Supplementary material for `Quantifying replicability and consistency in systematic reviews`

The package `metarep` is an extension to the package `meta`, which allows incorporating replicability-analysis tools to quantify consistency and replicability of treatment effect estimates in a meta-analysis. The tool was proposed by Jaljuli et. al. (submitted) for the fixed-effect and for the random-effects meta-analyses, whith or without the common-effect assumption.

Packages Instalation:

Currently, both `meta` and `metarep` packages can be downloaded from GitHub, therefore make sure that the package `devtools` is installed. `metarep` also requires the latest version of `meta` (>= 4.11-0, available on `guithub`)

Run the following commands in *console* to install the packages:

```
devtools::install_github( "guido-s/meta"    , force=T )
devtools::install_github( "IJaljuli/metarep", force=T )
```

Examples:

Here we demonstrate the approach implemented with `metrep` on several examples from the Systematic Review Cochrane Library. These examples are detailed in the paper as well, along with a demonstration of a way to incorporate our suggestions in standard meta-analysis reporting system.

We begin with an example based on fixed-effects meta-analysis from review number CD002943: the effect of mammogram invitation on attendance during the following 12 months.

1^{st} Example: Review CD002943

```
library(metarep)
data(CD002943_CMP001)

m2943 <- meta::metabin( event.e = N_EVENTS1, n.e = N_TOTAL1,
                     event.c = N_EVENTS2, n.c = N_TOTAL2,
                     studlab = STUDY, comb.fixed = T , comb.random = F,
                     method = 'Peto', sm = CD002943_CMP001$SM[1],
                     data = CD002943_CMP001)

m2943
                    OR            95%-CI %W(fixed)
  Sutton-1994   1.2836 [0.9933; 1.6589]      32.3
  Somkin-1997   1.8739 [1.5372; 2.2844]      54.2
  Turnbull-1991 3.5709 [1.9225; 6.6326]       5.5
  Mohler-1995   1.8764 [0.5272; 6.6788]       1.3
  Bodiya-1999   1.0758 [0.6110; 1.8941]       6.6
```

```
  Number of studies combined: k = 5

                          OR            95%-CI    z  p-value
  Fixed effect model 1.6564 [1.4317; 1.9164] 6.78 < 0.0001

  Quantifying heterogeneity:
   tau^2 = 0.0882 [0.0000; 1.5129]; tau = 0.2970 [0.0000; 1.2300];
   I^2 = 70.3% [24.4%; 88.3%]; H = 1.84 [1.15; 2.93]

  Test of heterogeneity:
       Q d.f. p-value
   13.47    4  0.0092

  Details on meta-analytical method:
  - Peto method
  - DerSimonian-Laird estimator for tau^2
  - Jackson method for confidence interval of tau^2 and tau

summary(m2943)
  Number of studies combined: k = 5

                          OR            95%-CI    z  p-value
  Fixed effect model 1.6564 [1.4317; 1.9164] 6.78 < 0.0001

  Quantifying heterogeneity:
   tau^2 = 0.0882 [0.0000; 1.5129]; tau = 0.2970 [0.0000; 1.2300];
   I^2 = 70.3% [24.4%; 88.3%]; H = 1.84 [1.15; 2.93]

  Test of heterogeneity:
       Q d.f. p-value
   13.47    4  0.0092

  Details on meta-analytical method:
  - Peto method
  - DerSimonian-Laird estimator for tau^2
  - Jackson method for confidence interval of tau^2 and tau
```

