```
# overall
fig2_data2 %>%
  select(year_split, EmpiricalStudyData, Type, n, prop, available) %>%
  print(., n = nrow(fig2_data2))
```

```
## # A tibble: 30 × 6
## # Groups:   year_split, EmpiricalStudyData [4]
##    year_split        EmpiricalStudyData Type                  n  prop availa…¹
##    <fct>             <fct>              <fct>             <int> <dbl> <fct>
##  1 Post-RC (2018-2019) Primary          "Materials "       106  75.7 " Not a…
##  2 Post-RC (2018-2019) Primary          "Materials "        34  24.3 " Avail…
##  3 Post-RC (2018-2019) Secondary        "Materials "        43  89.6 " Not a…
##  4 Post-RC (2018-2019) Secondary        "Materials "         5  10.4 " Avail…
##  5 Pre-RC (2008-09)    Primary          "Materials "       105  85.4 " Not a…
##  6 Pre-RC (2008-09)    Primary          "Materials "        18  14.6 " Avail…
##  7 Pre-RC (2008-09)    Secondary        "Materials "        48  98   " Not a…
##  8 Pre-RC (2008-09)    Secondary        "Materials "         1   2   " Avail…
##  9 Post-RC (2018-2019) Primary          "Raw Data "          5   3.6 " Avail…
## 10 Post-RC (2018-2019) Secondary        "Raw Data "         20  41.7 " Avail…
## 11 Pre-RC (2008-09)    Primary          "Raw Data "          8   6.5 " Avail…
## 12 Pre-RC (2008-09)    Secondary        "Raw Data "         17  32.7 " Avail…
## 13 Post-RC (2018-2019) Primary          "Raw Data "        135  96.4 " Not a…
## 14 Post-RC (2018-2019) Secondary        "Raw Data "         28  58.3 " Not a…
## 15 Pre-RC (2008-09)    Primary          "Raw Data "        116  93.5 " Not a…
## 16 Pre-RC (2008-09)    Secondary        "Raw Data "         35  67.3 " Not a…
## 17 Post-RC (2018-2019) Primary          "Processed Data "    5   3.6 " Avail…
## 18 Post-RC (2018-2019) Secondary        "Processed Data "    4   8.3 " Avail…
## 19 Pre-RC (2008-09)    Primary          "Processed Data "    1   0.8 " Avail…
## 20 Pre-RC (2008-09)    Secondary        "Processed Data "    2   3.8 " Avail…
## 21 Post-RC (2018-2019) Primary          "Processed Data "  135  96.4 " Not a…
## 22 Post-RC (2018-2019) Secondary        "Processed Data "   44  91.7 " Not a…
## 23 Pre-RC (2008-09)    Primary          "Processed Data "  123  99.2 " Not a…
## 24 Pre-RC (2008-09)    Secondary        "Processed Data "   50  96.2 " Not a…
## 25 Post-RC (2018-2019) Primary          "Analysis "          1   0.7 " Avail…
## 26 Post-RC (2018-2019) Secondary        "Analysis "          3   6.2 " Avail…
## 27 Post-RC (2018-2019) Primary          "Analysis "        139  99.3 " Not a…
```

```
## 28 Post-RC (2018-2019) Secondary        "Analysis "         45  93.8 " Not a…
## 29 Pre-RC (2008-09)     Primary          "Analysis "        124 100    " Not a…
## 30 Pre-RC (2008-09)     Secondary        "Analysis "         52 100    " Not a…
## # … with abbreviated variable name ¹available
```

```r
# summed by year
fig2_data2 %>%
  group_by(year_split, Type, available) %>%
  summarise(total_n = sum(n), .groups = "keep") %>%
  group_by(year_split, Type) %>%
  mutate(total_prop = total_n / sum(total_n) * 100)
```

```
