**Supplemental File S11: Cluster chi-squared tests and Fisher's tests.**

Title: Repetition and reproduction of preclinical medical studies: taking a leaf from the plant sciences.

Authors: Jeremy S.C. Clark, Anna Sałacka, Agnieszka Boroń, Thierry van de Wetering, Kamila Rydzewska, Konrad Podsiadło, Tomasz Prystacki, Leszek Domański, Mateusz Kurzawski, Andrzej Ciechanowicz.

**Rationale for analyses.**

The rationale for the analyses is to assess prevalence of various protocols (defined in text) in the plant and medical sciences. Fisher tests and Chi-squared tests with adjustments for clustering were used (because some journals from the plant sciences were sampled more than once).

**Contributions to knowledge.**

The analyses clearly show differences between the sciences. No similar such analyses have been found in the literature.

**1A. Chi-squared tests, with adjustments for journal clustering, between MEDICAL and PLANT triplicated and non-triplicated protocols, 2017 data.**

library(data.table)

library(aod)

library(vcd)

data2017 <- fread(".... path-to-file S12 Supplemental TABLE\_Clusters\_2017 S12...............")

data2017 <- as.data.frame(data2017)

data2017 <- data2017[c(1:143), c(1:5)]

donner(cbind(y, n-y) ~ group, data = data2017)

## Test of proportion homogeneity (Donner, 1989)

## ---------------------------------------------

## donner(formula = cbind(y, n - y) ~ group, data = data2017)

## N = 147 clusters, n = 160 subjects, y = 57 cases, I = 2 groups.

##

## Data and correction factors:

## group N n y p C

## 1 M 80 80 4 0.0500 1.0000

## 2 P 67 80 53 0.6625 0.7778

##

## Intra-cluster correlation (anova estimate): -0.6836

##

## Adjusted chi-squared test:

## X2 = 74.8, df = 1, P(> X2) = 0

pchisq(q = 74.8, df = 1, lower.tail = FALSE)

## p = 5.21 x 10^-18

**## confidence intervals for clustered data:**

**## plants:**

varbin(n, y, data = data2017[c(1:67), ], alpha = 0.05, R = 5000)

## varbin(n = n, y = y, data = data2017[c(1:67), ], alpha = 0.05,

## R = 5000)

## N = 67 clusters, n = 80 subjects, y = 53 cases.

##

## Proportion

## ----------

## p se lower upper

## Binomial 0.6625 0.0532 0.5582 0.7668

## Ratio 0.6625 0.0526 0.5594 0.7656

## Arithmetic 0.6567 0.0555 0.5480 0.7654

## Jackknife 0.6626 0.0527 0.5594 0.7659

## Bootstrap 0.6638 0.0525 0.5609 0.7667

##

## alpha = 0.05 for the CIs; R = 5000 samples for the bootstrap estimates.

**## medical:**

varbin(n, y, data = data2017[c(68:147), ], alpha = 0.05, R = 5000)

## N = 80 clusters, n = 80 subjects, y = 4 cases.

##

## Proportion

## ----------

## p se lower upper

## Binomial 0.050 0.0245 0.0019 0.0981

## Ratio 0.050 0.0245 0.0019 0.0981

## Arithmetic 0.050 0.0245 0.0019 0.0981

## Jackknife 0.050 0.0245 0.0019 0.0981

## Bootstrap 0.050 0.0245 0.0020 0.0980

##

## alpha = 0.05 for the CIs; R = 5000 samples for the bootstrap estimates.

**1B. Fisher's exact test between , between MEDICAL and PLANT protocols, 2017 (assumes no effect of journal clustering).**

|  |  |
| --- | --- |
| **Plants** |  |
| Triple result, | 7 |
| Triplicated study, global result | 40 |
| Double result | 11 |
| Duplicated study, global result | 6 |
| **­**Non-repeated study | 16 |
|  |  |
| **Medical science.** |  |
| Triple result, | 1 |
| Triplicated study, global result | 3 |
| Double result | 1 |
| Duplicated study, global result | 0 |
| **­**Non-repeated study | 75 |
|  |  |

Input =("  
Frequency Non-repeated Dup\_global Dup Tri\_global Tri   
Medical 75 0 1 3 1   
Plants 16 6 11 40 7     
")  
  
Matriz = as.matrix(read.table(textConnection(Input),  
                   header=TRUE,  
                   row.names=1))  
  
Matriz  
  
fisher.test(Matriz,  
            alternative="two.sided")

