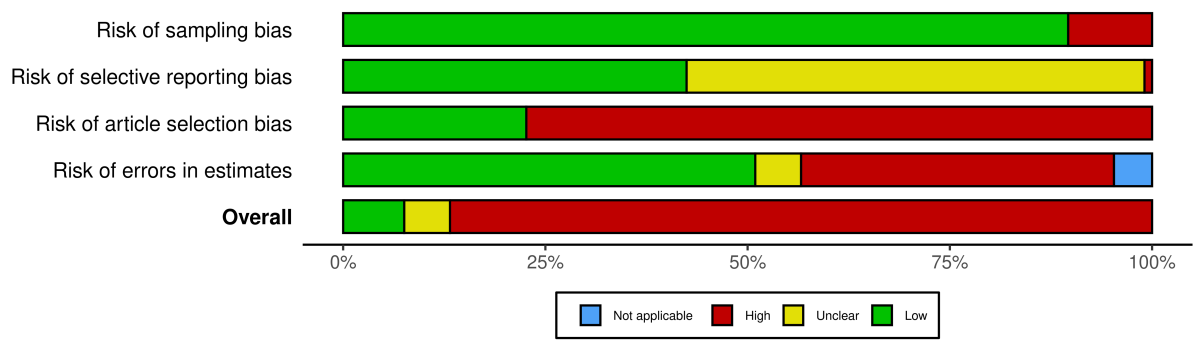


## SUPPLEMENTARY FIGURES AND TABLES

**Supplementary Figure 1. Summary of the results of the risk of bias assessments.**



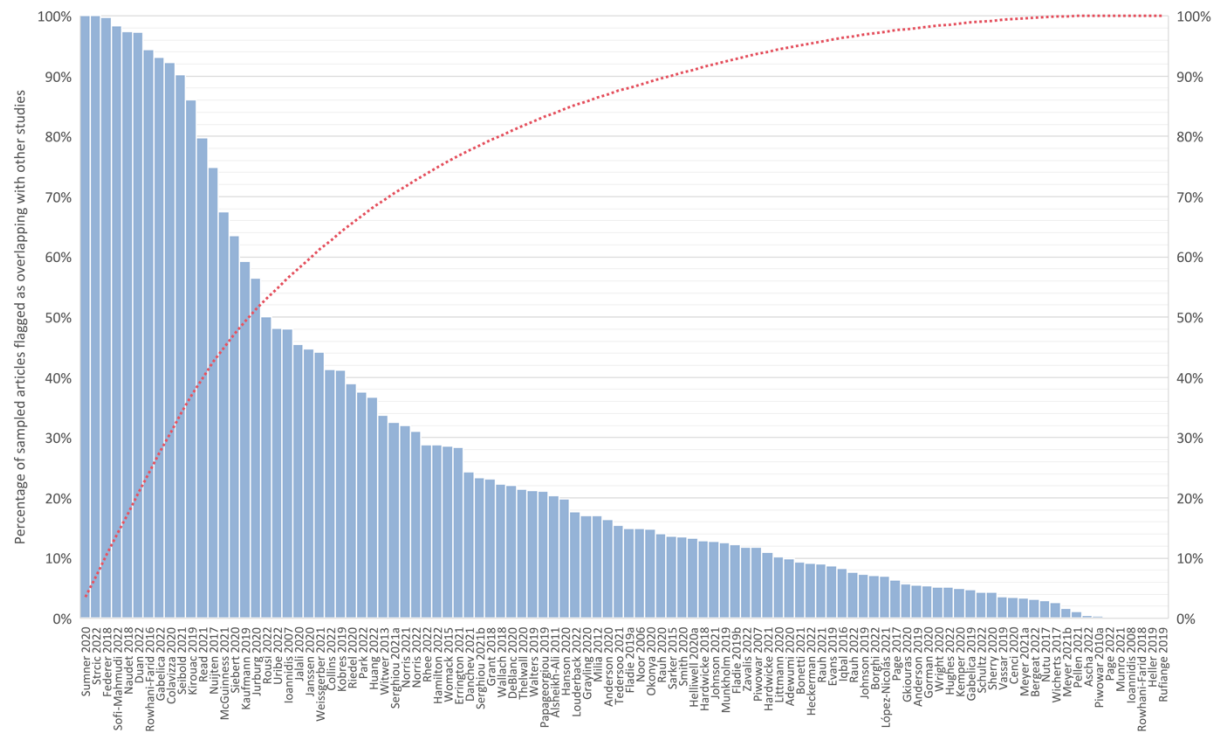
**Supplementary Figure 2. Risk of bias traffic light plot.**

