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
```

In this meta-analysis, the effect of sending invitation letters was examined in five studies. The authors main result is that: "The odds ratio in relation to the outcome, attendance in response to the mammogram invitation during the 12 months after the invitation, was 1.66 (95% CI 1.43 to 1.92)".

We suggest reporting the replicability-analysis results alongside: the r-value and lower confidence bounds on the number of studies. Results of complete replicability analysis can be added to the contents of a `meta` object or using the function `metarep(...)`, as well as to its summary using `summary( metarep(...) )`. To perform assumption-free replicability-analysis requiring replicability in at least `u = 2` (default) studies, we calculate $r(2)-value$ using truncated-Pearsons' test with truncation threshold `t=0.05` (default)

```
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.
```

The bottom two lines report the $r(2) - value$, lower bound on the number of studies with increased effect (u^L_{max}) and decreased effect (u^R_{max}) , respectively. The evidence towards an increased effect was replicable, with $r(2) - value = 0.0002$. Moreover, with 95% confidence, we can conclude that at least two studies had an increased effect. For higher replicability requirement, compute $r(u') - value$ for $u' > 2$ using `metarep(u = u' , ... )`.

The two-sided $r(u) - value$ of the model can be accessed via `r.value`:

```
m2943.ra$r.value
  [1] 0.0002067774
```

The replicability-analysis reported was performed with an assumption free test, based on truncated0-Pearsons' test with truncation level set at the nominal hypothesis testing level (i.e., `t=0.05`, default). For ordinary Pearsons' test, use `t=1`.

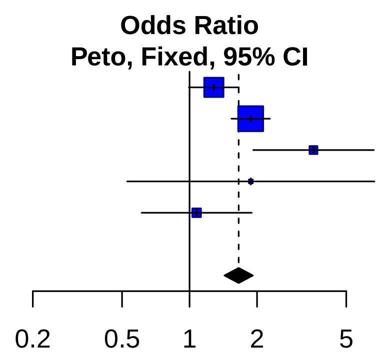
Although the fixed-effect model assumes that all studies are estimates of the same common effect θ, we recommend applying assumption-free replicability-analysis for protection against unsupported assumption. Despite that, we extend our suggested method with the common-effect incorporation in section 7. This analysis can be performed via `metarep( ... , common.effect = TRUE )`.

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

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

	Experimental		Control			Odds Ratio
Study	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto, Fixed, 95% CI**
Sutton−1994	576	977	167	316	32.3%	1.28 [0.99; 1.66]
Somkin−1997	310	1171	187	1171	54.2%	1.87 [1.54; 2.28]
Turnbull−1991	53	163	7	80	5.5%	3.57 [1.92; 6.63]
Mohler−1995	7	38	4	38	1.3%	1.88 [0.53; 6.68]
Bodiya−1999	36	102	37	110	6.6%	1.08 [0.61; 1.89]
Total (95% CI)		**2451**		**1715**	**100.0%**	**1.66 [1.43; 1.92]**

Heterogeneity: $Tau^2 = 0.0882$; $Chi^2 = 13.47$, df = 4 (P < 0.01); $I^2 = 70\%$
Test for overall effect: Z = 6.78 (P < 0.01)

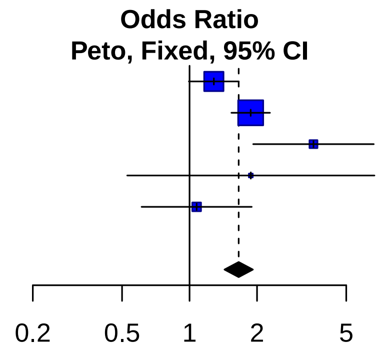


The computation of lower bounds on number of studies with replicability of increased & decreased effects can be suppressed using `metarep( ... , report.u.max = FALSE )`

