When and how should I combine patient-level data and literature data in a meta-analysis?

Jonathan L French, ScD
Patanjali Ravva, MS

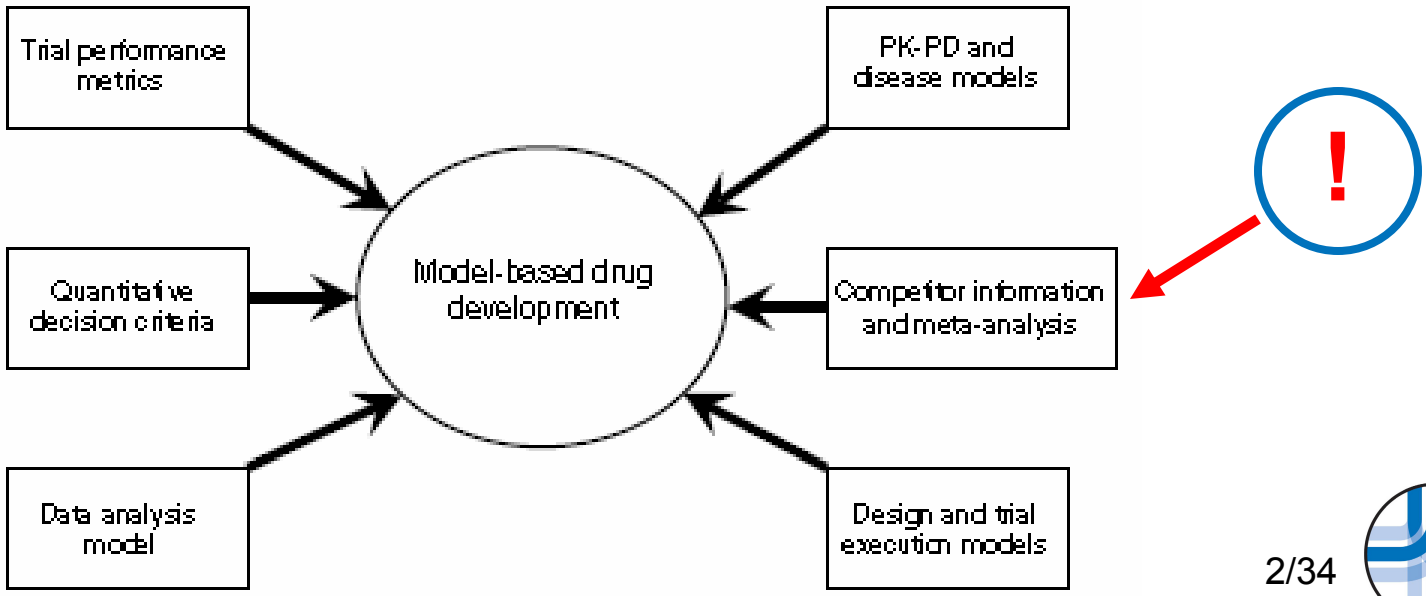
PAGE 2010, Berlin
10 June 2010

Pfizer

Global Pharmacometrics

Meta-analysis is one of the key pillars of model-based drug development

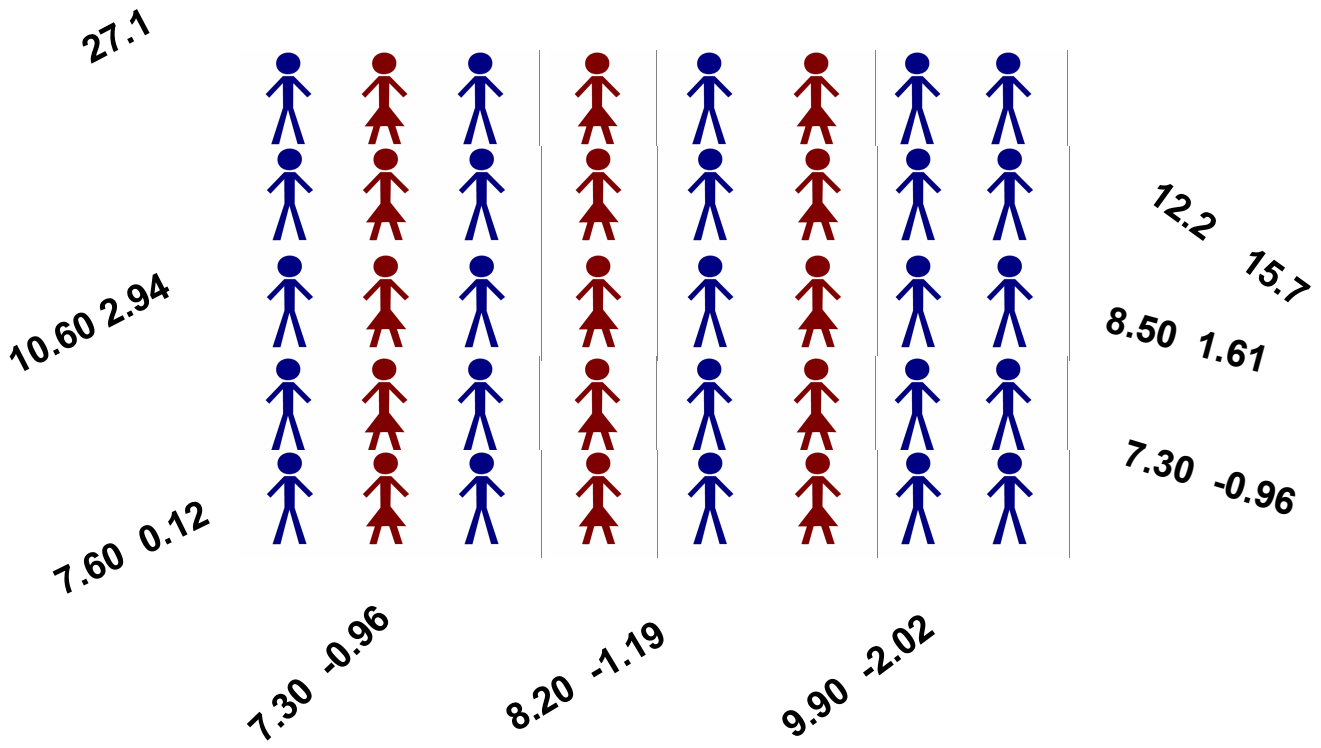
- "The statistical analysis of a large collection of [data] from individual studies for the purpose of integrating the findings." (Glass, 1976)
- Model-based meta-analysis has taken on an important role in drug development decision making



Lalonde et al. CPT 2007

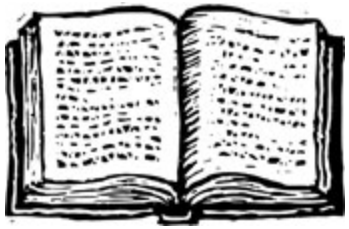
Meta-analysis of individual patient data (IPD) is the ‘gold standard

- We do this all the time with population models
- Sponsors have easy access to their own data but not other data



Meta-analysis of aggregate data (AD) is the norm for most traditional meta-analyses

- Typically, AD will be a measure of treatment effect (difference from control; log odds ratio; etc.)
- Can be an average response in a treatment arm
- Everyone has access to a large amount of AD through the published literature, SBAs, conference abstracts, etc.



