type_filtered$num_older_wom >= threshold, ]

# Re-run meta-analysis
small_studies_ma <- metaprop(event = num_hpv_pos,
                             n = num_older_wom,
                             studlab = author_year,
                             data = filtered_data,
                             sm = "PFT",
                             method.tau = "REML",
                             add = 0,
                             fixed = FALSE,
                             random = TRUE,
                             hakn = TRUE,
                             title = "Sensitivity Analysis - Excluding Small Studies")
summary(small_studies_ma)
```

```
## Review:     Sensitivity Analysis - Excluding Small Studies
##
##                               proportion           95%-CI %W(random)
## Demirci et al., 2019              0.2234 [0.1762; 0.2766]        1.7
## Dutta et al, 2012                 0.0965 [0.0492; 0.1661]        1.6
## Herrero et al., 2000              0.1655 [0.1311; 0.2047]        1.7
## Jin et al., 2019                  0.4025 [0.3800; 0.4252]        1.7
## Li, P. et al., 2021               0.1880 [0.1754; 0.2012]        1.7
## Li, XF. et al., 2021              0.1235 [0.1124; 0.1353]        1.7
## Richter et al., 2013              0.4215 [0.3672; 0.4773]        1.7
## Souho et al., 2016                0.5167 [0.4515; 0.5814]        1.7
## Maehama et al., 2002              0.1050 [0.0933; 0.1176]        1.7
## Brotherton et al., 2015           0.1704 [0.1112; 0.2446]        1.6
## Andujar et al., 2020              0.0806 [0.0677; 0.0950]        1.7
## Balanda et al., 2016              0.0774 [0.0516; 0.1106]        1.7
## Baloch et al., 2017               0.1667 [0.1019; 0.2506]        1.6
## Bansal et al., 2014               0.0718 [0.0475; 0.1035]        1.7
## Bi et al., 2015                   0.2866 [0.2618; 0.3124]        1.7
## Castellsague et al., 2012         0.0705 [0.0487; 0.0980]        1.7
## Castle et al., 2006               0.1222 [0.0967; 0.1516]        1.7
## Centurioni et al., 2005           0.1553 [0.1138; 0.2047]        1.7
## Chan et al., 2002                 0.0298 [0.0121; 0.0604]        1.7
## Chansaeroj et al., 2010           0.0678 [0.0471; 0.0938]        1.7
## Chen et al., 2015                 0.1473 [0.1226; 0.1747]        1.7
## Demers et al., 2012               0.0939 [0.0583; 0.1413]        1.7
## Gravitt et al., 2013              0.1533 [0.1145; 0.1992]        1.7
## Hamlin-Douglas et al., 2008       0.1983 [0.1314; 0.2806]        1.6
## Hernandez-Rosas et al., 2021      0.0607 [0.0369; 0.0932]        1.7
## Herrero et al., 2005              0.2845 [0.2625; 0.3072]        1.7
## Hong et al., 2015                 0.1415 [0.0976; 0.1958]        1.7
## Hooi et al., 2018                 0.1705 [0.1300; 0.2175]        1.7
## Jiang et al., 2011                0.1479 [0.1255; 0.1726]        1.7
## Shen et al., 2021                 0.1373 [0.0870; 0.2021]        1.6
## Xiao et al, 2016                  0.3458 [0.3242; 0.3679]        1.7
## Klug et al., 2007                 0.0254 [0.0185; 0.0339]        1.7
## Lee et al., 2012                  0.3278 [0.3211; 0.3346]        1.7
## Leinonen et al., 2013             0.0486 [0.0452; 0.0522]        1.7
## Levert et al., 2000               0.1485 [0.1221; 0.1782]        1.7
## Li et al., 2019                   0.2857 [0.2534; 0.3197]        1.7
## Li et al., 2011                   0.0625 [0.0337; 0.1045]        1.7
## Liu et al., 2014                  0.2724 [0.2378; 0.3092]        1.7
## López Rivera et al., 2012         0.0927 [0.0597; 0.1359]        1.7
## Moore et al., 2009                0.1160 [0.0980; 0.1360]        1.7
## Anh et al., 2003                  0.0422 [0.0275; 0.0616]        1.7
## Pista et al., 2011                0.0566 [0.0366; 0.0831]        1.7
## Schmitt et al., 2013              0.3367 [0.2710; 0.4075]        1.7
## Shakya et al., 2017               0.1293 [0.0890; 0.1794]        1.7
## Sukvirach et al., 2003            0.0447 [0.0301; 0.0635]        1.7
## Sun et al., 2014                  0.4094 [0.3645; 0.4554]        1.7
## Tang et al., 2017                 0.3341 [0.3165; 0.3520]        1.7
## Thomas et al., 2004               0.2491 [0.2007; 0.3028]        1.7
## Vu et al., 2013                   0.0857 [0.0687; 0.1052]        1.7
## Wang et al., 2018                 0.1885 [0.1758; 0.2017]        1.7
## Wei et al., 2014                  0.0230 [0.0206; 0.0256]        1.7
```

