METRUM
RESEARCH GROUP

Combining Summary-level and Patient-level Data for Longitudinal Dose-Response Models: Why and How

James A. Rogers, PhD

Metrum Research Group LLC

April 6, 2011

Acknowledgements

- Kaori Ito (Pfizer)
- Bill Gillespie (MetrumRG)
- Brian Corrigan (Pfizer)
- Marc Gastonguay (MetrumRG)

What Is the Question?

What Is the Question?

- Could I somehow support my wife and three children while living as a beach bum on Coronado Island?

What Is the Question?

- Could I somehow support my wife and three children while living as a beach bum on Coronado Island?
- How can we optimize clinical trials in Alzheimer's Disease with respect to, among other things:

What Is the Question?

- Could I somehow support my wife and three children while living as a beach bum on Coronado Island?
- How can we optimize clinical trials in Alzheimer's Disease with respect to, among other things:
  - trial duration
  - assessment schedule

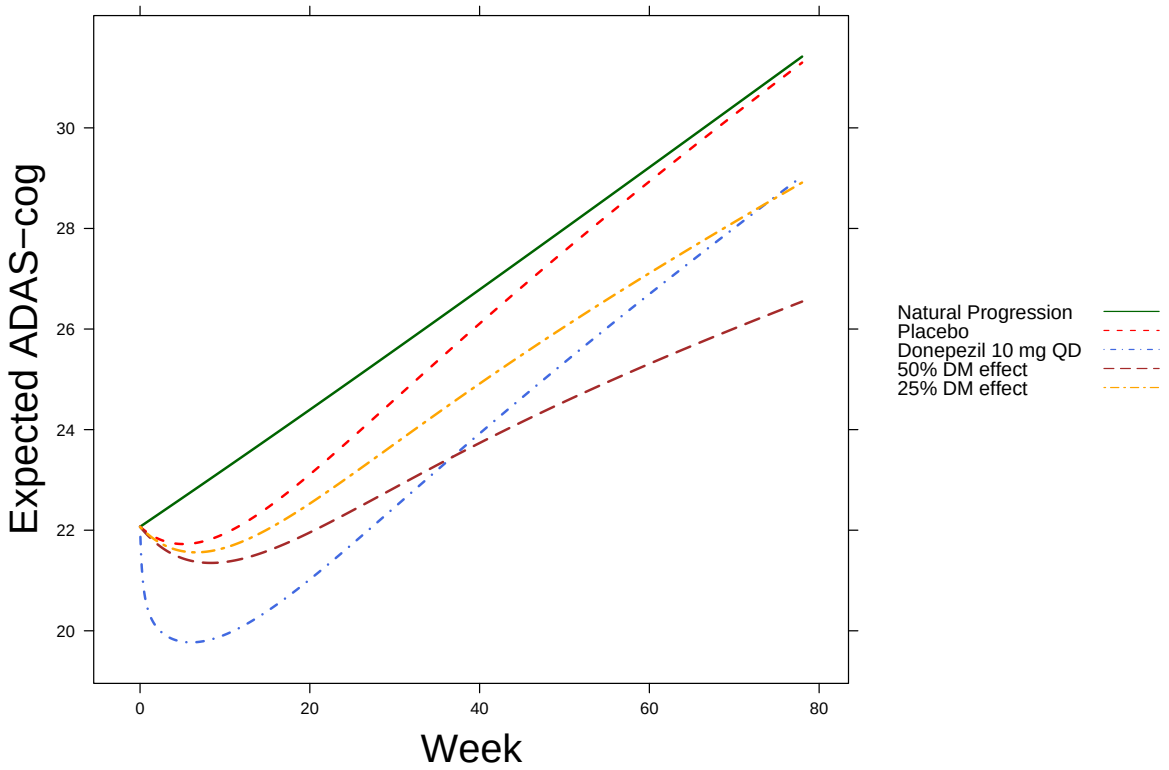
What Is the Question?

- Could I somehow support my wife and three children while living as a beach bum on Coronado Island?
- How can we optimize clinical trials in Alzheimer's Disease with respect to, among other things:
  - trial duration
  - assessment schedule
  - primary analysis methodology

What Is the Question?

- Could I somehow support my wife and three children while living as a beach bum on Coronado Island?
- How can we optimize clinical trials in Alzheimer's Disease with respect to, among other things:
  - trial duration
  - assessment schedule
  - primary analysis methodology
  - **inclusion / exclusion criteria**

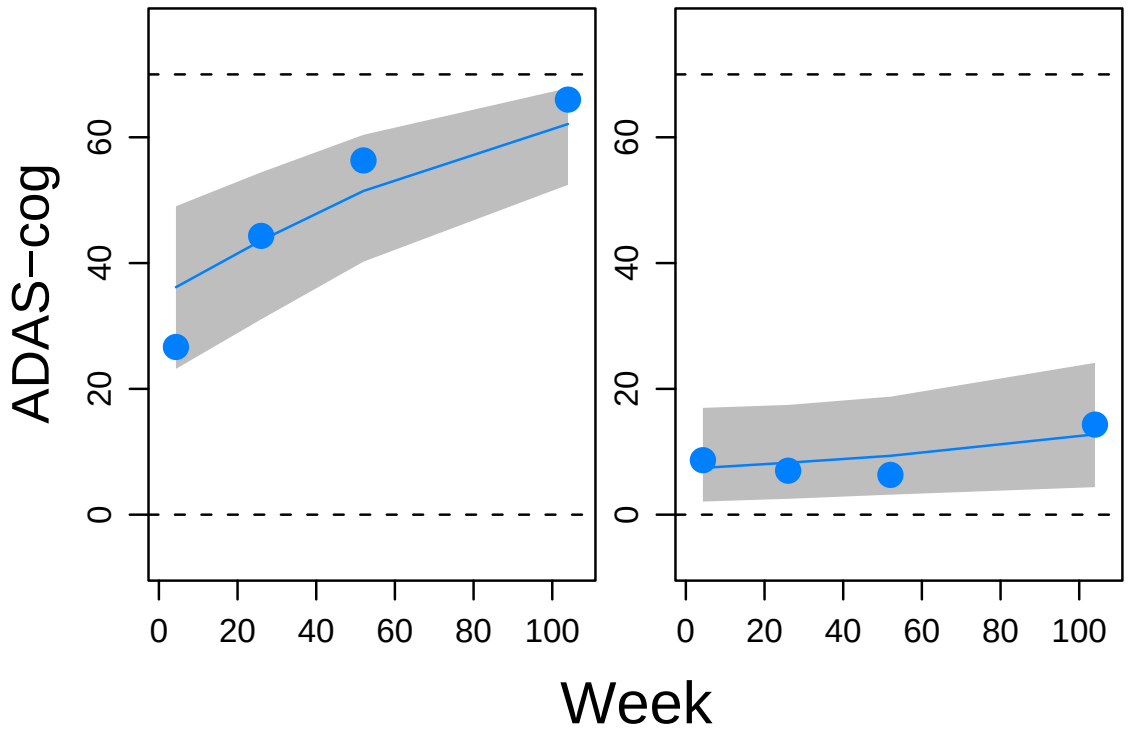
Structural Components of ADAS-cog Progression Model



Structural components (exposure-response version) originally proposed by Holford and Peace (1992).