The NEW ENGLAND JOURNAL of MEDICINE

THE LANCET

JOURNAL OF CLINICAL ONCOLOGY

JAMA®
The Journal of the American Medical Association

What might be the benefit of combining IPD and AD?

- Putting more information into our models must be a good thing, right?
  - Ultimately, we're interested in the IPD model but the AD part could be used to benchmark or to compare against or inform parts of the model not informed by the IPD
- Addition of AD may help to refine or add precision to parameter estimates that are based solely on a single study of IPD
  - Dose-response or disease progression models
- To yield a better model for clinical trial simulation
  - Allows us to account for between-study variability in drug effect, then this information is only available from multiple studies (hence including AD)
  - IPD can inform about the within- and between-subject variability
  - AD may be necessary for comparing effectiveness of two drugs / treatments
- Addition of IPD may help to inform about the correlation between observations over time in a model based solely on AD
  - This is typically missing from the reports that only give AD

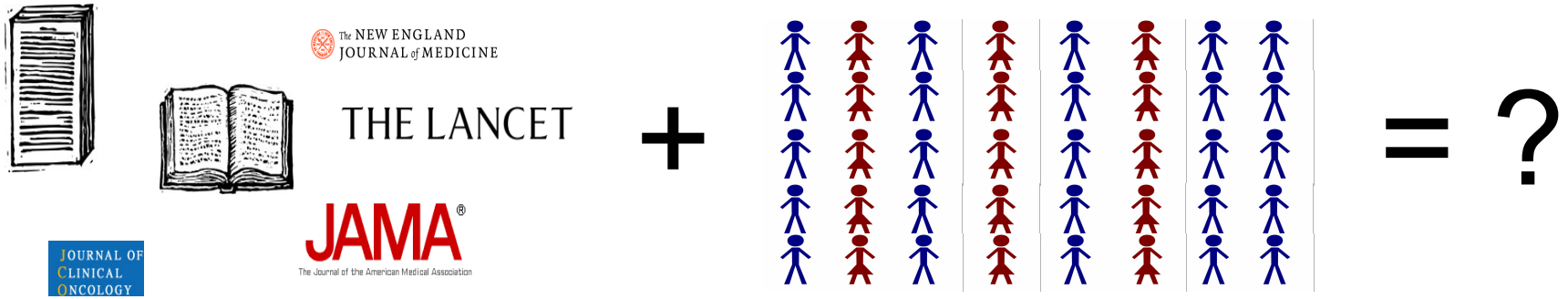
A quick poll

If you have IPD for your drug and AD for other compounds, when should you try to combine them into one meta-analysis model?

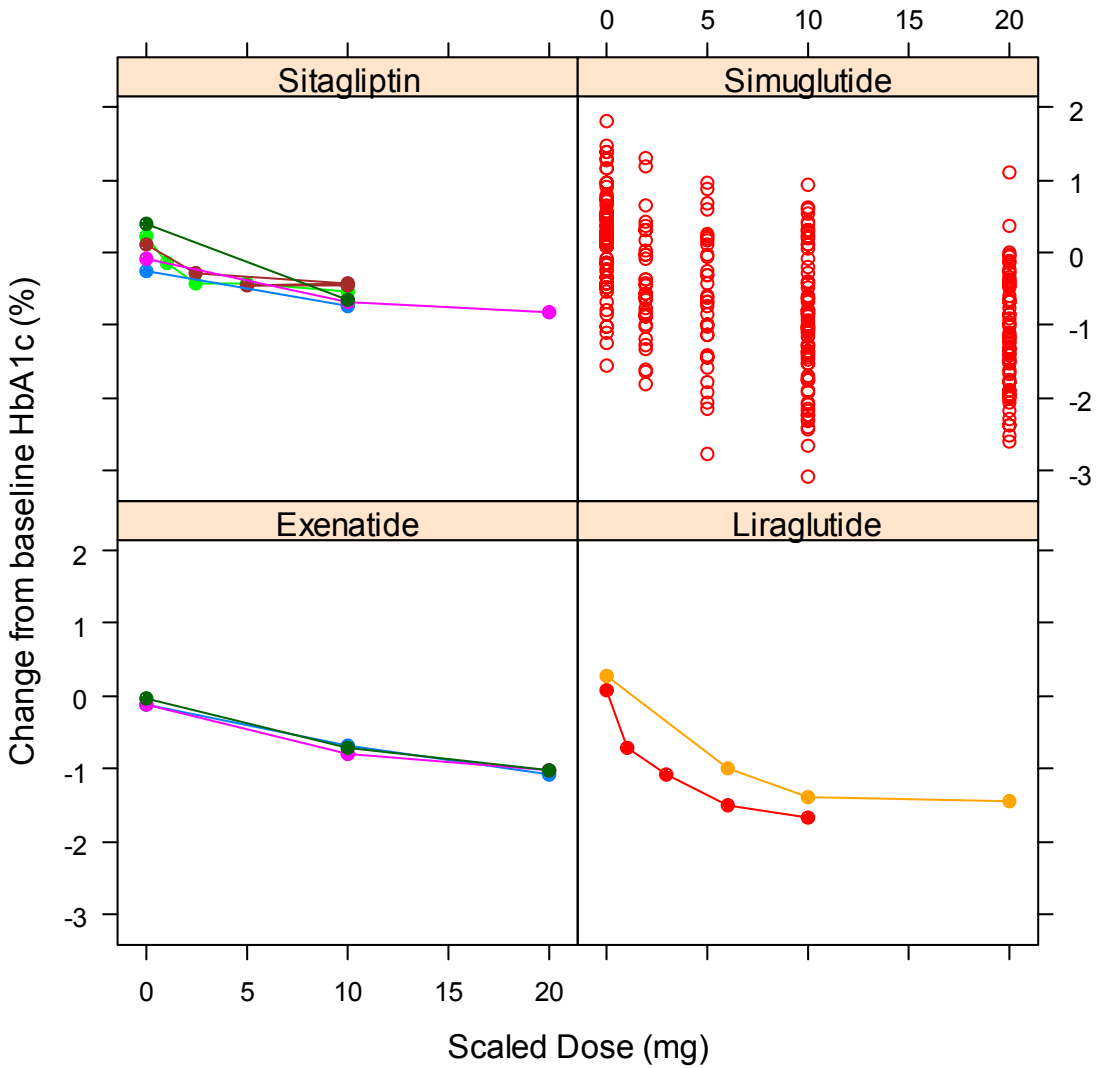
- ❑ Always – after all, that's what models are for, right?
- ❑ Sometimes – it depends on the situation
- ❑ Never – they're different types of data, from different studies - they're simply not combinable
- ❑ I have no idea
- ❑ Why are you bothering me with these questions? I'm here to listen not to think

How can we best combine AD and IPD in a single model?

