Some considerations about combining / borrowing evidence in clinical trials for rare disease

Hannover Medical School

Armin Koch. Kristina Weber
Institut für Biometrie
Medizinische Hochschule Hannover
Carl-Neuberg-Str. 1
30625 Hannover

Introduction:

“Patients suffering from rare conditions should be entitled to the same quality of treatment as other patients;”

“Patients with such conditions deserve the same quality, safety and efficacy in medicinal products as other patients; orphan medicinal products should therefore be submitted to the normal evaluation process;”

Regulation 141/2000 on orphan medicinal products

As it is not plausible to assume that

(i) drugs for rare diseases are on average more effective than drugs for normal disease

(ii) outcome in rare diseases is less variable than in normal disease

… it is highly unlikely that smaller trials can bring the same degree of confidence about efficacy and safety of new drugs for rare diseases than what would be requested / expected from drugs for frequent disease.

Introduction:

In case reduction of confidence is not considered 1st choice, combination or borrowing of information remain. Sufficient information for a classical MA not very likely available.

Tradition in drug regulation:

- Self standing data-based decision making
- Primary use of own data (class is of secondary interest)
- Pre-specified decision making process

Thus:

- In case data are available, preference is given to data (and not to expert opinion)
- In case information is borrowed, then this should be primarily “own” information
- Conclusions should be non-trivial

Combination of 2 (or a small number of) studies, unless homogeneous, is difficult:

For 2 and 6 studies the outcome is as follows:

highest: fixed effects
mid: standard rand. Eff.
lowest: H&K

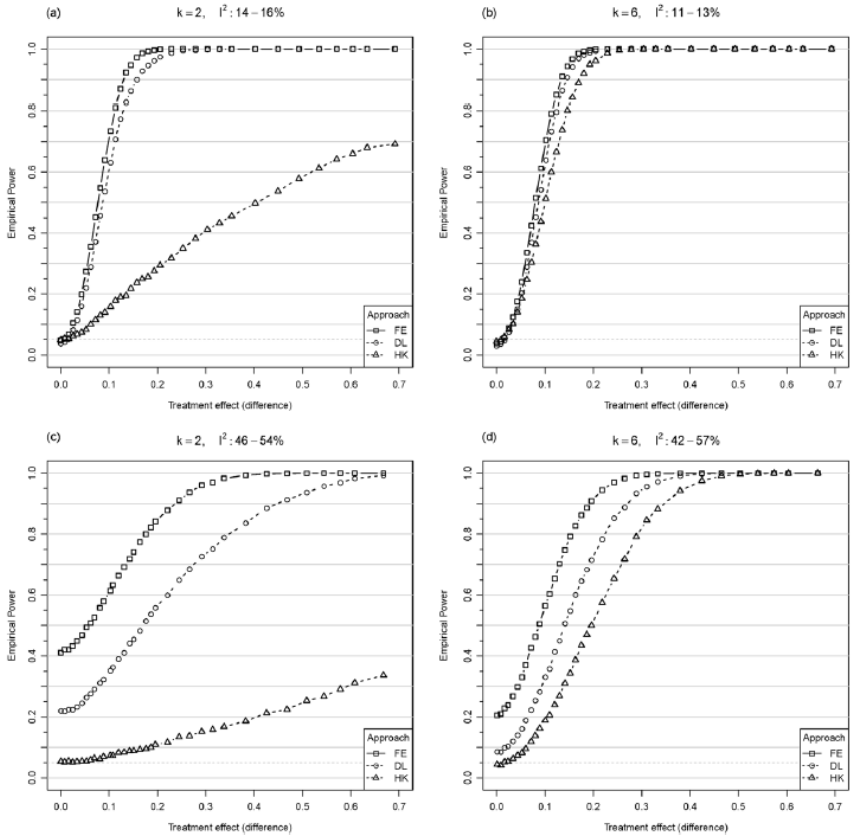


Figure 1. (a–d): Influence of heterogeneity in meta-analysis with two and six studies on empirical power. FE, fixed effects approach; DL, DerSimonian and Laird approach; HK, Hartung and Knapp approach. In the left column, simulation results with two studies are presented, whereas in the right column, situations with six studies are investigated. No heterogeneity is assumed in the top row, and in the bottom row, the impact of moderate heterogeneity is shown.

Formal detection of heterogeneity is difficult, as well:

Limitations of the “classical” heterogeneity test are well known:

Table 1: empirical type 1 error / power in a model with two FEMs in two strata

Situation	ratio	I^{2*}	Cochran’s Q	Q-Rule	G-Rule	E-Rule	KI-Rule
H_0	50:50	0.1489	0.1449	0.1427	0.0723	0.0179	0.0446
H_{1G}	50:50	0.4192	0.4804	0.4798	0.4580	0.2232	0.2998
H_{1E}	50:50	0.6136	0.7192	0.7172	0.7006	0.4395	0.5656
H_0	70:30	0.1495	0.1494	0.1541	0.2014	0.0725	0.2071
H_{1G}	70:30	0.3093	0.3472	0.3480	0.5110	0.3321	0.4299
H_{1E}	70:30	0.4374	0.5018	0.5027	0.6753	0.4927	0.6003
H_0	90:10	0.1486	0.1474	0.1459	0.4580	0.2237	0.5127
H_{1G}	90:10	0.2183	0.2356	0.2370	0.6176	0.4545	0.6100
H_{1E}	90:10	0.2668	0.2909	0.2943	0.6831	0.5496	0.6742