| Study              | Risk of bias |    |    |    |         | Study             | Risk of bias |    |    |    |         |
|--------------------|--------------|----|----|----|---------|-------------------|--------------|----|----|----|---------|
|                    | D1           | D2 | D3 | D4 | Overall |                   | D1           | D2 | D3 | D4 | Overall |
| Adewumi 2020       | +            | +  | -  | +  | -       | Sherry 2020       | +            | +  | -  | +  | -       |
| Alsheikh-Ali 2011  | -            | -  | +  | -  | -       | Siebert 2020      | +            | +  | +  | +  | +       |
| Anderson 2020      | +            | +  | -  | +  | -       | Smith 2020        | +            | +  | -  | +  | -       |
| Bonetti 2021       | +            | -  | +  | +  | -       | Sumner 2020       | +            | -  | +  | -  | -       |
| Cenci 2020         | +            | -  | -  | -  | -       | Thelwall 2020     | +            | -  | +  | +  | -       |
| Colavizza 2020     | +            | -  | -  | -  | -       | Vassar 2019       | +            | -  | -  | +  | -       |
| Danchev 2021       | +            | -  | -  | +  | -       | Wallach 2018      | +            | +  | -  | -  | -       |
| DeBlanc 2020       | +            | -  | -  | -  | -       | Walters 2019      | +            | +  | -  | +  | -       |
| Evans 2019         | +            | +  | -  | +  | -       | Wicherts 2017     | +            | -  | -  | ○  | -       |
| Federer 2018       | +            | -  | -  | -  | -       | Witwer 2013       | +            | -  | -  | -  | -       |
| Fladie 2019a       | +            | +  | -  | +  | -       | Womack 2015       | +            | -  | -  | -  | -       |
| Fladie 2019b       | +            | +  | -  | +  | -       | Wright 2020       | +            | +  | -  | +  | -       |
| Gabelica 2019      | +            | -  | -  | +  | -       | Anderson 2019     | +            | +  | -  | +  | -       |
| Gkiouras 2020      | +            | -  | -  | +  | -       | Ascha 2022        | +            | +  | -  | +  | -       |
| Gorman 2020        | -            | -  | +  | -  | -       | Bergeat 2022      | -            | +  | -  | +  | -       |
| Grant 2018         | +            | -  | -  | +  | -       | Borghi 2022       | +            | +  | -  | +  | -       |
| Grayling 2020      | +            | -  | +  | -  | -       | Collins 2022      | +            | -  | -  | -  | -       |
| Hanson 2020        | -            | -  | -  | -  | -       | Duan 2022         | +            | -  | -  | +  | -       |
| Hardwicke 2021     | +            | +  | +  | +  | +       | Errington 2021    | -            | +  | +  | +  | -       |
| Helliwell 2020a    | +            | -  | -  | -  | -       | Gabelica 2022     | +            | -  | -  | -  | -       |
| Ioannidis 2007     | +            | -  | -  | +  | -       | Hamilton 2022     | +            | +  | +  | -  | -       |
| Ioannidis 2008     | +            | -  | -  | +  | -       | Hardwicke 2018    | +            | -  | +  | -  | -       |
| Iqbal 2016         | +            | +  | +  | +  | +       | Heckerman 2022    | +            | +  | -  | +  | -       |
| Janssen 2020       | +            | -  | -  | -  | -       | Heller 2019       | +            | -  | -  | -  | -       |
| Johnson 2018       | -            | -  | -  | -  | -       | Huang 2022        | +            | -  | -  | -  | -       |
| Johnson 2019       | +            | +  | -  | +  | -       | Hughes 2022       | +            | +  | -  | +  | -       |
| Jurburg 2020       | +            | -  | -  | -  | -       | Jalali 2020       | -            | -  | -  | +  | -       |
| Kaufmann 2019      | +            | -  | -  | -  | -       | Johnson 2021      | +            | +  | -  | +  | -       |
| Kemper 2020        | -            | -  | -  | -  | -       | Kirouac 2019      | +            | -  | +  | -  | -       |
| Kobres 2019        | +            | -  | -  | -  | -       | Littmann 2020     | +            | -  | -  | +  | -       |
| López-Nicolás 2021 | +            | -  | -  | +  | -       | Louderback 2022   | +            | +  | -  | -  | -       |
| McGuinness 2021    | +            | +  | +  | +  | +       | Meyer 2021        | +            | -  | -  | -  | -       |
| Milia 2012         | +            | -  | -  | -  | -       | Miyakawa 2020     | +            | -  | -  | ○  | -       |
| Naudet 2018        | +            | +  | -  | +  | -       | Munkholm 2019     | +            | +  | -  | +  | -       |
| Noor 2006          | +            | -  | -  | -  | -       | Munro 2021        | +            | -  | -  | +  | -       |
| Nutu 2017          | +            | +  | -  | -  | -       | Norris 2022       | +            | +  | -  | +  | -       |
| Okonya 2020        | +            | +  | -  | +  | -       | Norris 2021       | +            | +  | -  | +  | -       |
| Papageorgiou 2019  | +            | +  | -  | +  | -       | Nuijten 2017      | +            | +  | +  | +  | +       |
| Pellen 2021        | +            | +  | -  | +  | -       | Page 2017         | +            | -  | -  | -  | -       |
| Piwowar 2007       | +            | -  | +  | -  | -       | Page 2022         | +            | +  | +  | +  | +       |
| Piwowar 2010a      | +            | -  | +  | -  | -       | Park 2022         | +            | -  | +  | +  | -       |
| Rauh 2020          | +            | +  | -  | +  | -       | Rauh 2021         | +            | +  | -  | +  | -       |
| Rauh 2022          | +            | +  | -  | +  | -       | Read 2021         | +            | +  | -  | -  | -       |
| Reidpath 2001      | +            | -  | -  | ○  | -       | Rhee 2022         | +            | -  | -  | +  | -       |
| Riedel 2020        | +            | -  | -  | -  | -       | Rousi 2022        | +            | -  | -  | -  | -       |
| Rowhani-Farid 2016 | +            | -  | -  | -  | -       | Schultz 2022      | +            | +  | +  | +  | +       |
| Rowhani-Farid 2018 | +            | -  | -  | -  | -       | Soft-Mahmudi 2022 | +            | +  | -  | -  | -       |
| Rufiange 2019      | -            | -  | -  | +  | -       | Stric 2022        | +            | +  | +  | +  | +       |
| Sarker 2015        | +            | -  | -  | -  | -       | Tedersoo 2021     | -            | -  | -  | -  | -       |
| Savage 2009        | -            | -  | -  | ○  | -       | Uribe 2022        | +            | +  | -  | -  | -       |
| Seibold 2021       | +            | -  | +  | -  | -       | Vanpaemel 2015    | +            | -  | -  | ○  | -       |
| Serghiou 2021a     | +            | +  | -  | +  | -       | Weissgerber 2021  | +            | -  | +  | -  | -       |
| Serghiou 2021b     | +            | -  | +  | -  | -       | Zavails 2022      | +            | +  | -  | -  | -       |

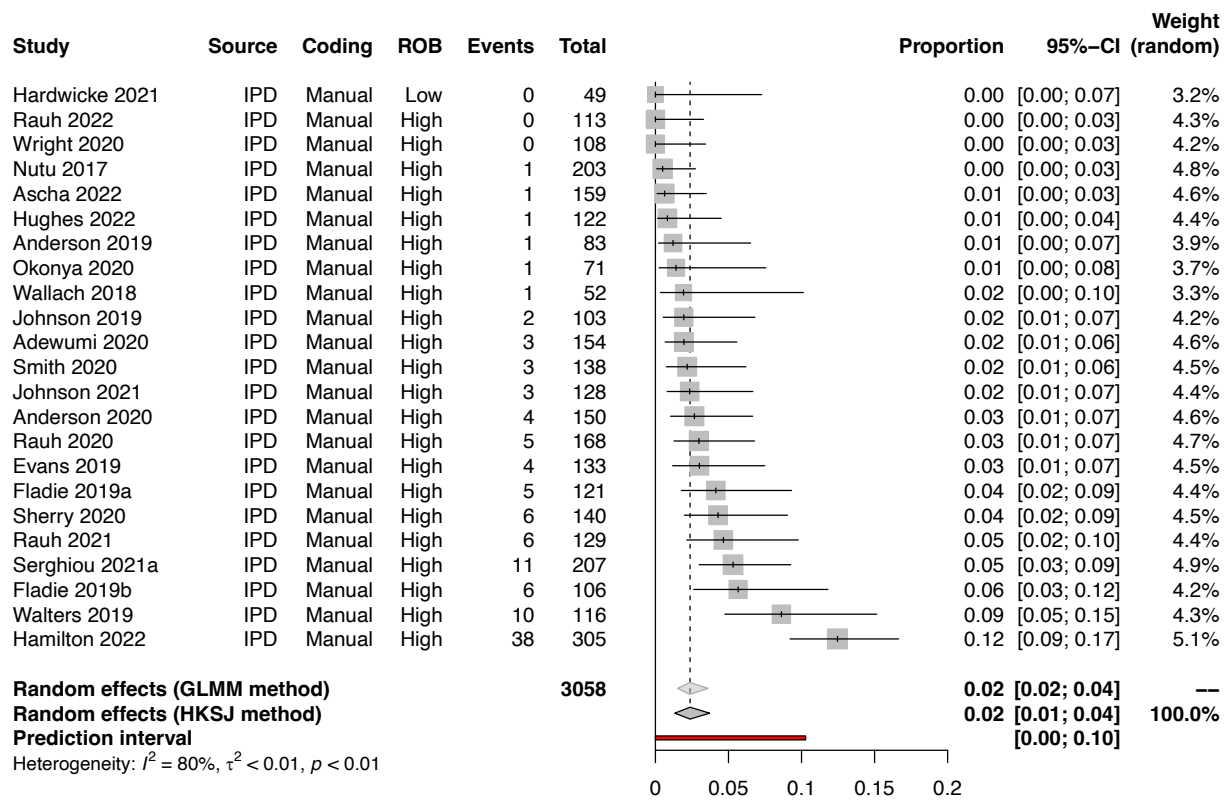
D1: Risk of sampling bias  
D2: Risk of selective reporting bias  
D3: Risk of article selection bias  
D4: Risk of errors in estimates