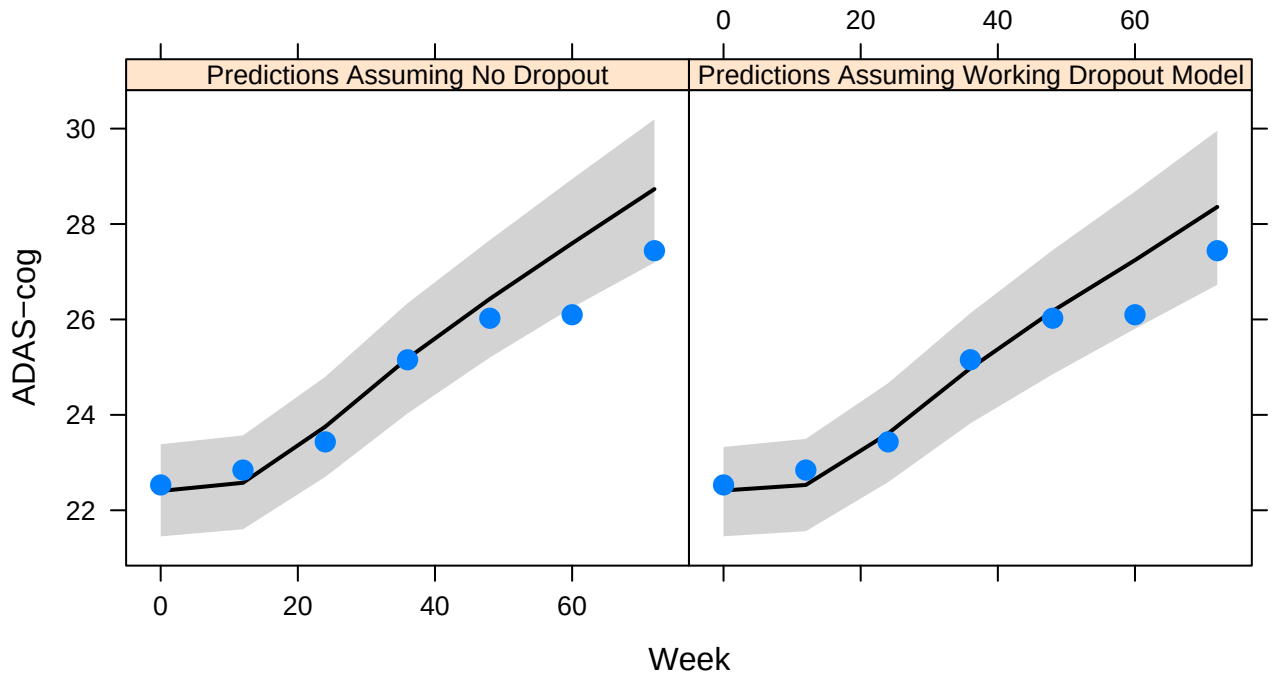
Adapted to longitudinal suitable for meta-analysis by Ito et al. (2008).

Beta-logit Approach to Respecting the 0–70 Constraints of the ADAS-cog



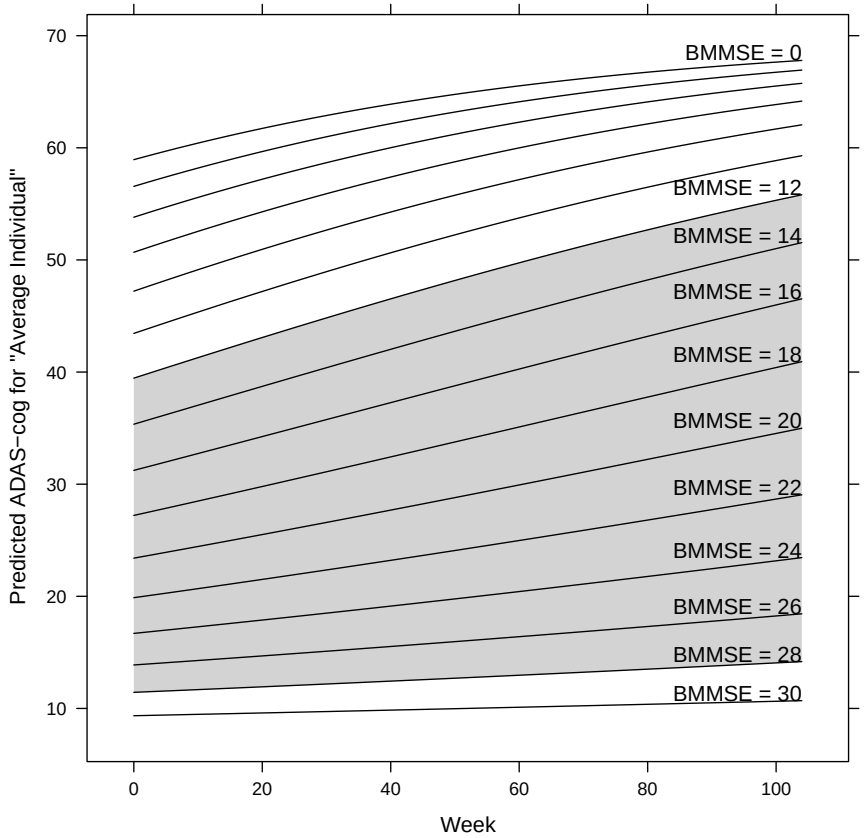
for more detail see http://www.metrumrg.com/images/stories/publications/adascog-bayesconf.pdf

Preliminary Consideration of Plausible Missing Data Mechanisms



for a review and discussion of some of the issues see Gastonguay *et al.* JCP 2010.

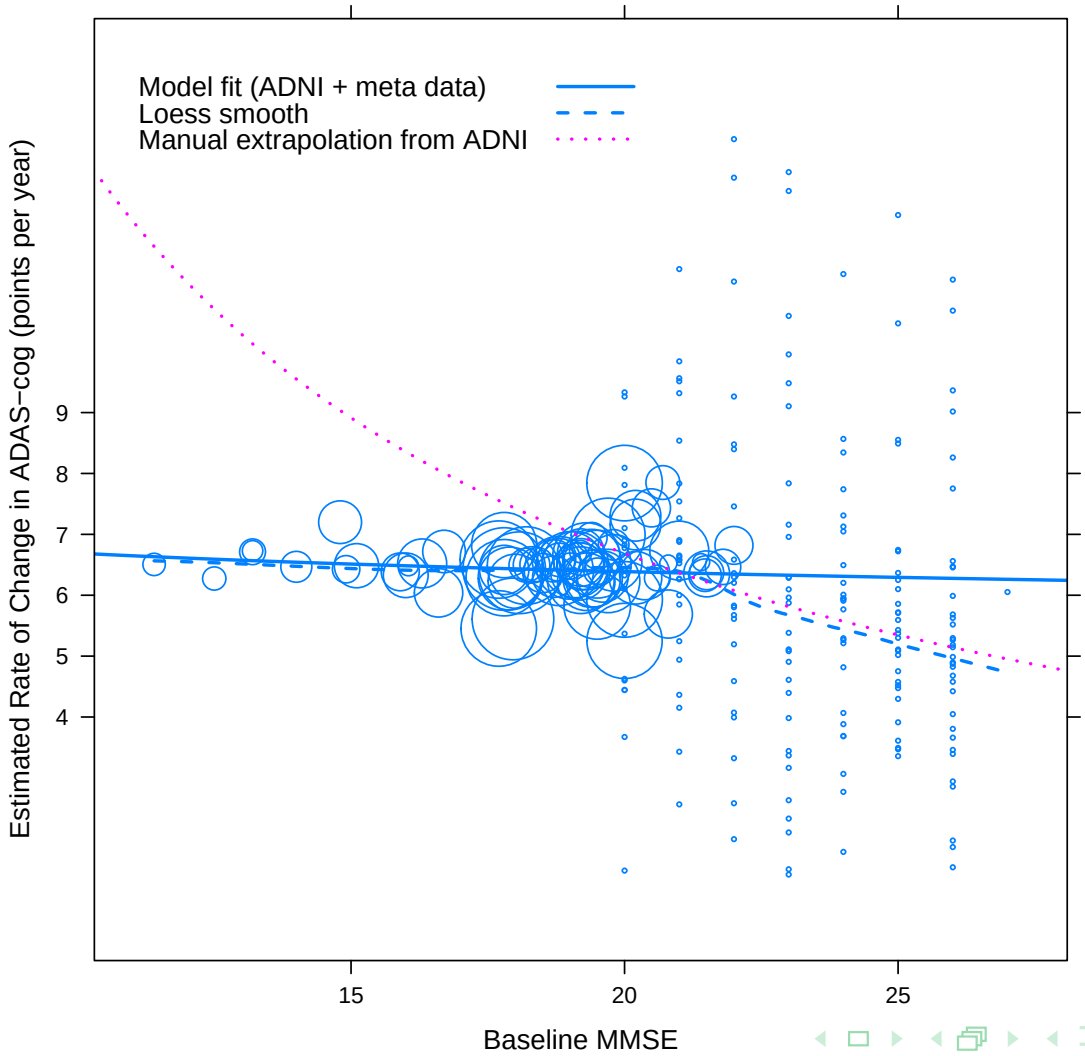
Baseline MMSE as a Covariate on Rate of Progression



