Bayesian Nonparametric and Meta Analyses of COVID-19 Studies

Haixin Yu

12/20/2020

Introduction

Background

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first case was reported in Wuhan, China, in December 2019. Then the disease rapidly spread across the world. On March 11, 2020, the World Health Organization (WHO) declared the outbreak a pandemic. As of today, a total of 76,381,409 and 17,702,516 confirmed cases have been recorded globally and in the United States. The rapid spread of the virus is largely attributed to the fact that infected patients may be asymptomatic. An asymptomatic patient is defined as a person with laboratory-confirmed COVID-19 infection that has no symptoms at the time of first clinical assessment and at the end of follow-up. Although asymptomatic COVID-19 infections have been reported in various studies, the proportion of asymptomatic infections varied widely between 5% and 95%. Thus, we conducted the Bayesian nonparametric and meta analyses to answer three questions:

1) What is the proportion of asymptomatic infections among people who infected with COVID-19?
2) What are the factors of asymptomatic infections in COVID-19?
3) Most importantly, is it reasonable to synthesize all the studies to produce a single estimate?

Data Source

The file listing all the studies used for my analysis is available from the Harvard Dataverse Database. Four online databases including PubMed, Embase, bioRxiv and medRxiv were searched through 31 July 2020. All included studies reported a specific number of asymptomatic COVID-19 infections. There was no language restriction. Risk of bias may be included in studies.

Analysis and Results

Data Preprocessing

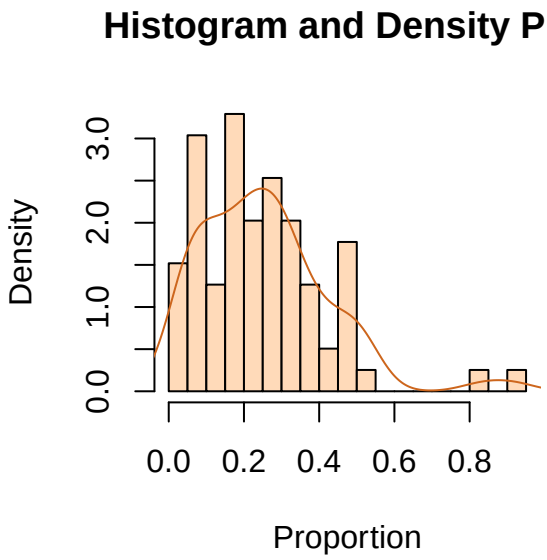
All analyses were conducted using R version 4.0.2. A total of 6,602 patients from 79 studies were included. The original dataset had 7 variables. I separated the variable `'setting'` into two variables: `'code'` and `'setting'`. I added a new variable `'proportion'` that calculated by dividing the number of asymptomatic infections by the number of total COVID-19 infections. Thus, there are 79 observations of 9 variables in our dataset. The nine variables are:

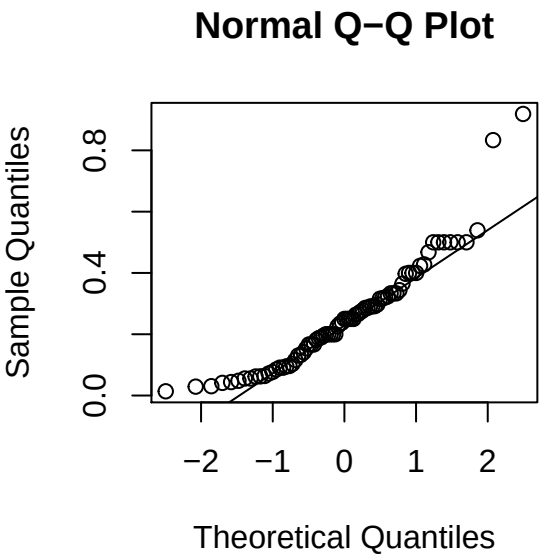
- **id**: record id
- **author**: first author
- **collection**: setting of data collection
- **code**: setting code
- **setting** (seven study settings):
  - contact investigation
  - contact investigation, aggregated

  - outbreak investigation
  - screening
  - hospitalized adults
  - hospitalized children
  - hospitalized children and adults
- **cases**: the number of asymptomatic infections
- **total**: the number of total infections
- **proportion**: cases/total
- **source** (four databases):
  - PubMed
  - Embase
  - bioRxiv
  - medRxiv

For the seven categories of the variable `'setting'`: contact investigation included 16 studies, contact investigation, aggregated included 9 studies, outbreak investigation included 12 studies, screening included 7 studies, hospitalized adults included 15 studies, hospitalized children included 10 studies, and hospitalized children and adults included 10 studies.

The histogram and density plot below gave us a general idea of how the data was distributed. The normal Q-Q plot below examined the normality of the data. I performed two types of analysis on the dataset. The first one was meta-analysis using `meta` and `metafor` packages. The second one was Bayesian nonparametric analysis using `DPpackage`.