Judgement  
- High (Red circle with X)  
- Unclear (Yellow circle with -)  
- Low (Green circle with +)  
- Not applicable (Grey circle with ○)

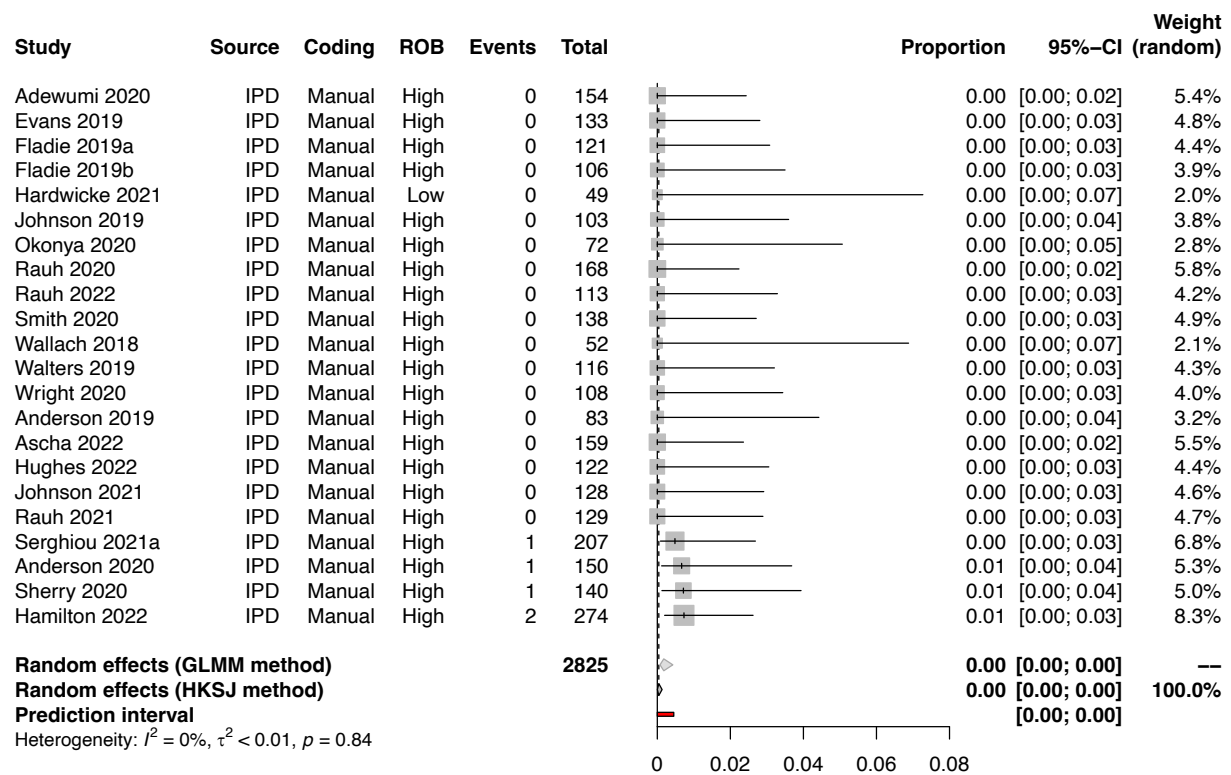
**Supplementary Figure 3. Proportion of primary articles assessed by eligible meta-research studies flagged as potential duplicates.**



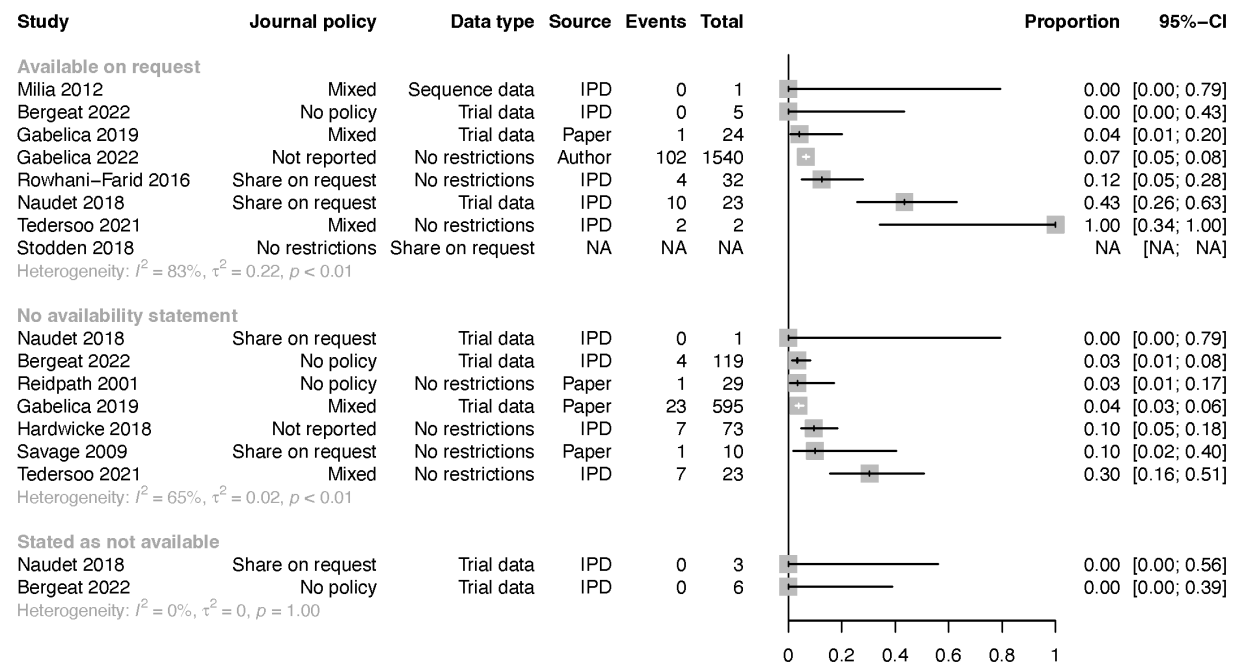
Supplementary Figure 4. Declared private data sharing rates since 2016.