## > Matriz

## Non.repeated Dup\_global Dup Tri\_global Tri

## Medical 75 0 1 3 1

## Plants 16 6 11 40 7

## >

## > fisher.test(Matriz,

## + alternative="two.sided")

##

## Fisher's Exact Test for Count Data

##

## data: Matriz

## p-value < 2.2e-16

## alternative hypothesis: two.sided

##

assocstats(Matriz)

## X^2 df P(> X^2)

## Likelihood Ratio 102.503 4 0

## Pearson 88.923 4 0

##

## Phi-Coefficient : NA

## Contingency Coeff.: 0.598

## Cramer's V : 0.746

##

**2A. Chi-squared tests, with adjustments for journal clustering, between MEDICAL and PLANT triplicated and non-triplicated protocols, 2019 data.**

data2019 <- fread("...........path to file S13 Supplemental TABLE\_Clusters\_2019 S13")

data2019 <- as.data.frame(data2019)

data2019 <- data2019[c(1:143), c(1:5)]

donner(cbind(y, n-y) ~ group, data = data2019)

> donner(cbind(y,n-y) ~group, data = data2019)

## Test of proportion homogeneity (Donner, 1989)

## ---------------------------------------------

## donner(formula = cbind(y, n - y) ~ group, data = data2019)

## N = 143 clusters, n = 160 subjects, y = 52 cases, I = 2 groups.

##

## Data and correction factors:

## group N n y p C

## 1 M 80 80 4 0.05 1.0000

## 2 P 63 80 48 0.60 0.8345

##

## Intra-cluster correlation (anova estimate): -0.3484

##

## Adjusted chi-squared test:

## X2 = 60.6, df = 1, P(> X2) = 0

pchisq(q = 60.6, df = 1, lower.tail = FALSE)

## p = 7.00 x 10^-15

**## confidence intervals for clustered data:**

**## plants:**

varbin(n, y, data = data2019[c(1:63), ], alpha = 0.05, R = 5000)

## N = 63 clusters, n = 80 subjects, y = 48 cases.

##

## Proportion

## ----------

## p se lower upper

## Binomial 0.6000 0.0551 0.4920 0.7080

## Ratio 0.6000 0.0559 0.4904 0.7096

## Arithmetic 0.5714 0.0598 0.4542 0.6887

## Jackknife 0.6005 0.0560 0.4907 0.7103

## Bootstrap 0.5994 0.0547 0.4921 0.7067

##

## alpha = 0.05 for the CIs; R = 5000 samples for the bootstrap estimates.

**## medical:**

varbin(n, y, data = data2019[c(64:143), ], alpha = 0.05, R = 5000)

## N = 80 clusters, n = 80 subjects, y = 4 cases.

##

## Proportion

## ----------

## p se lower upper

## Binomial 0.0500 0.0245 0.0019 0.0981

## Ratio 0.0500 0.0245 0.0019 0.0981

## Arithmetic 0.0500 0.0245 0.0019 0.0981

## Jackknife 0.0500 0.0245 0.0019 0.0981

## Bootstrap 0.0495 0.0242 0.0020 0.0971

##

## alpha = 0.05 for the CIs; R = 5000 samples for the bootstrap estimates.

**2B. Fisher's exact test between MEDICAL and PLANT protocols 2019 (assumes no effect of journal clustering).**

|  |  |
| --- | --- |
| **Plant sciences.** |  |
| Triple result, | 29 |
| Triplicated study, global result | 19 |
| Double result | 6 |
| Duplicated study, global result | 0 |
| **­**Non-repeated study | 26 |
|  |  |
| **Medical science.** |  |
| Triple result, | 2 |
| Triplicated study, global result | 2 |
| Double result | 1 |
| Duplicated study, global result | 0 |
| **­**Non-repeated study | 75 |
|  |  |

Input =("

Frequency Non-repeated Dup\_global Dup Tri\_global Tri

Medical 75 0 1 2 2

Plants 26 0 6 19 29

")

Matriz = as.matrix(read.table(textConnection(Input),

header=TRUE,

row.names=1))

Matriz

fisher.test(Matriz,

alternative="two.sided")

## Non.repeated Dup\_global Dup Tri\_global Tri

## Medical 75 0 1 2 2

## Plants 26 0 6 19 29

## >

## > fisher.test(Matriz,

## + alternative="two.sided")

##

## Fisher's Exact Test for Count Data

##

## data: Matriz

## p-value < 2.2e-16

## alternative hypothesis: two.sided

##

assocstats(Matriz)

## X^2 df P(> X^2)

## Likelihood Ratio 72.815 4 5.7732e-15

## Pearson NaN 4 NaN

##

## Phi-Coefficient : NA

## Contingency Coeff.: NaN

## Cramer's V : NaN

##