```
meta::forest( metarep(x = m2943, report.u.max = F ),
              layout='revman5',digits.pval = 2 , test.overall = T )
```

Study	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio Peto, Fixed, 95% CI
Sutton−1994	576	977	167	316	32.3%	1.28 [0.99; 1.66]
Somkin−1997	310	1171	187	1171	54.2%	1.87 [1.54; 2.28]
Turnbull−1991	53	163	7	80	5.5%	3.57 [1.92; 6.63]
Mohler−1995	7	38	4	38	1.3%	1.88 [0.53; 6.68]
Bodiya−1999	36	102	37	110	6.6%	1.08 [0.61; 1.89]
Total (95% CI)		**2451**		**1715**	**100.0%**	**1.66 [1.43; 1.92]**

Heterogeneity: $Tau^2 = 0.0882$; $Chi^2 = 13.47$, df = 4 (P < 0.01); $I^2 = 70\%$
Test for overall effect: Z = 6.78 (P < 0.01)



The lower bounds u^L_{max} and u^R_{max} are calculated with $1 - \alpha = 95\%$ confidence level (default), meaning that each of the null hypotheses

$$H^{u^L_{max}/n}(L) \text{ and } H^{u^L_{max}/n}(R)$$

is tested at level $\alpha/2 = 2.5\%$, resulting in bounds in overall type error rate 5%. Type I error rate can be controlled for any desired α using the argument `confidence = 1 -`α.

The calculation of u^L_{max} and u^R_{max} can also be calculated directly using the function `find_umax()` with the option to specify one-sided alternative, confidence level, truncation threshold and common-effect assumption. For example, let's compute u^L_{max} with the same confidence level as produced by `m2943.ra.bounds`.

```
find_umax(x = m2943 , common.effect = F,alternative = 'less',t = 0.05,confidence = 0.975)
  $worst.case
  [1] "Sutton-1994"   "Somkin-1997"   "Turnbull-1991" "Mohler-1995"
  [5] "Bodiya-1999"

  $side
  Direction of the stronger signal
                            "less"

  $u_max
  u^L
    0

  $r.value
  r^R
    1

  $Replicability_Analysis
  [1] "out of 5 studies, 0 with decreased effect."
```

Note that this function produces 2 main types of results:

1. Worst-case scenario studies: A list of $n - u^L_{max} + 1$ studies names yielding the maximum

$$\max_{\forall\{i_1,\ldots,i_{n-u+1}\}\subset\{1,\ldots,n\}} \{\, p^L_{i_1,\ldots,i_{n-u+1}} \,\}$$

2. Replicability-analysis results, including:

   - u^L_{max} or u^R_{max}. If `alternative='two-sided'`, then $u_{max} = \max\{u^L_{max}\,,\, u^R_{max}\}$ is also reported.

- $r(u^L_{max}) - value$ or $r(u^R_{max}) - value$ if setting `alternative='less'` or `alternative='greater'`, respectively. If `alternative='two-sided'`, then

$$rvalue = r(u_{max}) = 2 \cdot \min\{r^R(u^R_{max}), r^L(u^L_{max})\}$$

is also reported.

For demonstration, see the following example.

```
find_umax(x = m2943 , common.effect = F,alternative = 'two-sided',t = 0.05,confidence = 0.95)
  $worst.case
  [1] "Sutton-1994"   "Turnbull-1991" "Mohler-1995"   "Bodiya-1999"

  $side
  Direction of the stronger signal
                         "greater"

  $u_max
  u_max   u^L   u^R
      2     0     2

  $r.value
  r.value     r^L     r^R
    2e-04   1e+00   1e-04

  $Replicability_Analysis
  [1] "out of 5 studies: 0 with decreased effect, and 2 with increased effect."
```

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
```

The meta-analysis overall effect is significant, with a 95% CI entirely to the right of the null value, despite the fact that the two largest studies report conflicting significant effects. The authors write as a main result that “Cosmesis (physician-reported) appeared worse with PBI/APBI (odds ratio (OR) 1.51, 95% CI 1.17 to 1.95, five studies, 1720 participants, low-quality evidence)”.