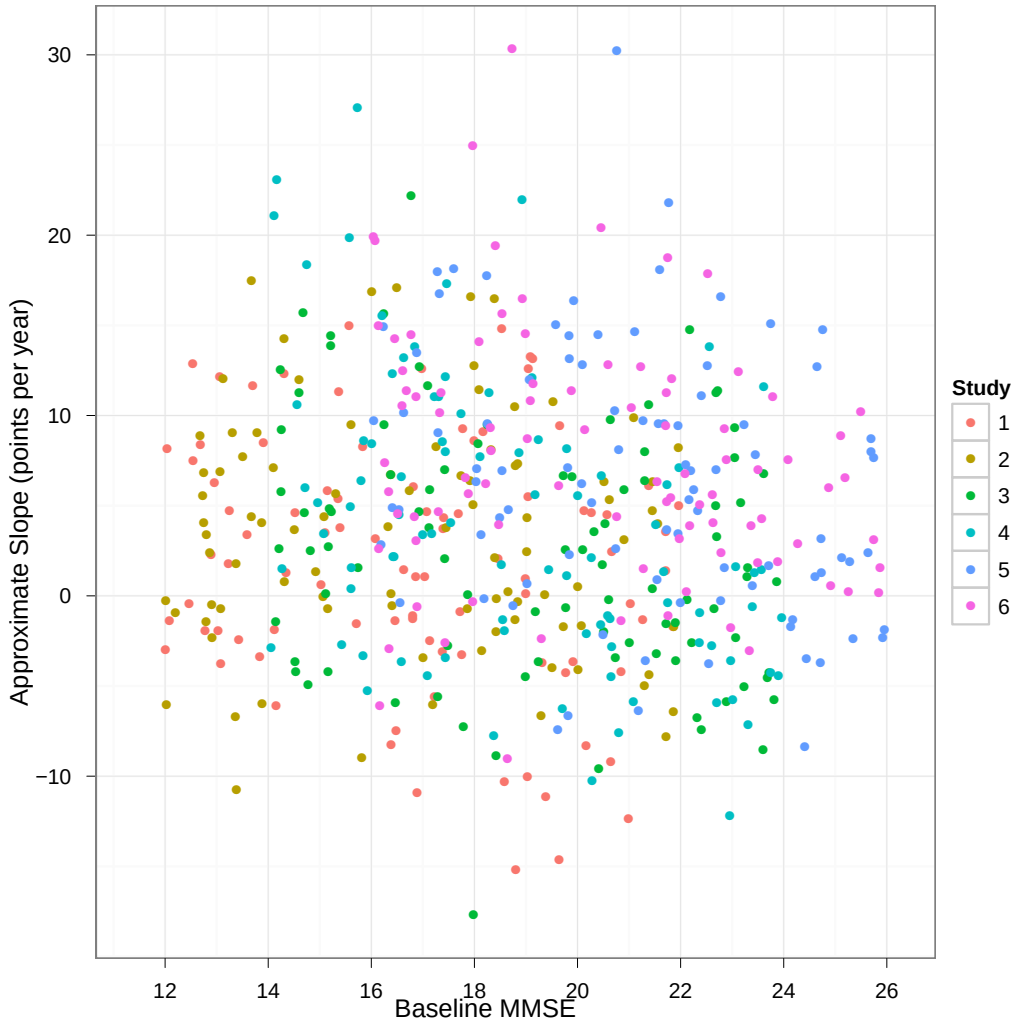
Baseline MMSE	Posterior Median	95% Credible Interval
12	9.69	8.16–11.31
14	8.75	7.53–10.08
16	7.84	6.85–8.93
18	6.92	6.03–7.82
20	6.01	5.05–6.84
22	5.09	3.96–6.02
24	4.17	2.78–5.3
26	3.24	1.58–4.69

Covariate-adjusted slopes for “average individuals” (i.e. with all random effects set to zero). We use baseline MMSE because this is entry criteria.

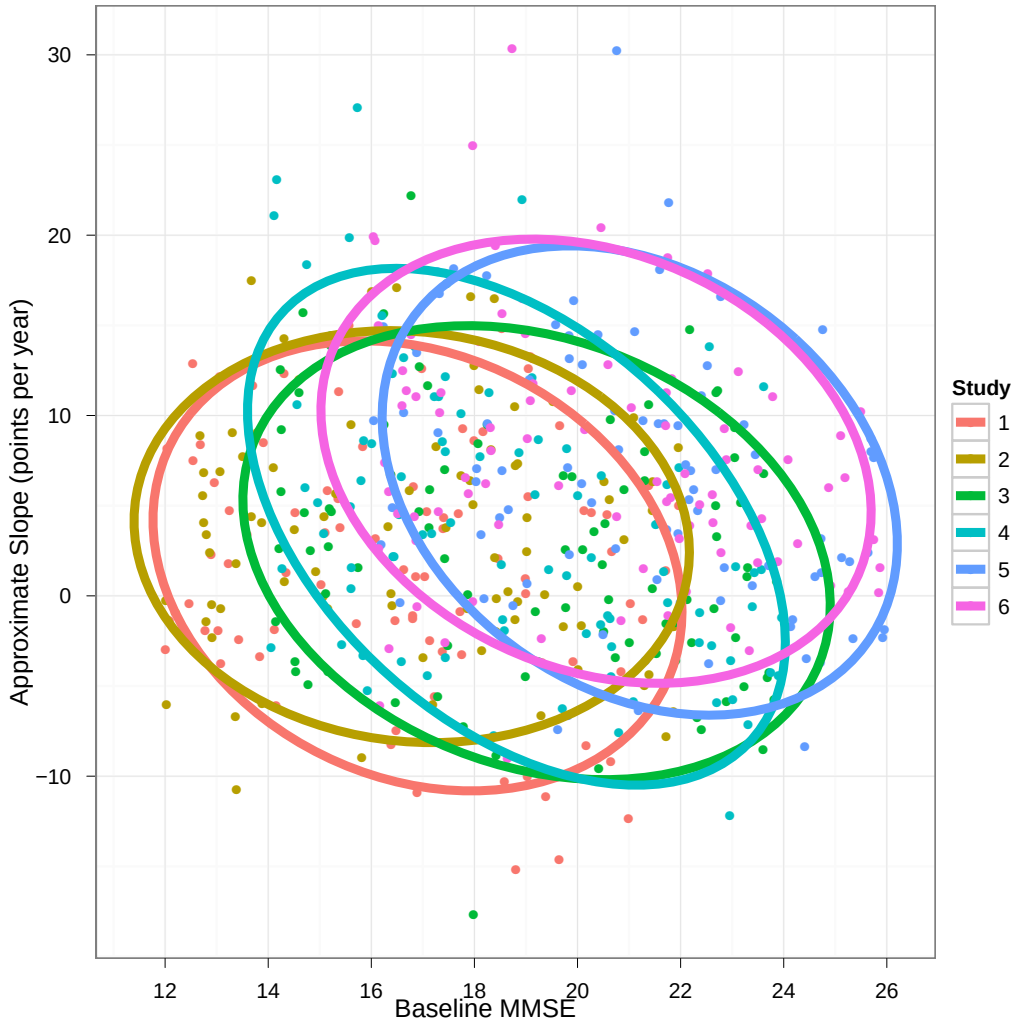
Covariate Relationship Appears Different in Summary-level Versus Patient-level Data



