APPENDIX

Formula 1 Prediction interval

In order to calculate the 95% prediction interval, the summary meta-analysis estimate M, the two sided critical t-value $t_{1\text{-}0.05/2,\,k\text{-}1}$ and the standard deviation for the prediction interval SD_{PI} are needed. Here, *t* is the two-sided critical t-value that can be calculated via http://www.danielsoper.com/statcalc3/calc.aspx?id=10. Fill in DF=k-1 and probability level 0.025, with *k* the number of studies in the meta-analysis. SD_{PI} is the standard deviation of the prediction interval: $SD_{PI} = \sqrt{(\tau^2+SE^2)}$,where τ^2 is the estimated heterogeneity and SE is the standard error of M[1 18]. If the SE was not reported, it can be approximated by dividing the distance between the limits of the 95% CI of the SMD by 3.92. The lower and upper limits of the 95% prediction interval are equal to $M \pm t_{1\text{-}0.05/2,\,k\text{-}1} \times SD_{PI}$. Of course it is possible to estimate prediction intervals with a different coverage, e.g. an 80% prediction interval would be based on $t_{1\text{-}0.20/2,\,6}$. Note that the interval is calculated under the assumption that the value of τ^2 is known (and not estimated).

Estimations for ORs, risk ratios and hazard ratios are generally performed on the natural logarithm scale. As an example we take the calculation of a 95% prediction interval for an OR of 2.28 with a 95% CI from 1.05 to 4.96, τ^2= 0.353 and *k*=7. The prediction interval will first be estimated on log scale. Note that the reported τ^2 is in general already the heterogeneity for log OR, not for OR, and can thus be used directly in the calculations. The SE of the log OR is calculated by dividing the distance between the log of the limits of the 95% CI of the OR by 3.92. This results in SE=0.318. The lower and upper limits of the 95% prediction interval for the log OR are $\log(2.28) \pm 2.45\sqrt{(0.353 + 0.318^2)}$. The value 2.45 results from the $t_{1\text{-}0.05/2}$ distribution with 6 DF. Finally, we exponentiate the limits to return to the OR scale. The resulting prediction interval ranges from 0.44 to 11.86, and can be interpreted as the 95% range of true ORs to be expected in similar studies.

Formula 2 Probability that effect is larger than threshold D

The probability P that the true effect in a new study will be below a threshold D (e.g. the null effect) can be calculated with the left-tail cumulative t-distribution with k-1 degrees of freedom. The probability that the effect is above D equals 1 – P.

In our example on nasal polyps the probability that the SMD ≥ 0 can be estimated as follows:

1. Start to calculate the probability P that a true SMD ≤ 0. This is equivalent to the probability that a t-value $\leq$ T, where T is equal to $(D - M)/SD_{PI}$, with summary treatment effect M= -0.51, SD_{PI} = 0.425 and D=0. This results in T = 1.207, with 6 degrees of freedom (DF).
2. The probability P can be calculated online at http://www.danielsoper.com/statcalc3/calc.aspx?id=41. Fill in t value = 1.207 and DF = 6. The one-tailed probability $P(t \leq 1.207)$= 0.864.
3. We want the probability that the SMD ≥ 0, this is 1 – P= 0.136.

In the example on the OR (see formula 1), if we are interested in the probability of a null or negative effect, we are interested in the probability that a true OR $\leq$ 1. For ORs, calculations must be based on the ln OR, with M= ln(2.28)=0.824, SD_{PI} = 0.674, and DF=6. A true OR $\leq$ 1 corresponds to a true ln OR $\leq$ 0. Fill in T = (0-0.824)/0.674 = -1.223 and DF=6. The probability that a true OR $\leq$ 1 is equal to 0.134.

Formula 3 Prediction interval starting with I^2

In order to calculate prediction intervals starting with an assumed I^2 value (as percentage), we first calculated the corresponding τ^2value:

$$\tau^2 = s^2 \frac{I^2}{100 - I^2}$$

with s^2 the typical study variance, equal to $\frac{\sum \mathrm{w_i}\ (\mathrm{k}-1)}{(\sum \mathrm{W_i})^2 - \sum w_i^2}$, and w_i equal to the inverse of the study variance of study i (i=1..k) and k the number of studies.[23] Subsequently formula 1 can be applied.

Formula 4 Power of a future study

Usually sample size calculations are performed without consideration of the heterogeneity. If we do take into account the heterogeneity, the expected power, i.e. the probability that a new study with N patients will have a positive result at significance level α, given values for the standard error *s* of the new study and μ and τ^2 as above, can be approximated with the delta method if τ^2 is not too large:

$$E(power) = g(\mu) + 0.5\, \tau^2\, g''(\mu)$$

where g is the power at the meta-analysis summary estimate μ, and $g''(\mu)$ is the second derivative of g at μ. For $g''(\mu)$ we can take the second derivative of the normal cumulative distribution function if N is sufficiently large.

This results in $g''(\mu) = \frac{z_\mu e^{-0.5 z_\mu^2}}{s^2\sqrt{2\pi}}$, with $z_\mu = \frac{1.96\mathrm{s} - \mu}{\mathrm{s}}$.

If the sample size N of the new study is such that the power for an effect of size μ is 80%, the expected power of the study will be smaller than 80% if τ^2 is positive, because the corresponding value of z_μ is negative.