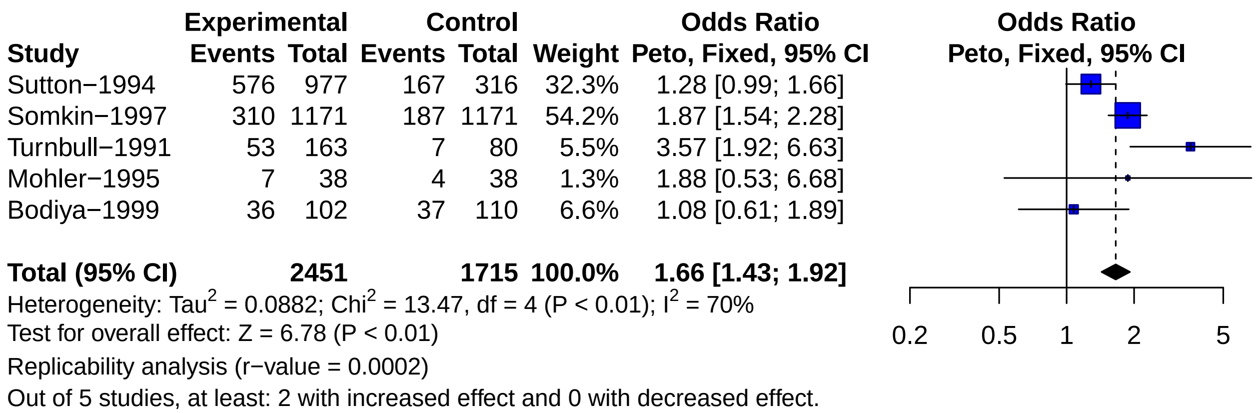
The replicability-analysis results follow.

```
m2943.ra <- metarep(x = m2943 , u = 2 , common.effect = F ,t = 0.05 ,report.u.max = T)
m2943.ra
                     OR            95%-CI %W(fixed)
  Sutton-1994   1.2836 [0.9933; 1.6589]      32.3
  Somkin-1997   1.8739 [1.5372; 2.2844]      54.2
  Turnbull-1991 3.5709 [1.9225; 6.6326]       5.5
  Mohler-1995   1.8764 [0.5272; 6.6788]       1.3
  Bodiya-1999   1.0758 [0.6110; 1.8941]       6.6

  Number of studies combined: k = 5

                          OR            95%-CI    z  p-value
  Fixed effect model 1.6564 [1.4317; 1.9164] 6.78 < 0.0001

  Quantifying heterogeneity:
```

```
  tau^2 = 0.0882 [0.0000; 1.5129]; tau = 0.2970 [0.0000; 1.2300];
  I^2 = 70.3% [24.4%; 88.3%]; H = 1.84 [1.15; 2.93]

 Test of heterogeneity:
      Q d.f. p-value
  13.47    4  0.0092

 Details on meta-analytical method:
 - Peto method
 - DerSimonian-Laird estimator for tau^2
 - Jackson method for confidence interval of tau^2 and tau
 - replicability analysis (r-value = 2e-04)
 - out of 5 studies, at least: 2 with increased effect and 0 with decreased effect.
```

metarep allows adding replicability results to the conventional forest plots by meta. This can be done by simply applying `meta::forest()` on a metarep object.

```
metarep::forest(m2943.ra, layout='revman5',digits.pval = 2 , test.overall = T )
```



2^{nd} Example: Review CD007077

The second example is based on a fixed-effects meta analysis in review CD007077. The main objective of this review is to determine whether PBI/APBI is equivalent to or better than conventional or hypo-fractionated whole breast radiotherapy (WBRT) after breast-conservation therapy for early-stage breast cancer. The primary outcome was Cosmesis.

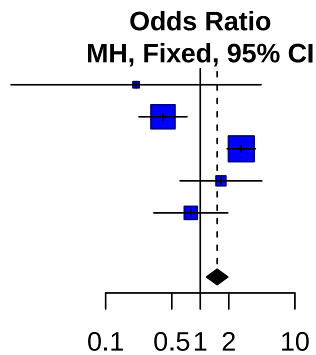
```
data(CD007077_CMP001)

m7077 <- meta::metabin( event.e = N_EVENTS1, n.e = N_TOTAL1,
                     event.c = N_EVENTS2, n.c = N_TOTAL2,
                     studlab = STUDY, comb.fixed = T , comb.random = F,
                     method = 'MH', sm = CD007077_CMP001$SM[1],
                     data = CD007077_CMP001)

summary(m7077)
 Number of studies combined: k = 5

                         OR           95%-CI    z p-value
 Fixed effect model 1.5078 [1.1673; 1.9476] 3.14  0.0017

 Quantifying heterogeneity:
```

```
  tau^2 = 1.0240; tau = 1.0119; I^2 = 88.4% [75.5%; 94.5%]; H = 2.94 [2.02; 4.26]

 Test of heterogeneity:
      Q d.f.  p-value
  34.47    4 < 0.0001

 Details on meta-analytical method:
 - Mantel-Haenszel method
 - DerSimonian-Laird estimator for tau^2
 - Mantel-Haenszel estimator used in calculation of Q and tau^2 (like RevMan 5)
 - Continuity correction of 0.5 in studies with zero cell frequencies
```