```
## Wu et al., 2013                 0.1156 [0.0914; 0.1436]        1.7
## Xue et al., 2015                0.1152 [0.0972; 0.1353]        1.7
## Yip et al., 2010                0.1275 [0.0964; 0.1642]        1.7
## Zhao et al., 2009               0.0489 [0.0327; 0.0699]        1.7
## Zhu et al., 2021                0.2349 [0.2049; 0.2670]        1.7
## Giorgi Rossi et al., 2010       0.0465 [0.0300; 0.0684]        1.7
## Giuliano et al., 2005           0.0604 [0.0280; 0.1116]        1.6
## Jin et al., 2010                0.0905 [0.0576; 0.1339]        1.7
##
## Number of studies: k = 59
## Number of observations: o = 92985
## Number of events: e = 15389
##
##                       proportion           95%-CI
## Random effects model     0.1473 [0.1203; 0.1765]
##
## Quantifying heterogeneity:
##  tau^2 = 0.0224 [0.0158; 0.0335]; tau = 0.1496 [0.1256; 0.1830]
##  I^2 = 99.5% [99.5%; 99.6%]; H = 14.63 [14.12; 15.17]
##
## Test of heterogeneity:
##         Q d.f. p-value
##  12419.83   58       0
##
## Details on meta-analytical method:
## - Inverse variance method
## - Restricted maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 58)
## - Freeman-Tukey double arcsine transformation
## - Clopper-Pearson confidence interval for individual studies
```

Results:

The sensitivity analysis conducted by excluding studies with small sample sizes (less than 100 participants) from the meta-analysis of HPV prevalence among women aged 50 and older with predominantly normal cytology yielded the following: The analysis included 59 studies with 92,985 observations and 15,389 events. The estimated HPV prevalence (proportion) is 14.73% with a 95% confidence interval (CI) of [12.03%; 17.65%]. The heterogeneity remains high, indicated by an I^2 statistic of 99.5% and a tau^2 value of 0.0224. The Q test for heterogeneity is significant ($Q = 12419.83$, $df = 58$, $p < 0.001$).

Explanation:

Proportion of 14.73%: This reduced prevalence (compared to the initial analysis including smaller studies) suggests that the overall HPV prevalence estimate becomes slightly lower when focusing on larger studies. This might reflect more stable prevalence estimates in larger samples. High I^2 value: Despite excluding smaller studies, the heterogeneity remains very high, indicating substantial differences in prevalence rates across studies. This suggests that factors other than sample size are contributing to this variability. Tau^2 value: This provides an estimate of between-study variance in true effect sizes, which remains significant, further confirming the presence of heterogeneity. Q test's significance: This reinforces the presence of heterogeneity in the meta-analysis, as the variability in effect sizes is greater than what would be expected by chance.

Interpretation:

The slightly lower prevalence rate after excluding small studies might indicate that smaller studies could be showing higher variability or bias in prevalence estimates. The persistent high heterogeneity after excluding small studies suggests the need for further exploration into other sources of heterogeneity, such as differences in study methodologies, population demographics, or regional factors. The findings imply that while sample size is a factor, it is not the sole contributor to the variability in HPV prevalence rates observed across studies.