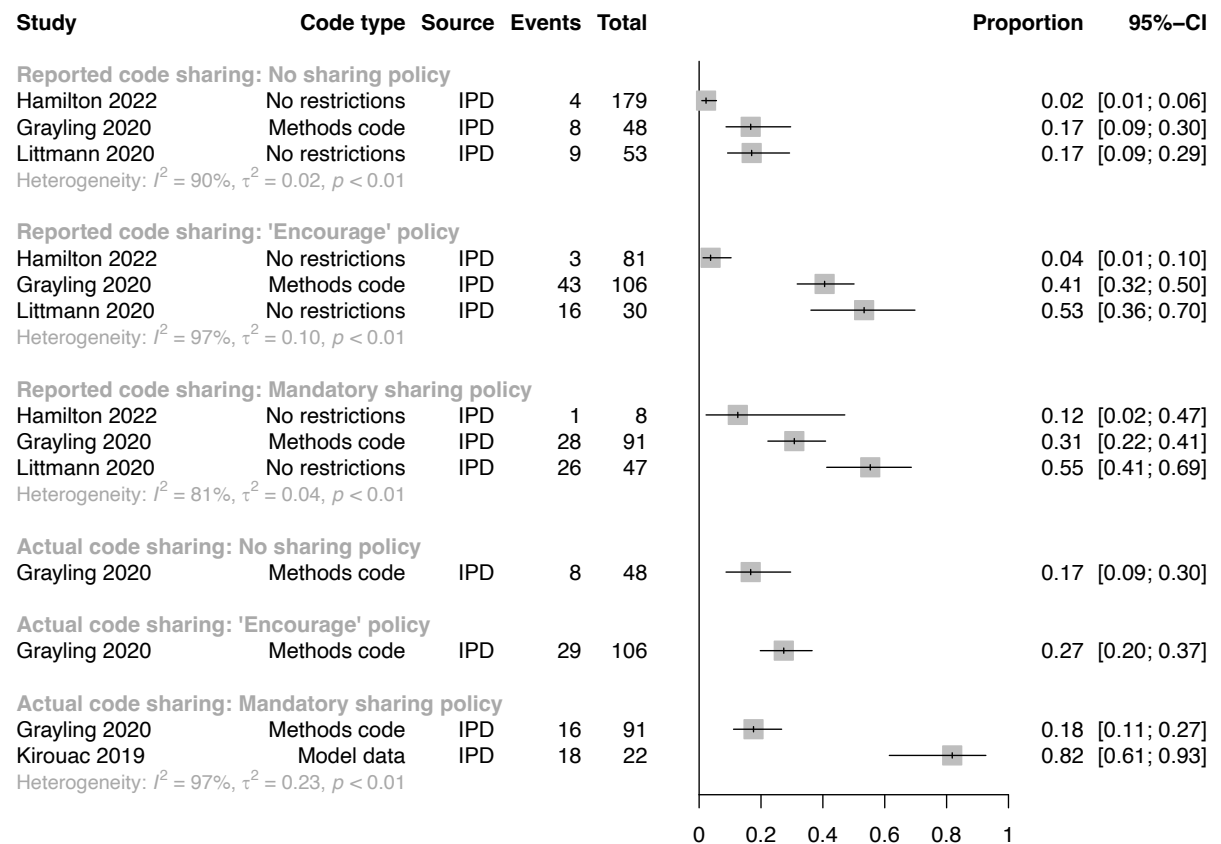
Supplementary Figure 5. Declared private code sharing rates since 2016.



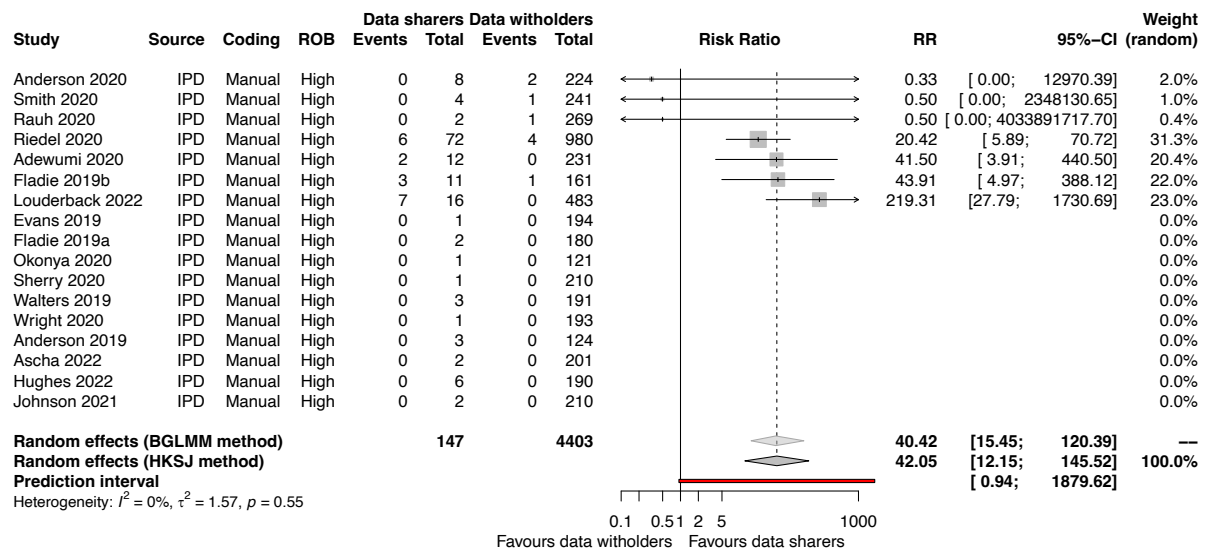
**Supplementary Figure 6. Success rates of private requests for data from published medical research by declaration type.**