- A drug developer will have some IPD for their drug and AD for others (including placebo)
- Intuitively, it makes sense to combine these into one model, but how best to do it?



Aggregate and (hypothetical) individual patient data for four diabetes compounds



- 10 studies with AD from 3 drugs (N=40 – 400)
- 1 study with IPD (n~35 or 70 / dose)
- All 4 drugs have a related mechanism of action.
- Can we combine these data into one model to make comparisons between the drugs?
- If so, how?
- What if we also have baseline HbA1c to consider as a covariate…

The diabetes data look like this…

Study	drug	bl	n	CHG	dose
1	Exenatide	8.20	113	-0.11	0
1	Exenatide	8.26	110	-0.67	10
1	Exenatide	8.18	113	-1.09	20
2	Exenatide	8.70	123	-0.12	0
2	Exenatide	8.48	125	-0.81	10
2	Exenatide	8.59	129	-1.02	20

Aggregate data

. < Additional aggregate data >

24	Simuglutide	10.60	1	2.94	0
24	Simuglutide	8.20	1	-1.19	2
24	Simuglutide	9.90	1	-2.02	5
24	Simuglutide	8.50	1	1.61	2
24	Simuglutide	7.60	1	0.12	5
24	Simuglutide	7.30	1	-0.96	10

Individual patient data

. < Additional patient-level data>

Assumptions for the rest of the talk

- Endpoint measured at a single, landmark time
- Continuous response data
  - Although similar principles can be followed for categorical data
- One covariate

- AD consists of mean response, N, mean covariate value (either at the study or treatment-arm level)
  - Not using observed standard error but could easily be incorporated

- IPD consists of individual-level response and covariate values

- Intentionally starting simply
  - The same basic approach should generalize to more complicated situations (but that is still work in progress)

What are some possible ways to combine the data?

- A two-stage approach
  - Convert the IPD to AD and fit an AD-only model
  - Doesn't allow us to realize the benefits of having IPD
- Reconstruct the IPD from the AD
  - Only applicable in limited situations (e.g., binary response data with no covariates)
- Fit a hierarchical/multilevel model
  - View the IPD as nested within a study and build a model for both levels
  - Can be fit using maximum likelihood or with a Bayesian model
- Fit a Bayesian model with informative priors
  - Use the AD to form a prior distribution for the model of the IPD

The two-stage approach for the diabetes data would look like this…

Combined AD and IPD

Study	drug	bl	n	CHG	dose
1	Exenatide	8.20	113	-0.11	0
1	Exenatide	8.26	110	-0.67	10
1	Exenatide	8.18	113	-1.09	20
2	Exenatide	8.70	123	-0.12	0
2	Exenatide	8.48	125	-0.81	10
2	Exenatide	8.59	129	-1.02	20
.					
.	< Lots of other data >				
.					
24	Simuglutide	10.60	1	2.94	0
24	Simuglutide	8.20	1	-1.19	2
24	Simuglutide	9.90	1	-2.02	5
24	Simuglutide	8.50	1	1.61	2
24	Simuglutide	7.60	1	0.12	5
24	Simuglutide	7.30	1	-0.96	10
.					
.					
.					

IPD converted to AD

Study	drug	bl	n	CHG	dose
1	Exenatide	8.20	113	-0.11	0
1	Exenatide	8.26	110	-0.67	10
1	Exenatide	8.18	113	-1.09	20
2	Exenatide	8.70	123	-0.12	0
2	Exenatide	8.48	125	-0.81	10
2	Exenatide	8.59	129	-1.02	20
.					
.	< Lots of other data >				
.					
24	Simuglutide	8.34	70	0.12	0
24	Simuglutide	7.94	36	-0.23	2
24	Simuglutide	8.36	35	-0.95	5
24	Simuglutide	8.35	74	-1.22	10
24	Simuglutide	8.34	69	-1.38	20

The two-stage approach may be adequate in some situations

- Depending on the data at-hand and how you want to use your model, this may be entirely satisfactory
- There should be little loss in information if
  - There are no covariates
  - No need to use individual-level data to inform about certain aspects of the model (e.g., residual error variance)
- Some, possibly large, loss in information if
  - There are covariates to incorporate into the model
  - You need to describe correlations of observations over time and/or residual error variance
- Because we're typically in the latter setting, this approach is not ideal
  - However, it is certainly the easiest approach to implement

Hierarchical/multilevel model approach

- View the IPD as nested within a study and build a model for both levels

- Goldstein et al. (2000) describe this approach for a linear mixed effects model
  - Describe effects of class size on achievement in schools
- Sutton et al. (2008) do the same for a linear logistic regression model

- A related method builds a model for the IPD and derives the corresponding AD model
  - Hierarchical related regression (Jackson et al., 2006, 2008) in an ecological regression (using a logistic regression model)
  - Gillespie et al. (2009) demonstrate this approach in constructing a disease progression model in Alzheimer's disease (ADAS-cog)

A naïve approach is to use the same structural model for the AD and IPD

For the IPD, let's consider the model

$$Y_{ijk} = \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta\left(baseline_{ijk} - 8\right)\right] \bullet dose_{ijk}}{\mathrm{ED50}_{\mathrm{drug}} + dose_{ijk}} + \delta_{ijk} \qquad (1)$$

$$Var\left(\delta_{ijk}\right) = \sigma^2$$

where Y_{ijk} is the change from baseline HbA1c in the k[th] subject at the j[th] dose in the i[th] study and $baseline_{ijk}$ is the corresponding baseline HbA1c.