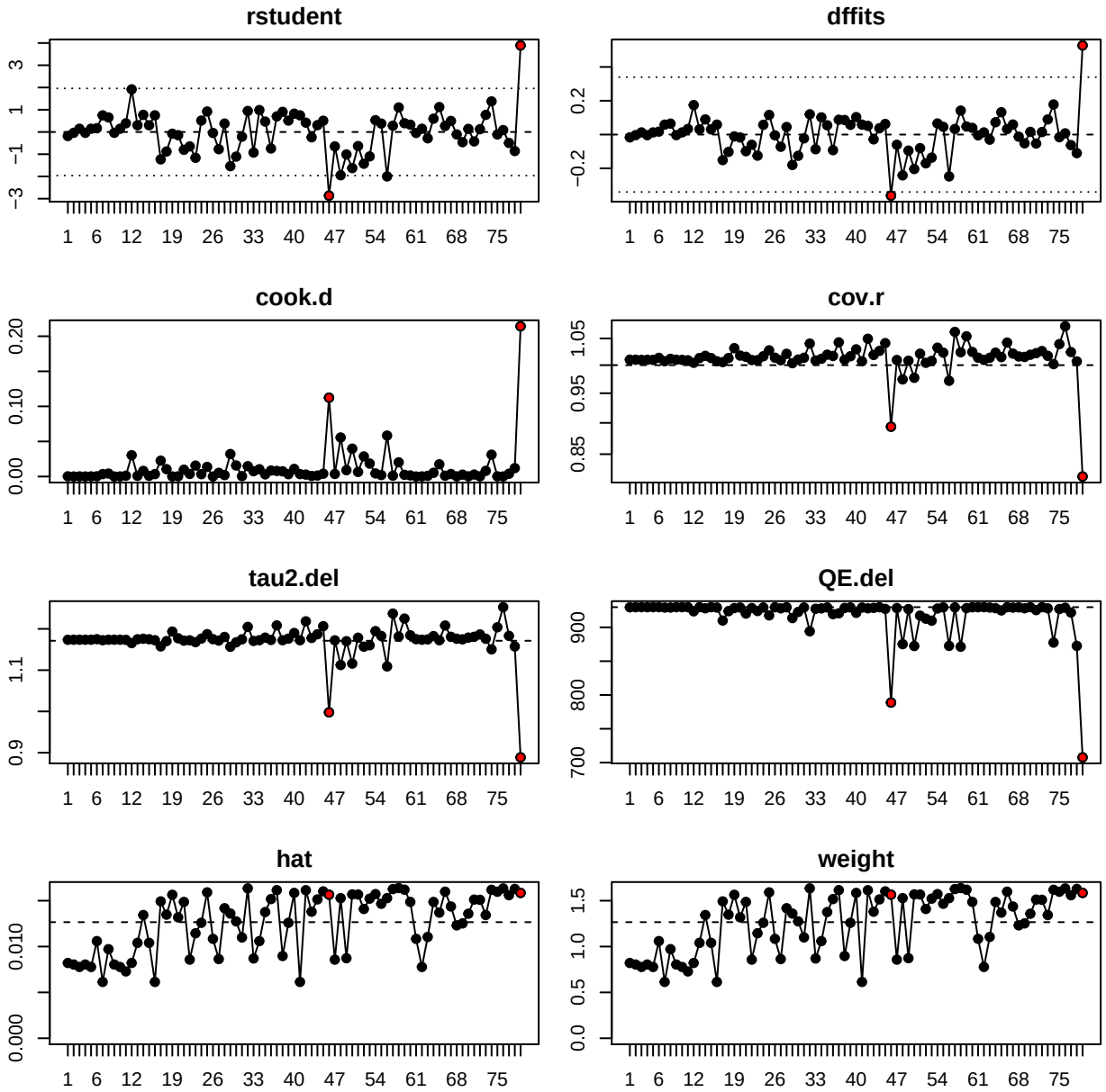


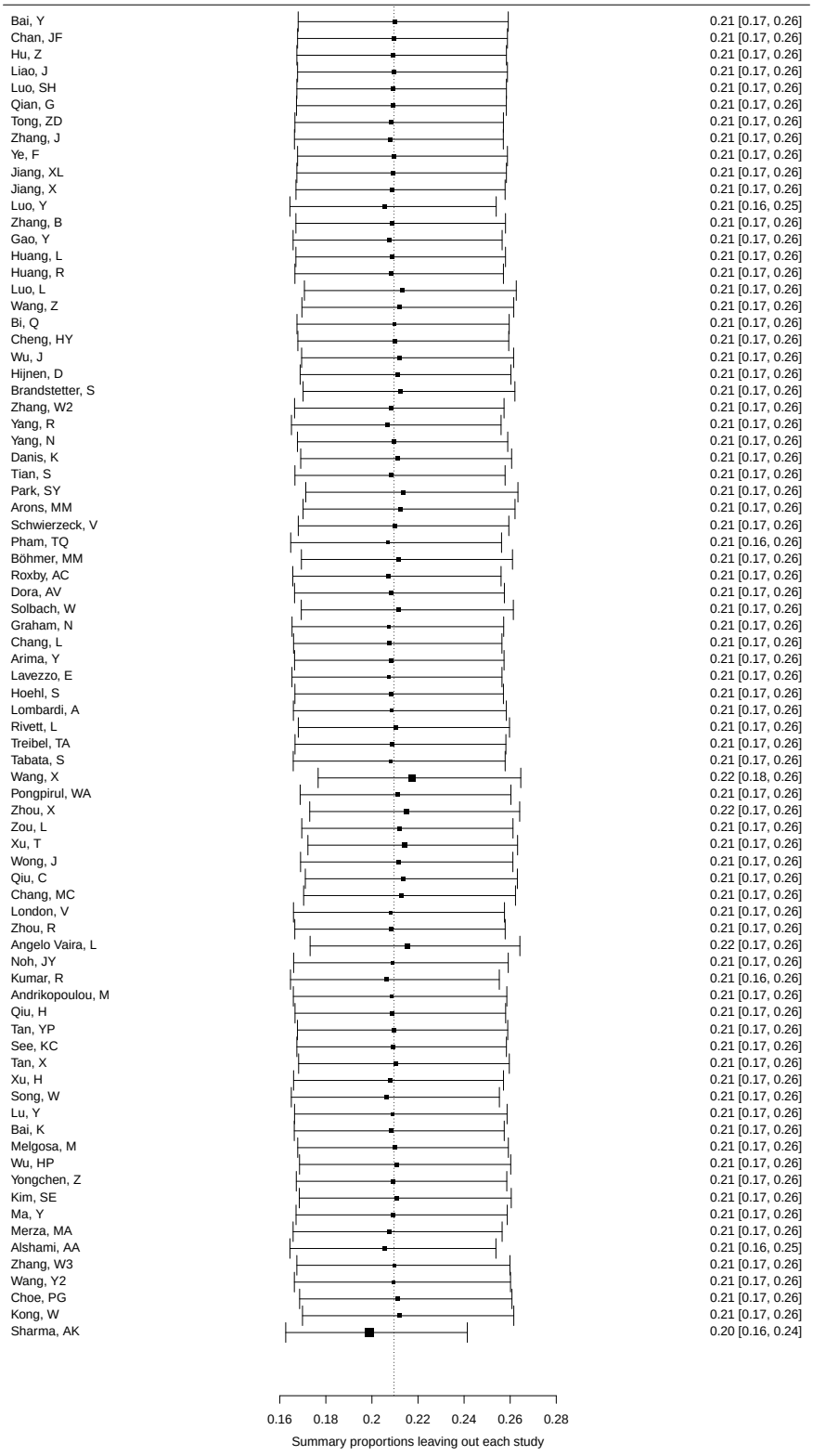
Meta Analysis Using meta and metafor Packages

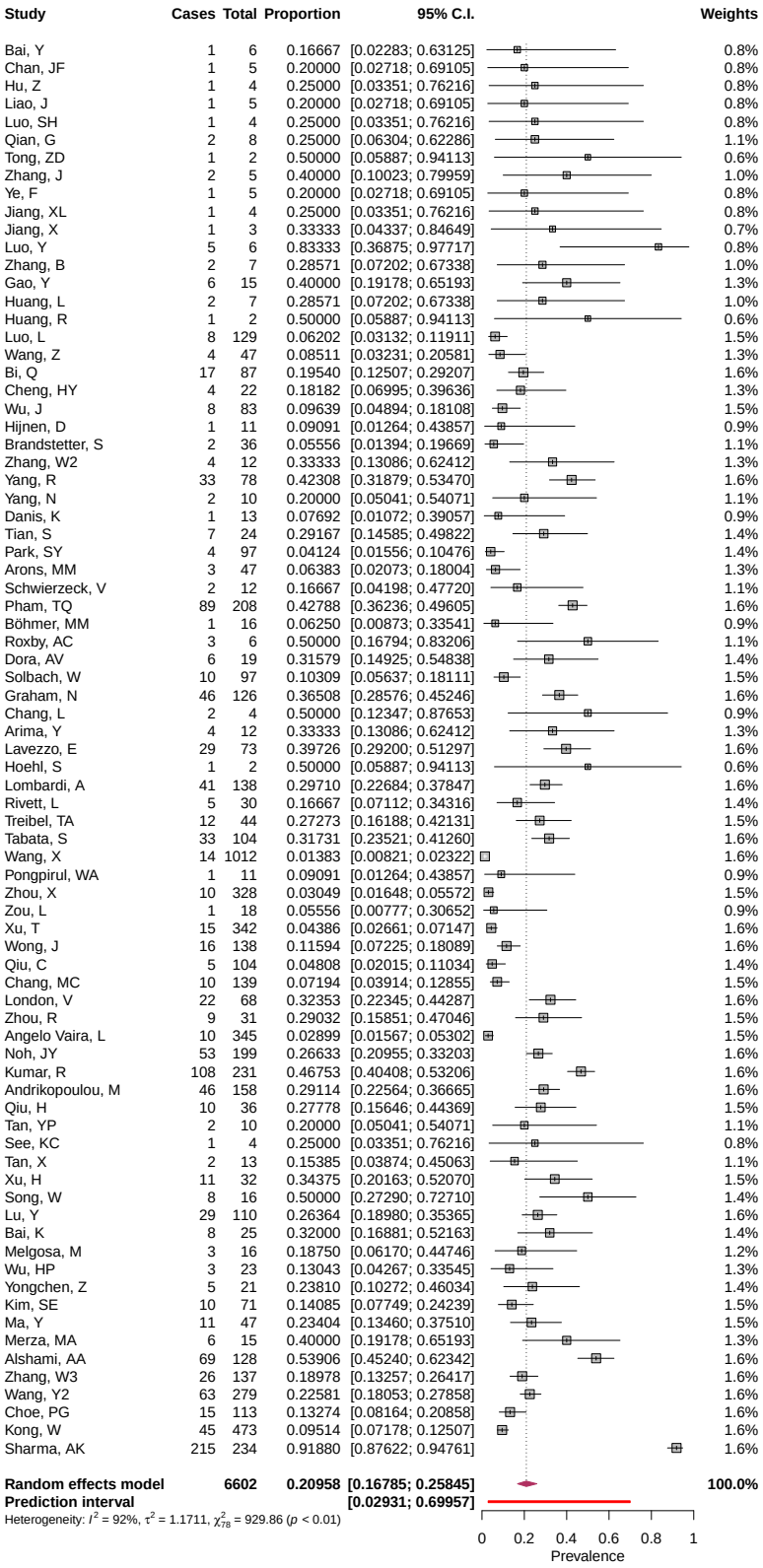
Based on the QQ plot above, the proportion of asymptomatic infections was transformed using the logit transformation to make it conform to the normal distribution. The `escalc()` function in the `metafor` package was used to calculate individual effect sizes (yi) and its sampling variances (vi). The overall effect size and its 95% confidence interval was based on the random-effects-model using DerSimonian-Laird estimator and normal approximation interval respectively. Unlike the fixed-effects-model, which assumes all studies along with their effect sizes stem from a single homogeneous population, the random-effects-model assumes a distribution of true effect sizes from a "universe" of populations, which takes both within- and between-study variances into account.

In the forest plot below, we displayed the overall estimate of asymptomatic COVID-19 infections through the `metaprop()` and `forest()` function in the `meta` package. Of 6,602 confirmed cases, 1,273 were defined as asymptomatic infections. The overall estimate of the proportion of people who become infected with COVID-19 and remain asymptomatic throughout the course of infection was 20.96%, with a 95% confidence interval of 16.8% to 25.8% and a prediction interval of 3.0% to 70.0%. Apparently, the estimates of asymptomatic infections were quite different from study to study in the forest plot. The high heterogeneity was also confirmed by the tau^2 and I^2 statistic (tau^2 = 1.17, I^2 = 92%, $p < 0.01$).

There are three methods to quantify heterogeneity: 1) the test for heterogeneity (Q), 2) the estimate of between-study variance (tau^2), 3) the estimate for the proportion of the observed variability that reflects the between-study variance (I^2). Together, they can inform us if the effects are consistent. According to our results, it is not appropriate to combine all the studies to make a summary estimate of the proportion of asymptomatic COVID-19 infections. To identify and visually inspect the influential studies, leave-one-out analysis and diagnostic test were used.