Subgroup Analysis by Geographic Region

```
subgroup_analysis <- metaprop(event = num_hpv_pos,
                              n = num_older_wom,
                              studlab = author_year,
                              data = any_type_filtered,
                              sm = "PFT",
                              method.tau = "REML",
                              byvar = world_subregion,
                              title = "Subgroup Analysis - By Geographic Region")
summary(subgroup_analysis)
```

```
## Review:     Subgroup Analysis - By Geographic Region
##
##                               proportion           95%-CI %W(common) %W(random)
## Tsedenbal et al., 2018            0.3600 [0.1797; 0.5748]        0.0        1.0
## Çolakoğlu et al., 2017            0.4423 [0.3047; 0.5867]        0.1        1.1
## Demirci et al., 2019              0.2234 [0.1762; 0.2766]        0.3        1.3
## Dutta et al, 2012                 0.0965 [0.0492; 0.1661]        0.1        1.2
## Herrero et al., 2000              0.1655 [0.1311; 0.2047]        0.4        1.3
## Jin et al., 2019                  0.4025 [0.3800; 0.4252]        2.0        1.3
## Li, P. et al., 2021               0.1880 [0.1754; 0.2012]        3.8        1.3
## Li, XF. et al., 2021              0.1235 [0.1124; 0.1353]        3.5        1.3
## Nuñez-Troconis et al., 2009       0.2400 [0.0936; 0.4513]        0.0        1.0
## Richter et al., 2013              0.4215 [0.3672; 0.4773]        0.3        1.3
## Souho et al., 2016                0.5167 [0.4515; 0.5814]        0.3        1.3
## Tezcan et al., 2014               0.2453 [0.1376; 0.3828]        0.1        1.1
## Donkoh et al., 2022               0.3662 [0.2550; 0.4890]        0.1        1.2
## Maehama et al., 2002              0.1050 [0.0933; 0.1176]        2.7        1.3
## Brotherton et al., 2015           0.1704 [0.1112; 0.2446]        0.1        1.2
## Ahmadi et al., 2020               0.1500 [0.0571; 0.2984]        0.0        1.1
## Andujar et al., 2020              0.0806 [0.0677; 0.0950]        1.7        1.3
## Ardhaoui et al., 2016             0.1702 [0.0765; 0.3081]        0.1        1.1
## Balanda et al., 2016              0.0774 [0.0516; 0.1106]        0.4        1.3
## Baloch et al., 2017               0.1667 [0.1019; 0.2506]        0.1        1.2
## Bansal et al., 2014               0.0718 [0.0475; 0.1035]        0.4        1.3
## Becker et al., 1991               0.0000 [0.0000; 0.1058]        0.0        1.0
## Bell et al., 2007                 0.1667 [0.0474; 0.3738]        0.0        0.9
## Bi et al., 2015                   0.2866 [0.2618; 0.3124]        1.3        1.3
## Castellsague et al., 2012         0.0705 [0.0487; 0.0980]        0.5        1.3
## Castellsague et al., 2001         0.2195 [0.1056; 0.3761]        0.0        1.1
## Castle et al., 2006               0.1222 [0.0967; 0.1516]        0.6        1.3
## Cathro et al., 2009               0.0682 [0.0143; 0.1866]        0.0        1.1
## Centurioni et al., 2005           0.1553 [0.1138; 0.2047]        0.3        1.3
## Chan et al., 2002                 0.0298 [0.0121; 0.0604]        0.3        1.3
## Chansaeroj et al., 2010           0.0678 [0.0471; 0.0938]        0.5        1.3
```

```
## Chen et al., 2015                0.1473 [0.1226; 0.1747]        0.8        1.3
## Chong et al., 2010               0.3913 [0.1971; 0.6146]        0.0        0.9
## Coser et al., 2013               0.2449 [0.1636; 0.3421]        0.1        1.2
## Dai et al., 2006                 0.1429 [0.0783; 0.2319]        0.1        1.2
## DeVuyst et al., 2003             0.2500 [0.0549; 0.5719]        0.0        0.7
## Debrah et al., 2021              0.2857 [0.0367; 0.7096]        0.0        0.6
## Demers et al., 2012              0.0939 [0.0583; 0.1413]        0.2        1.3
## Dufit et al., 2016               0.4396 [0.3356; 0.5475]        0.1        1.2
## Eren et al., 2010                0.0000 [0.0000; 0.4593]        0.0        0.5
## Foliaki et al., 2014             0.1857 [0.1028; 0.2966]        0.1        1.2
## Gravitt et al., 2013             0.1533 [0.1145; 0.1992]        0.3        1.3
## Hamlin-Douglas et al., 2008      0.1983 [0.1314; 0.2806]        0.1        1.2
## Hernandez-Rosas et al., 2021     0.0607 [0.0369; 0.0932]        0.3        