METRUM INSTITUTE

MI260: Model-based Meta-analysis to Support Decision-Making in Clinical Drug Development

Bill Gillespie, Jonathan French and other Metrum Institute scientists

Metrum Institute

Spring 2013

Course Introduction

MI260: Model-based Meta-analysis to Support Decision-Making in Clinical Drug Development provides an introduction to model-based meta-analysis of summary data or a combination of summary and individual data from clinical trials to support decision-making in clinical drug development. The course duration and content is equivalent to a single semester 3 credit course at a typical institution of higher learning. Each week's topic will consist of a lecture (two hours) followed by a hands-on lab (one hour). The general plan will be as follows:

- Lectures will be on Wednesdays at 2 PM EDT.
- Hands-on labs will be on Mondays at 2 PM EDT (in some cases, the lecture may finish during the first part of the lab on Monday).

Student Expectations and Requirements for Certificate

- All students are expected to attempt the hands-on exercises prior to the lab session. Instructors will not grade homework assignments, but will review solutions with the entire class on a weekly basis.
- A midterm take-home exam will be assigned at the midpoint of the course.
- A final take-home exam will be assigned at the end of the course (due one week after it is posted).
- Students will be required to complete and submit a modeling project before the end of the course. This project will be based on a real-world (or similar) problem, and will include components of data assembly, model development and evaluation, and a brief report. More details to follow...
- Course grade will be based on the midterm (25%), final (25%) and modeling project (50%)

Course Content Management Site

All students should already have an account to access the main course website. Here's the link: (http://training.metruminstitute.org). Postings to this site will automatically generate an email message to your own e-mail accounts. This site is intended to be the primary repository for all course resources including:

- **News Forum:** Here you'll find updates about class schedule, assignments, etc.
- **Discussion Forum:** Direct your questions about course content to instructors or other students here. You can contribute to ongoing discussions or start a new thread.
- **Link** to the GoToWebinar webcast registration form.
- **Course Materials:** You'll find course notes, examples, and links to recorded lectures under each weekly class heading.

Computer resources

Use of your own computer

- The initial course materials, including the software we will use throughout the course, are contained in the (rather large) zip file named MI260USB.zip available on the course website.
- For the course we recommend you run the software and course examples directly from a USB flash drive via the shortcuts provided. Install the contents of MI260USB.zip according to:
  1. Download MI260USB.zip to your computer.
  2. Unzip MI260USB.zip. This will create a folder named MI260USB.
  3. Obtain and insert a USB flash drive with $\geq$ 1 GB capacity.
  4. Copy the *contents* of MI260USB (not the MI260USB folder itself) to the flash drive.
  5. Check the software installation by double-clicking on the shortcuts named "R 2.15.2.cmd" and "WinBUGS14.cmd". In each case the corresponding program (R or WinBUGS) should be successfully launched. If not, please report it via the Technical Support forum on the course website.

Model-based Meta-analysis to Support Decision-making in Clinical Drug Development

Objective

Provide an introduction to model-based meta-analysis of summary data or a combination of summary and individual data from clinical trials to support decision-making in clinical drug development.

Primary intended audience: pharmacometricians with biological or statistical modeling skills

Background assumed

- PK/PD or statistical modeling
- Some familiarity and hands-on experience with nonlinear regression, mixed effects modeling, Bayesian modeling using WinBUGS and use of R (or S-PLUS).

Course Outline

- Introduction
  - Rationale and role of model-based meta-analysis in clinical drug development
    - Why do it?
    - What decisions benefit from meta-analysis and model-based meta-analysis in particular?
  - Motivating examples
- The systematic review and planning your meta-analysis
  - Analysis plan
  - Database construction
    - Data sources
    - Data types, e.g., mean, mean change from baseline, percent change from baseline, standard deviation, standard error, ...

Course Outline (cont.)

- Traditional meta-analysis
  - What is it?
  - Fixed effects meta-analysis
- Random effects meta-analysis and meta-regression
  - Measures of heterogeneity
  - What the traditional random effects model is and how it differs from fixed effects
  - Meta-regression
- Selection bias and missing data
- Combining different types of data
- Network Meta-analysis
  - Relationship to random effects meta-analysis
  - Assumptions
  - Fitting the models

Course Outline (cont.)