*

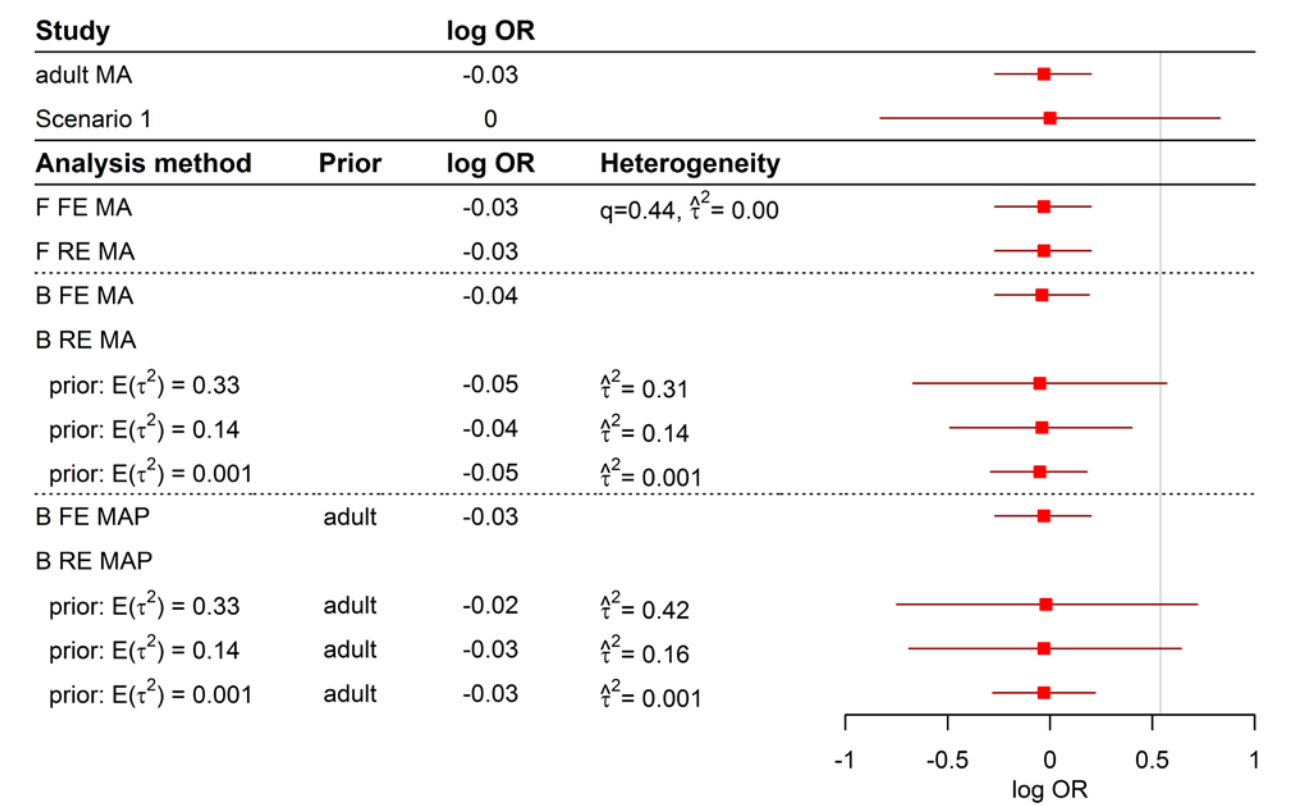
Why not use it in a way medics would use a blood-count?

False positives should be no problem, if there is an agreement that this is signal detection.

Borrowing information (irrespective of methodological approach) is a weighting problem:

Pediatric extrapolation, substantial evidence from adults available, equivalence to be shown:

Scenario 1: homogeneity is assumed



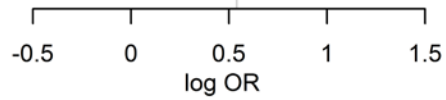
Borrowing information (irrespective of methodological approach) is a weighting problem:

Pediatric extrapolation, substantial evidence from adults available, equivalence to be shown:

Scenario 2:
outcome is different in the target population.

In the Bayesian framework the conclusion depends on the assumed degree of heterogeneity.