1.3
## Herrero et al., 2005             0.2845 [0.2625; 0.3072]        1.7        1.3
## Hong et al., 2015                0.1415 [0.0976; 0.1958]        0.2        1.3
## Hooi et al., 2018                0.1705 [0.1300; 0.2175]        0.3        1.3
## Jiang et al., 2011               0.1479 [0.1255; 0.1726]        1.0        1.3
## Shen et al., 2021                0.1373 [0.0870; 0.2021]        0.2        1.2
## Xiao et al, 2016                 0.3458 [0.3242; 0.3679]        2.0        1.3
## Klug et al., 2007                0.0254 [0.0185; 0.0339]        1.8        1.3
## Kobetz et al., 2012              0.1852 [0.0925; 0.3143]        0.1        1.1
## Lee et al., 2012                 0.3278 [0.3211; 0.3346]       20.1        1.3
## Leinonen et al., 2013            0.0486 [0.0452; 0.0522]       15.6        1.3
## Levert et al., 2000              0.1485 [0.1221; 0.1782]        0.7        1.3
## Li et al., 2019                  0.2857 [0.2534; 0.3197]        0.8        1.3
## Li et al., 2011                  0.0625 [0.0337; 0.1045]        0.2        1.3
## Liu et al., 2014                 0.2724 [0.2378; 0.3092]        0.7        1.3
## López Rivera et al., 2012        0.0927 [0.0597; 0.1359]        0.3        1.3
## Moore et al., 2009               0.1160 [0.0980; 0.1360]        1.2        1.3
## Mudderis et al., 2019            0.1781 [0.0984; 0.2853]        0.1        1.2
## Anh et al., 2003                 0.0422 [0.0275; 0.0616]        0.6        1.3
## Pista et al., 2011               0.0566 [0.0366; 0.0831]        0.5        1.3
## Rahmat et al., 2021              0.0000 [0.0000; 0.2849]        0.0        0.7
## Schmitt et al., 2013             0.3367 [0.2710; 0.4075]        0.2        1.3
## Shakya et al., 2017              0.1293 [0.0890; 0.1794]        0.2        1.3
## Sukvirach et al., 2003           0.0447 [0.0301; 0.0635]        0.7        1.3
## Sun et al., 2014                 0.4094 [0.3645; 0.4554]        0.5        1.3
## Tang et al., 2017                0.3341 [0.3165; 0.3520]        2.9        1.3
## Thomas et al., 2004              0.2491 [0.2007; 0.3028]        0.3        1.3
## Vu et al., 2013                  0.0857 [0.0687; 0.1052]        1.0        1.3
## Wang et al., 2018                0.1885 [0.1758; 0.2017]        3.8        1.3
## Wei et al., 2014                 0.0230 [0.0206; 0.0256]       15.2        1.3
## Wu et al., 2013                  0.1156 [0.0914; 0.1436]        0.7        1.3
## Xue et al., 2015                 0.1152 [0.0972; 0.1353]        1.2        1.3
## Yip et al., 2010                 0.1275 [0.0964; 0.1642]        0.4        1.3
## Zhao et al., 2009                0.0489 [0.0327; 0.0699]        0.6        1.3
## Zhu et al., 2021                 0.2349 [0.2049; 0.2670]        0.8        1.3
## Zoa Assoumou et al., 2016        0.6667 [0.5045; 0.8043]        0.0        1.1
## Giorgi Rossi et al., 2010        0.0465 [0.0300; 0.0684]        0.5        1.3
## Giuliano et al., 2005            0.0604 [0.0280; 0.1116]        0.2        1.2
## Giuliano et al., 2001            0.0441 [0.0092; 0.1236]        0.1        1.2
## Jin et al., 2010                 0.0905 [0.0576; 0.1339]        0.3        1.3
##                                        world_subregion
## Tsedenbal et al., 2018                    Central Asia
```

```
## Çolakoğlu et al., 2017                    Western Asia
## Demirci et al., 2019                      Western Asia
## Dutta et al, 2012                        Southern Asia
## Herrero et al., 2000                    Central America
## Jin et al., 2019                           Eastern Asia
## Li, P. et al., 2021                        Eastern Asia
## Li, XF. et al., 2021                       Eastern Asia
## Nuñez-Troconis et al., 2009              South America
## Richter et al., 2013                    Southern Africa
## Souho et al., 2016                      Northern Africa
## Tezcan et al., 2014                        Western Asia
## Donkoh et al., 2022                      Western Africa
## Maehama et al., 2002                       Eastern Asia
## Brotherton et al., 2015      Australia and New Zealand
## Ahmadi et al., 2020                      Southern Asia
## Andujar et al., 2020                    Southern Europe
## Ardhaoui et al., 2016                   