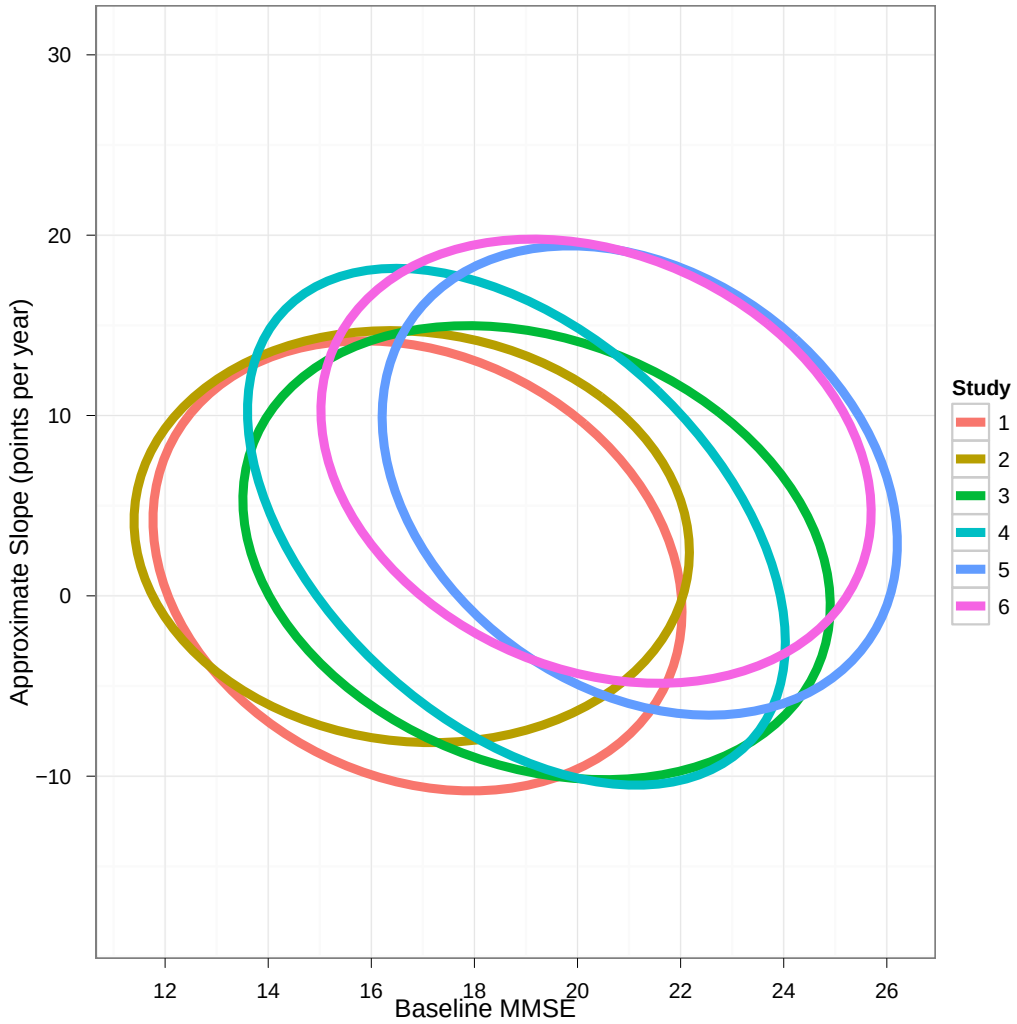
Recreating the Problem With Simulated Data



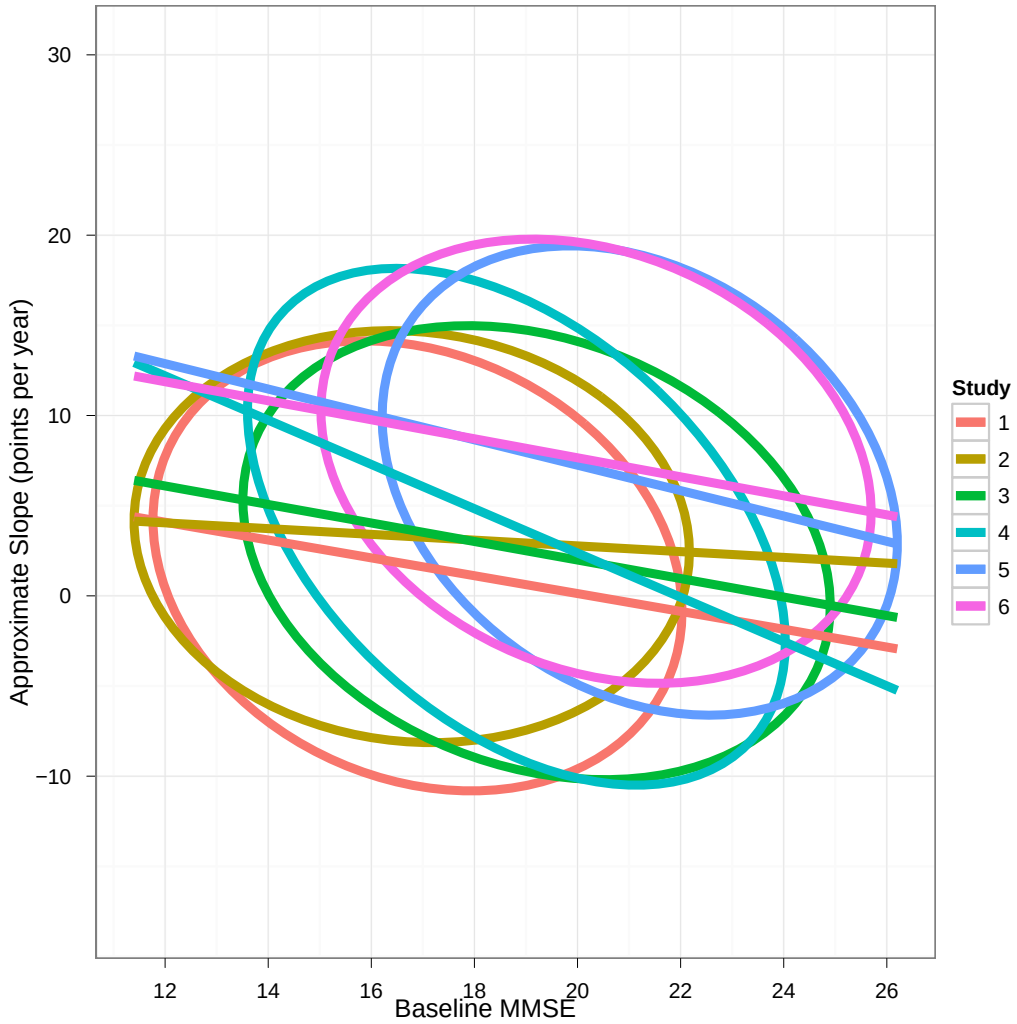
Higher Baseline MMSE → Slower Rate of Progression