For the AD, we will consider the model

$$\overline{Y}_{ij} = \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta\left(\overline{baseline}_{ij} - 8\right)\right] \bullet dose_{ij}}{\mathrm{ED50}_{\mathrm{drug}} + dose_{ij}} + \varepsilon_{ij} \qquad (2)$$

$$Var\left(\varepsilon_{ij}\right) = \sigma^2 / n_{ij}$$

where $\overline{Y}_{ij}$ is the mean change from baseline HbA1c in the j[th] group in the i[th] study and $\overline{baseline}_{ij}$ is the corresponding mean baseline HbA1c

Are the parameters in these two models actually describing the same effects?

To answer this question, we need to view the model as a function of the *covariate*

For aggregate data, we observe $\overline{Y}_{ij}$ which is an estimate of

$$E(Y \mid dose) = \int E(Y \mid dose, x)\, p(x \mid dose)\, dx$$

For the IPD model(1), the covariate is baseline HbA1c and the mean response is

$$E(Y \mid dose) = E\left[\mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta(X - 8)\right] \bullet dose}{\mathrm{ED50}_{\mathrm{drug}} + dose} + \delta \right]$$

$$= \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta(E(X) - 8)\right] \bullet dose}{\mathrm{ED50}_{\mathrm{drug}} + dose}$$

We can approximate this by replacing $E(X)$ with $\overline{x}_{ij}$.

In general, when the covariate enters the model linearly, the IPD and AD structural models are the same. Thus, we can pool the types of data relatively easily.

When the model is non-linear in the covariates, then things are not as simple

Imagine, instead, the IPD model was

$$Y_{ijk} = \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet dose_{ijk}}{\mathrm{ED50}_{\mathrm{drug}} \left(\frac{x_{ijk}}{8} \right)^{\theta} + dose_{ijk}} + \delta_{ijk} \qquad (3)$$

In this case, the aggregate data model does not collapse as nicely

$$E\left(Y \mid dose\right) = \mathrm{E0}_i + \mathrm{Emax} \bullet dose_{ij} \bullet E\left[\left(\mathrm{ED50}_{\mathrm{drug}} \left(\frac{X}{8} \right)^{\theta} + dose \right)^{-1} \right]$$

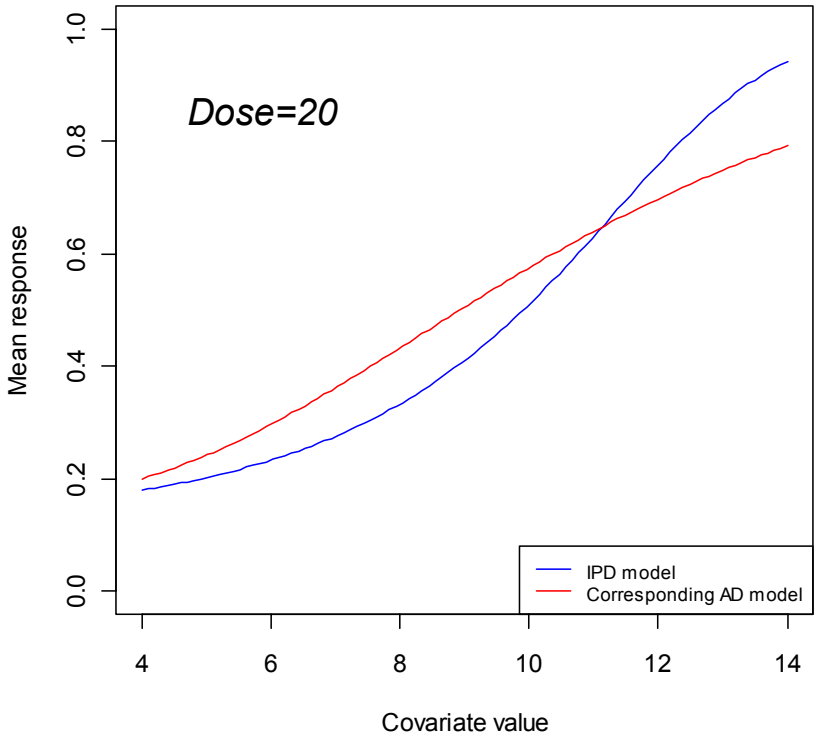
$$\neq \mathrm{E0}_i + \mathrm{Emax} \bullet dose \left(\mathrm{ED50}_{\mathrm{drug}} \left(\frac{E(X)}{8} \right)^{\theta} + dose \right)^{-1}$$

In general, when the covariate enters the model non-linearly, the IPD and AD models are not the same. Whether or not the parameters have a similar interpretation will depend on the degree of non-linearity in the model (as a function of the covariate).

This is important because of the potential for aggregation bias

When the effect of the covariate at the AD level is different from that at the IPD level we are prone to Aggregation (aka Ecological) bias

- This has been recognized as a problem for a long time in the ecological regression field (Wakefield 2008) and more recently in meta-analysis (Berlin et al. 2002)



IPD Model :

$$\text{logit}\left(P\left(response \mid x\right)\right) = \text{logit}(.2) + \frac{3\,\text{e}^{0.2(x-8)}\,d}{20+d}$$

$$\ln\left(x\right) \sim N\left(\mu, 0.4^2\right)$$

Corresponding AD Model :