With the replicability analysis:

```
m7077.ra <- metarep(x = m7077 , u = 2, t = 0.05 , report.u.max = T )

metarep::forest(m7077.ra, layout='revman5',digits.pval = 2 , test.overall = T )
```

	Experimental		Control			Odds Ratio
Study	**Events**	**Total**	**Events**	**Total**	**Weight**	**MH, Fixed, 95% CI**
Livi−2015	0	246	2	260	2.5%	0.21 [0.01; 4.39]
Polg_x00e1_r−2007	24	125	43	116	37.4%	0.40 [0.23; 0.72]
RAPID	140	399	61	367	42.8%	2.71 [1.92; 3.82]
Rodriguez	12	51	8	51	6.3%	1.65 [0.61; 4.47]
TARGIT	12	55	13	50	11.0%	0.79 [0.32; 1.95]
Total (95% CI)		**876**		**844**	**100.0%**	**1.51 [1.17; 1.95]**



Heterogeneity: $Tau^2 = 1.0240$; $Chi^2 = 34.47$, df = 4 (P < 0.01); $I^2 = 88\%$
Test for overall effect: Z = 3.14 (P < 0.01)
Replicability analysis (r−value = 1)
Out of 5 studies, at least: 1 with increased effect and 1 with decreased effect.

3^{rd} Example:

Based on a random-effects meta-analysis in review CD006823, where the meta-analysis finding was statistically significant. The authors examine the effects of wound drainage after axillary dissection for breast carcinoma on the incidence of post-operative Seroma formation.

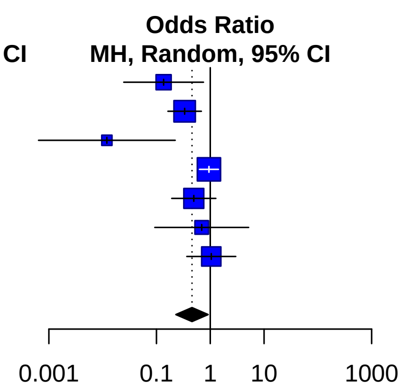
```
data(CD006823_CMP001)

m6823 <- meta::metabin( event.e = N_EVENTS1, n.e = N_TOTAL1,
                     event.c = N_EVENTS2, n.c = N_TOTAL2,
                     studlab = STUDY, comb.fixed = F , comb.random = T,
                     method = 'MH', sm = CD006823_CMP001$SM[1],
                     data = CD006823_CMP001)

m6823
                       OR            95%-CI %W(random)
  Cameron-1988      0.1358 [0.0247; 0.7478]       10.0
  Somers-1992       0.3341 [0.1633; 0.6837]       20.1
  Zavotksy-1998     0.0120 [0.0006; 0.2230]        4.6
  Purushotham-2002 0.9457 [0.6305; 1.4183]       23.6
  Jain-2004         0.4942 [0.1922; 1.2704]       17.4
  Soon-2005         0.6939 [0.0931; 5.1693]        8.1
  Classe-2006       1.0446 [0.3665; 2.9774]       16.2
```

```
Number of studies combined: k = 7

                         OR            95%-CI      z p-value
Random effects model 0.4576 [0.2293; 0.9130] -2.22  0.0266

Quantifying heterogeneity:
 tau^2 = 0.4833; tau = 0.6952; I^2 = 67.7% [28.3%; 85.4%]; H = 1.76 [1.18; 2.62]

Test of heterogeneity:
     Q d.f. p-value
 18.58    6  0.0049

Details on meta-analytical method:
- Mantel-Haenszel method
- DerSimonian-Laird estimator for tau^2
- Mantel-Haenszel estimator used in calculation of Q and tau^2 (like RevMan 5)
- Continuity correction of 0.5 in studies with zero cell frequencies
```

With replicability-analysis:

```
m6823.ra <- metarep(x = m6823 , u = 2, t = 0.05, report.u.max = T )

metarep::forest(m6823.ra, layout='revman5',digits.pval = 2 , test.overall = T )
```