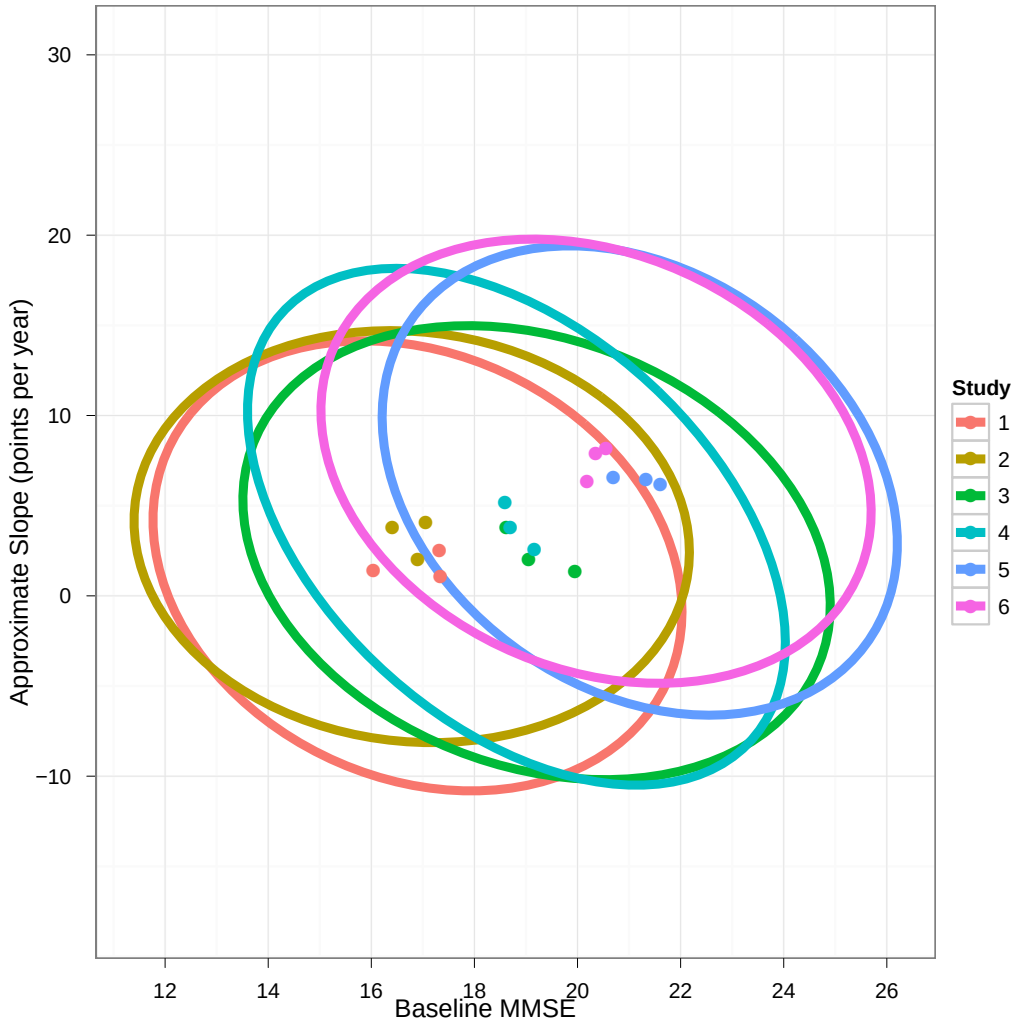
Higher Baseline MMSE → Slower Rate of Progression



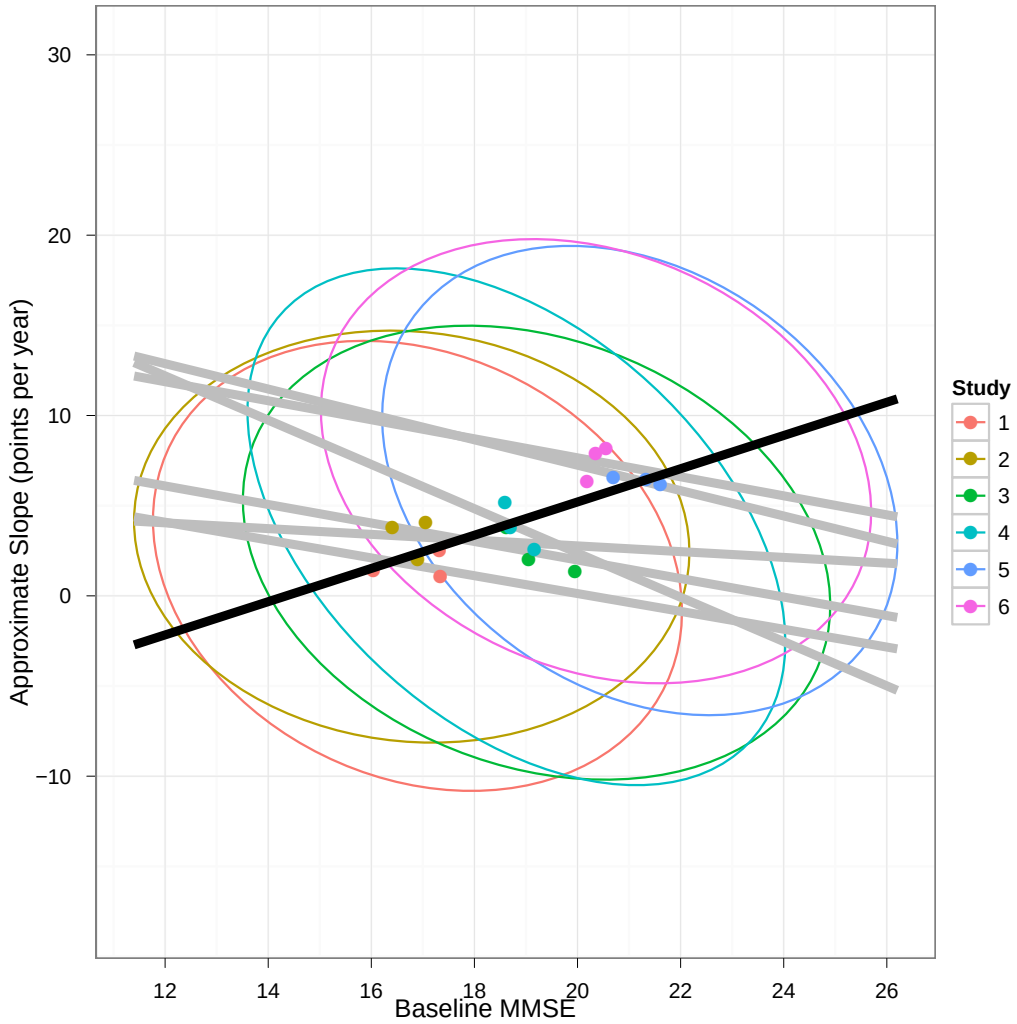
Higher Baseline MMSE → Slower Rate of Progression



If We Analyze the Group Means Naively, . . .



. . . Apparent Direction of Association May be Wrong



A Reminder of the Obvious

We randomize within studies, we don't randomize between studies.

For example, no one randomly assigned studies in more severe populations to the 1980s and studies of less severe populations to the 2000s.