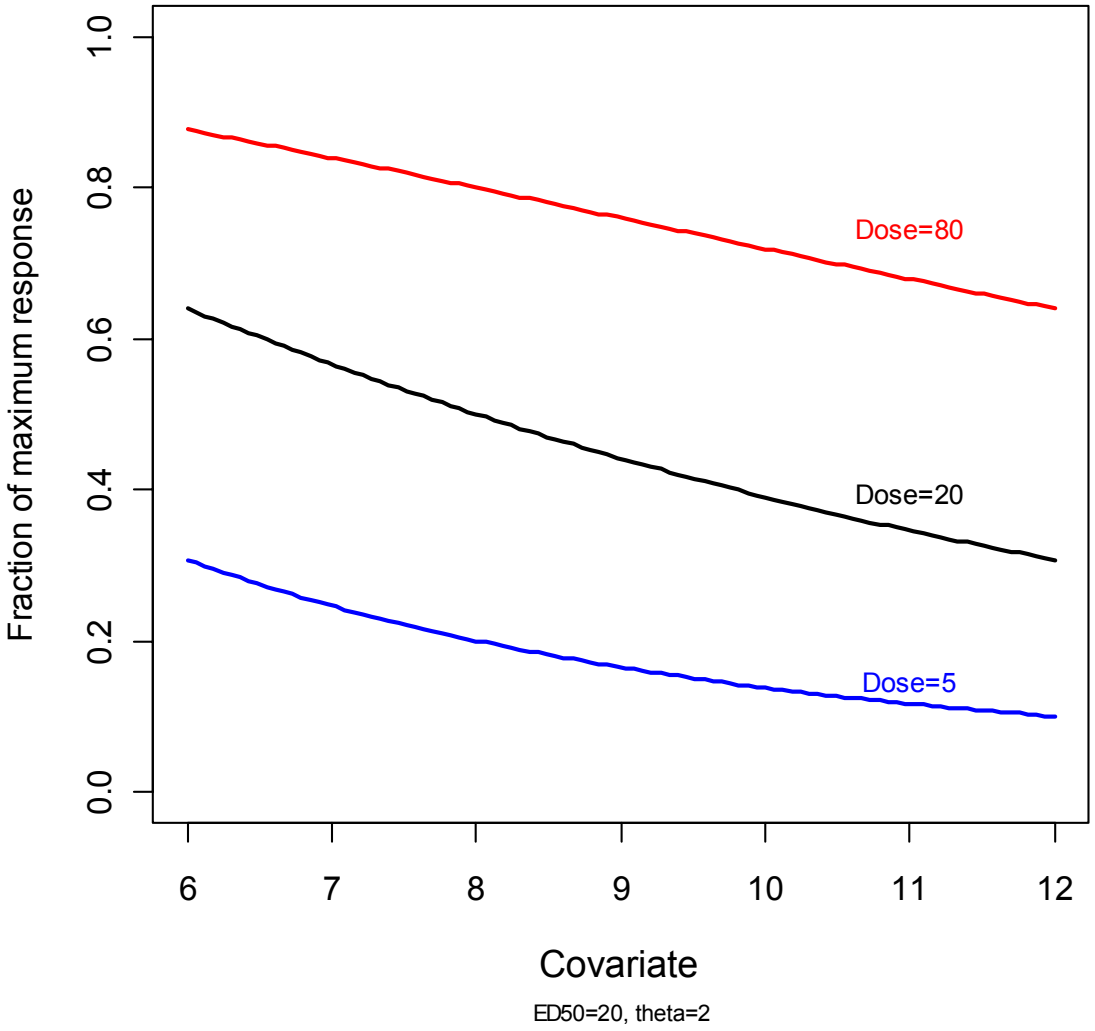
$$P\left(response \mid \mu\right) = \int P\left(response \mid x\right) p\left(x \mid \mu\right) dx$$

The covariate-effect relationship may differ between the IPD and AD models

This can have a major impact when we're trying to combine AD and IPD using the naïve model

- This is particular to models in which a covariate enter the model non-linearly
- In this situation, the naïve approach to combining the IPD and AD model will lead to biased parameter values (because the AD model is incorrectly specified)
- The bias will depend on how non-linear the function is in the covariate

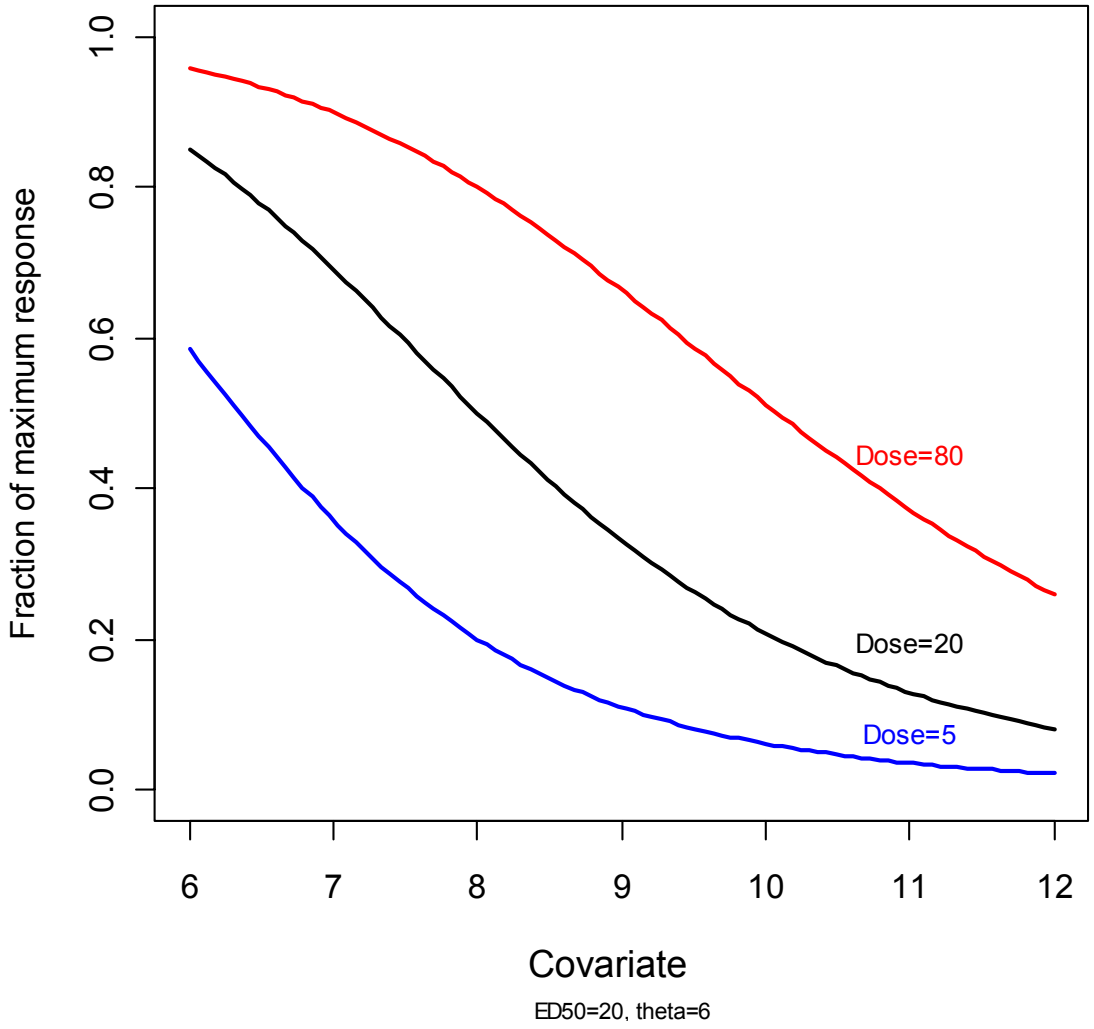
Let's view this as a function of a continuous covariate for fixed values of ED50 and θ



This relationship is reasonably linear across a range of doses, despite the fact that the covariate enters the model non-linearly.