Northern Africa
## Balanda et al., 2016                     South America
## Baloch et al., 2017                        Eastern Asia
## Bansal et al., 2014                        Western Asia
## Becker et al., 1991                       North America
## Bell et al., 2007                       Northern Africa
## Bi et al., 2015                            Eastern Asia
## Castellsague et al., 2012                Western Europe
## Castellsague et al., 2001                Eastern Africa
## Castle et al., 2006                       North America
## Cathro et al., 2009                     Central America
## Centurioni et al., 2005                 Southern Europe
## Chan et al., 2002                          Eastern Asia
## Chansaeroj et al., 2010               Southeastern asia
## Chen et al., 2015                          Eastern Asia
## Chong et al., 2010                    Southeastern asia
## Coser et al., 2013                        South America
## Dai et al., 2006                           Eastern Asia
## DeVuyst et al., 2003                     Eastern Africa
## Debrah et al., 2021                      Western Africa
## Demers et al., 2012                       North America
## Dufit et al., 2016                        South America
## Eren et al., 2010                          Western Asia
## Foliaki et al., 2014                          Melanesia
## Gravitt et al., 2013                      North America
## Hamlin-Douglas et al., 2008               North America
## Hernandez-Rosas et al., 2021            Central America
## Herrero et al., 2005                    Central America
## Hong et al., 2015                          Eastern Asia
## Hooi et al., 2018                             Caribbean
## Jiang et al., 2011                        North America
## Shen et al., 2021                          Eastern Asia
## Xiao et al, 2016                           Eastern Asia
## Klug et al., 2007                        Western Europe
## Kobetz et al., 2012                       North America
## Lee et al., 2012                           Eastern Asia
## Leinonen et al., 2013                   Northern Europe
## Levert et al., 2000                      Western Europe
```

```
## Li et al., 2019                            Eastern Asia
## Li et al., 2011                            Eastern Asia
## Liu et al., 2014                           Eastern Asia
## López Rivera et al., 2012                Central America
## Moore et al., 2009                         North America
## Mudderis et al., 2019                      Western Asia
## Anh et al., 2003                        Southeastern asia
## Pista et al., 2011                        Western Europe
## Rahmat et al., 2021                     Southeastern asia
## Schmitt et al., 2013                      Western Europe
## Shakya et al., 2017                        Southern Asia
## Sukvirach et al., 2003                  Southeastern asia
## Sun et al., 2014                           Eastern Asia
## Tang et al., 2017                          Eastern Asia
## Thomas et al., 2004                       Western Africa
## Vu et al., 2013                         Southeastern asia
## Wang et al., 2018                          Eastern Asia
## Wei et al., 2014                           Eastern Asia
## Wu et al., 2013                            Eastern Asia
## Xue et al., 2015                           Eastern Asia
## Yip et al., 2010                           Eastern Asia
## Zhao et al., 2009                          Eastern Asia
## Zhu et al., 2021                           Eastern Asia
## Zoa Assoumou et al., 2016                 Middle Africa
## Giorgi Rossi et al., 2010               Southern Europe
## Giuliano et al., 2005                     North America
## Giuliano et al., 2001                     North America
## Jin et al., 2010                           Eastern Asia
##
## Number of studies: k = 83
## Number of observations: o = 94086
## Number of events: e = 15654
##
##                       proportion           95%-CI
## Common effect model       0.1390 [0.1367; 0.1413]
## Random effects model      0.1603 [0.1342; 0.1881]
##
## Quantifying heterogeneity:
##  tau^2 = 0.0255 [0.0187; 0.0369]; tau = 0.1596 [0.1369; 0.1922]
##  I^2 = 99.4% [99.3%; 99.4%]; H = 12.41 [11.99; 12.83]
##