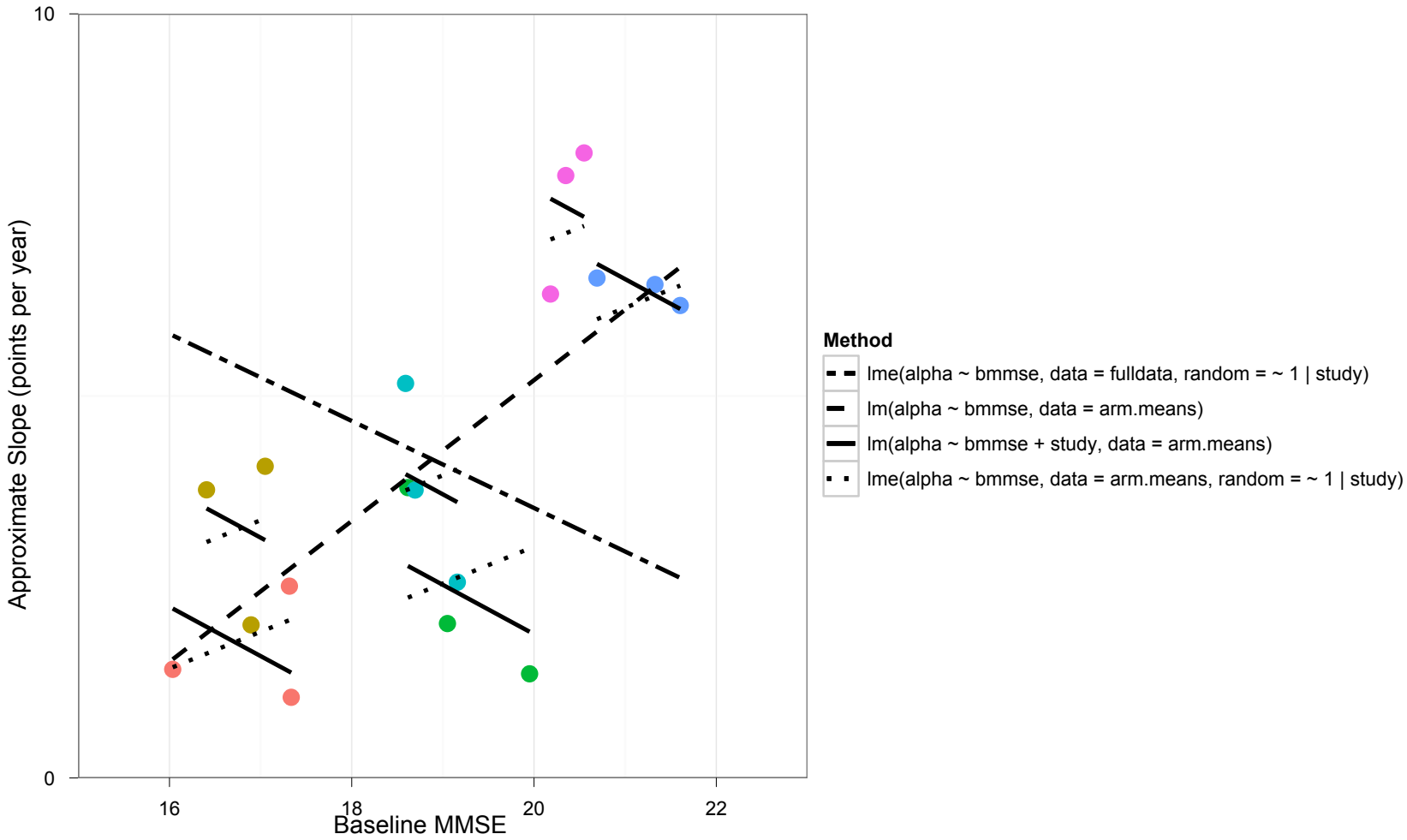
Much greater opportunity for confounding!

Solution? Incorporate fixed or random study effects?

The problem is that we often have fairly little within-study information

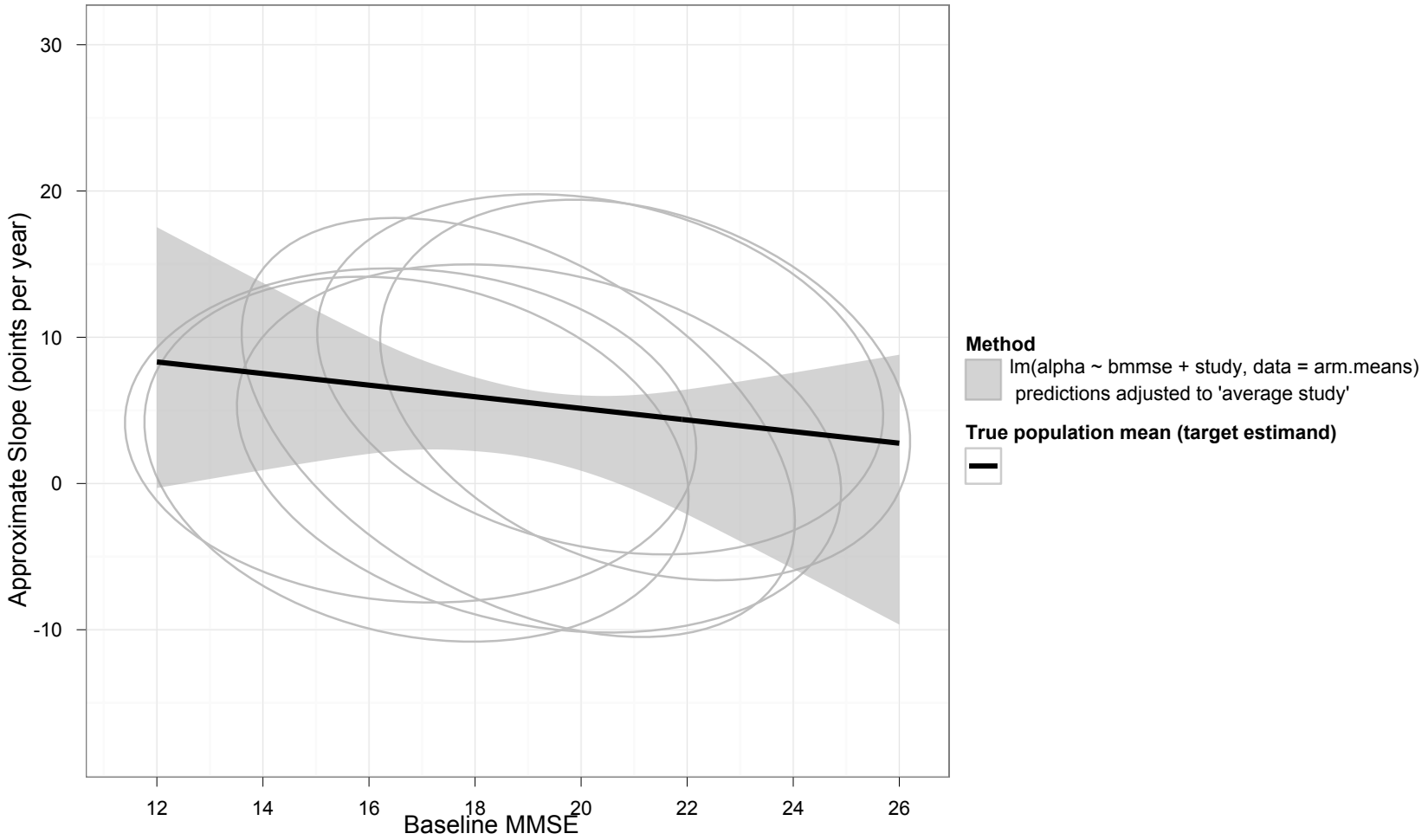
- Generally relatively few treatment arms within each study.
- Generally very little dispersion in the “X” variable within each study. In fact, studies are generally randomized to ensure this!

The Problem With Random Study Effects



Overshrinkage!