- Model-based meta-analysis (MBMA)
  - What is it?
  - Role of MBMA: Objectives not adequately addressed by traditional meta-analysis
  - Why Bayesian? / Why BUGS?
- Modeling sample mean data
- Modeling sample standard deviations: why and how

Course Outline (cont.)

- Population simulations
  - Simulating probable ranges of population estimands, e.g., population mean, probability of an event, etc.
  - Using simulation results to support decision-making in a competitive market environment
- Issues arising from analysis of summary data
  - Applying models developed to describe responses in individuals to summary data
  - Analysis of longitudinal data
    - Pitfalls of treating treatment arms as “super-patients”
    - Within-arm correlation
    - Approaches for addressing these issues

Course Outline (cont.)

- Modeling other types of summary statistics:
  - Number or fraction of patients with a particular outcome or that experience an event
  - Number or fraction of patients within each level of an ordinal scale
  - Number of events per patient
  - Summary statistics for time-to-event measurements
- Issues arising from use of LOCF and OC data
- Combining summary and individual data

Course Outline (cont.)

- Incorporating a broader range of data and knowledge
  - Leveraging the Bayesian framework to incorporate additional quantitative knowledge via informative prior distributions
  - Integrating preclinical, biomarker and clinical outcome data to improve prediction and decision-making in early clinical development
- Miscellaneous topics
- Closing discussion

Why do it?

- The primary rationale for model-based meta-analysis (MBMA) is to improve decision-making by better leveraging prior information from multiple sources.
- Decision-makers generally attempt to consider such prior information, but it is usually done in a relatively qualitative manner, and each individual decision-maker is usually aware of only a subset of the prior information.
- MBMA seeks to make the process more quantitative and comprehensive.
- The process and results of MBMA should be made visible (aka transparent) to the decision-makers.
- The end result is that the decision-makers are better informed, and they can contribute their knowledge to the modeling process leading to better, more trusted models and model-based inferences.

Why do it?

- Leverage prior knowledge from multiple sources
  - Data on the NCE of interest
    - Preclinical, Phase I safety & biomarkers, clinical safety & efficacy
  - Knowledge about the target disease & affected physiologic systems
  - Knowledge/data on related compounds
    - From proprietary or public sources
    - Data for analogs that share the same MOA and possibly chemical properties
    - Data for competitors should be included to provide benchmarks for comparison.
- Comparison of competing drugs/treatments
  - Information about analogs may permit information borrowing
  - Shared model parameters are more precisely estimated
  - Comparative inferences are also more precise
- Prediction of unobserved clinical outcomes
  - Build models for preclinical-to-clinical or biomarker-to-clinical relationships
  - Use those models to predict clinical outcomes for NCEs for which no clinical data is yet available

What decisions benefit from model-based meta-analysis?

- MBMA can enhance dose-selection and PoC decisions in at least 2 ways.
  1. MBMA can improve quantitative comparisons with competing treatments. This may permit better selection of a dosing regimen that performs comparable to or better than the competing treatment.
  2. Alternatively, the MBMA results may demonstrate that no dosing regimen of the new drug performs favorably relative to competitors. If so, MBMA may support a better and earlier decision to terminate development.

What decisions benefit from model-based meta-analysis?

- In some cases models developed via MBMA may be suitable for simulating clinical trials and thereby improving decisions regarding the designs for such trials.
  - MBMA is particularly valuable in cases where no clinical efficacy data is yet available for the new treatment, but quantitative predictions for efficacy-related measurements are possible by leveraging data for related compounds.
  - Even if those predictions are highly uncertain, the clinical trial simulations provides the means for optimizing the trial design in the presence of such uncertainty.
- MBMA may also be useful for decisions related to market differentiation and choice or prioritization of indications.