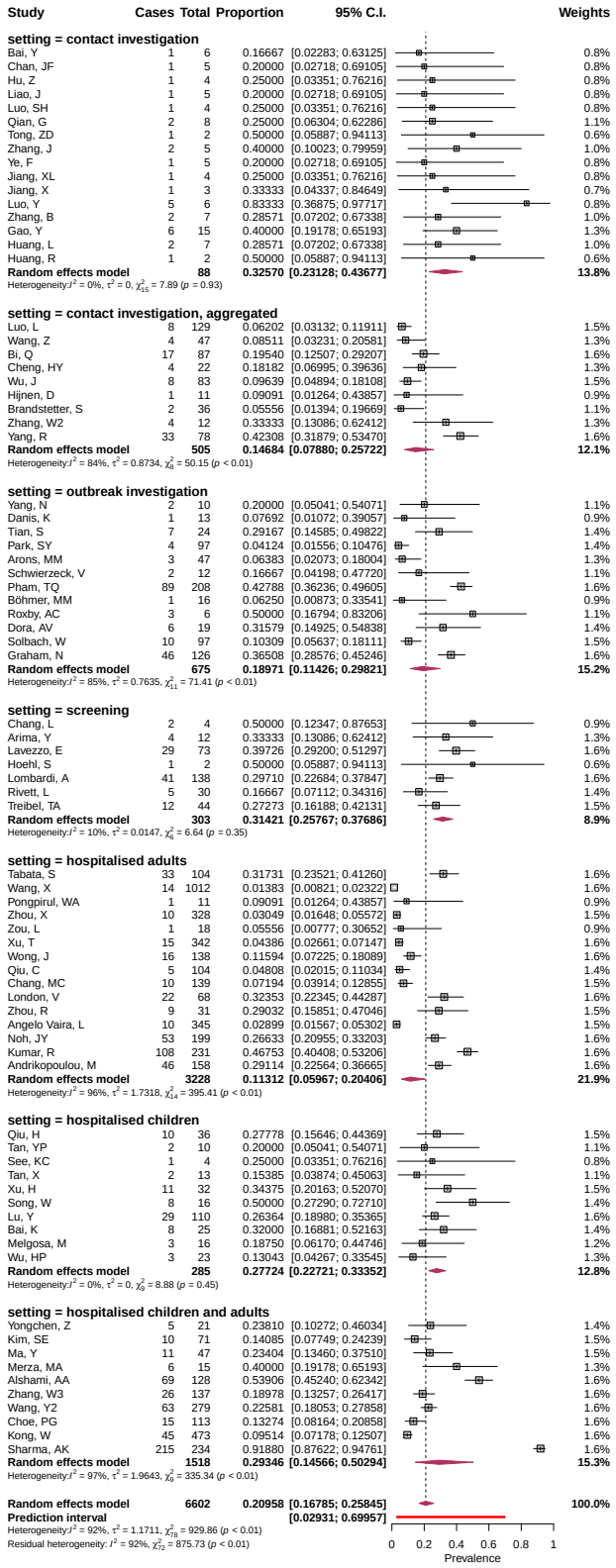

Bai, Y 0.21 [0.17, 0.26]
Chan, JF 0.21 [0.17, 0.26]
Hu, Z 0.21 [0.17, 0.26]
Liao, J 0.21 [0.17, 0.26]
Luo, SH 0.21 [0.17, 0.26]
Qian, G 0.21 [0.17, 0.26]
Tong, ZD 0.21 [0.17, 0.26]
Zhang, J 0.21 [0.17, 0.26]
Ye, F 0.21 [0.17, 0.26]
Jiang, XL 0.21 [0.17, 0.26]
Jiang, X 0.21 [0.17, 0.26]
Luo, Y 0.21 [0.16, 0.25]
Zhang, B 0.21 [0.17, 0.26]
Gao, Y 0.21 [0.17, 0.26]
Huang, L 0.21 [0.17, 0.26]
Huang, R 0.21 [0.17, 0.26]
Luo, L 0.21 [0.17, 0.26]
Wang, Z 0.21 [0.17, 0.26]
Bi, Q 0.21 [0.17, 0.26]
Cheng, HY 0.21 [0.17, 0.26]
Wu, J 0.21 [0.17, 0.26]
Hijnen, D 0.21 [0.17, 0.26]
Brandstetter, S 0.21 [0.17, 0.26]
Zhang, W2 0.21 [0.17, 0.26]
Yang, R 0.21 [0.17, 0.26]
Yang, N 0.21 [0.17, 0.26]
Danis, K 0.21 [0.17, 0.26]
Tian, S 0.21 [0.17, 0.26]
Park, SY 0.21 [0.17, 0.26]
Arons, MM 0.21 [0.17, 0.26]
Schwierzeck, V 0.21 [0.17, 0.26]
Pham, TQ 0.21 [0.16, 0.26]
Böhmer, MM 0.21 [0.17, 0.26]
Roxby, AC 0.21 [0.17, 0.26]
Dora, AV 0.21 [0.17, 0.26]
Solbach, W 0.21 [0.17, 0.26]
Graham, N 0.21 [0.17, 0.26]
Chang, L 0.21 [0.17, 0.26]
Arima, Y 0.21 [0.17, 0.26]
Lavezzo, E 0.21 [0.17, 0.26]
Hoehl, S 0.21 [0.17, 0.26]
Lombardi, A 0.21 [0.17, 0.26]
Rivett, L 0.21 [0.17, 0.26]
Treibel, TA 0.21 [0.17, 0.26]
Tabata, S 0.21 [0.17, 0.26]
Wang, X 0.22 [0.18, 0.26]
Pongpirul, WA 0.21 [0.17, 0.26]
Zhou, X 0.22 [0.17, 0.26]
Zou, L 0.21 [0.17, 0.26]
Xu, T 0.21 [0.17, 0.26]
Wong, J 0.21 [0.17, 0.26]
Qiu, C 0.21 [0.17, 0.26]
Chang, MC 0.21 [0.17, 0.26]
London, V 0.21 [0.17, 0.26]
Zhou, R 0.21 [0.17, 0.26]
Angelo Vaira, L 0.22 [0.17, 0.26]
Noh, JY 0.21 [0.17, 0.26]
Kumar, R 0.21 [0.16, 0.26]
Andrikopoulou, M 0.21 [0.17, 0.26]
Qiu, H 0.21 [0.17, 0.26]
Tan, YP 0.21 [0.17, 0.26]
See, KC 0.21 [0.17, 0.26]
Tan, X 0.21 [0.17, 0.26]
Xu, H 0.21 [0.17, 0.26]
Song, W 0.21 [0.17, 0.26]
Lu, Y 0.21 [0.17, 0.26]
Bai, K 0.21 [0.17, 0.26]
Melgosa, M 0.21 [0.17, 0.26]
Wu, HP 0.21 [0.17, 0.26]
Yongchen, Z 0.21 [0.17, 0.26]
Kim, SE 0.21 [0.17, 0.26]
Ma, Y 0.21 [0.17, 0.26]
Merza, MA 0.21 [0.17, 0.26]
Alshami, AA 0.21 [0.16, 0.25]
Zhang, W3 0.21 [0.17, 0.26]
Wang, Y2 0.21 [0.17, 0.26]
Choe, PG 0.21 [0.17, 0.26]
Kong, W 0.21 [0.17, 0.26]
Sharma, AK 0.20 [0.16, 0.24]
0.16 0.18 0.2 0.22 0.24 0.26 0.28
Summary proportions leaving out each study


Study Cases Total Proportion 95% C.I. Weights
Bai, Y 1 6 0.16667 [0.02283; 0.63125] 0.8%
Chan, JF 1 5 0.20000 [0.02718; 0.69105] 0.8%
Hu, Z 1 4 0.25000 [0.03351; 0.76216] 0.8%
Liao, J 1 5 0.20000 [0.02718; 0.69105] 0.8%
Luo, SH 1 4 0.25000 [0.03351; 0.76216] 0.8%
Qian, G 2 8 0.25000 [0.06304; 0.62286] 1.1%
Tong, ZD 1 2 0.50000 [0.05887; 0.94113] 0.6%
Zhang, J 2 5 0.40000 [0.10023; 0.79959] 1.0%
Ye, F 1 5 0.20000 [0.02718; 0.69105] 0.8%
Jiang, XL 1 4 0.25000 [0.03351; 0.76216] 0.8%
Jiang, X 1 3 0.33333 [0.04337; 0.84649] 0.7%
Luo, Y 5 6 0.83333 [0.36875; 0.97717] 0.8%
Zhang, B 2 7 0.28571 [0.07202; 0.67338] 1.0%
Gao, Y 6 15 0.40000 [0.19178; 0.65193] 1.3%
Huang, L 2 7 0.28571 [0.07202; 0.67338] 1.0%
Huang, R 1 2 0.50000 [0.05887; 0.94113] 0.6%
Luo, L 8 129 0.06202 [0.03132; 0.11911] 1.5%
Wang, Z 4 47 0.08511 [0.03231; 0.20581] 1.3%
Bi, Q 17 87 0.19540 [0.12507; 0.29207] 1.6%
Cheng, HY 4 22 0.18182 [0.06995; 0.39636] 1.3%
Wu, J 8 83 0.09639 [0.04894; 0.18108] 1.5%
Hijnen, D 1 11 0.09091 [0.01264; 0.43857] 0.9%
Brandstetter, S 2 36 0.05556 [0.01394; 0.19669] 1.1%
Zhang, W2 4 12 0.33333 [0.13086; 0.62412] 1.3%
Yang, R 33 78 0.42308 [0.31879; 0.53470] 1.6%
Yang, N 2 10 0.20000 [0.05041; 0.54071] 1.1%
Danis, K 1 13 0.07692 [0.01072; 0.39057] 0.9%
Tian, S 7 24 0.29167 [0.14585; 0.49822] 1.4%
Park, SY 4 97 0.04124 [0.01556; 0.10476] 1.4%
Arons, MM 3 47 0.06383 [0.02073; 0.18004] 1.3%
Schwierzeck, V 2 12 0.16667 [0.04198; 0.47720] 1.1%
Pham, TQ 89 208 0.42788 [0.36236; 0.49605] 1.6%
Böhmer, MM 1 16 0.06250 [0.00873; 0.33541] 0.9%
Roxby, AC 3 6 0.50000 [0.16794; 0.83206] 1.1%
Dora, AV 6 19 0.31579 [0.14925; 0.54838] 1.4%
Solbach, W 10 97 0.10309 [0.05637; 0.18111] 1.5%
Graham, N 46 126 0.36508 [0.28576; 0.45246] 1.6%
Chang, L 2 4 0.50000 [0.12347; 0.87653] 0.9%
Arima, Y 4 12 0.33333 [0.13086; 0.62412] 1.3%
Lavezzo, E 29 73 0.39726 [0.29200; 0.51297] 1.6%
Hoehl, S 1 2 0.50000 [0.05887; 0.94113] 0.6%
Lombardi, A 41 138 0.29710 [0.22684; 0.37847] 1.6%
Rivett, L 5 30 0.16667 [0.07112; 0.34316] 1.4%
Treibel, TA 12 44 0.27273 [0.16188; 0.42131] 1.5%
Tabata, S 33 104 0.31731 [0.23521; 0.41260] 1.6%
Wang, X 14 1012 0.01383 [0.00821; 0.02322] 1.6%
Pongpirul, WA 1 11 0.09091 [0.01264; 0.43857] 0.9%
Zhou, X 10 328 0.03049 [0.01648; 0.05572] 1.5%
Zou, L 1 18 0.05556 [0.00777; 0.30652] 0.9%
Xu, T 15 342 0.04386 [0.02661; 0.07147] 1.6%
Wong, J 16 138 0.11594 [0.07225; 0.18089] 1.6%
Qiu, C 5 104 0.04808 [0.02015; 0.11034] 1.4%
Chang, MC 10 139 0.07194 [0.03914; 0.12855] 1.5%
London, V 22 68 0.32353 [0.22345; 0.44287] 1.6%
Zhou, R 9 31 0.29032 [0.15851; 0.47046] 1.5%
Angelo Vaira, L 10 345 0.02899 [0.01567; 0.05302] 1.5%
Noh, JY 53 199 0.26633 [0.20955; 0.33203] 1.6%
Kumar, R 108 231 0.46753 [0.40408; 0.53206] 1.6%
Andrikopoulou, M 46 158 0.29114 [0.22564; 0.36665] 1.6%
Qiu, H 10 36 0.27778 [0.15646; 0.44369] 1.5%
Tan, YP 2 10 0.20000 [0.05041; 0.54071] 1.1%
See, KC 1 4 0.25000 [0.03351; 0.76216] 0.8%
Tan, X 2 13 0.15385 [0.03874; 0.45063] 1.1%
Xu, H 11 32 0.34375 [0.20163; 0.52070] 1.5%
Song, W 8 16 0.50000 [0.27290; 0.72710] 1.4%
Lu, Y 29 110 0.26364 [0.18980; 0.35365] 1.6%
Bai, K 8 25 0.32000 [0.16881; 0.52163] 1.4%
Melgosa, M 3 16 0.18750 [0.06170; 0.44746] 1.2%
Wu, HP 3 23 0.13043 [0.04267; 0.33545] 1.3%
Yongchen, Z 5 21 0.23810 [0.10272; 0.46034] 1.4%
Kim, SE 10 71 0.14085 [0.07749; 0.24239] 1.5%
Ma, Y 11 47 0.23404 [0.13460; 0.37510] 1.5%
Merza, MA 6 15 0.40000 [0.19178; 0.65193] 1.3%
Alshami, AA 69 128 0.53906 [0.45240; 0.62342] 1.6%
Zhang, W3 26 137 0.18978 [0.13257; 0.26417] 1.6%
Wang, Y2 63 279 0.22581 [0.18053; 0.27858] 1.6%
Choe, PG 15 113 0.13274 [0.08164; 0.20858] 1.6%
Kong, W 45 473 0.09514 [0.07178; 0.12507] 1.6%
Sharma, AK 215 234 0.91880 [0.87622; 0.94761] 1.6%
Random effects model 6602 0.20958 [0.16785; 0.25845] 100.0%
Prediction interval [0.02931; 0.69957]
Heterogeneity: I^2 = 92%, τ^2 = 1.1711, χ^2_{78} = 929.86 (p < 0.01)
0 0.2 0.4 0.6 0.8 1
Prevalence

