ph1861_hw2_ygu5

Yue Gu

2023-11-21

Meta-Analysis Problem #1 (50 points) The following paper reviewed and meta-analyzed the proportion of Campylobacter cases that develop chronic sequelae, based on studies published prior to July 2011. (Keithlin 2014, "Systematic review and meta-analysis of the proportion of Campylobacter cases that develop chronic sequelae"). Import the number of cases that developed reactive arthritis (ReA) from page 7 of the article and follow the steps below to conduct a meta-analysis of the proportion/prevalence.

Please input data in the exact order from the paper. Calculate the proportion of people with Campylobacter who developed a ReA for each study. Add a count of 0.5 both to those reporting ReA outcome of 0% and to those totals. Calculate the logit of the outcome and logit of the standard error using the equation from Lipsey & Wilson. [Eq 3.5, p40].

```
# read data
keith_data = read.csv("./data/Keithlin2014_ReA.csv") %>%
  janitor::clean_names() %>%
  mutate(prop_seq = round(number_developing_sequelae/number_of_people_with_campylobacter,4)) %>%
  mutate(number_of_people_with_campylobacter_add = ifelse(prop_seq==0.0000, number_of_people_with_campyl
         number_developing_sequelae_add = ifelse(prop_seq==0.0000, number_developing_sequelae+0.5, numbe
  mutate(prop_seq_add = number_developing_sequelae_add/number_of_people_with_campylobacter_add,
         logit_outcome = log(prop_seq/(1-prop_seq)),
         se_outcome = sqrt(1/(number_of_people_with_campylobacter_add*prop_seq_add)+1/(number_of_people_

# check total number of cases
s1=sum(keith_data$number_developing_sequelae_add)
# check total number
s2=sum(keith_data$number_of_people_with_campylobacter)
# check the proportion developing ReA
p1 = s1/s2
# the logit outcome for the K=25 studies
knitr::kable(keith_data %>% select(first_author_year_reference_number, logit_outcome))
```

first_author__year_reference_number	logit_outcome
Ternhag, [52]	-8.111428
Schoenberg-Norio, jejuni [51]	-3.183275
Townes, [47]	-4.269191
Kosunen, [39]	-3.731341
Petersen, [27]	-Inf
Short, [25]	-Inf
Hannu, [43]	-2.528269
Hannu, [43]	-2.371141
Hannu, [43]	-2.599044
Hannu, [43]	-1.891268

first_author_year_reference_number	logit_outcome
Hannu, [43]	-Inf
Melby, [32]	-5.060886
Pitkanen, [23]	-2.871116
Locht, [37]	-1.687537
Schiellerup, [41]	-1.895664
Hannu, [54]	-3.635228
Pitkanen, [29]	-2.989555
Helms, [38]	-6.724233
Ponka, [40]	-3.832326
Doorduyn, [44]	-3.029746
Eastmond, [35]	-4.462676
Eastmond [35]	-Inf
Melby, [55]	-3.623315
Gumpel, [49]	-1.139566
Bremell, [33]	-4.171143

```r
# the standard error for the K=25 studies
knitr::kable(keith_data %>% select(first_author_year_reference_number, se_outcome))
```

first_author_year_reference_number	se_outcome
Ternhag, [52]	0.2582326
Schoenberg-Norio, jejuni [51]	0.3608065
Townes, [47]	0.1752951
Kosunen, [39]	0.3577625
Petersen, [27]	1.4228107
Short, [25]	1.4375906
Hannu, [43]	0.1549041
Hannu, [43]	0.1450038
Hannu, [43]	0.1703968
Hannu, [43]	0.3792993
Hannu, [43]	1.4411534
Melby, [32]	1.0031596
Pitkanen, [23]	0.5934655
Locht, [37]	0.2094907
Schiellerup, [41]	0.0937037
Hannu, [54]	0.3377035
Pitkanen, [29]	0.3416105
Helms, [38]	0.2133312
Ponka, [40]	0.4126461
Doorduyn, [44]	0.2289442
Eastmond, [35]	1.0057307
Eastmond [35]	1.4226066
Melby, [55]	0.7164728
Gumpel, [49]	0.4062019
Bremell, [33]	1.0076629

Now that the data are set, answer following questions with corresponding STATA or R code: (a) List the number of cases, total number, the proportion developing ReA, the logit outcome, and the logit standard error for the K=25 studies.

The number of cases for the K=25 studies is 448, total number is 83178, the proportion developing ReA is

0.005386, the logit outcome and standard error for each studies could be found above.
(b) Run a fixed effects meta-analysis using the exponential form (ie, exponentiating the result back to the original scale). Attach the forest plot.

(c) What is the pooled estimate? Is it significant? Please interpret the result.

(d) If necessary, run a random effect meta-analysis using exponential form.

(e) Why don't we have to convert the proportion to Cohen's D?

Write a short methods and results paragraph as if you were reporting these two sections in a journal article (refer to our in-class case studies).