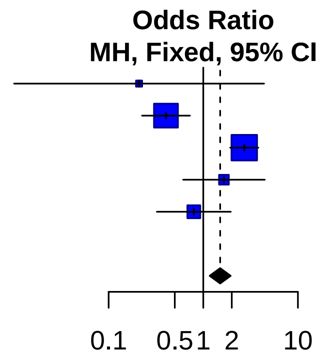
With the replicability analysis:

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

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

Study	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio 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: $Z = 3.14$ ($P < 0.01$)



we suggest adding the r-value and lower confidence bounds on the number of studies, as follows: “We cannot rule out the possibility that this result is critically based on a single study ($r-value = 1$). Moreover, the results are inconsistent, since with 95% confidence, we conclude that at least one study had an increased effect and at least one study had a 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
```

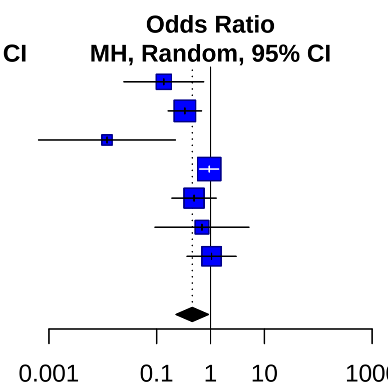
The authors write "The OR for Seroma formation was 0.46 (95% CI 0.23 to 0.91, P = 0.03) in favor of a reduced incidence of Seroma in participants with drains inserted." With replicability-analysis:

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

meta::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)



we conclude that the evidence towards a decreased effect was replicable (r-value = 0.0002). Moreover, with 95% confidence, we can conclude that at least two studies had an decreased effect (and zero studies had an increased 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_ 2.1913 [1.6974;  2.8290]        4.1
  EU_x002d_93011             1.4340 [1.1970;  1.7179]        4.7
  _x0033_06-Study-Group      1.0824 [1.0196;  1.1490]        5.4
  AGO                        0.7525 [0.5883;  0.9625]        4.2
  Blohmer                    1.0938 [0.9511;  1.2579]        5.0
  Bonneterre                 1.3286 [0.8736;  2.0205]        2.9
  Bontenbal                  1.0568 [0.9505;  1.1749]        5.2
  CECOG-BM1                  1.1047 [1.0090;  1.2094]        5.3
  EORTC-10961                1.0919 [0.9882;  1.2066]        5.3
  HERNATA                    1.9171 [1.3090;  2.8078]        3.2
  Jassem                     1.3576 [1.1833;  1.5576]        5.0
  Lyman                      0.9903 [0.7529;  1.3025]        4.0
  Nabholtz                   1.1618 [1.0839;  1.2452]        5.4
  Rugo                       1.4062 [0.3031;  6.5255]        0.4
  TRAVIOTA                   0.1708 [0.0651;  0.4483]        1.0
  _x0033_03-Study-Group      1.0319 [0.9833;  1.0828]        5.5
  _x0033_04-Study-Group      0.9988 [0.9499;  1.0501]        5.5
  ANZ-TITG                   0.4362 [0.3151;  0.6040]        3.6
  Dieras                     4.1053 [0.9311; 18.1008]        0.5
  ECOG-E1193-_x0028_B_x0029_ 2.4365 [1.8921;  3.1377]        4.2
  EORTC-10923                0.4719 [0.3875;  0.5747]        4.6
  JCOG9802                   1.6019 [1.0899;  2.3545]        3.1
  Meier                      0.2467 [0.1314;  0.4633]        1.8
  Sjostrom                   4.8162 [3.2206;  7.2023]        3.0
  Talbot                     5.7895 [1.4441; 23.2103]        0.5
  TOG                        0.6048 [0.3018;  1.2119]        1.6
  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 1.1278 [1.0165; 1.2513] 2.27  0.0233

  Quantifying heterogeneity:
   tau^2 = 0.0509; tau = 0.2255; I^2 = 92.3% [90.0%; 94.1%]; H = 3.61 [3.16; 4.11]

  Test of heterogeneity:
        Q d.f.  p-value
   351.17   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)
```