## Test of heterogeneity:
##         Q d.f. p-value
##  12620.36   82       0
##
## Results for subgroups (common effect model):
##                                              k proportion           95%-CI
## world_subregion = Central Asia               1     0.3600 [0.1812; 0.5598]
## world_subregion = Western Asia               6     0.1446 [0.1198; 0.1710]
## world_subregion = Southern Asia              3     0.1194 [0.0881; 0.1544]
## world_subregion = Central America            5     0.2068 [0.1914; 0.2227]
## world_subregion = Eastern Asia              26     0.1794 [0.1763; 0.1824]
## world_subregion = South America              4     0.1517 [0.1225; 0.1834]
## world_subregion = Southern Africa            1     0.4215 [0.3683; 0.4757]
```

```
## world_subregion = Northern Africa             3     0.4308 [0.3752; 0.4873]
## world_subregion = Western Africa              3     0.2611 [0.2149; 0.3098]
## world_subregion = Australia and New Zealand   1     0.1704 [0.1112; 0.2389]
## world_subregion = Southern Europe             3     0.0790 [0.0684; 0.0902]
## world_subregion = North America              10     0.1216 [0.1109; 0.1327]
## world_subregion = Western Europe              5     0.0629 [0.0550; 0.0713]
## world_subregion = Eastern Africa              2     0.2219 [0.1145; 0.3494]
## world_subregion = Southeastern asia           6     0.0541 [0.0451; 0.0637]
## world_subregion = Melanesia                   1     0.1857 [0.1022; 0.2863]
## world_subregion = Caribbean                   1     0.1705 [0.1302; 0.2149]
## world_subregion = Northern Europe             1     0.0486 [0.0452; 0.0522]
## world_subregion = Middle Africa               1     0.6667 [0.5159; 0.8024]
##                                                Q   I^2
## world_subregion = Central Asia              0.00    --
## world_subregion = Western Asia             60.42 91.7%
## world_subregion = Southern Asia             1.16  0.0%
## world_subregion = Central America         149.81 97.3%
## world_subregion = Eastern Asia           8766.76 99.7%
## world_subregion = South America            64.70 95.4%
## world_subregion = Southern Africa           0.00    --
## world_subregion = Northern Africa          29.80 93.3%
## world_subregion = Western Africa            3.84 47.9%
## world_subregion = Australia and New Zealand 0.00    --
## world_subregion = Southern Europe          24.64 91.9%
## world_subregion = North America            38.02 76.3%
## world_subregion = Western Europe          219.27 98.2%
## world_subregion = Eastern Africa            0.09  0.0%
## world_subregion = Southeastern asia        34.09 85.3%
## world_subregion = Melanesia                 0.00    --
## world_subregion = Caribbean                 0.00    --
## world_subregion = Northern Europe           0.00    --
## world_subregion = Middle Africa             0.00    --
##
## Test for subgroup differences (common effect model):
##                       Q d.f. p-value
## Between groups 3227.75   18       0
## Within groups  9392.61   64       0
##
## Results for subgroups (random effects model):
##                                              k proportion           95%-CI
## world_subregion = Central Asia               1     0.3600 [0.1812; 0.5598]
## world_subregion = Western Asia               6     0.1876 [0.0867; 0.3127]
## world_subregion = Southern Asia              3     0.1194 [0.0881; 0.1544]
## world_subregion = Central America            5     0.1294 [0.0612; 0.2174]
## world_subregion = Eastern Asia              26     0.1737 [0.1317; 0.2202]
## world_subregion = South America              4     0.2337 [0.0939; 0.4105]
## world_subregion = Southern Africa            1     0.4215 [0.3683; 0.4757]
## world_subregion = Northern Africa            3     0.2859 [0.0899; 0.5345]
## world_subregion = Western Africa             3     0.2864 [0.1898; 0.3928]
## world_subregion = Australia and New Zealand  1     0.1704 [0.1112; 0.2389]
## world_subregion = Southern Europe            3     0.0881 [0.0379; 0.1563]
## world_subregion = North America             10     0.1093 [0.0769; 0.1463]
## world_subregion = Western Europe             5     0.1080 [0.0314; 0.2222]
## world_subregion = Eastern Africa             2     0.2219 [0.1145; 0.3494]
```

```