The Problem With Fixed Study Effects



Overparameterization → Imprecise Estimates

Summary of the Problem

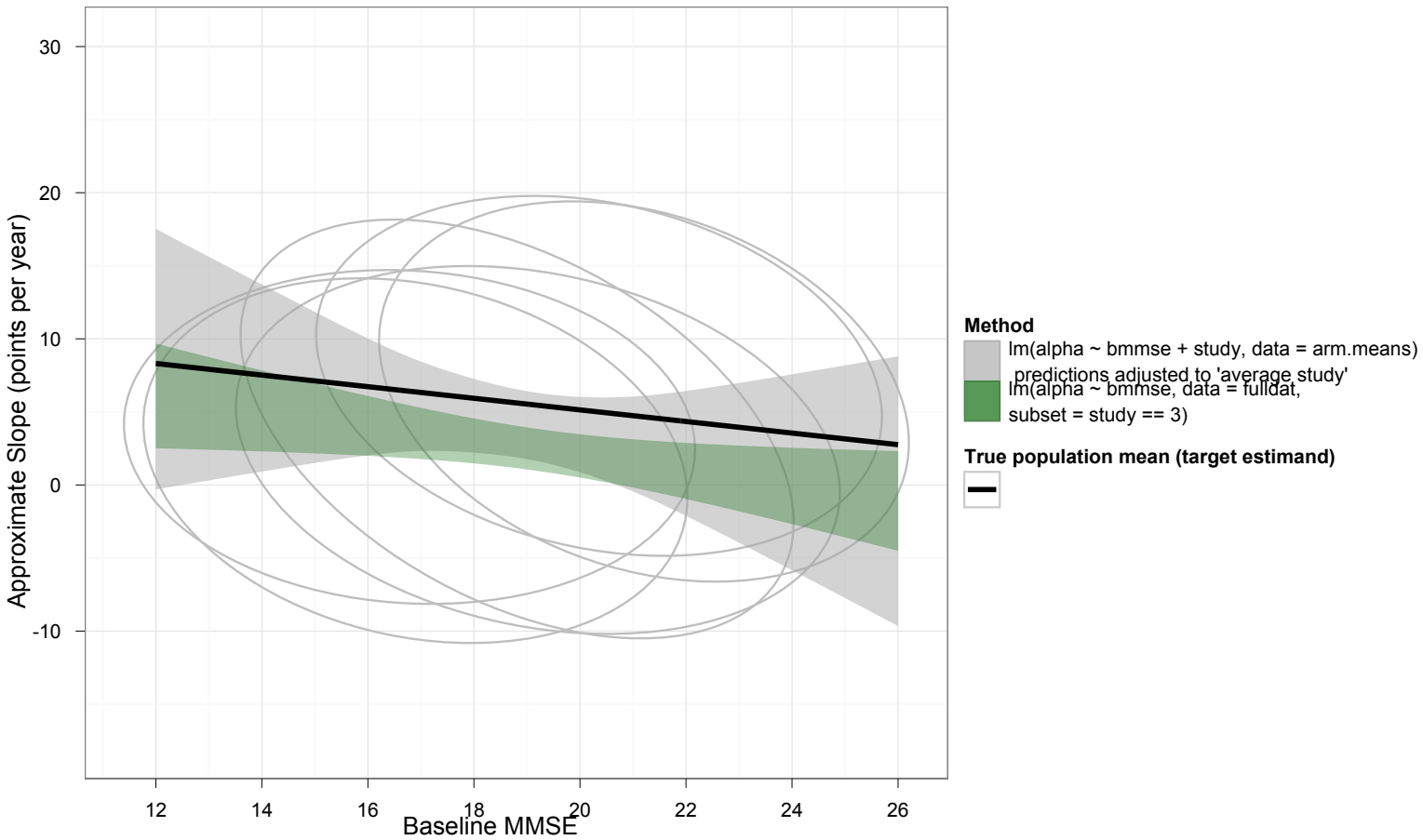
When using only summary-level data, you might be out of luck:

A. Failing to model inter-study variation $\longrightarrow$ potentially totally misleading results.

B. Modeling inter-study variation using random effects $\longrightarrow$ over-shrinkage $\longrightarrow$ Almost equivalent to option A.

C. Modeling inter-study variation using fixed effects $\longrightarrow$ highly variable estimates, and no rigorous framework for predicting the "next study".

Tradeoffs Between Summary-level Analysis and Patient-level Analysis



The Reese's Peanut Butter Cup Approach

Known fact: two great tastes can taste great together

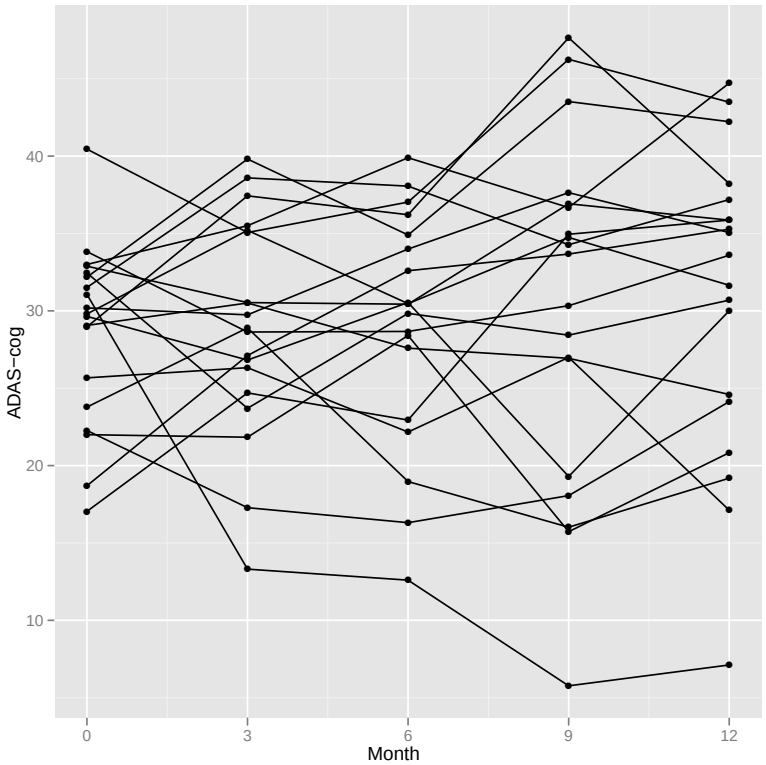
Summary-Level Data

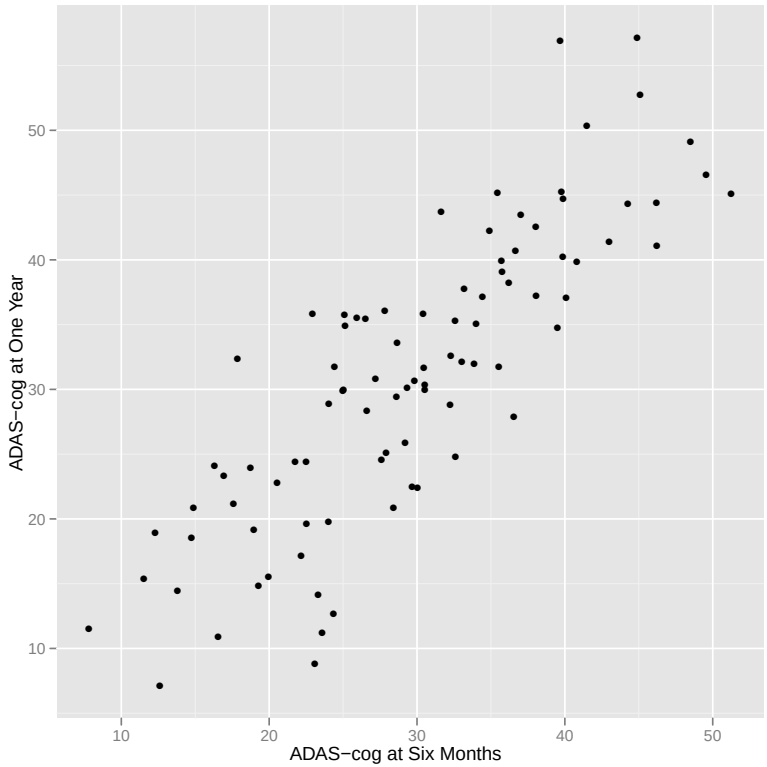
- More balanced, comprehensive view.
- Estimates of competitor drug effects.

Patient-level Data

- Better estimation of covariate relationships.
- Better estimation of variance components.