	Experimental		Control			Odds Ratio
Study	**Events**	**Total**	**Events**	**Total**	**Weight**	**MH, Random, 95% CI**
Cameron−1988	2	20	9	20	10.0%	0.14 [0.02; 0.75]
Somers−1992	79	108	106	119	20.1%	0.33 [0.16; 0.68]
Zavotksy−1998	0	24	14	22	4.6%	0.01 [0.00; 0.22]
Purushotham−2002	98	190	98	185	23.6%	0.95 [0.63; 1.42]
Jain−2004	15	58	12	29	17.4%	0.49 [0.19; 1.27]
Soon−2005	34	36	49	51	8.1%	0.69 [0.09; 5.17]
Classe−2006	9	51	8	47	16.2%	1.04 [0.37; 2.98]
Total (95% CI)		**487**		**473**	**100.0%**	**0.46 [0.23; 0.91]**



Heterogeneity: $Tau^2 = 0.4833$; $Chi^2 = 18.58$, df = 6 ($P < 0.01$); $I^2 = 68\%$
Test for overall effect: $Z = -2.22$ ($P = 0.03$)
Replicability analysis (r−value = 0.0056)
Out of 7 studies, at least: 0 with increased effect and 2 with decreased effect.

4^{th} Example: Review CD003366

Based on a random-effects meta-analysis in review CD003366. The authors compare chemotherapy regimens on overall effect in Leukopaenia. Pooling 28 studies, the random-effects meta-analysis fails to declare any significant difference between regimens, due to the highly-significant yet contradicting results.

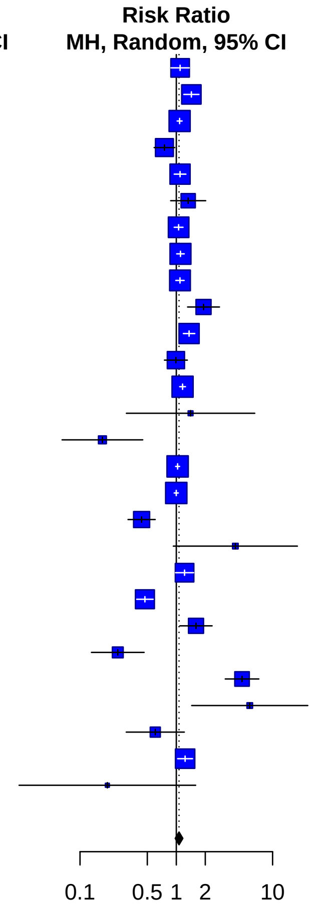
```
data(CD003366_CMP005)

m3366 <- meta::metabin( event.e = N_EVENTS1, n.e = N_TOTAL1,
                     event.c = N_EVENTS2, n.c = N_TOTAL2,
                     studlab = STUDY, comb.fixed = F , comb.random = T,
                     method = 'MH', sm = CD003366_CMP005$SM[1],
                     data = CD003366_CMP005)


m3366
                                 RR             95%-CI %W(random)
```

```
ECOG-E1193-_x0028_A_x0029_ 1.0957 [0.8799;  1.3643]        4.4
EU_x002d_93011             1.4340 [1.1970;  1.7179]        4.8
_x0033_06-Study-Group      1.0824 [1.0196;  1.1490]        5.7
AGO                        0.7525 [0.5883;  0.9625]        4.1
Blohmer                    1.0938 [0.9511;  1.2579]        5.2
Bonneterre                 1.3286 [0.8736;  2.0205]        2.6
Bontenbal                  1.0568 [0.9505;  1.1749]        5.4
CECOG-BM1                  1.1047 [1.0090;  1.2094]        5.6
EORTC-10961                1.0919 [0.9882;  1.2066]        5.5
HERNATA                    1.9171 [1.3090;  2.8078]        2.9
Jassem                     1.3576 [1.1833;  1.5576]        5.2
Lyman                      0.9903 [0.7529;  1.3025]        3.8
Nabholtz                   1.1618 [1.0839;  1.2452]        5.7
Rugo                       1.4062 [0.3031;  6.5255]        0.3
TRAVIOTA                   0.1708 [0.0651;  0.4483]        0.8
_x0033_03-Study-Group      1.0319 [0.9833;  1.0828]        5.8
_x0033_04-Study-Group      0.9988 [0.9499;  1.0501]        5.8
ANZ-TITG                   0.4362 [0.3151;  0.6040]        3.3
Dieras                     4.1053 [0.9311; 18.1008]        0.4
ECOG-E1193-_x0028_B_x0029_ 1.2183 [0.9812;  1.5125]        4.4
EORTC-10923                0.4719 [0.3875;  0.5747]        4.6
JCOG9802                   1.6019 [1.0899;  2.3545]        2.9
Meier                      0.2467 [0.1314;  0.4633]        1.5
Sjostrom                   4.8162 [3.2206;  7.2023]        2.7
Talbot                     5.7895 [1.4441; 23.2103]        0.4
TOG                        0.6048 [0.3018;  1.2119]        1.3
TXT                        1.2342 [1.0326;  1.4752]        4.8
Yardley                    0.1923 [0.0233;  1.5888]        0.2

Number of studies combined: k = 28

                         RR            95%-CI    z p-value
Random effects model 1.0673 [0.9747; 1.1688] 1.41  0.1596

Quantifying heterogeneity:
 tau^2 = 0.0365; tau = 0.1911; I^2 = 89.7% [86.3%; 92.2%]; H = 3.11 [2.70; 3.59]

Test of heterogeneity:
      Q d.f.  p-value
 261.41   27 < 0.0001

Details on meta-analytical method:
- Mantel-Haenszel method
- DerSimonian-Laird estimator for tau^2
- Mantel-Haenszel estimator used in calculation of Q and tau^2 (like RevMan 5)
```

