

# On the road to clinical extrapolation

Kristina Weber, Armin Koch



Hannover Medical School

# Application of Bayesian methodology

- Proposed for situations with limited options to recruit patients into studies (rare disease, pediatric trials)  
or potential limited need (extrapolation from adult to pediatric indications)
- Use of „expert opinion“ to justify certain assumptions about priors that interlink some sort of pathophysiological or pharmacological plausibility with a response parameter
- In rare disease some pre-specified expert opinion may be the only option to reduce the burden of evidence needed for “proof” of efficacy
- In extrapolation, however, data in adults are available to inform about prior knowledge regarding a drug in a certain context (e.g. immunosuppression in organ transplantation)

# Bayesian extrapolation (and regulatory context)

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 (e.g. the prior completely determines the evaluation of the new experiment)

# Paediatric extrapolation

In contrast to other situations:

- Available data have been sufficient for licensing a new drug
- PK/PD and mechanism of action are usually well understood
- PK/PD in paediatric patients available (or can be generated “easily”)

Why then clinical data in paediatric patients?

- Low belief that similar PK/PD leads to the same clinical efficacy
- No reliable PD endpoint
- Puzzling outcome in previous steps of the extrapolation exercise

Drug regulation clarifies the need-to-knows and not the nice-to-knows. To have “at least some paediatric data” would be neither ethical nor scientific as a motivation to do a human experiment.

## Regulatory question

Going for an extrapolation exercise assumes an agreement that there is no need for formal (self-standing) proof of efficacy in the paediatric population. Instead, the following questions need to be addressed:

- A. Which paediatric experiment is needed to detect with good probability relevant deviations from adult expectations regarding the treatment effect?
- B. How to define and assess “relevant deviations”?

To be presented here:

1. Play-games with differing amounts of information (e.g. a lot of information in adults and only a few children),
2. Discussing the EVR case-study in the light of this.

## EVR case-study

Adult studies in de novo kidney transplants with EVR (NIM(log(OR))): 0.54)

study	EVR events/treated	MPA events/treated	Log(OR) 95% CI P-value
B201 Vitko 2004	58/194 (29.9%)	61/196 (31.1%)	-0.05 (-0.48, 0.38) 0.793
B251 Lorber 2005	48/193 (24.9%)	54/196 (27.6%)	-0.13 (-0.58, 0.32) 0.548
A2309 Tedesco 2010	70/277 (25.3%)	67/277 (24.2%)	0.06 (-0.33, 0.45) 0.844
Meta-Analysis (FEM & REM)			-0.035 (-0.28, 0.21) 0.776

Studies investigated different comparators, but demonstration of equivalence was felt relevant in all instances.

B201 (Vitko 2004): **CS+CsA(s)+EVR vs. CS+CsA(s)+MMF,**  
 B251 (Lorber 2005): **CS+CsA(s)+EVR vs. CS+CsA(s)+MMF,**  
 A2309 (Tedesco 2010): **CS+B+CsA(r)+EVR vs. CS+B+CsA(s)+MPA.**

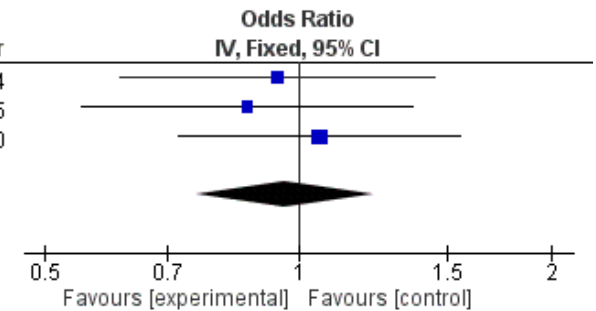
# EVR case-study

Aim: extrapolation to the paediatric population with one clinical study

Investigation of two different scenarios:

study	EVR events/treated	MPA events/treated	log (OR) 95% CI P-value
Scenario 1	16/53 30.2%	16/53 30.2%	0.00 (-0.83; 0.83) 1.00
Scenario 2	22/53 41.5%	16/53 30.2%	0.50 (-0.31; 1.30) 0.33