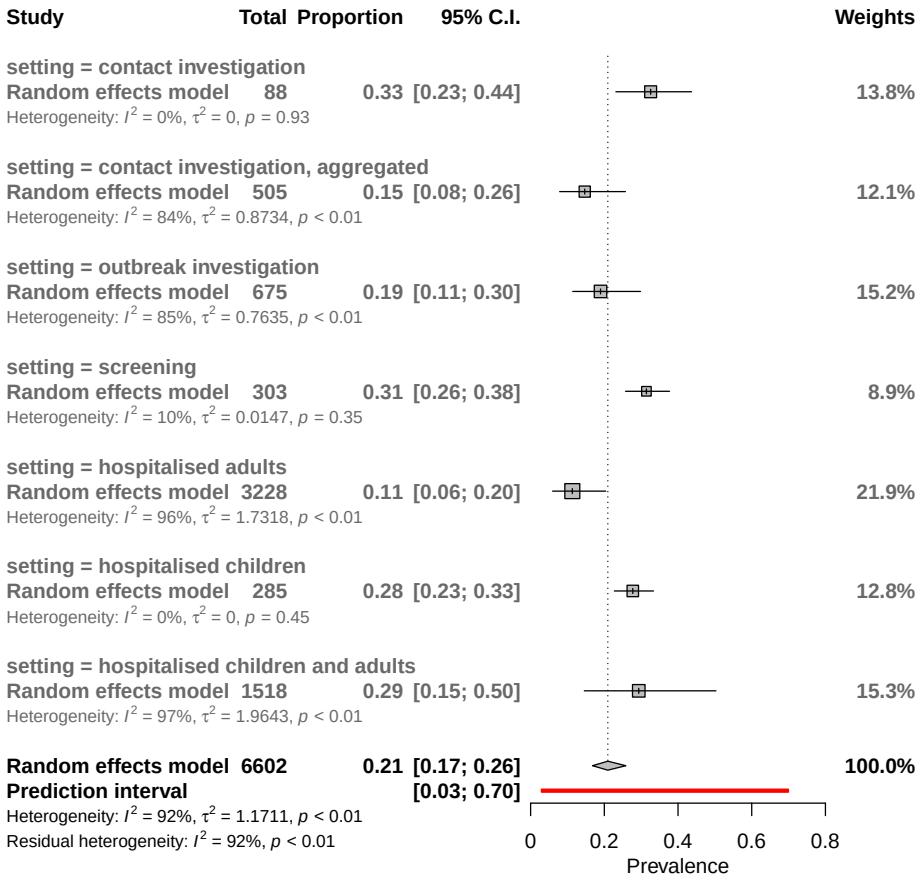
Subgroup Analysis

Subgroup analysis were conducted to explain potential factors contributing to heterogeneity. There are seven defined subgroups based on the method of selection of asymptomatic status. As shown in the forest plot in presence of subgroups, the summary estimate of the proportion of people with asymptomatic COVID-19 infections varied from setting to setting. The first three subgroups involved studies that reported on different types of contact investigation. In 16 studies of single-family contact investigation, the summary estimate was 33% (95% CI 23%-44%). In 9 studies of aggregated contact investigation from both asymptomatic and symptomatic people, the summary estimate was 15% (95% CI 8%-26%). There were 12 studies that reported on outbreak investigation arising from a single symptomatic case in nursing homes or occupational settings. The summary estimate of these studies was 19% (95% CI 11%-30%). People with asymptomatic COVID-19 infections that detected through screening in defined populations were reported in 7 studies. The screened populations included healthcare workers, evacuees from transmission sites, and the entire population of an Italian village. The summary estimate of these studies was 31% (95% CI 26%-38%). The remaining studies were done in hospital settings, included adult only (15 studies), children only (10 studies), and children and adults (10 studies). The summary estimates of these subgroups were 11% (95% CI 6%-20%), 28% (95% CI 23%-33%), and 29% (95% CI 15%-50%) respectively.

```
forest(pes.summary,
       xlim=c(0,1),
       leftcols=c("studlab", "event", "n", "effect", "ci"),
       leftlabs=c("Study", "Cases", "Total", "Proportion", "95% C.I."),
       rightcols=c("w.random"),
       rightlabs=c("Weights"),
       xlab="Prevalence",
       prediction=TRUE,
       col.predict.lines="red",
       fs.xlab=12,
       fs.study=12,
       fs.study.lables=12,
       fs.heading=14,
       squaresize=0.5,
       col.square="grey",
       col.square.lines="black",
       col.diamond="maroon",
       col.diamond.lines="maroon",
       col.by="black",
       pooled.totals=TRUE,
       comb.fixed=FALSE,
       lty.random=2,
       type.study="square",
       type.random="diamond",
       ff.random="bold",
       hetlab="Heterogeneity:",
       fs.hetstat=10,
       smlab="",
       print.tau2=TRUE,
       print.Q=TRUE,
       print.pval.Q=TRUE,
       print.I2=TRUE,
       digits.Q=2,
       digits=5)
```


Study Cases Total Proportion 95% C.I. Weights
setting = contact investigation
Bai, Y 1 6 0.16667 [0.02283; 0.63125] 0.8%
Chan, JF 1 5 0.20000 [0.02718; 0.69105] 0.8%
Hu, Z 1 4 0.25000 [0.03351; 0.76216] 0.8%
Liao, J 1 5 0.20000 [0.02718; 0.69105] 0.8%
Luo, SH 1 4 0.25000 [0.03351; 0.76216] 0.8%
Qian, G 2 8 0.25000 [0.06304; 0.62286] 1.1%
Tong, ZD 1 2 0.50000 [0.05887; 0.94113] 0.6%
Zhang, J 2 5 0.40000 [0.10023; 0.79959] 1.0%
Ye, F 1 5 0.20000 [0.02718; 0.69105] 0.8%
Jiang, XL 1 4 0.25000 [0.03351; 0.76216] 0.8%
Jiang, X 1 3 0.33333 [0.04337; 0.84649] 0.7%
Luo, Y 5 6 0.83333 [0.36875; 0.97717] 0.8%
Zhang, B 2 7 0.28571 [0.07202; 0.67338] 1.0%
Gao, Y 6 15 0.40000 [0.19178; 0.65193] 1.3%
Huang, L 2 7 0.28571 [0.07202; 0.67338] 1.0%
Huang, R 1 2 0.50000 [0.05887; 0.94113] 0.6%
Random effects model 88 0.32570 [0.23128; 0.43677] 13.8%
Heterogeneity: I^2 = 0%, τ^2 = 0, χ^2_{15} = 7.89 (p = 0.93)
setting = contact investigation, aggregated
Luo, L 8 129 0.06202 [0.03132; 0.11911] 1.5%
Wang, Z 4 47 0.08511 [0.03231; 0.20581] 1.3%
Bi, Q 17 87 0.19540 [0.12507; 0.29207] 1.6%
Cheng, HY 4 22 0.18182 [0.06995; 0.39636] 1.3%
Wu, J 8 83 0.09639 [0.04894; 0.18108] 1.5%
Hijnen, D 1 11 0.09091 [0.01264; 0.43857] 0.9%
Brandstetter, S 2 36 0.05556 [0.01394; 0.19669] 1.1%
Zhang, W2 4 12 0.33333 [0.13086; 0.62412] 1.3%
Yang, R 33 78 0.42308 [0.31879; 0.53470] 1.6%
Random effects model 505 0.14684 [0.07880; 0.25722] 12.1%
Heterogeneity: I^2 = 84%, τ^2 = 0.8734, χ^2_{8} = 50.15 (p < 0.01)
setting = outbreak investigation
Yang, N 2 10 0.20000 [0.05041; 0.54071] 1.1%
Danis, K 1 13 0.07692 [0.01072; 0.39057] 0.9%
Tian, S 7 24 0.29167 [0.14585; 0.49822] 1.4%
Park, SY 4 97 0.04124 [0.01556; 0.10476] 1.4%
Arons, MM 3 47 0.06383 [0.02073; 0.18004] 1.3%
Schwierzeck, V 2 12 0.16667 [0.04198; 0.47720] 1.1%
Pham, TQ 89 208 0.42788 [0.36236; 0.49605] 1.6%
Böhmer, MM 1 16 0.06250 [0.00873; 0.33541] 0.9%
Roxby, AC 3 6 0.50000 [0.16794; 0.83206] 1.1%
Dora, AV 6 19 0.31579 [0.14925; 0.54838] 1.4%
Solbach, W 10 97 0.10309 [0.05637; 0.18111] 1.5%
Graham, N 46 126 0.36508 [0.28576; 0.45246] 1.6%
Random effects model 675 0.18971 [0.11426; 0.29821] 15.2%
Heterogeneity: I^2 = 85%, τ^2 = 0.7635, χ^2_{11} = 71.41 (p < 0.01)
setting = screening
Chang, L 2 4 0.50000 [0.12347; 0.87653] 0.9%
Arima, Y 4 12 0.33333 [0.13086; 0.62412] 1.3%
Lavezzo, E 29 73 0.39726 [0.29200; 0.51297] 1.6%
Hoehl, S 1 2 0.50000 [0.05887; 0.94113] 0.6%
Lombardi, A 41 138 0.29710 [0.22684; 0.37847] 1.6%
Rivett, L 5 30 0.16667 [0.07112; 0.34316] 1.4%
Treibel, TA 12 44 0.27273 [0.16188; 0.42131] 1.5%
Random effects model 303 0.31421 [0.25767; 0.37686] 8.9%
Heterogeneity: I^2 = 10%, τ^2 = 0.0147, χ^2_{6} = 6.64 (p = 0.35)
setting = hospitalised adults
Tabata, S 33 104 0.31731 [0.23521; 0.41260] 1.6%
Wang, X 14 1012 0.01383 [0.00821; 0.02322] 1.6%
Pongpirul, WA 1 11 0.09091 [0.01264; 0.43857] 0.9%
Zhou, X 10 328 0.03049 [0.01648; 0.05572] 1.5%
Zou, L 1 18 0.05556 [0.00777; 0.30652] 0.9%
Xu, T 15 342 0.04386 [0.02661; 0.07147] 1.6%
Wong, J 16 138 0.11594 [0.07225; 0.18089] 1.6%
Qiu, C 5 104 0.04808 [0.02015; 0.11034] 1.4%
Chang, MC 10 139 0.07194 [0.03914; 0.12855] 1.5%
London, V 22 68 0.32353 [0.22345; 0.44287] 1.6%
Zhou, R 9 31 0.29032 [0.15851; 0.47046] 1.5%
Angelo Vaira, L 10 345 0.02899 [0.01567; 0.05302] 1.5%
Noh, JY 53 199 0.26633 [0.20955; 0.33203] 1.6%
Kumar, R 108 231 0.46753 [0.40408; 0.53206] 1.6%
Andrikopoulou, M 46 158 0.29114 [0.22564; 0.36665] 1.6%
Random effects model 3228 0.11312 [0.05967; 0.20406] 21.9%
Heterogeneity: I^2 = 96%, τ^2 = 1.7318, χ^2_{14} = 395.41 (p < 0.01)
setting = hospitalised children
Qiu, H 10 36 0.27778 [0.15646; 0.44369] 1.5%
Tan, YP 2 10 0.20000 [0.05041; 0.54071] 1.1%
See, KC 1 4 0.25000 [0.03351; 0.76216] 0.8%
Tan, X 2 13 0.15385 [0.03874; 0.45063] 1.1%
Xu, H 11 32 0.34375 [0.20163; 0.52070] 1.5%
Song, W 8 16 0.50000 [0.27290; 0.72710] 1.4%
Lu, Y 29 110 0.26364 [0.18980; 0.35365] 1.6%
Bai, K 8 25 0.32000 [0.16881; 0.52163] 1.4%
Melgosa, M 3 16 0.18750 [0.06170; 0.44746] 1.2%
Wu, HP 3 23 0.13043 [0.04267; 0.33545] 1.3%
Random effects model 285 0.27724 [0.22721; 0.33352] 12.8%
Heterogeneity: I^2 = 0%, τ^2 = 0, χ^2_{9} = 8.88 (p = 0.45)
setting = hospitalised children and adults
Yongchen, Z 5 21 0.23810 [0.10272; 0.46034] 1.4%
Kim, SE 10 71 0.14085 [0.07749; 0.24239] 1.5%
Ma, Y 11 47 0.23404 [0.13460; 0.37510] 1.5%
Merza, MA 6 15 0.40000 [0.19178; 0.65193] 1.3%
Alshami, AA 69 128 0.53906 [0.45240; 0.62342] 1.6%
Zhang, W3 26 137 0.18978 [0.13257; 0.26417] 1.6%
Wang, Y2 63 279 0.22581 [0.18053; 0.27858] 1.6%
Choe, PG 15 113 0.13274 [0.08164; 0.20858] 1.6%
Kong, W 45 473 0.09514 [0.07178; 0.12507] 1.6%
Sharma, AK 215 234 0.91880 [0.87622; 0.94761] 1.6%
Random effects model 1518 0.29346 [0.14566; 0.50294] 15.3%
Heterogeneity: I^2 = 97%, τ^2 = 1.9643, χ^2_{9} = 335.34 (p < 0.01)
Random effects model 6602 0.20958 [0.16785; 0.25845] 100.0%
Prediction interval [0.02931; 0.69957]
Heterogeneity: I^2 = 92%, τ^2 = 1.1711, χ^2_{29} = 929.86 (p < 0.01)
Residual heterogeneity: I^2 = 92%, χ^2_{72} = 875.73 (p < 0.01)
0 0.2 0.4 0.6 0.8 1
Prevalence