**Supplementary Figure 7. Declared and actual public code sharing rates by journal code sharing policy.**

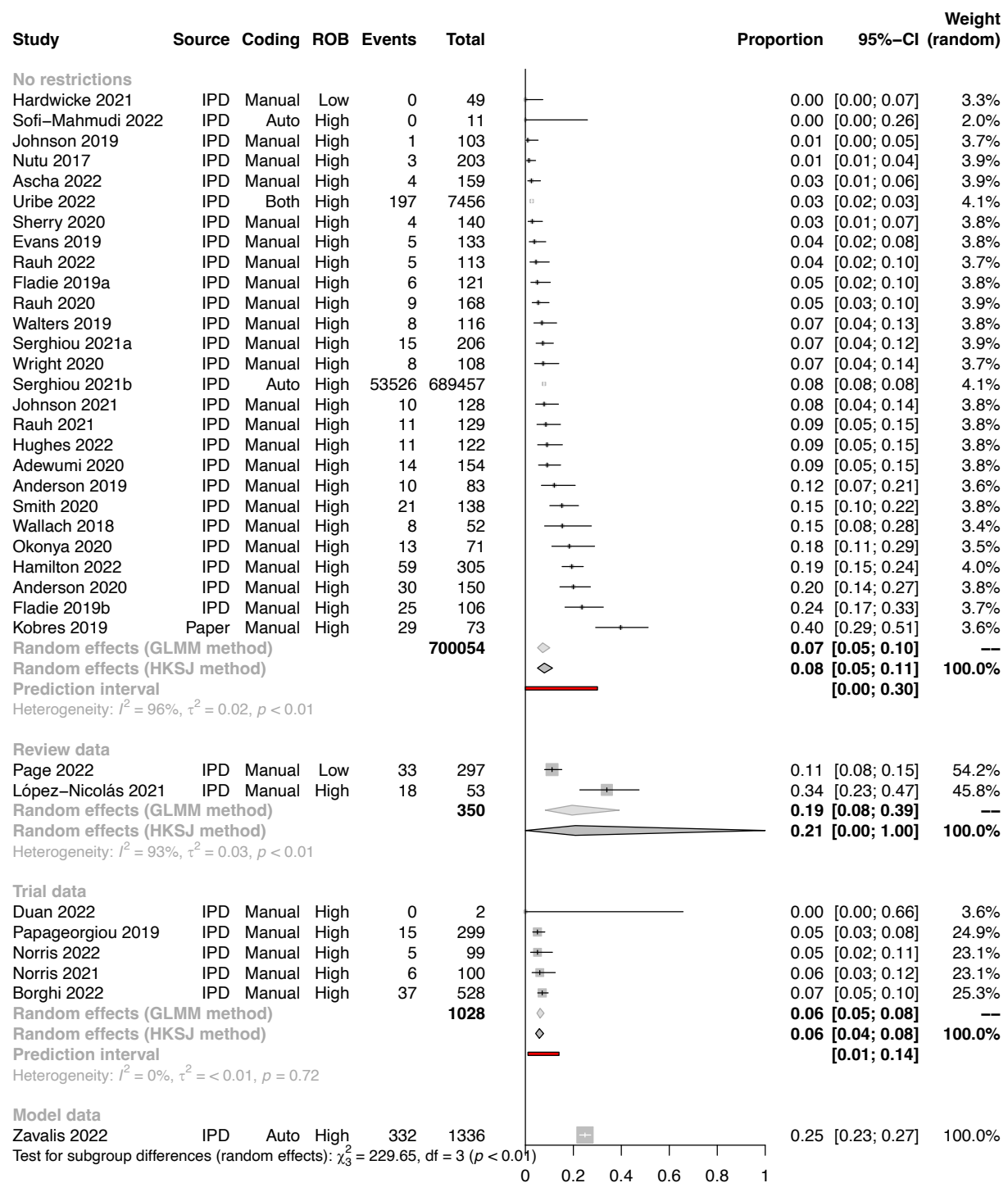


**Supplementary Figure 8. Association between data sharing and code sharing (actual availability).**

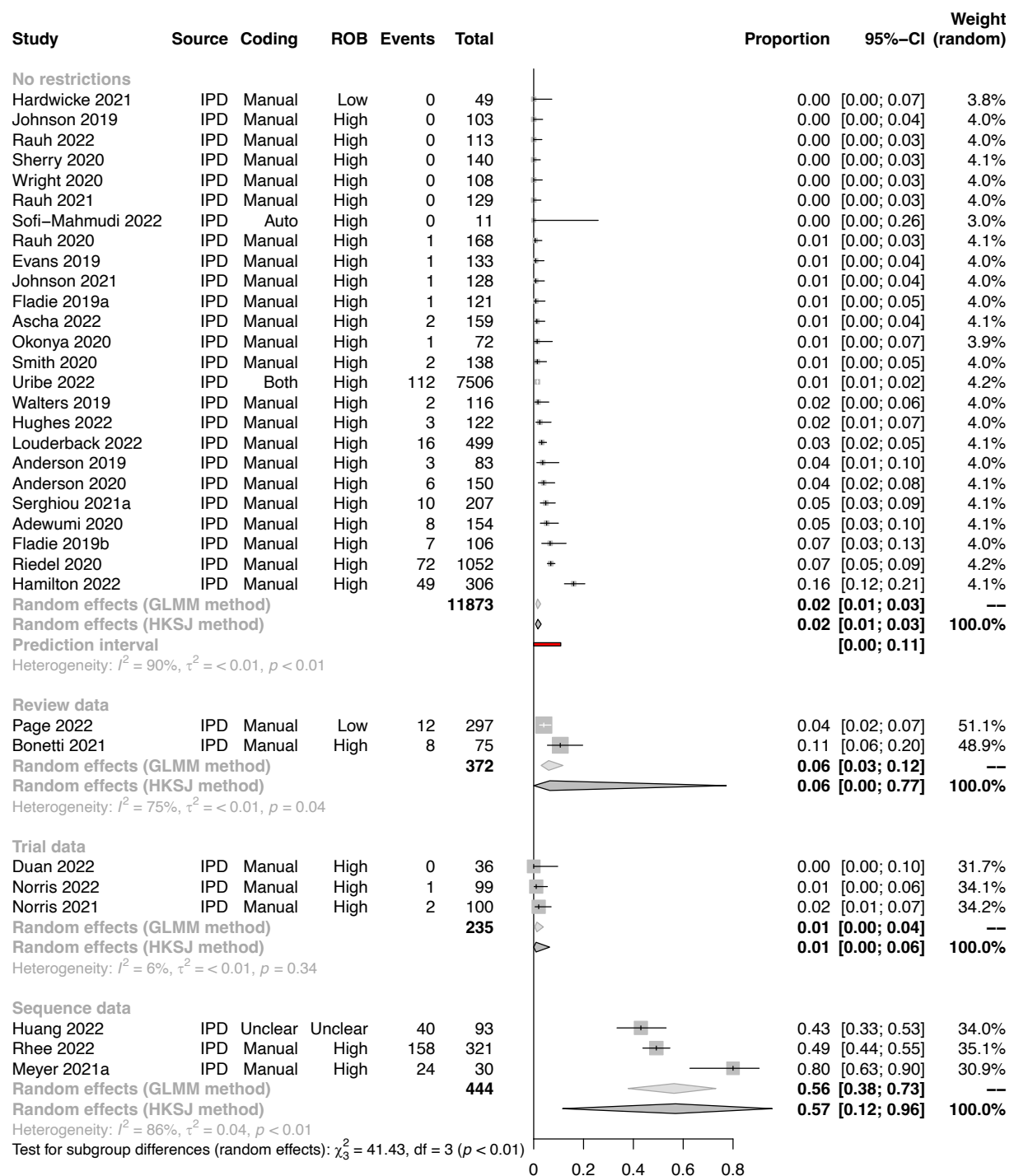




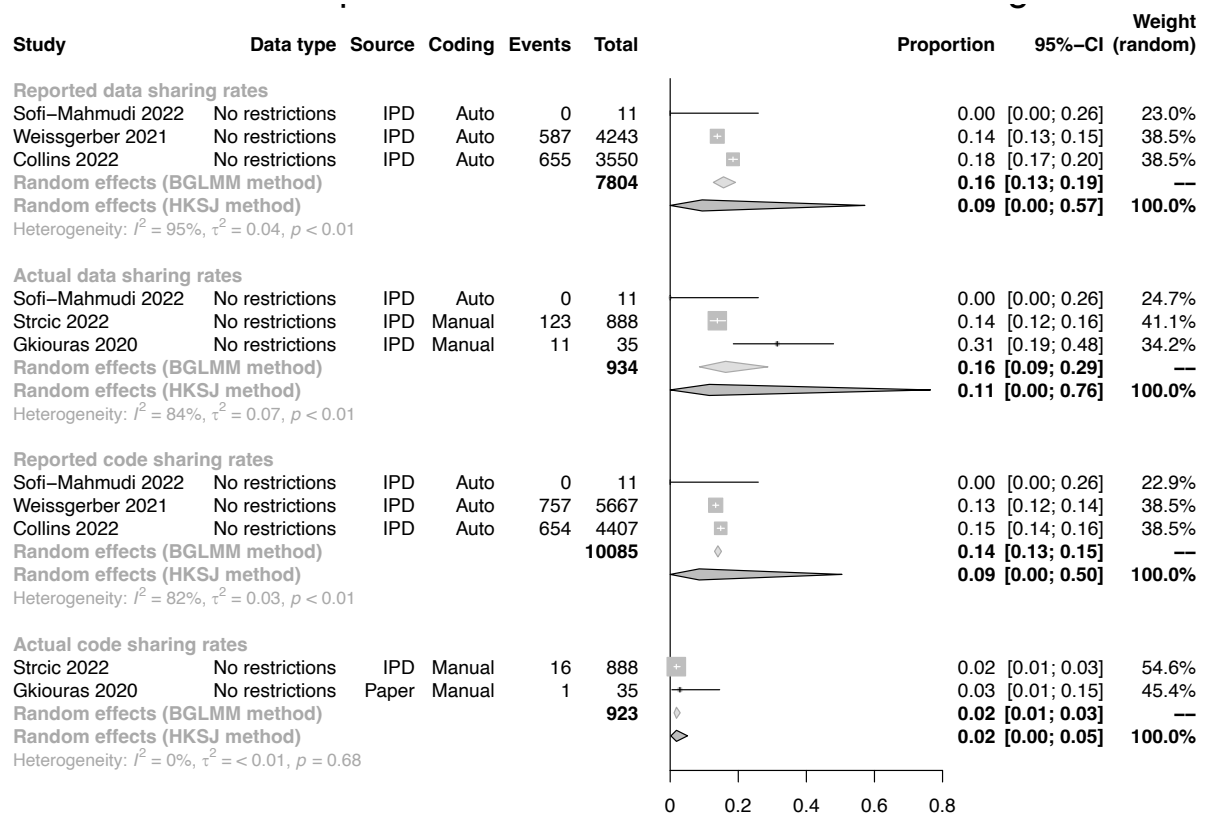
Supplementary Figure 9. Declared public data sharing rates since 2016 by data type.



**Supplementary Figure 10. Actual public data sharing rates since 2016 by data type.**



**Supplementary Figure 11. Data and code sharing rates among studies investigating COVID-19.**



**Supplementary Table 1. Deviations from the original review protocol.**

| Original plan  | Revised plan   | Reason for modification  |
|--|--|--|
| <p>Please refer to Table 1 in the review protocol (<a href="https://f1000research.com/articles/10-491/v2-T1">https://f1000research.com/articles/10-491/v2-T1</a>).</p> | <p>Two extra options were added to the ‘risk of sampling bias’ (“Sampled whole population of interest”) and ‘risk of article selection bias’ categories (“No article screening performed”). Both items were considered low risk of bias.</p> | <p>The two options were added to account for meta-research articles that assessed all articles within a population of interest (e.g., assessed all articles indexed in PubMed Central a la Serghiou et al. (2021) [14]) or did not perform article screening respectively.</p>   |
| <p>Prevalence rates will be transformed using the <u>Freeman-Tukey double arcsine transformation</u> and combined using standard inverse variance methods.</p>         | <p>We pooled prevalence estimates by first stabilising the variances of the raw proportions using <u>arcsine square root transformations</u>, then applied random-effects models using the Hartung-Knapp-Sidik-Jonkman method.