with replicability-analysis:

```
m3366.ra <- metarep(x = m3366 , u = 2 , t = 0.05, report.u.max = T )

metarep::forest(m3366.ra, layout='revman5',digits.pval = 2 , test.overall = T )
```

Study	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio MH, Random, 95% CI
ECOG−E1193−_x0028_A_x0029_	126	230	56	112	4.4%	1.10 [0.88; 1.36]
EU_x002d_93011	76	85	53	85	4.8%	1.43 [1.20; 1.72]
_x0033_06−Study−Group	202	213	184	210	5.7%	1.08 [1.02; 1.15]
AGO	69	204	89	198	4.1%	0.75 [0.59; 0.96]
Blohmer	101	125	82	111	5.2%	1.09 [0.95; 1.26]
Bonneterre	31	70	24	72	2.6%	1.33 [0.87; 2.02]
Bontenbal	96	108	90	107	5.4%	1.06 [0.95; 1.17]
CECOG−BM1	113	122	109	130	5.6%	1.10 [1.01; 1.21]
EORTC−10961	121	136	110	135	5.5%	1.09 [0.99; 1.21]
HERNATA	56	139	29	138	2.9%	1.92 [1.31; 2.81]
Jassem	119	134	87	133	5.2%	1.36 [1.18; 1.56]
Lyman	31	45	32	46	3.8%	0.99 [0.75; 1.30]
Nabholtz	224	238	192	237	5.7%	1.16 [1.08; 1.25]
Rugo	3	32	3	45	0.3%	1.41 [0.30; 6.53]
TRAVIOTA	4	40	24	41	0.8%	0.17 [0.07; 0.45]
_x0033_03−Study−Group	154	159	153	163	5.8%	1.03 [0.98; 1.08]
_x0033_04−Study−Group	188	200	176	187	5.8%	1.00 [0.95; 1.05]
ANZ−TITG	31	105	67	99	3.3%	0.44 [0.32; 0.60]
Dieras	8	38	2	39	0.4%	4.11 [0.93; 18.10]
ECOG−E1193−_x0028_B_x0029_	137	229	55	112	4.4%	1.22 [0.98; 1.51]
EORTC−10923	66	164	139	163	4.6%	0.47 [0.39; 0.57]
JCOG9802	50	147	31	146	2.9%	1.60 [1.09; 2.35]
Meier	9	58	39	62	1.5%	0.25 [0.13; 0.46]
Sjostrom	105	136	21	131	2.7%	4.82 [3.22; 7.20]
Talbot	10	19	2	22	0.4%	5.79 [1.44; 23.21]
TOG	11	97	18	96	1.3%	0.60 [0.30; 1.21]
TXT	65	79	60	90	4.8%	1.23 [1.03; 1.48]
Yardley	1	52	5	50	0.2%	0.19 [0.02; 1.59]
Total (95% CI)		**3404**		**3160**	**100.0%**	**1.07 [0.97; 1.17]**



Heterogeneity: $Tau^2 = 0.0365$; $Chi^2 = 261.41$, df = 27 ($P < 0.01$); $I^2 = 90\%$
Test for overall effect: $Z = 1.41$ ($P = 0.16$)
Replicability analysis (r−value = < 0.0001)
Out of 28 studies, at least: 10 with increased effect and 3 with decreased effect.