Based on the subgroup results showing in the forest plot, it is more reasonable to make summary estimates of the proportion of asymptomatic COVID-19 infections in different settings. Contact investigation and screening were both at a relatively high asymptomatic infection rate of 33% and 31% respectively. The summary estimate of the proportion of asymptomatic infections in children (28%) is much higher than in adults (11%). However, we can still observe some overlap in the confidence intervals of these subgroups, which was confirmed by the significant heterogeneity among studies conducted in aggregated contact investigation (I^2 = 84%, p < 0.01), outbreak investigation (I^2 = 85%, p < 0.01), hospitalized adults (I^2 = 96%, p < 0.01) and hospitalized children and adults (I^2 = 97%, p < 0.01). That is, when we grouped the included studies according to different study settings, there was no significant difference in effect sizes between the seven subgroups. This was also supported by the significant unexplained heterogeneity left in the data (I^2 = 92%, p < 0.01), which implied that there might be one or more important missing factors could better explain the heterogeneity.



Bayesian Nonparametric Analysis Using DPmeta {DPpackage}

Bayesian nonparametric analysis was conducted using `DPpackage`. This package contains a series of Bayesian nonparametric models under Dirichlet process. A Bayesian nonparametric model is a Bayesian model on an infinite dimensional parameter space. Its high flexibility and robustness allow for better data modeling than parametric models. A Dirichlet process is a probability distribution whose range is itself a set of probability distributions. It means that instead of generating a single parameter, a single draw from the Dirichlet process outputs another distribution. Therefore, the `DPmeta()` function in the `DPpackage` will generate a posterior density sample for a semiparametric linear mixed effects meta-analysis model using a Dirichlet process or a mixture of Dirichlet process prior for the distribution of the random effects. The highest posterior density (HPD) interval will be used to represent the Bayesian credible intervals. It is the shortest interval among all of the Bayesian credible intervals.

The `DPmeta()` function includes five arguments:

- **formula**: a two-sided linear formula object.
- **prior**: a list giving the prior information.
  - alpha (the value of the precision parameter)
  - mu (the value of the mean of the centering distribution)
  - sigma2 (the value of the variance of the centering distribution)
- **mcmc**: a list giving the MCMC parameters.
  - nburn (the number of burn-in scans)
  - nskip (the thinning interval)
  - nsave (the total number of scans to be saved)
  - ndisplay (the number of saved scans to be displayed on screen)
- **state**: a list giving the current value of the parameters.
- **status**: a logical variable.
  - the run is new (TRUE)
  - the continuation of a previous analysis (FALSE)

`formula` describes how to predict the proportion of asymptomatic COVID-19 infections (y). `prior` gives the prior information (alpha, mu, tau1 and tau2). `mcmc` gives the MCMC parameters (nburn, nsave, nskip, ndisplay). `state` is set to NULL. `status` indicates the run is new. Since we know the distribution of the included studies, we can fairly estimate the value of the mean of the centering distribution (mu) and the hyperparameters for the prior distribution of the variance of the centering distribution (tau1 and tau2). Then by tuning the precision parameter (alpha), the minimum number of clusters was obtained when alpha was 0.01 (See alpha Tuning Plot). The model summary indicated that the overall estimate of the proportion of people who become infected with COVID-19 and remain asymptomatic throughout the course of infection was 22.15%, with a 95% HPD interval of 18.1% to 26.3%. Five clusters were generated. To extract the predictive information of random effects, we used the `DPrandom()` function in the `DPpackage`. It showed that the prediction of the proportion of asymptomatic COVID-19 infection was 22.19%, along with a 95% HPD interval of 1.5% to 46.0%. The trace and density of model parameters are shown in the model summary plots below.

```
set.seed(1234)
# Prior information
prior = list(alpha=0.01, mu=mean(effects), tau1=0.01, tau2=0.01)

# Initial state
state = NULL

# MCMC parameters
nburn = 20000    # the number of burn-in scans
nskip = 20       # the thinning interval
nsave = 10000    # the total number of scans to be saved
ndisplay = 400   # the number of saved scans to be displayed on screen
```

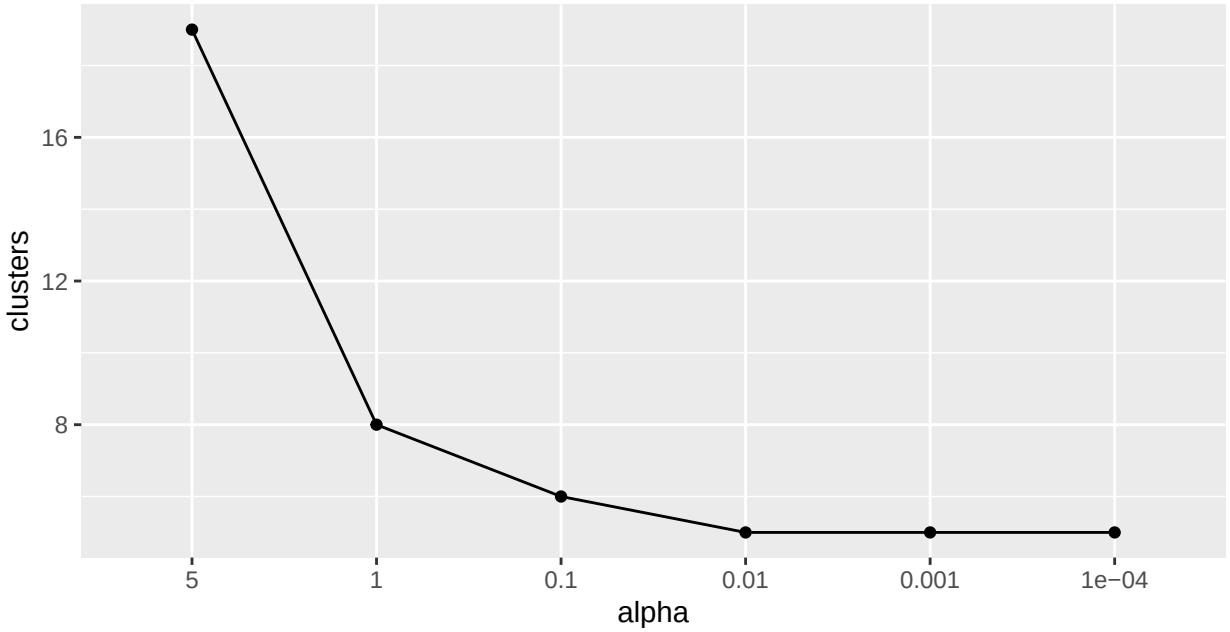
```r
mcmc = list(nburn=nburn, nsave=nsave, nskip=nskip, ndisplay=ndisplay)

# Fit the model: First run
fit = DPmeta(formula=y~1, prior=prior, mcmc=mcmc, state=state, status=TRUE)
```