As θ gets larger, this relationship becomes more non-linear.

For a larger value of θ, the relationship is no longer (approximately) linear



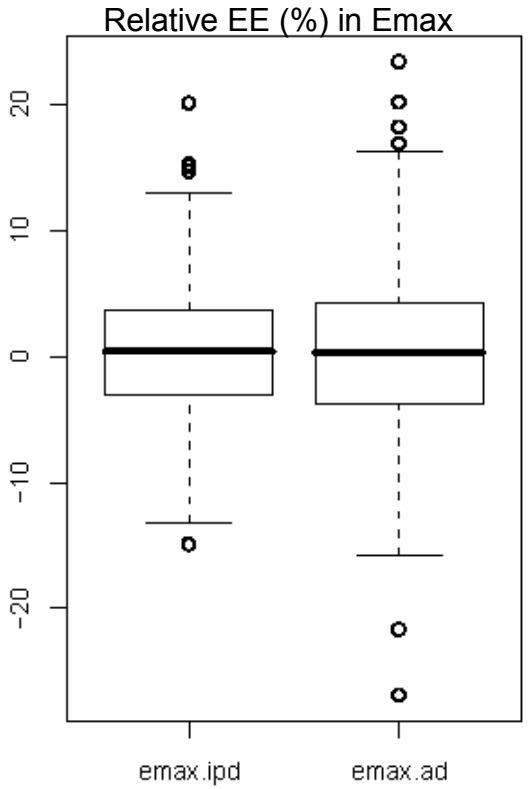
For a larger value of θ, this relationship is now rather non-linear across a range of doses.

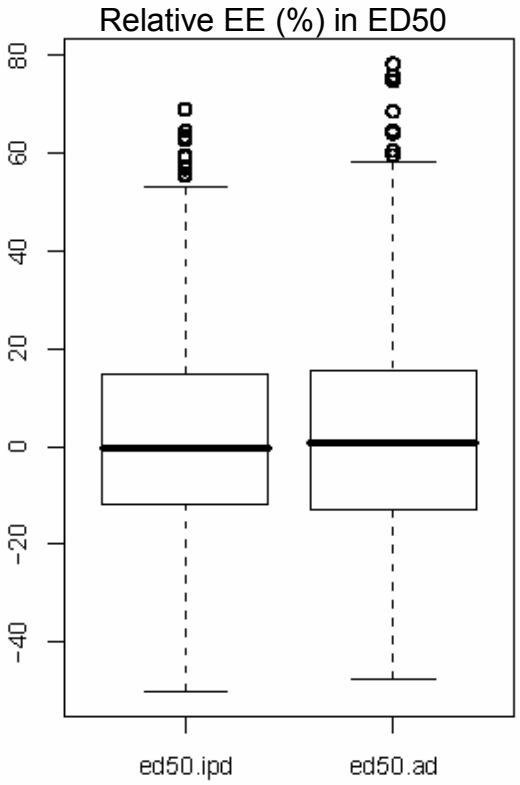
We can see similar behavior for other types of models.

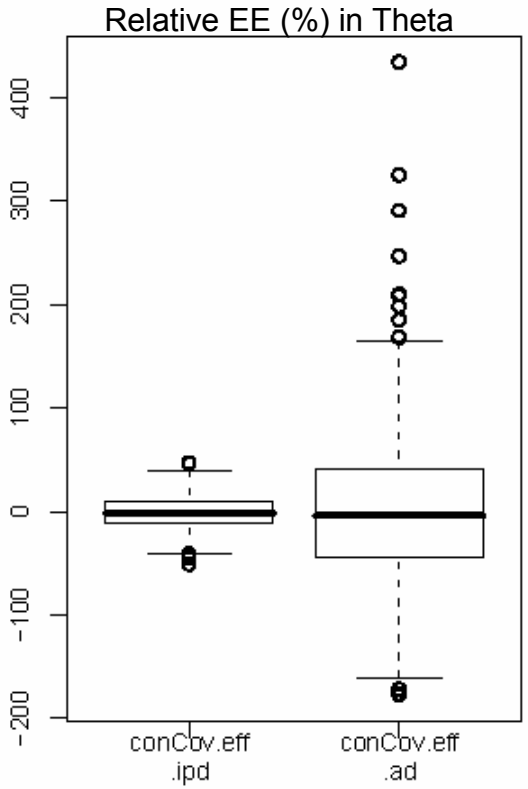
A small simulation study demonstrates that the naive approach is okay for linear effect of the covariate

We simulated IPD data from the final diabetes model (11 studies in total), then fitted the model separately to the IPD data from all studies and the AD model from all studies.

No appreciable bias from the AD only model shows that the IPD and AD models are estimating the same parameters. The IPD model is more efficient for the covariate effect, though.