Study		log OR	
adult MA		-0.04	
Scenario 2		0.5	
Analysis method	**Prior**	**log OR**	**Heterogeneity**
F FE MA		0.01	q=0.44, $\hat{\tau}^2$= 0.00
F RE MA		0.01	
B FE MA		0	
B RE MA			
prior: E(τ^2) = 0.33		0.05	$\hat{\tau}^2$= 0.32
prior: E(τ^2) = 0.14		0.04	$\hat{\tau}^2$= 0.15
prior: E(τ^2) = 0.001		-0.01	$\hat{\tau}^2$= 0.001
B FE MAP	adult	0.01	
B RE MAP			
prior: E(τ^2) = 0.33	adult	0.38	$\hat{\tau}^2$= 0.43
prior: E(τ^2) = 0.14	adult	0.31	$\hat{\tau}^2$= 0.16
prior: E(τ^2) = 0.001	adult	0.01	$\hat{\tau}^2$= 0.001



Summary:

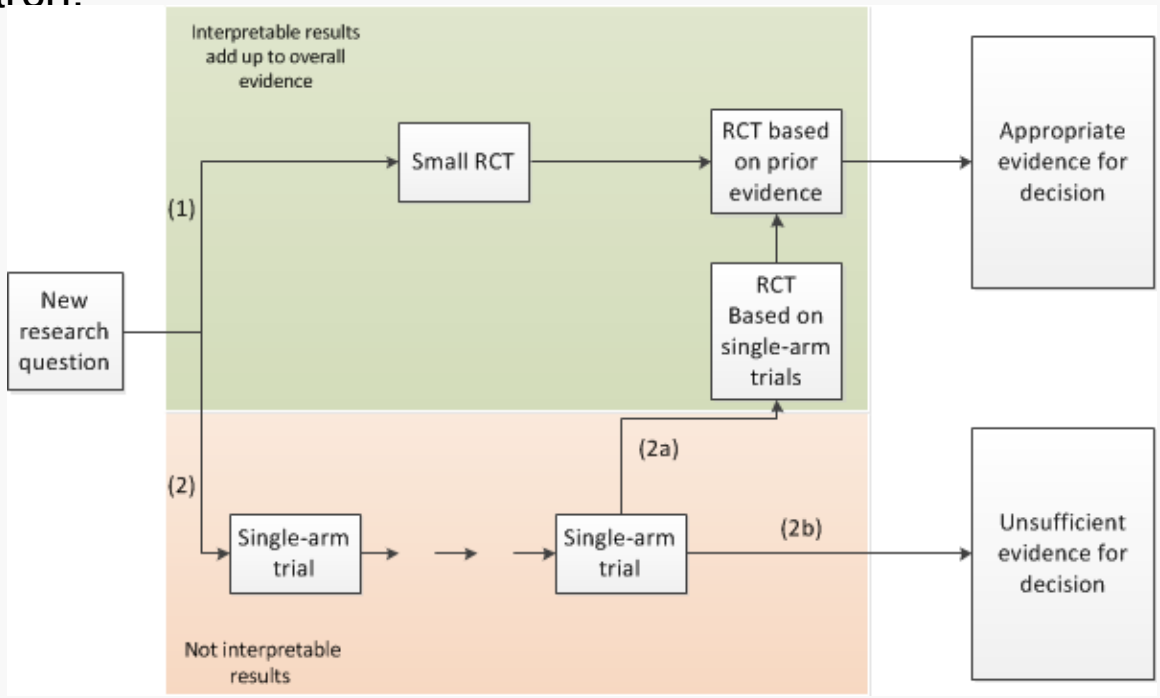
MA-thinking helps to identify problems with combining / borrowing:

- Both, a MA and a borrowing exercise require careful upfront justification.
- Observed heterogeneity / discrepancy may be an indicator that something is not well understood, or is just a chance finding.
- Consequences of the two states are different (ignorable, usable to improve treatment).
- Best methodology (if it exists) can eventually cope with the first, unlikely with the latter.
- To think about liberal behaviour (FEM, REM), poor power (H&K, Cochran-test) and weighting issues is illustrative: What if statistics is right and there is simply not enough information?

Summary:

- Many (most) rare diseases are heterogeneous. Implications of heterogeneity apply for analysis of the individual trial, when combining it, and when borrowing information.
- Arguments for Lego®-research: Chalmers revisited

  In the presence of non-homogeneous data, single-arm trials are more fragile than randomized trials



Hannover Medical School

Finis:

Amount of evidence from a trial is determined by

- number of patients
- duration of observation
- quality of study design

I doubt that there are shortcuts

Moral?

Plan and conduct the best trial that can be done in a reasonable time-frame and then assess it “in the light of” various assumptions.

References:

1. Gonnermann, A., Framke, T., Großhennig, A., & Koch, A. (2015). No solution yet for combining two independent studies in the presence of heterogeneity. *Statistics in Medicine*, *34*(16), 2476–2480. http://doi.org/10.1002/sim.6473
2. Gonnermann, A., Kottas, M., & Koch, A. (2015). Biometrische Entscheidungsunterstützung in Zulassung und Nutzenbewertung am Beispiel der Implikationen von heterogenen Ergebnissen in Untergruppen der Studienpopulation. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 1–9. http://doi.org/10.1007/s00103-014-2105-2
3. Hasford, J., & Koch, A. (2017). Ethische Aspekte der klinischen Prüfung bei seltenen Erkrankungen. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*. http://doi.org/10.1007/s00103-017-2537-6