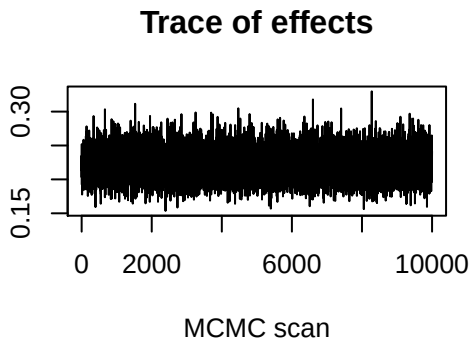
```
##
## MCMC scan 400 of 10000 (CPU time: 1.578 s)
## MCMC scan 800 of 10000 (CPU time: 2.062 s)
## MCMC scan 1200 of 10000 (CPU time: 2.547 s)
## MCMC scan 1600 of 10000 (CPU time: 3.031 s)
## MCMC scan 2000 of 10000 (CPU time: 3.516 s)
## MCMC scan 2400 of 10000 (CPU time: 3.984 s)
## MCMC scan 2800 of 10000 (CPU time: 4.500 s)
## MCMC scan 3200 of 10000 (CPU time: 4.984 s)
## MCMC scan 3600 of 10000 (CPU time: 5.469 s)
## MCMC scan 4000 of 10000 (CPU time: 5.953 s)
## MCMC scan 4400 of 10000 (CPU time: 6.438 s)
## MCMC scan 4800 of 10000 (CPU time: 6.922 s)
## MCMC scan 5200 of 10000 (CPU time: 7.406 s)
## MCMC scan 5600 of 10000 (CPU time: 7.906 s)
## MCMC scan 6000 of 10000 (CPU time: 8.406 s)
## MCMC scan 6400 of 10000 (CPU time: 8.875 s)
## MCMC scan 6800 of 10000 (CPU time: 9.375 s)
## MCMC scan 7200 of 10000 (CPU time: 9.844 s)
## MCMC scan 7600 of 10000 (CPU time: 10.328 s)
## MCMC scan 8000 of 10000 (CPU time: 10.812 s)
## MCMC scan 8400 of 10000 (CPU time: 11.312 s)
## MCMC scan 8800 of 10000 (CPU time: 11.797 s)
## MCMC scan 9200 of 10000 (CPU time: 12.281 s)
## MCMC scan 9600 of 10000 (CPU time: 12.781 s)
## MCMC scan 10000 of 10000 (CPU time: 13.250 s)
```

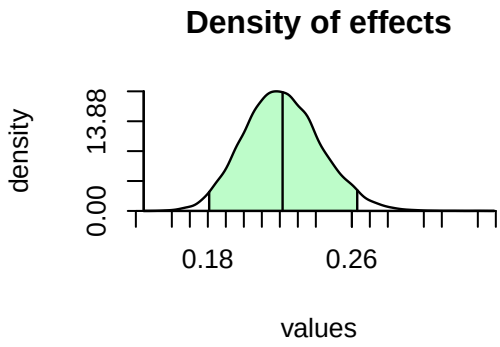
```r
fit
```

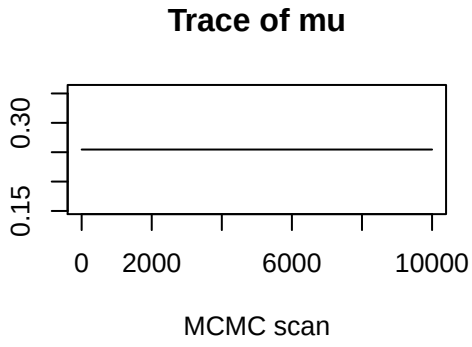
```
##
## Bayesian semiparametric linear mixed effects meta-analysis
##
## Call:
## DPmeta.default(formula = y ~ 1, prior = prior, mcmc = mcmc, state = state,
##     status = TRUE)
##
## Posterior Inference of Parameters:
##  effects        mu    sigma2  ncluster
##   0.2215    0.2545    0.1882    5.0671
##
## Number of Studies: 79
```

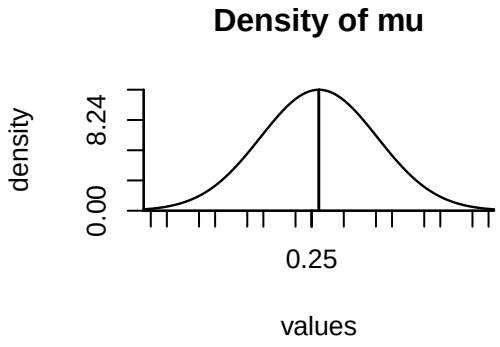


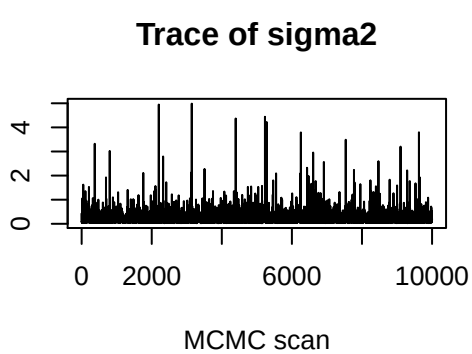

```
##
## Bayesian semiparametric linear mixed effects meta-analysis
##
## Call:
## DPmeta.default(formula = y ~ 1, prior = prior, mcmc = mcmc, state = state,
##     status = TRUE)
##
## Posterior Predictive Distributions (log):
##    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -2.8917   0.3107   0.7064   0.7165   1.1084   3.2312
##
## Model's performance:
##    Dbar     Dhat     pD       DIC      LPML
##    -189.15  -245.36    56.21  -132.94    56.60
##
## Regression coefficients:
##          Mean       Median     Std. Dev.  Naive Std.Error  95%HPD-Low
## effects  0.2215233  0.2205796  0.0213156  0.0002132        0.1807334
##          95%HPD-Upp
## effects  0.2629920
##
## Baseline distribution:
##         Mean      Median    Std. Dev.  Naive Std.Error  95%HPD-Low  95%HPD-Upp
## mu      0.254521  0.254521  0.000000   0.000000         0.254521    0.254521
## sigma2  0.188225  0.128675  0.229249   0.002292         0.025506    0.513459
##
## Precision parameter:
##           Mean      Median    Std. Dev.  Naive Std.Error  95%HPD-Low
## ncluster  5.067100  5.000000  0.256913   0.002569         5.000000
##           95%HPD-Upp
## ncluster  6.000000
##
## Number of Studies: 79
```