## world_subregion = Southeastern asia           6     0.0740 [0.0190; 0.1549]
## world_subregion = Melanesia                   1     0.1857 [0.1022; 0.2863]
## world_subregion = Caribbean                   1     0.1705 [0.1302; 0.2149]
## world_subregion = Northern Europe             1     0.0486 [0.0452; 0.0522]
## world_subregion = Middle Africa               1     0.6667 [0.5159; 0.8024]
##                                            tau^2    tau
## world_subregion = Central Asia                --     --
## world_subregion = Western Asia            0.0244 0.1563
## world_subregion = Southern Asia                0      0
## world_subregion = Central America         0.0160 0.1264
## world_subregion = Eastern Asia            0.0224 0.1498
## world_subregion = South America           0.0329 0.1814
## world_subregion = Southern Africa             --     --
## world_subregion = Northern Africa         0.0436 0.2087
## world_subregion = Western Africa          0.0045 0.0670
## world_subregion = Australia and New Zealand     --     --
## world_subregion = Southern Europe         0.0080 0.0895
## world_subregion = North America           0.0060 0.0772
## world_subregion = Western Europe          0.0308 0.1755
## world_subregion = Eastern Africa               0      0
## world_subregion = Southeastern asia       0.0196 0.1399
## world_subregion = Melanesia                   --     --
## world_subregion = Caribbean                   --     --
## world_subregion = Northern Europe             --     --
## world_subregion = Middle Africa               --     --
##
## Test for subgroup differences (random effects model):
##                      Q d.f.  p-value
## Between groups 628.20   18 < 0.0001
##
## Details on meta-analytical method:
## - Inverse variance method
## - Restricted maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Freeman-Tukey double arcsine transformation
## - Clopper-Pearson confidence interval for individual studies
```

Results:

Overall Findings: The subgroup analysis by world subregion included 83 studies, with 94,086 observations and 15,654 events. The random-effects model estimated an HPV prevalence of 16.03% with a 95% CI of [13.42%; 18.81%]. Heterogeneity: High heterogeneity is indicated with an I^2 of 99.4% and tau^2 of 0.0255, suggesting substantial variation in prevalence across studies. Subgroup Results: The prevalence varied significantly across regions, with the highest prevalence in Middle Africa (66.67%) and the lowest in Southeastern Asia (5.41% in the common effect model and 7.40% in the random effects model). Subgroup Differences: Significant differences were found between groups ($Q = 628.20$, $p < 0.0001$), indicating substantial variation in HPV prevalence across different world subregions.

Explanation:

Variation in Prevalence: The wide variation in HPV prevalence among different subregions could be due to differences in demographic factors, HPV screening practices, and prevalence of high-risk HPV types. High Heterogeneity: The high heterogeneity within and between subgroups suggests that prevalence rates are influenced by diverse factors, including study methodologies and population characteristics. Subgroup Variability: The substantial variability between subregions highlights the importance of considering geographical

differences when interpreting HPV prevalence data.

Interpretation of Results:

Global Variability: These findings underscore the global variability in HPV prevalence among women aged 50 and older with predominantly normal cytology, emphasizing the need for region-specific public health strategies. Public Health Implications: The significant differences in prevalence rates across regions suggest that public health interventions and HPV screening programs need to be tailored to specific regional contexts.

Additional Analyses:

a. Genotype-Specific Analysis: Where data on specific HPV genotypes are available, conduct separate analyses for high-risk types (e.g., HPV-16, HPV-18) and other types.

b. Age-Subgroup Analysis: Explore HPV prevalence in different age brackets within the over-50 demographic.

c. Meta-Regression: If data permits, explore factors (e.g., study year, testing methods) that might explain heterogeneity.

If there are X numbers of studies with different prevalence rates 3.2

Abstract