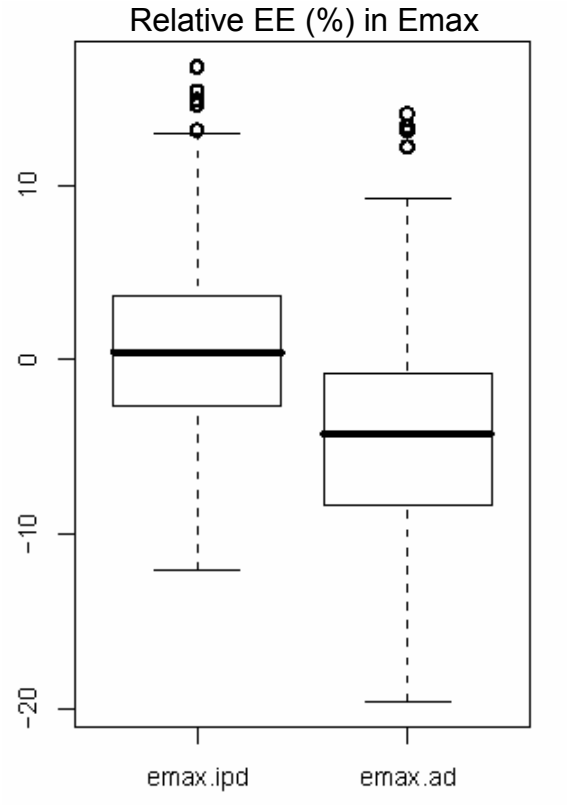


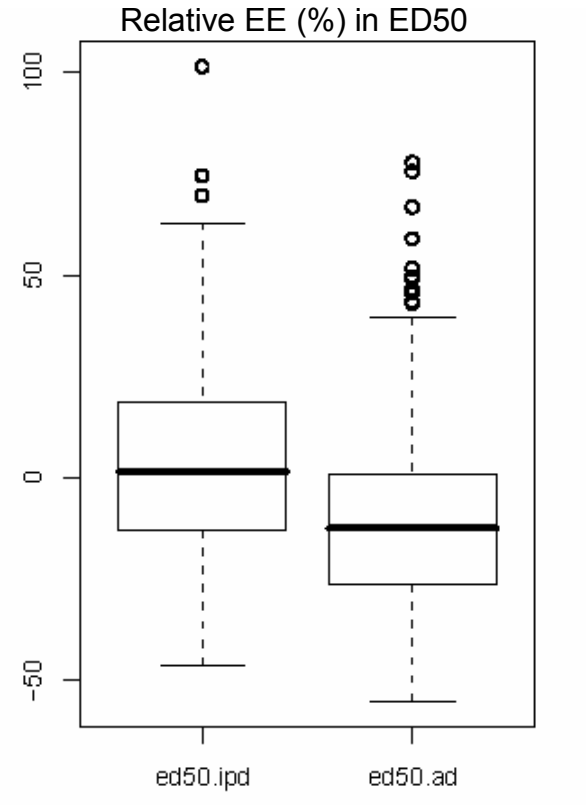


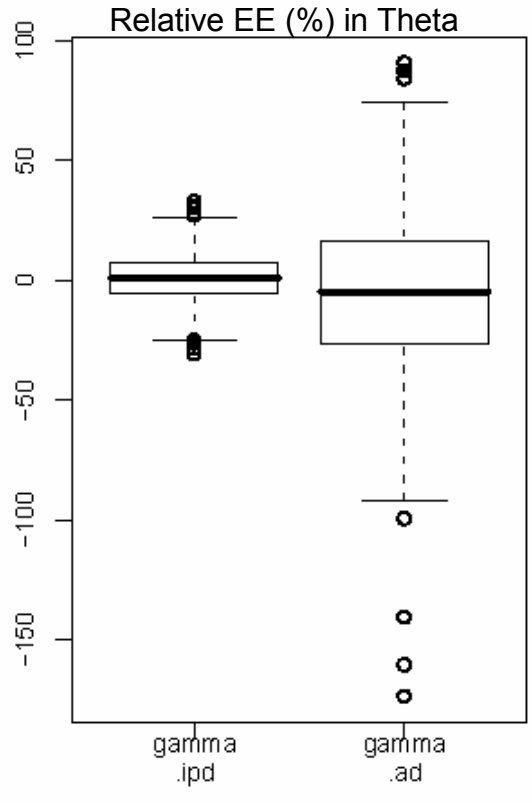
However, it would not be okay for non-linear effect of the covariate

We performed a similar simulation using the a model similar to that on slide 21.

Approximately 5% bias in Emax and ~10% bias in ED50 and covariate effect shows that the IPD and AD models are not estimating the same parameters in these models.







When can we use the same structural model for the IPD and AD?

If the structural model (viewed as a function of the covariate) can be reasonably approximated by a first-order Taylor series approximation, then we can replace the individual-level covariate values in the IPD model with the average covariate values in the AD model.

That is, if we have the model

$$Y_{ijk} = f(\text{cov}_{ijk}) + e_{ijk}$$

and

$$f(\text{cov}_{ijk}) \simeq f\left(\mu_{\text{cov}}\right) + f'\left(\mu_{\text{cov}}\right)\left(\text{cov}_{ijk} - \mu_{\text{cov}}\right)$$

then

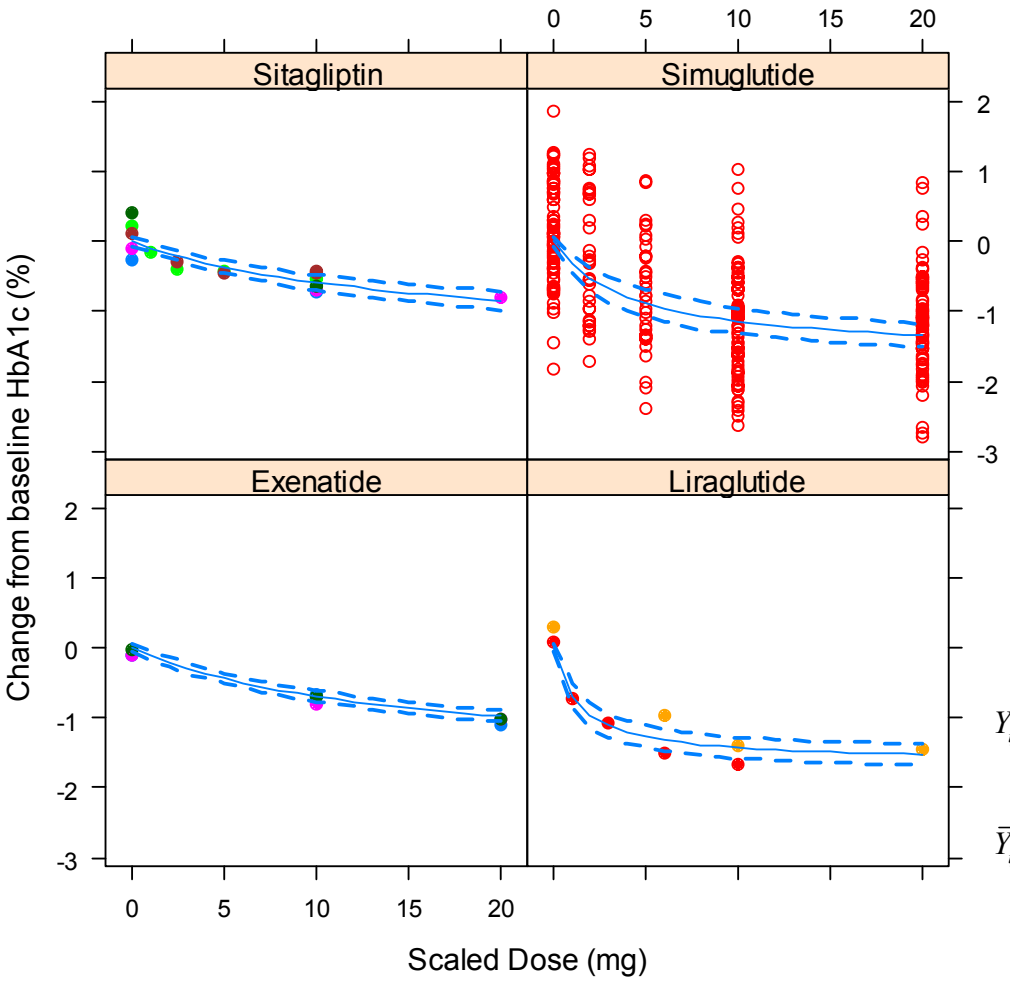
$$E_{\text{cov}}\left(Y_{ijk}\right) = E_{\text{cov}}\left(\overline{Y}_{ij}\right) = f\left(\mu_{\text{cov}}\right)$$

and the IPD and AD models have (approximately) the same structural form and we can use the "naive" approach.