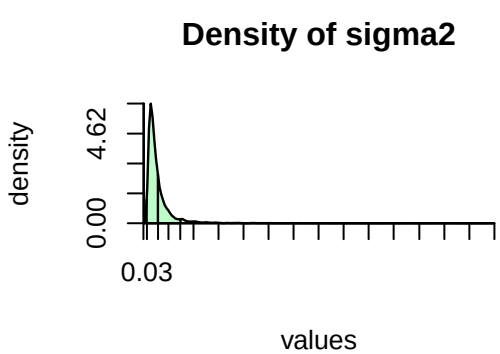

Trace of effects
0.15
0.30
0
2000
6000
10000
MCMC scan

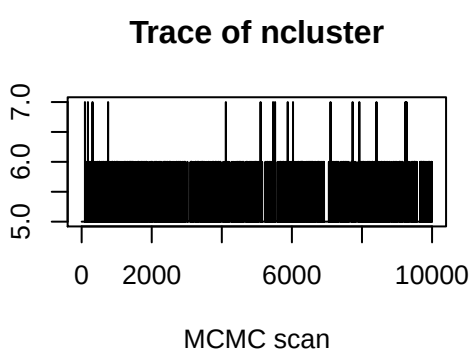

Density of effects
density
0.00
13.88
0.18
0.26
values

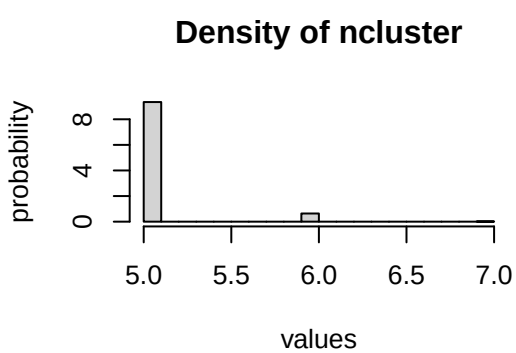

Trace of mu
0.15
0.30
0
2000
6000
10000
MCMC scan

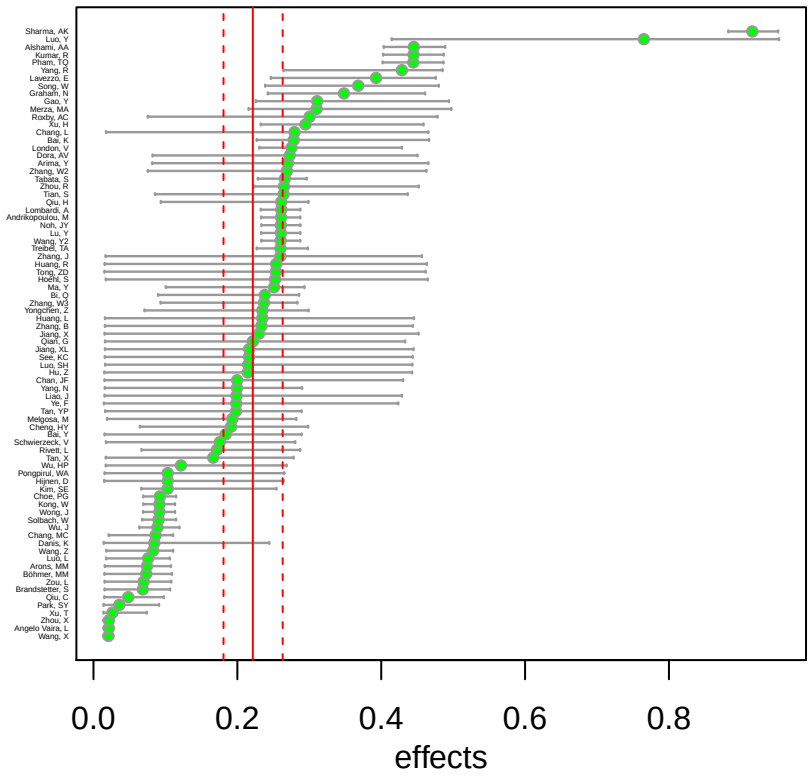

Density of mu
density
0.00
8.24
0.25
values


Trace of sigma2
0
2
4
0
2000
6000
10000
MCMC scan

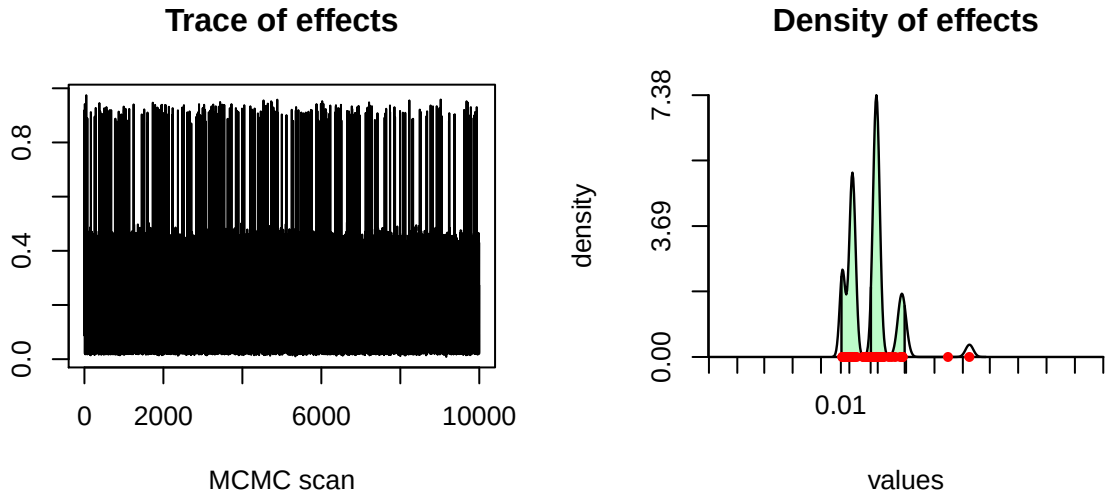

Density of sigma2
density
0.00
4.62
0.03
values


Trace of ncluster
5.0
6.0
7.0
0
2000
6000
10000
MCMC scan


Density of ncluster
probability
0
4
8
5.0
5.5
6.0
6.5
7.0
values

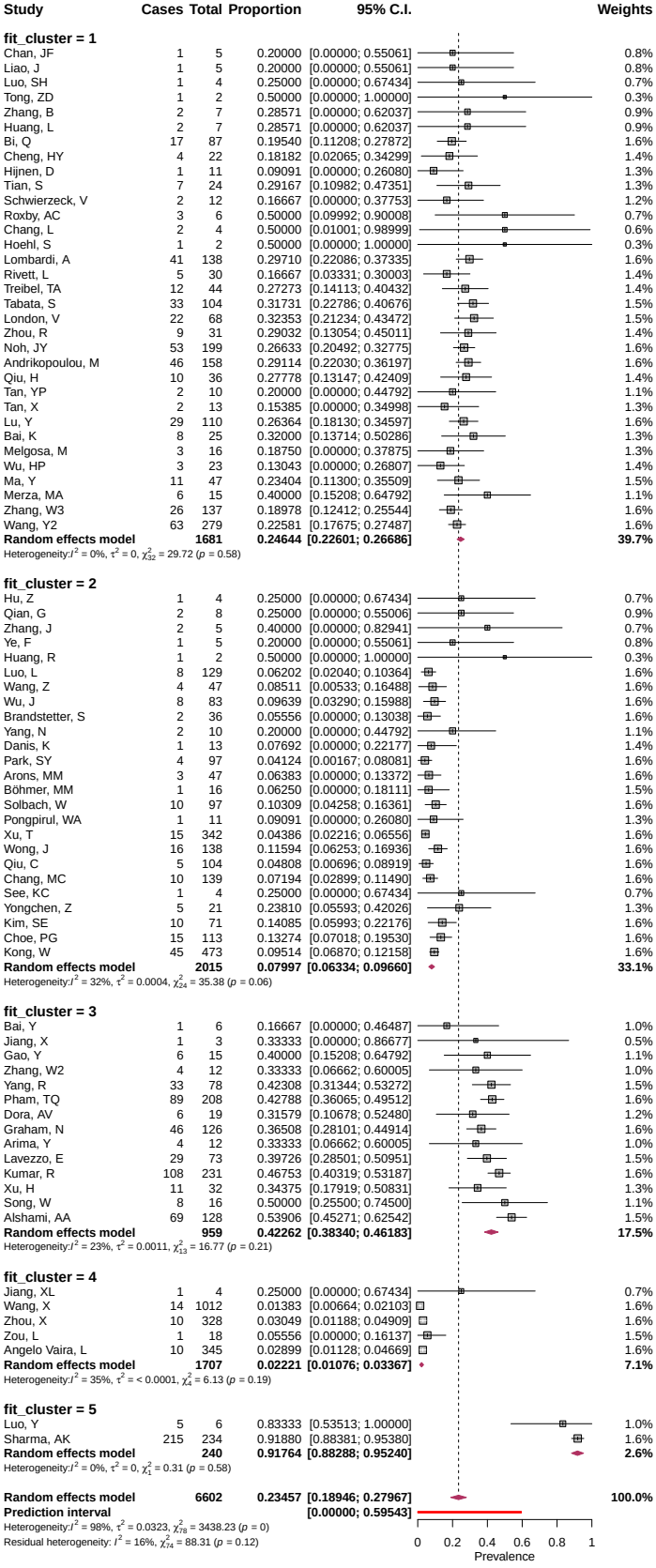



```
##
## Random effect information for the DP object:
##
## Call:
## DPmeta.default(formula = y ~ 1, prior = prior, mcmc = mcmc, state = state,
##     status = TRUE)
##
##
## Predictive distribution:
##         Mean      Median    Std. Dev.  Naive Std.Error  95%HPD-Low  95%HPD-Upp
## theta  0.221928  0.249146  0.167730   0.001677         0.014302    0.459616
```

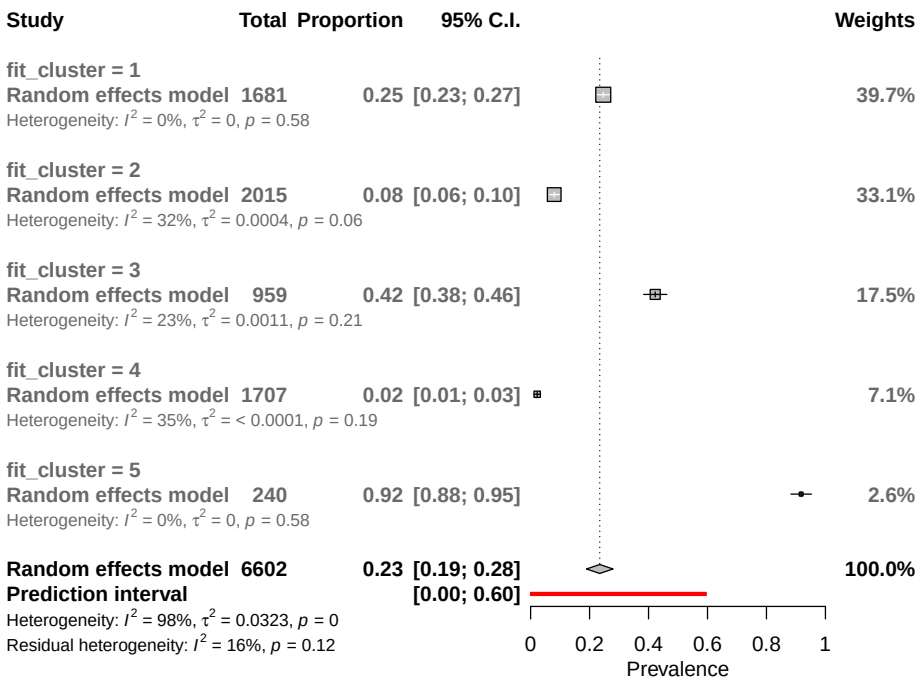


Cluster Analysis

The `DPmeta` model divided the included studies into five clusters. For further study and comparison, I also conducted a subgroup analysis on the five clusters. As shown in the forest plot in presence of clusters, the summary estimates of the proportion of people with asymptomatic COVID-19 infections varied significantly from cluster to cluster. Cluster 2 and cluster 4 were to the left of the overall summary estimate. Cluster 3 and cluster 5 were to the right of the overall summary estimate. Cluster 1 was on the overall summary estimate. There were 33 studies in cluster 1. The summary estimate of these studies was 25% (95% CI 23%-27%). It included studies from all seven settings. In cluster 2, there were 25 studies. The summary estimate of them was 8% (95% CI 6%-10%). It included studies under six settings other than screening. In 14 studies of cluster 3, the summary estimate of the proportion with asymptomatic infections was 42% (95% CI 38%-46%). It also included studies from all seven settings. Cluster 4 had 5 studies with a summary estimate of 2% (95% CI 1%-3%). Two settings were included: contact investigation and hospitalized adults. The remaining two studies were in cluster 5. They came from contact investigation and hospitalized children and adults. The summary estimate of these two studies was 92% (95% CI 88%-95%).