Study or Subgroup	Experimental		Control		Weight	Odds Ratio IV, Fixed, 95% CI	Year
	Events	Total	Events	Total			
Vitko 2004	58	194	61	196	31.7%	0.94 [0.61, 1.45]	2004
Lorber 2005	48	193	54	196	28.8%	0.87 [0.55, 1.37]	2005
Tedesco 2010	70	277	67	277	39.5%	1.06 [0.72, 1.56]	2010
<b>Total (95% CI)</b>		<b>664</b>		<b>669</b>	<b>100.0%</b>	<b>0.97 [0.76, 1.23]</b>	
Total events	176		182				
Heterogeneity: Chi <sup>2</sup> = 0.44, df = 2 (P = 0.80); I <sup>2</sup> = 0%							
Test for overall effect: Z = 0.28 (P = 0.78)							



# Approaches to a summary evaluation of individual sources of information

- **Frequentist Meta-Analysis**

- Joint analysis of existing and new trial (eventually looking into heterogeneity) in a fixed (FEM) or a random (REM) effects model

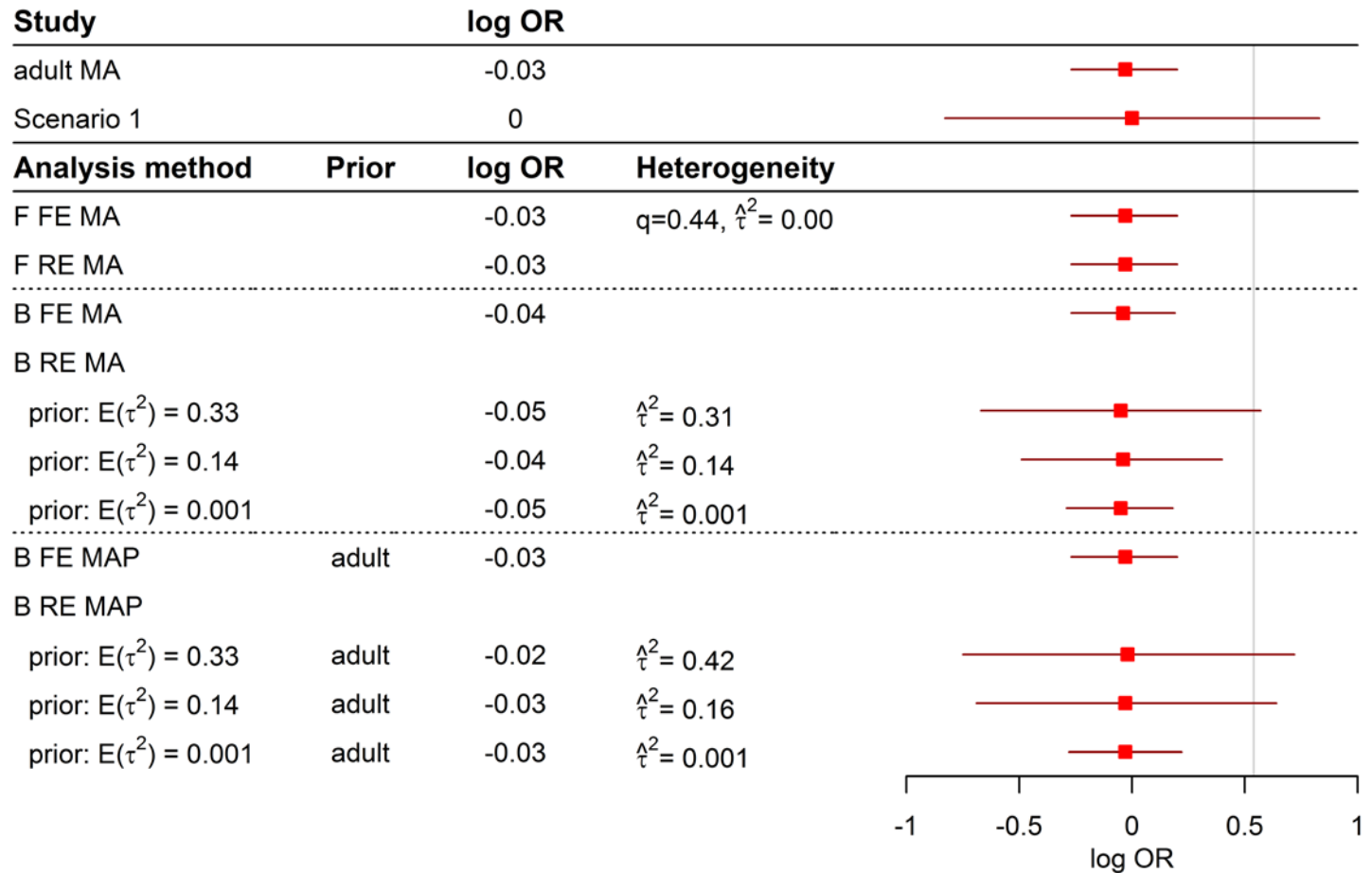
- **Bayesian Meta-Analysis**

- Joint analysis of existing and new trial in a FEM or a REM (Smith et al., 1995)

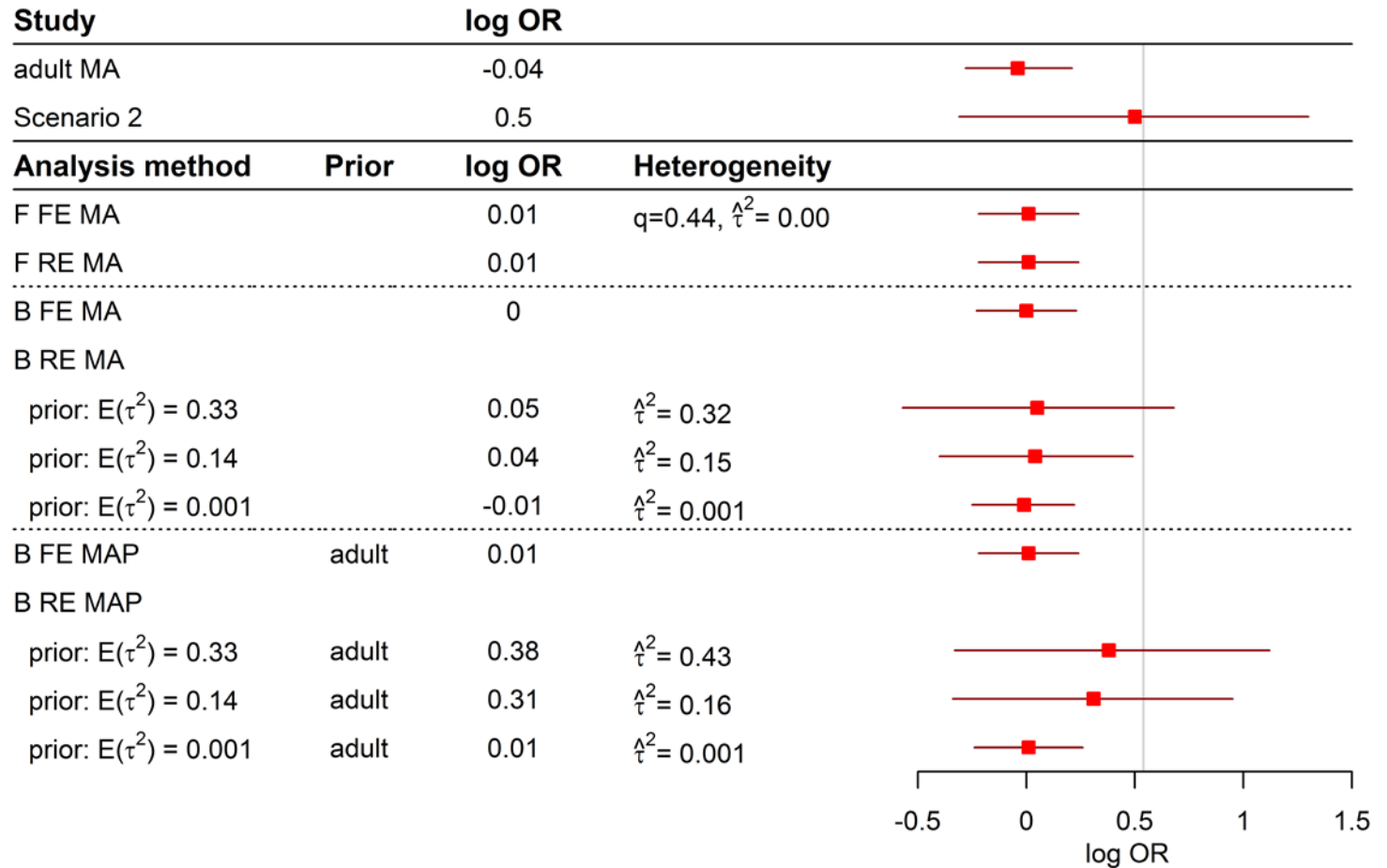
- **Bayesian meta-analytic predictive approach**