## # A tibble: 15 × 5
## # Groups:   year_split, Type [8]
##    year_split        Type            available      total_n total_prop
##    <fct>             <fct>           <fct>            <int>     <dbl>
##  1 Pre-RC (2008-09)  "Materials "    " Available"        19      11.0
##  2 Pre-RC (2008-09)  "Materials "    " Not available"   153      89.0
##  3 Pre-RC (2008-09)  "Raw Data "     " Available"        25      14.2
##  4 Pre-RC (2008-09)  "Raw Data "     " Not available"   151      85.8
##  5 Pre-RC (2008-09)  "Processed Data " " Available"       3       1.70
##  6 Pre-RC (2008-09)  "Processed Data " " Not available" 173      98.3
##  7 Pre-RC (2008-09)  "Analysis "     " Not available"   176     100
##  8 Post-RC (2018-2019) "Materials "    " Available"        39      20.7
##  9 Post-RC (2018-2019) "Materials "    " Not available"   149      79.3
## 10 Post-RC (2018-2019) "Raw Data "     " Available"        25      13.3
## 11 Post-RC (2018-2019) "Raw Data "     " Not available"   163      86.7
## 12 Post-RC (2018-2019) "Processed Data " " Available"       9       4.79
## 13 Post-RC (2018-2019) "Processed Data " " Not available" 179      95.2
## 14 Post-RC (2018-2019) "Analysis "     " Available"         4       2.13
## 15 Post-RC (2018-2019) "Analysis "     " Not available"   184      97.9
```

```r
# summed by study data
fig2_data2 %>%
  group_by(EmpiricalStudyData, Type, available) %>%
  summarise(total_n = sum(n), .groups = "keep") %>%
  group_by(EmpiricalStudyData, Type) %>%
  mutate(total_prop = total_n / sum(total_n) * 100)
```

```
## # A tibble: 16 × 5
```

```
## # Groups:   EmpiricalStudyData, Type [8]
##    EmpiricalStudyData Type             available          total_n total_prop
##    <fct>              <fct>            <fct>                <int>      <dbl>
##  1 Secondary          "Materials "     " Available"             6       6.19
##  2 Secondary          "Materials "     " Not available"        91      93.8
##  3 Secondary          "Raw Data "      " Available"            37      37
##  4 Secondary          "Raw Data "      " Not available"        63      63
##  5 Secondary          "Processed Data " " Available"            6       6
##  6 Secondary          "Processed Data " " Not available"       94      94
##  7 Secondary          "Analysis "      " Available"             3       3
##  8 Secondary          "Analysis "      " Not available"        97      97
##  9 Primary            "Materials "     " Available"            52      19.8
## 10 Primary            "Materials "     " Not available"       211      80.2
## 11 Primary            "Raw Data "      " Available"            13       4.92
## 12 Primary            "Raw Data "      " Not available"       251      95.1
## 13 Primary            "Processed Data " " Available"            6       2.27
## 14 Primary            "Processed Data " " Not available"      258      97.7
## 15 Primary            "Analysis "      " Available"             1       0.379
## 16 Primary            "Analysis "      " Not available"       263      99.6
```

```r
# completely overall

fig2_data2 %>%
  group_by(Type, available) %>%
  summarise(total_n = sum(n), .groups = "keep") %>%
  group_by(Type) %>%
  mutate(total_prop = total_n / sum(total_n) * 100)
```

```
## # A tibble: 8 × 4
## # Groups:   Type [4]
##   Type              available          total_n total_prop
##   <fct>             <fct>                <int>      <dbl>
## 1 "Materials "      " Available"            58      16.1
## 2 "Materials "      " Not available"       302      83.9
## 3 "Raw Data "       " Available"            50      13.7
## 4 "Raw Data "       " Not available"       314      86.3
## 5 "Processed Data " " Available"            12       3.30
## 6 "Processed Data " " Not available"       352      96.7
```

```
## 7 "Analysis "      " Available"        4     1.10
## 8 "Analysis "      " Not available"  360     98.9
```