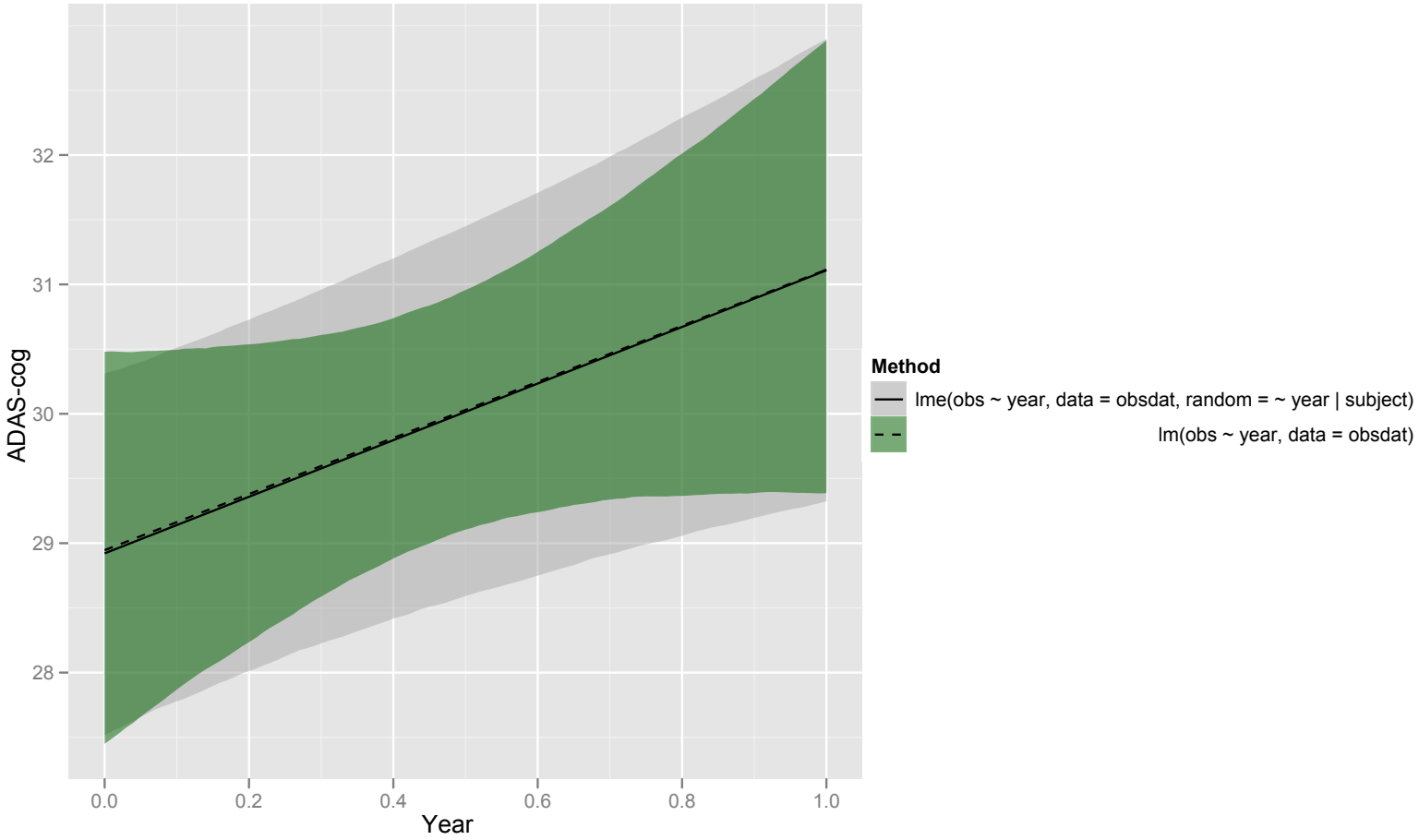
Recall How We Handle Patient-level Longitudinal Data





Generally use Mixed Effects Modeling or some other methodology (e.g. generalized least squares) that reflects stochastic dependence at different time points.

Treating Postively Correlated Obs'ns As If They Were Independent → Too-Narrow Conf. Intervals



Basic Sampling Theory For Mixed Effects Models

Let *i* be the index on patients and *j* be the index on timepoints.

- **If:**

$$\begin{aligned} Y_{ij} &= \mu + S_i + \epsilon_{ij} \\ S_i \sim N(0, \sigma_S^2) \quad &; \quad \epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \\ \text{implying that } \mathrm{Corr}(Y_{11}, Y_{12}) &= \sigma_S^2/(\sigma_S^2 + \sigma_\epsilon^2) \end{aligned}$$

- **Then:**

$$\begin{aligned} \overline{Y}_{\cdot j} &= \mu + \overline{S}_{\cdot} + \overline{\epsilon}_{\cdot j} \\ \overline{S}_{\cdot} \sim N(0, \sigma_S^2/n) \quad &; \quad \overline{\epsilon}_{\cdot j} \sim N(0, \sigma_\epsilon^2/n) \\ \text{implying that } \mathrm{Corr}(\overline{Y}_{\cdot 1}, \overline{Y}_{\cdot 2}) &= \sigma_S^2/(\sigma_S^2 + \sigma_\epsilon^2) \end{aligned}$$

The Moral

- Longitudinal models still need random effects[1] even when based on meta-data.
- It does not suffice to simply weight residuals by $\sqrt{n}$.
- $\sqrt{n}$ is not a magic talisman: basic sampling theory dictates where and when to use it.

[1] or other means of reflect dependence between time points

WinBUGS Model Specification

```
...
for(j in 1:nUnits) {

  ## Between-unit variation, conditional on study :
  muEtaAdj[j] <- muEta[ studyUnit[j] ] + lambdaEta * (bmmse[j] - meanBMMSE)
  tauEtaAdj[j] <- tauEta * nPatUnit[j]
  eta[j] ~ dnorm(muEtaAdj[j], tauEtaAdj[j])

  muAlphaAdj[j] <- muAlpha[ studyUnit[j] ] + lambdaAlpha * (bmmse[j] - meanBMMSE)
  tauAlphaAdj[j] <- tauAlpha * nPatUnit[j]
  alpha[j] ~ dnorm(muAlphaAdj[j], tauAlphaAdj[j])

}

...
```

WinBUGS Model Specification

```
...
  for(i in 1:nObs) {

    ePlacebo[i] <- beta * ( exp( -kel * time[i] ) - exp( -keq * time[i] ) )

    iDrug[i] <- max(1, drug[i])   # prevents syntax error when treatment = placebo
    eDrug[i] <- (1 - equals(drug[i], 0)) *
      pow(dose[i] / doseRef[ iDrug[i] ], gamma[ iDrug[i] ]) *
        eDelta[ iDrug[i] ] * time[i] / (et50.use[ iDrug[i] ] + time[i])

    logitTheta[i] <- eta[ unit[i] ] + alpha[ unit[i] ] * time[i] + ePlacebo[i] + eDrug[i]
    theta[i] <- 1 / (1 + exp(- logitTheta[i]) )

    ## Wihin-unit variation (conditional on unit and study)
    tauAdj[i] <- tau * nPat[i] + nPat[i]  - 1
    aAdas[i] <-  (theta[i] + 0.001) * tauAdj[i]
    bAdas[i] <- (1 + 0.001 - theta[i]) * tauAdj[i]
    normAdas[i] ~ dbeta(aAdas[i], bAdas[i])
}
...
```

Other MBMA Challenges Not Addressed Here

- Assessing publication bias in model-based meta-analyses.
- Within-study biases due to missing data.
- Unavailability of complete longitudinal sample sizes.
- Non-collapsibility of nonlinear relationships (do Wikipedia Jensen's Inequality if you're not familiar with it!)

Thank You!