What if we can't use the naïve approach?

- Use a higher order (e.g., second order) approximation to the structural model
  - In this case the AD mean response will the depend on the variance of the covariate
- Introduce both study-level (AD) and subject-level (IPD) effects of the covariate (Goldstein et al. 2000)

Results of fitting the model to the combined diabetes data



Parameter	Estimate	95% CI
E0	-0.012	-0.08, 0.06
Emax	-1.56	-1.76, -1.35
Theta	0.18	0.06, 0.30
ED50.exe	14.5	10.5, 19.9
ED50.lira	1.43	0.74, 2.73
ED50.sita	16.4	12.3, 21.8
ED50.simu	6.37	3.62, 11.2

$$Y_{ijk} = \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta\left(baseline_{ijk} - 8\right)\right] \bullet dose_{ijk}}{\mathrm{ED50}_{\mathrm{drug}} + dose_{ijk}} + \delta_{ijk}, \quad Var\left(\delta_{ijk}\right) = \sigma^2$$

$$\overline{Y}_{ij} = \mathrm{E0}_i + \frac{\mathrm{Emax} \bullet \left[1 + \theta\left(\overline{baseline}_{ij} - 8\right)\right] \bullet dose_{ij}}{\mathrm{ED50}_{\mathrm{drug}} + dose_{ij}} + \varepsilon_{ij}, \quad Var\left(\varepsilon_{ij}\right) = \sigma^2 / n_{ij}$$

Benefit on standard errors of using combined data

Parameter	SE from Combined data	Relative Standard Error	
		AD-only dataset	IPD-only dataset
E0	0.0347	1.10	2.53
Emax	0.103	1.96	1.89
Theta	0.060	4.56	0.99
ED50.exe	0.162	1.75	NA
ED50.lira	0.332	1.73	NA
ED50.sita	0.146	2.00	NA
ED50.simu	0.289	NA	1.56

The basic idea extends to a model with more than one covariate

In this case we observe $\overline{Y}_{ij}$ which is an estimate of

$$E\left(Y \mid dose\right) = \int E\left(Y \mid dose, x_1, x_2, x_3\right) p\left(x_1, x_2, x_3 \mid dose\right) dx_1 dx_2 dx_3$$

- Again, if linear in all covariates, then no problem using the naïve approach
- If not, then problematic just like with a single covariate
  - But now we need to consider the relationship between the covariates
  - Problem becomes (much) simple if the covariates are independent

Bayesian model with informative priors

- Derive (moderately) informative prior distributions for IPD model based on the AD
- If using the same structural model for IPD and AD, this approach is very similar to the hierarchical model approach
  - Because of the way Bayes theorem works
  - Thus the same issues arise with regard to ecological bias
  - That is, if the AD and IPD models are estimating different parameters, then it's not right to use a prior for the IPD parameters based on the AD model.
- Allows the modeler to down-weight the influence of the AD
  - By increasing the prior variance
  - Not necessarily easy to choose what the prior should be in this case, though.
- Relatively easily implemented in WinBUGS and NONMEM
- Need to be willing to do a Bayesian analysis (☺)

What about more complex situations ?

- This is still an open problem, particularly for
  - Non-linear mixed effects models
  - Residual models other than additive (e.g., exponential)
  - Longitudinal measures for odd-type data with covariates
  - Multiple, dependent covariates

- The basic framework is the same, though
  - Viewing the observed mean values as an expectation of a conditional model over covariate and/or random effects distributions

- Consider that for these models, the variance will now depend on the covariates and model parameters other than just omega and sigma.

- Gillespie et al. (2009) have started some nice work in this area for longitudinal models for ADAS-cog

In summary…

- There are several ways to combine AD and IPD in a meta-analysis
- The two-stage approach is simple, but does not address the primary reasons for wanting to pool IPD and AD
- The most promising is to use the hierarchical model / Bayesian approach
  - Conceptually simple and a comfortable fit for modelers who normally work with multilevel models
  - However, this has not been thoroughly evaluated for linear or non-linear models
  - Potential limitations due to aggregation bias for models that are (highly) non-linear in the covariates
  - Easy to fit in R/S-PLUS, NONMEM, etc.
  - The Bayesian approach is conceptually similar to the hierarchical model with the added flexibility of being able to down-weight the AD

….and finally

- The conceptual framework for both of these is
  - Specify the IPD model
  - Derive the corresponding AD model
  - Assess if the model is (approximately) linear in covariates
    - If so, the naïve hierarchical/Bayesian model should be fine
    - If not, be aware of the potential for aggregation bias and consider alternative approaches
  - If you're comfortable with the assumptions, fit the combined model

Acknowledgements

- Kevin Sweeney, PhD
- Caroline French

References

- Berlin JA, Santanna J, Schmid CH et al. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statist. Med*. 2002; 21: 589-624.
- Gillespie WR, Rogers JA, Ito K, and Gastonguay MR. Population dose-response model for ADAS-cog scores in patients with Alzheimer's disease by meta-analysis of a mixture of summary and individual data. American Conference on Pharmacometrics. Mashantucket, CT. October 4-7, 2009.
- Glass GV. Primary, secondary and meta-analysis of research. *Educational Researcher*. 1976; 5: 3-8.
- Goldstein H, Yang M, Omar R et al. Meta-analysis using multilevel models with an application to the study of class size effects. *Appl Statist*. 2000; 49: 399-412.
- Jackson C, Best N, and Richardson S. Improving ecological inference using individual-level data. *Statist. Med.* 2006; 25: 2136-59.
- Jackson C, Best N, and Richardson S. Hierarchical related regression for combining aggregate and individual data in studies of socio-economic disease risk factors. *J.R. Statist. Soc. A.* 2008; 171: 159-78.
- Lalonde RL, Kowalski KG, Hutmacher MM et al. Model-based drug development. *Clin Pharmacol Ther*. 2007; 82: 21-32.
- Sutton AJ, Kendrick D and Coupland CAC. Meta-analysis of individual- and aggregate-level data. *Statist. Med*. 2008; 27: 651-69.
- Wakefield J. Ecological studies revisited. *Annu. Rev. Public Health*. 2008; 29: 75-90.