</p>           | <p>Due to large sample size imbalances between included studies, and the negative influence such skewed ranges of sample sizes can have on the harmonic mean which is used to back-transform meta-analytic estimates transformed with the Freeman-Tukey double arcsine method [Schwarzer 2019], we decided to use the arcsine square root transformation instead.</p>  |
| <p>We did not plan to conduct sensitivity analyses to investigate differences in <u>pooled risk ratios</u> when using generalised linear mixed models.</p>             | <p>We examined differences in <u>pooled risk ratios</u> when using generalised linear mixed models to aggregate findings [32,33].</p>  | <p>We decided post-hoc to check the robustness of the meta-analyses of risk ratios when using bivariate generalised linear mixed effects models as proposed by Chu et al. (2012) [33]. Like the meta-analyses of proportions, we chose this method as it has been specifically recommended in situations when the event of interest is rare, and individual study sample sizes are small and circumvent the need to add arbitrary continuity corrections allowing the analysis of both single-zero and double-zero events [27; 165].</p> |
| <p>We did not plan to collect information on <u>data type</u>, nor perform a subgroup analysis to explore its effects on the study’s findings.</p>                     | <p>We collected data on data type and conducted a sub-group analysis to investigate whether prevalence estimates of public data sharing differed depending on <u>the data type</u>, or ...</p>   | <p>We decided to collect data on data type and perform a subgroup analysis exploring its effects on reported findings based on feedback from colleagues on the protocol.</p>   |