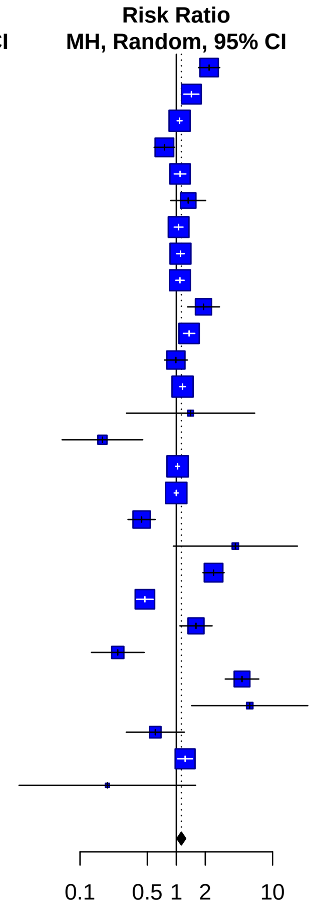
The authors write: "Overall, there was no difference in the risk of Leukopaenia (RR 1.07; 95% CI 0.97 to 1.17; P = 0.16; participants = 6564; Analysis 5.2) with significant heterogeneity across the studies (I2 = 90%; P < 0.00001)". Although, with replicability-analysis:

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

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

	Experimental		Control			Risk Ratio
Study	Events	Total	Events	Total	Weight	MH, Random, 95% CI
ECOG−E1193−_x0028_A_x0029_	126	230	56	224	4.1%	2.19 [1.70; 2.83]
EU_x002d_93011	76	85	53	85	4.7%	1.43 [1.20; 1.72]
_x0033_06−Study−Group	202	213	184	210	5.4%	1.08 [1.02; 1.15]
AGO	69	204	89	198	4.2%	0.75 [0.59; 0.96]
Blohmer	101	125	82	111	5.0%	1.09 [0.95; 1.26]
Bonneterre	31	70	24	72	2.9%	1.33 [0.87; 2.02]
Bontenbal	96	108	90	107	5.2%	1.06 [0.95; 1.17]
CECOG−BM1	113	122	109	130	5.3%	1.10 [1.01; 1.21]
EORTC−10961	121	136	110	135	5.3%	1.09 [0.99; 1.21]
HERNATA	56	139	29	138	3.2%	1.92 [1.31; 2.81]
Jassem	119	134	87	133	5.0%	1.36 [1.18; 1.56]
Lyman	31	45	32	46	4.0%	0.99 [0.75; 1.30]
Nabholtz	224	238	192	237	5.4%	1.16 [1.08; 1.25]
Rugo	3	32	3	45	0.4%	1.41 [0.30; 6.53]
TRAVIOTA	4	40	24	41	1.0%	0.17 [0.07; 0.45]
_x0033_03−Study−Group	154	159	153	163	5.5%	1.03 [0.98; 1.08]
_x0033_04−Study−Group	188	200	176	187	5.5%	1.00 [0.95; 1.05]
ANZ−TITG	31	105	67	99	3.6%	0.44 [0.32; 0.60]
Dieras	8	38	2	39	0.5%	4.11 [0.93; 18.10]
ECOG−E1193−_x0028_B_x0029_	137	229	55	224	4.2%	2.44 [1.89; 3.14]
EORTC−10923	66	164	139	163	4.6%	0.47 [0.39; 0.57]
JCOG9802	50	147	31	146	3.1%	1.60 [1.09; 2.35]
Meier	9	58	39	62	1.8%	0.25 [0.13; 0.46]
Sjostrom	105	136	21	131	3.0%	4.82 [3.22; 7.20]
Talbot	10	19	2	22	0.5%	5.79 [1.44; 23.21]
TOG	11	97	18	96	1.6%	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**		**3384**	**100.0%**	**1.13 [1.02; 1.25]**

Heterogeneity: $Tau^2 = 0.0509$; $Chi^2 = 351.17$, df = 27 ($P < 0.01$); $I^2 = 92\%$
Test for overall effect: Z = 2.27 (P = 0.02)



there is inconsistent evidence for the direction of effect: an increased effect in at least two studies, and a decreased effect in at least ten studies (with 95% confidence).