- Analysis of new trial „in light of“ the already existing trial in a FEM or a REM (Viele et al., 2014 and Spiegelhalter et al., 2004)

# Results with Scenario 1 (assumed homogeneity)



# Results with Scenario 2 (log OR = 0.50, at the margin)



# Assessment of the exemplary analyses

Many approaches and ...

- ... many different conclusions about the same data possible
- If meta-analysis is used as a tool to arrive at an overall conclusion, no difference between a frequentist approach or a Bayesian approach can be detected: actually summary estimates will always be dominated by adult data.
- Using the predictive approach might allow that the pediatric data stand against the adult data (in case a prior is chosen that will allow for heterogeneity), however then even in case of homogeneity nothing can be concluded with the current sample-size.
- If heterogeneity is restricted, the impact of the adult data is increased (similar to frequentist MA).
- Precise pre-specification of the assumptions is required / recommended.

## Assessment of the exemplary analyses (ctd)

- Such considerations could be used to determine sample-size for a pediatric trial.
- Power priors have been proposed (e.g. Goodman and Sladky, 2005), but result in an arbitrary down weighing and add another screw that may impact on conclusions.

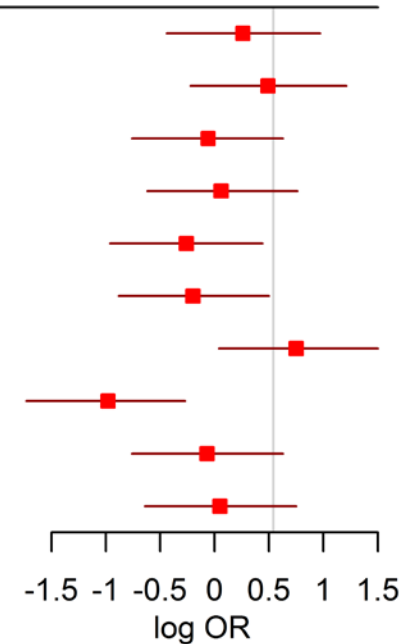
What could be done?

- Avoiding “overweight” in the MA-approach with content-wise selection of adult patients (e.g. only use data from young adults to weigh in for the assessment of adolescent pediatric patients)
- Be precise about the weight of the prior information
- Change of emphasis from “Does it work?” towards “Is there evidence for differential effects?”