**Supplementary Table 1 continued. Deviations from the original review protocol.**

| Original plan  | Revised plan   | Reason for modification   |
|--|--|---|
| We did not plan to conduct sensitivity analyses to investigate differences in pooled prevalence rates when <u>excluding studies that used automated coding strategies</u> .  | We conducted sensitivity analyses to investigate differences in pooled prevalence rates when <u>excluding studies that used automated coding strategies</u> .  | We decided to include this sensitivity analysis to evaluate whether pooled prevalence rates deviated when the results of meta-research studies which used automated coding strategies (methods have been shown to have inferior accuracy, positive predictive value (PPV), and negative predictive value (NPV) when compared to manual coding strategies) were removed. |
| We originally planned to perform subgroup analyses to explore differences in public data sharing frequencies between primary articles reporting the results of <u>clinical trials or not</u> , as well as articles reporting the results of studies using <u>human participants versus not</u> . | We conducted analyses of association for these outcomes rather than a subgroup analysis comparing pooled proportions between groups. Consequently, these two subgroup analyses have been included as secondary outcomes. | We changed the analysis plan for these outcomes due to the availability of data that allowed us to directly explore associations between both of these factors (i.e. calculation of risk ratios)  |
| We originally planned to perform a subgroup analysis to explore differences in public data sharing frequencies between primary articles studying <u>COVID-19 or not</u> .  | We conducted a sensitivity analysis to examine how data sharing frequencies changed when analyses were restricted to meta-research studies examining COVID-19.   | We decided not to compare data and code sharing rates between COVID and non-COVID research because of the large amount of methodological heterogeneity in meta-research studies examining non-COVID research.   |

**Supplementary Table 2. Risk of bias criteria.**

| <b>Item</b>  | <b>Low risk of bias</b>  | <b>High risk of bias</b>   | <b>Unclear risk of bias</b>  |
|--|--|--|--|
| Risk of sampling bias                                | The meta-research study evaluated a random sample of primary articles or sampled the entire population of interest.  | The meta-research study included a non- or pseudorandom sample of primary articles.  | The sampling frame for the sample of primary articles was unclear.   |
| Risk of selective reporting bias                     | Eligible outcomes and associations reported in the protocol for the meta-research study were fully reported in the results section of the publication.   | Not all eligible outcomes and associations reported in the protocol for the meta-research study were reported in the results section of the publication.               | It was unclear if all eligible outcomes and associations were fully reported in the results section of the publication (e.g., because a study protocol for the meta-research study was unavailable). |
| Risk of article selection bias                       | Details about which studies were excluded from the study and why have been shared and match the criteria described in the methods, or no article screening needed to be performed (e.g., because all articles identified by a literature search were analysed) | Details about which studies were excluded and why were not reported.   | Details about the eligibility criteria and study selection process was unclear.  |
| Risk of errors in the accuracy of reported estimates | All outcome data were either manually coded by at least two people independently in parallel or coded by one person and checked in full by another.  | Outcome data were manually coded by one researcher, an automated algorithm, or according to another methodology different from that outlined in the Low Risk category. | The method used to extract data from the included primary studies was unclear.   |

**Supplementary Table 3. Findings of eligible meta-research studies where summary data were not available for the review (N=9).**

| Study           | Year | Discipline          | Journals examined | Primary study date range | Data types          | Sample size     | IPD available | Exclusion reason   |
|-----------------|------|---------------------|-------------------|--------------------------|---------------------|-----------------|---------------|--|
| Helliwell 2020b | 2020 | COVID-19, MERS      | Multiple          | 2019-2020, 2018-2019     | Any                 | 398<br>55       | Partial       | Reported prevalence estimates could not be coded in accordance with the study codebook   |
| Hemkens 2016    | 2016 | General Medical     | Multiple          | 2012                     | Clinical data       | 124             | No            | Reported prevalence estimates could not be coded in accordance with the study codebook   |
| Jiao 2022       | 2022 | Multidisciplinary   | PLOS One          | 2014-2020                | Any                 | 127,935*        | No            | Prevalence estimates not reported separately for medical articles  |
| McDonald 2017   | 2017 | General Medical     | BMJ               | 2015-2017                | Clinical data       | 237             | Partial       | Reported prevalence estimates could not be coded in accordance with the study codebook   |
| Ramke 2018      | 2018 | Ophthalmology       | Multiple          | 2000-2014                | Clinical data       | 153             | No            | Prevalence estimates not reported  |
| Rustici 2021    | 2021 | Biomedicine         | Multiple          | 2009-2013, 2012          | RNA-Seq, Microarray | 1,114*,<br>347* | No            | Prevalence estimates not reported separately for medical articles  |
| Stodden 2018    | 2018 | Multidisciplinary   | Science           | 2009-2010                | Any                 | 204*            | No            | Reported prevalence estimates could not be coded in accordance with the study codebook and are also not reported separately for medical articles |
| Towse 202       | 2020 | Clinical Psychology | Multiple          | 2014-2017                | Any                 | 1,900*          | Partial       | Prevalence estimates not reported separately for medical articles  |
| Zhao 2017       | 2017 | Multidisciplinary   | PLOS One          | 2014-2015                | Any                 | 50*             | No            | Prevalence estimates not reported separately for medical articles  |

Supplementary Table 4. Meta-regression results.