Study Cases Total Proportion 95% C.I. Weights
fit_cluster = 1
Chan, JF 1 5 0.20000 [0.00000; 0.55061] 0.8%
Liao, J 1 5 0.20000 [0.00000; 0.55061] 0.8%
Luo, SH 1 4 0.25000 [0.00000; 0.67434] 0.7%
Tong, ZD 1 2 0.50000 [0.00000; 1.00000] 0.3%
Zhang, B 2 7 0.28571 [0.00000; 0.62037] 0.9%
Huang, L 2 7 0.28571 [0.00000; 0.62037] 0.9%
Bi, Q 17 87 0.19540 [0.11208; 0.27872] 1.6%
Cheng, HY 4 22 0.18182 [0.02065; 0.34299] 1.4%
Hijnen, D 1 11 0.09091 [0.00000; 0.26080] 1.3%
Tian, S 7 24 0.29167 [0.10982; 0.47351] 1.3%
Schwierzeck, V 2 12 0.16667 [0.00000; 0.37753] 1.2%
Roxby, AC 3 6 0.50000 [0.09992; 0.90008] 0.7%
Chang, L 2 4 0.50000 [0.01001; 0.98999] 0.6%
Hoehl, S 1 2 0.50000 [0.00000; 1.00000] 0.3%
Lombardi, A 41 138 0.29710 [0.22086; 0.37335] 1.6%
Rivett, L 5 30 0.16667 [0.03331; 0.30003] 1.4%
Treibel, TA 12 44 0.27273 [0.14113; 0.40432] 1.4%
Tabata, S 33 104 0.31731 [0.22786; 0.40676] 1.5%
London, V 22 68 0.32353 [0.21234; 0.43472] 1.5%
Zhou, R 9 31 0.29032 [0.13054; 0.45011] 1.4%
Noh, JY 53 199 0.26633 [0.20492; 0.32775] 1.6%
Andrikopoulou, M 46 158 0.29114 [0.22030; 0.36197] 1.6%
Qiu, H 10 36 0.27778 [0.13147; 0.42409] 1.4%
Tan, YP 2 10 0.20000 [0.00000; 0.44792] 1.1%
Tan, X 2 13 0.15385 [0.00000; 0.34998] 1.3%
Lu, Y 29 110 0.26364 [0.18130; 0.34597] 1.6%
Bai, K 8 25 0.32000 [0.13714; 0.50286] 1.3%
Melgosa, M 3 16 0.18750 [0.00000; 0.37875] 1.3%
Wu, HP 3 23 0.13043 [0.00000; 0.26807] 1.4%
Ma, Y 11 47 0.23404 [0.11300; 0.35509] 1.5%
Merza, MA 6 15 0.40000 [0.15208; 0.64792] 1.1%
Zhang, W3 26 137 0.18978 [0.12412; 0.25544] 1.6%
Wang, Y2 63 279 0.22581 [0.17675; 0.27487] 1.6%
Random effects model 1681 0.24644 [0.22601; 0.26686] 39.7%
Heterogeneity: I^2 = 0%, τ^2 = 0, χ^2_{32} = 29.72 (p = 0.58)
fit_cluster = 2
Hu, Z 1 4 0.25000 [0.00000; 0.67434] 0.7%
Qian, G 2 8 0.25000 [0.00000; 0.55006] 0.9%
Zhang, J 2 5 0.40000 [0.00000; 0.82941] 0.7%
Ye, F 1 5 0.20000 [0.00000; 0.55061] 0.8%
Huang, R 1 2 0.50000 [0.00000; 1.00000] 0.3%
Luo, L 8 129 0.06202 [0.02040; 0.10364] 1.6%
Wang, Z 4 47 0.08511 [0.00533; 0.16488] 1.6%
Wu, J 8 83 0.09639 [0.03290; 0.15988] 1.6%
Brandstetter, S 2 36 0.05556 [0.00000; 0.13038] 1.6%
Yang, N 2 10 0.20000 [0.00000; 0.44792] 1.1%
Danis, K 1 13 0.07692 [0.00000; 0.22177] 1.4%
Park, SY 4 97 0.04124 [0.00167; 0.08081] 1.6%
Arons, MM 3 47 0.06383 [0.00000; 0.13372] 1.6%
Böhmer, MM 1 16 0.06250 [0.00000; 0.18111] 1.5%
Solbach, W 10 97 0.10309 [0.04258; 0.16361] 1.6%
Pongpirul, WA 1 11 0.09091 [0.00000; 0.26080] 1.3%
Xu, T 15 342 0.04386 [0.02216; 0.06556] 1.6%
Wong, J 16 138 0.11594 [0.06253; 0.16936] 1.6%
Qiu, C 5 104 0.04808 [0.00696; 0.08919] 1.6%
Chang, MC 10 139 0.07194 [0.02899; 0.11490] 1.6%
See, KC 1 4 0.25000 [0.00000; 0.67434] 0.7%
Yongchen, Z 5 21 0.23810 [0.05593; 0.42026] 1.3%
Kim, SE 10 71 0.14085 [0.05993; 0.22176] 1.6%
Choe, PG 15 113 0.13274 [0.07018; 0.19530] 1.6%
Kong, W 45 473 0.09514 [0.06870; 0.12158] 1.6%
Random effects model 2015 0.07997 [0.06334; 0.09660] 33.1%
Heterogeneity: I^2 = 32%, τ^2 = 0.0004, χ^2_{24} = 35.38 (p = 0.06)
fit_cluster = 3
Bai, Y 1 6 0.16667 [0.00000; 0.46487] 1.0%
Jiang, X 1 3 0.33333 [0.00000; 0.86677] 0.5%
Gao, Y 6 15 0.40000 [0.15208; 0.64792] 1.1%
Zhang, W2 4 12 0.33333 [0.06662; 0.60005] 1.0%
Yang, R 33 78 0.42308 [0.31344; 0.53272] 1.5%
Pham, TQ 89 208 0.42788 [0.36065; 0.49512] 1.6%
Dora, AV 6 19 0.31579 [0.10678; 0.52480] 1.2%
Graham, N 46 126 0.36508 [0.28101; 0.44914] 1.6%
Arima, Y 4 12 0.33333 [0.06662; 0.60005] 1.0%
Lavezzo, E 29 73 0.39726 [0.28501; 0.50951] 1.5%
Kumar, R 108 231 0.46753 [0.40319; 0.53187] 1.6%
Xu, H 11 32 0.34375 [0.17919; 0.50831] 1.3%
Song, W 8 16 0.50000 [0.25500; 0.74500] 1.1%
Alshami, AA 69 128 0.53906 [0.45271; 0.62542] 1.5%
Random effects model 959 0.42262 [0.38340; 0.46183] 17.5%
Heterogeneity: I^2 = 23%, τ^2 = 0.0011, χ^2_{13} = 16.77 (p = 0.21)
fit_cluster = 4
Jiang, XL 1 4 0.25000 [0.00000; 0.67434] 0.7%
Wang, X 14 1012 0.01383 [0.00664; 0.02103] 1.6%
Zhou, X 10 328 0.03049 [0.01188; 0.04909] 1.6%
Zou, L 1 18 0.05556 [0.00000; 0.16137] 1.5%
Angelo Vaira, L 10 345 0.02899 [0.01128; 0.04669] 1.6%
Random effects model 1707 0.02221 [0.01076; 0.03367] 7.1%
Heterogeneity: I^2 = 35%, τ^2 = < 0.0001, χ^2_4 = 6.13 (p = 0.19)
fit_cluster = 5
Luo, Y 5 6 0.83333 [0.53513; 1.00000] 1.0%
Sharma, AK 215 234 0.91880 [0.88381; 0.95380] 1.6%
Random effects model 240 0.91764 [0.88288; 0.95240] 2.6%
Heterogeneity: I^2 = 0%, τ^2 = 0, χ^2_1 = 0.31 (p = 0.58)
Random effects model 6602 0.23457 [0.18946; 0.27967] 100.0%
Prediction interval [0.00000; 0.59543]
Heterogeneity: I^2 = 98%, τ^2 = 0.0323, χ^2_{76} = 3438.23 (p = 0)
Residual heterogeneity: I^2 = 16%, χ^2_{74} = 88.31 (p = 0.12)
0 0.2 0.4 0.6 0.8 1
Prevalence

As shown in the forest plot of cluster results below, there were significant differences in the proportion of asymptomatic COVID-19 infections among the five clusters generated by the `DPmeta` model. Cluster 5 had an extremely high asymptomatic infection rate of 92% (95% CI 88%-95%). Cluster 4 had the lowest asymptomatic infection rate of 2% (95% CI 1%-3%). It is reasonable to say that the summary estimates of the proportion of asymptomatic COVID-19 infections based on the five clusters are more appropriate than based on the seven study settings. There was only a small amount of heterogeneity in cluster 2 (32%), cluster3 (23%), and cluster 4 (35%). The residual heterogeneity in the data was also only 16%. However, due to the lack of individual level data on COVID-19, we cannot explain the underlying factors behind the clusters at this stage.



Comparison

The overall summary estimate of the proportion of asymptomatic COVID-19 infections by Bayesian nonparametric analysis using the `DPpackage` (22.15%, 95% HPD 18.1%-26.3%) is slightly higher than by meta-analysis using the `meta` and `metafor` packages (20.96%, 95% CI 16.8%-25.8%). However, the prediction interval of Bayesian nonparametric analysis (3.0%-70%) is much narrower than that of meta-analysis (1.5%-46.0%), which enables a more precise proportion estimate of the asymptomatic COVID-19 infections. Also, cluster analysis based on the `DPmeta` model can better explain the heterogeneity among included studies than subgroup analysis based on study settings. However, the factors behind the clusters cannot be summarized since individual level data on COVID-19 is not available.

Conclusion

We conducted the Bayesian nonparametric and meta analyses to address three questions: 1) What is the proportion of asymptomatic infections among people who infected with COVID-19? 2) What are the factors of asymptomatic infection in COVID-19? 3) And most importantly, is it reasonable to synthesize all the studies to produce a single estimate? The overall summary estimate of the proportion of asymptomatic COVID-19 infections in both methods were around 20%. However, it is not appropriate to synthesize the proportions of all the studies into a single estimate. Subgroup analysis can be conducted to explain the heterogeneity between studies, but not all factors can be summarized in the available data. The variable `'setting'` can only be part of the contributing factors. Further analysis is needed as more studies and features become available. Therefore, preventive measures such as masks and social distancing will continue to be needed to reduce transmission.

References

WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/

Müller P, Quintana FA, Jara A, Hanson (2015) Bayesian Nonparametric Data Analysis. Springer, New York

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing Meta-Analysis in R: A Hands-on Guide. DOI: 10.5281/zenodo.2551803. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011 Feb 10;342:d549. doi: 10.1136/bmj.d549. PMID: 21310794.

He W, Yi GY, Zhu Y. Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. J Med Virol. 2020 May 29:10.1002/jmv.26041. doi: 10.1002/jmv.26041. Epub ahead of print. PMID: 32470164; PMCID: PMC7283745.

Zheng B, Wang H, Yu C. An increasing public health burden arising from children infected with SARS-CoV2: A systematic review and meta-analysis. Pediatr Pulmonol. 2020 Dec;55(12):3487-3496. doi: 10.1002/ppul.25008. Epub 2020 Sep 25. PMID: 32757374; PMCID: PMC7436588.

He J, Guo Y, Mao R, Zhang J. Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. J Med Virol. 2020;1–11. https://doi.org/10.1002/jmv.26326

Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998 Apr 30;17(8):857-72. doi: 10.1002/(sici)1097-0258(19980430)17:8<857::aid-sim777>3.0.co;2-e. PMID: 9595616.

Heneghan C, Brassey J, Jefferson T. Covid-19: What Proportion Are Asymptomatic? Oxford: Centre for Evidence Based Medicine; 2020 [cited 2020 Jul 27]. Available from: https://www.cebm.net/covid-19/covid-19-what-proportion-are-asymptomatic/

Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. (2020) Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and metaanalysis. PLoS Med 17(9): e1003346. https://doi.org/10.1371/journal.pmed.1003346

Jara, A., Hanson, T. E., Quintana, F. A., Müller, P., & Rosner, G. L. (2011). DPpackage: Bayesian semi- and nonparametric modeling in R. Journal of Statistical Software, 40(5), 1-30. https://doi.org/10.18637/jss.v040.i05