# “Simulation” to reduce optimism

Some random draws under the assumption of homogeneity;

Analysis method	prior	log OR	est. Heterogeneity
B RE MA prior: $E(\tau^2) = 0.33$	adult	0.26	0.43
B RE MA prior: $E(\tau^2) = 0.33$	adult	0.49	0.44
B RE MA prior: $E(\tau^2) = 0.33$	adult	-0.06	0.41
B RE MA prior: $E(\tau^2) = 0.33$	adult	0.06	0.42
B RE MA prior: $E(\tau^2) = 0.33$	adult	-0.26	0.42
B RE MA prior: $E(\tau^2) = 0.33$	adult	-0.2	0.41
B RE MA prior: $E(\tau^2) = 0.33$	adult	0.75	0.48
B RE MA prior: $E(\tau^2) = 0.33$	adult	-0.98	0.5
B RE MA prior: $E(\tau^2) = 0.33$	adult	-0.07	0.41
B RE MA prior: $E(\tau^2) = 0.33$	adult	0.05	0.42



# Everolimus – Pediatric extrapolation exercise

## Recruitment of pediatric patients difficult:

PDCO agreed to a M&S approach for EVR:

## Extrapolation

- Use Bayesian meta-regression analysis of the 51 studies to determine a predictive distribution for EVR event-rate in the pediatric trials.
- Calculate the credible interval for the EVR-rate based on this predictive distribution (specifying the treatments under investigation)
- Validation: check, whether each EVR event-rate estimate from the pediatric trials is in the prediction interval.
- Applicant states that this cannot be made precise, because the number of pediatric patients is not yet available.

# Everolimus - Pediatric extrapolation exercise

Completeness vs. best fitting studies:

- 12 months results often not available → use of 6 months data
- Composite endpoint often not available → use of part of the composite as a surrogate endpoint
- Wide range of doses in the treatment arms modelled partly
- Not much discussion about heterogeneity
- Availability of small single-arm studies could be taken as an argument **not** to base decision strategy on best comparative evidence

Nothing in life is free:

Having the “best” control effect may come at the price of uncertainties that are difficult to quantify.

Do we need all or sufficient information?

# Summary and conclusions

- EVR-case is a no-brainer from the medical perspective: everybody believes that this is working
- Extrapolation <-> self standing evidence
- Idea exists that extrapolation is an iterative process (model -> collect data -> check fit -> evaluate -> eventually redo)
- This may be feasible in PK/PD in general, but I feel that this is not true in the field extrapolation:
  - All knowledge has been used-up for the best prediction of pediatric outcome.
  - If then reality doesn't fit our plans – isn't this evidence that extrapolation from adult to pediatric is (too) limited / not possible?
  - Re-do in the world of clinical trials would be extremely costly

# Summary and conclusions

- Clinical extrapolation could be seen as a descriptive exercise (w/o need for confirmatory decision making), but how then to justify sample-size?
- One may decide that no pediatric clinical trial is needed (PK or PK/PD is sufficient), but if one is done, it needs to have an objective to be achieved.
- The value of confirmatory (pre-planned) decision making:
  - a chance to discuss the required amount of information upfront
  - avoid unethical / costly collection of data that is difficult to use
- A lot can be done (relax T1E, increase NI-margin, meta-analyze, pep-up your control group or just omit it).
- Methodological problems exist, but not in the field of whether Bayesian or Frequentist statistics is more appropriate.
- It is more important to precisely define the research question and get the metrics clear to make maximum out of the fact that formal proof of efficacy in adults is already available.

## Summary and conclusions:

F. de Antres-Trelles has made the most important point:

The need for extrapolation should not come as a surprise but should be well reflected in the adult development program.



# References

Steven N Goodman and John T Sladky. A Bayesian approach to randomized controlled trials in children utilizing information from adults: the case of Guillain-Barré syndrome. *Clinical trials (London, England)*, 2(4):305–310; discussion 364–378, 2005.

Smith, T.C., Spiegelhalter, D.J., Thomas, A. (1995). Bayesian Approaches to Random-Effects Meta-Analysis: A Comparative Study. *Statistics in Medicine*, 14: 2685-2699

David J. Spiegelhalter, Keith R. Abrams, and Jonathan P. Myles. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, 2004.

Kert Viele, Scott Berry, Beat Neuenschwander, Billy Amzal, Fang Chen, Nathan Enas, Brian Hobbs, Joseph G. Ibrahim, Nelson Kinnersley, Stacy Lindborg, Sandrine Micallef, Satrajit Roychoudhury, and Laura Thompson. Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics*.