|                              | Model Coefficients |         |           |        |                |          | Level 3 (Between-study) |                       |          | Level 2 (Within-study) |                       |          | AIC     | BIC     |
|------------------------------|--------------------|---------|-----------|--------|----------------|----------|-------------------------|-----------------------|----------|------------------------|-----------------------|----------|---------|---------|
|                              | Intercept          | SE      | $\beta^*$ | SE     | 95% CI         | <i>p</i> | $\tau^2$                | <i>I</i> <sup>2</sup> | <i>k</i> | $\tau^2$               | <i>I</i> <sup>2</sup> | <i>o</i> |         |         |
| <b>Declared data sharing</b> |                    |         |           |        |                |          |                         |                       |          |                        |                       |          |         |         |
| - Three-level (All)          | -18.4290           | 1.4417  | 0.0093    | 0.0007 | 0.0078-0.0107  | <0.0001  | 0.0124                  | 90.66%                | 27       | 0.0013                 | 9.11%                 | 155      | -278.92 | -266.80 |
| - Three-level (Manual)       | -34.0273           | 10.1446 | 0.0170    | 0.0050 | 0.0070-0.0270  | 0.0010   | 0.0117                  | 55.06%                | 25       | 0.0032                 | 15.05%                | 118      | -149.50 | -138.48 |
| - Two-level (All)            | -17.9399           | 2.4845  | 0.0090    | 0.0012 | 0.0066-0.0115  | <0.0001  | 0.0084                  | 99.62%                | 155      | -                      | -                     | -        | -195.29 | -186.20 |
| - Two-level (Manual)         | -42.7527           | 8.9317  | 0.0213    | 0.0044 | 0.0126-0.0301  | <0.0001  | 0.0132                  | 67.64%                | 118      | -                      | -                     | -        | -114.37 | -106.10 |
| <b>Actual data sharing</b>   |                    |         |           |        |                |          |                         |                       |          |                        |                       |          |         |         |
| - Three-level (All)          | -9.0520            | 6.5014  | 0.0045    | 0.0032 | -0.0018-0.0109 | 0.1615   | 0.0088                  | 76.38%                | 26       | 0.0002                 | 2.06%                 | 125      | -230.02 | -218.77 |
| - Three-level (Manual)       | -16.0684           | 8.5307  | 0.0080    | 0.0042 | -0.0004-0.0164 | 0.0604   | 0.0087                  | 61.08%                | 25       | 0.0004                 | 2.54%                 | 119      | -205.38 | -194.33 |
| - Two-level (All)            | -20.4047           | 6.9122  | 0.0102    | 0.0034 | 0.0034-0.0170  | 0.0036   | 0.0057                  | 69.65%                | 125      | -                      | -                     | -        | -197.37 | -188.93 |
| - Two-level (Manual)         | -21.0863           | 7.5362  | 0.0105    | 0.0037 | 0.0031-0.0179  | 0.0058   | 0.0065                  | 55.57%                | 119      | -                      | -                     | -        | -178.70 | -170.42 |
| <b>Declared code sharing</b> |                    |         |           |        |                |          |                         |                       |          |                        |                       |          |         |         |
| - Three-level (All)          | -7.1880            | 0.6262  | 0.0036    | 0.0003 | 0.0030-0.0042  | <0.0001  | 0.0010                  | 82.70%                | 24       | 0.0002                 | 15.03%                | 139      | -428.82 | -417.14 |
| - Three-level (Manual)       | -1.5145            | 10.6961 | 0.0008    | 0.0053 | -0.0098-0.0113 | 0.8852   | 0.0014                  | 18.54%                | 22       | 0.0000                 | 0%                    | 102      | -236.20 | -225.78 |
| - Two-level (All)            | -4.4004            | 0.7886  | 0.0022    | 0.0004 | 0.0014-0.0030  | <0.0001  | 0.0005                  | 94.92%                | 139      | -                      | -                     | -        | -400.39 | -391.63 |
| - Two-level (Manual)         | -5.4033            | 11.2150 | 0.0027    | 0.0056 | -0.0083-0.0137 | 0.6290   | 0.0009                  | 13.19%                | 102      | -                      | -                     | -        | -235.61 | -227.80 |
| <b>Actual code sharing</b>   |                    |         |           |        |                |          |                         |                       |          |                        |                       |          |         |         |
| - Three-level (All)          | -5.1997            | 10.5109 | 0.0026    | 0.0052 | -0.0078-0.0129 | 0.6199   | 0.0009                  | 15.15%                | 21       | 0.0000                 | 0%                    | 99       | -242.96 | -232.66 |
| - Three-level (Manual)       | NA                 | NA      | NA        | NA     | NA             | NA       | NA                      | NA                    | NA       | NA                     | NA                    | NA       | NA      | NA      |
| - Two-level (All)            | -11.2948           | 10.6347 | 0.0056    | 0.0053 | -0.0048-0.0161 | 0.2894   | 0.0006                  | 10.26%                | 99       | -                      | -                     | -        | -241.18 | -233.45 |
| - Two-level (Manual)         | NA                 | NA      | NA        | NA     | NA             | NA       | NA                      | NA                    | NA       | -                      | -                     | -        | NA      | NA      |



**Supplementary Table 5. Sensitivity analyses for secondary outcomes and subgroup analyses.**

|   | Declared public sharing |                        |             |    |                | Actual public sharing |              |              |    |                |
|---|-------------------------|------------------------|-------------|----|----------------|-----------------------|--------------|--------------|----|----------------|
|   | RR                      | 95% CI                 | 95% PI      | k  | I <sup>2</sup> | RR                    | 95% CI       | 95% PI       | k  | I <sup>2</sup> |
| <b>Association between data and code sharing</b>  |                         |                        |             |    |                |                       |              |              |    |                |
| - HKSJ method   | 8.03                    | 2.86-22.53             | 0.33-194.43 | 12 | 32%            | 42.05                 | 12.15-145.52 | 0.94-1879.62 | 7  | 0%             |
| - BGLMM method (SZC)  | 7.88                    | 2.44-18.01             | NA          | 12 | NA             | 40.42                 | 15.45-120.39 | NA           | 7  | NA             |
| - BGLMM method (DZC)  | 10.51                   | 3.00-18.01             | NA          | 23 | NA             | 52.85                 | 9.46-132.52  | NA           | 17 | NA             |
| - Low ROB   | -                       | -                      | -           | -  | -              | -                     | -            | -            | -  | -              |
| - FAIR studies  | 11.84                   | 0-1.33x10 <sup>7</sup> | NA          | 2  | 82%            | -                     | -            | -            | -  | -              |
| - IPD only  | -                       | -                      | -           | -  | -              | -                     | -            | -            | -  | -              |
| - Manual coding   | 4.52                    | 1.38-14.86             | 0.20-101.05 | 10 | 0%             | -                     | -            | -            | -  | -              |
| <b>Trial versus non-trial</b>   |                         |                        |             |    |                |                       |              |              |    |                |
| - HKSJ method   | 0.69                    | 0.45-1.07              | 0.12-4.13   | 23 | 0%             | 0.96                  | 0.53-1.72    | 0.15-5.95    | 19 | 0%             |
| - BGLMM method (SZC)  | 0.55                    | 0.35-0.77              | NA          | 23 | NA             | 0.67                  | 0.26-1.39    | NA           | 19 | NA             |
| - BGLMM method (DZC)  | 0.56                    | 0.37-0.79              | NA          | 25 | NA             | 0.69                  | 0.27-1.52    | NA           | 24 | NA             |
| <b>Human versus non-human</b>   |                         |                        |             |    |                |                       |              |              |    |                |
| - HKSJ method   | 0.65                    | 0.42-0.99              | 0.12-3.61   | 19 | 57%            | 0.44                  | 0.24-0.81    | 0.05-3.57    | 16 | 28%            |
| - BGLMM method (SZC)  | 0.69                    | 0.46-1.01              | NA          | 19 | NA             | 0.58                  | 0.29-1.00    | NA           | 16 | NA             |
| - BGLMM method (DZC)  | 0.69                    | 0.48-1.00              | NA          | 20 | NA             | 0.59                  | 0.30-0.97    | NA           | 20 | NA             |
| SZC – Did not include studies with no events in both groups in analyses, DZC – Included studies with no events in both groups in analyses |                         |                        |             |    |                |                       |              |              |    |                |