Indication	Endpoints modeled	Drugs
Restless Legs Syndrome	RLS rating scale	placebo, ropinirole, pramipexole, internal compounds
Insomnia	LPS	eszopiclone, zolpidem, indiplon, zaleplon, zolpidem, placebo, internal compounds
Generalized anxiety disorder	HAM-A	Pregabalin, paroxetine, placebo
Chronic Obstructive Pulmonary disease	FEV1	beclomethasone, montelukast, Roflumilast, placebo
Diabetes	Fasting plasma glucose, glycosolated hemoglobin, weight	troglitazone, placebo, internal compounds
Depression	HAM-D, MADRS, dropout rate	citalopram, duloxetine, fluoxetine, fluvoxamine, nefazadone, paroxetine, sertraline, venlafaxine, venlafaxine xr, placebo
Hypercholesterolemia		
obesity	Weight loss, glycosolated hemoglobin, blood pressure, heart rate, HDL,	Orlistat, rimonababnt, sibutramine, internal candidates.
Pain (acute)	dental pain, WOMAC pain, WOMAC composite, PGIC	celcoxib, rofecoxib, naproxen/na, valdecoxib, placebo and internal compounds
Pain (neuropathic)	Pain Likert scale, adverse events	gabapentin, pregabalin, internal compounds
Alzheimers		donepezil, placebo,
Schizophrenia	PANSS, D2/D3 receptor occupancy	Ziprasidone, olanzepine, risperidone, aripiprazole, haloperidol, placebo
Venous thromboembolism	VTE, %severe bleeding, aPTT	melagatran, enoxaparin, PD0348292
HIV	Viral RNA copies	various compounds

Table 1. List of Indications, Endpoints, and Drugs Currently Being Modeled Using a Model-based Meta-Analyses Approach

Table from Corrigan et al. AAPS NEWSMAGAZINE September 2007 pp 26-27 illustrates the increasing use of MBMA in clinical drug development

The AAPS Journal 2005; 7 (3) Article 52 (http://www.aapsj.org).

Themed Issue: Population Pharmacokinetics - In Memory of Lewis Sheiner
Guest Editors - Peter Bonate and Diane Mould

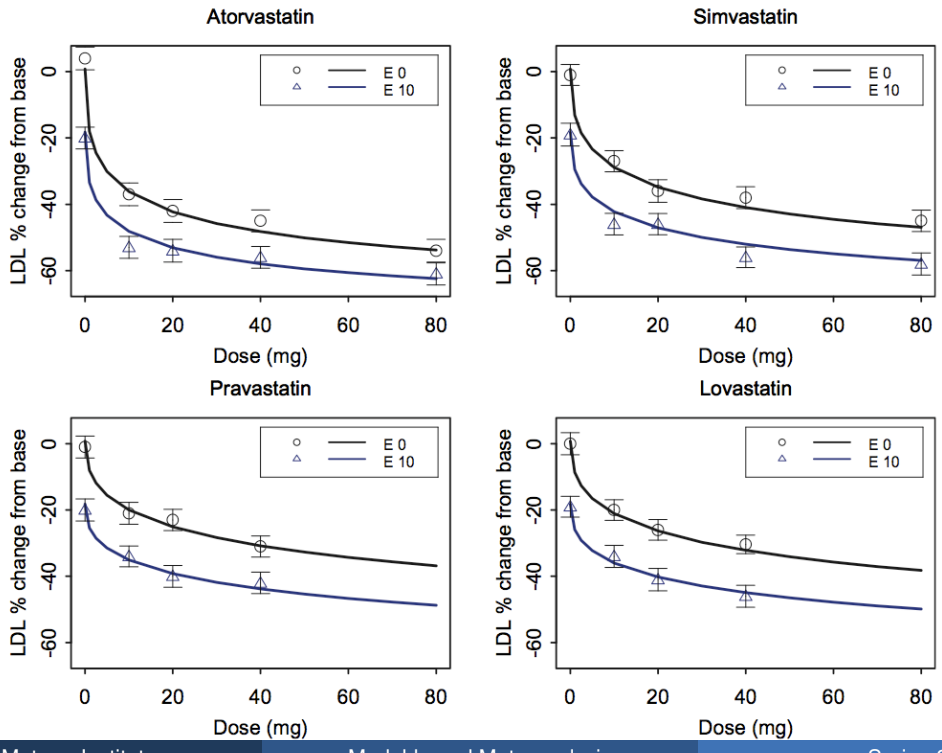
Model-Based Development of Gemcabene, a New Lipid-Altering Agent

Submitted: May 3, 2005; Accepted: May 4, 2005; Published: October 7, 2005

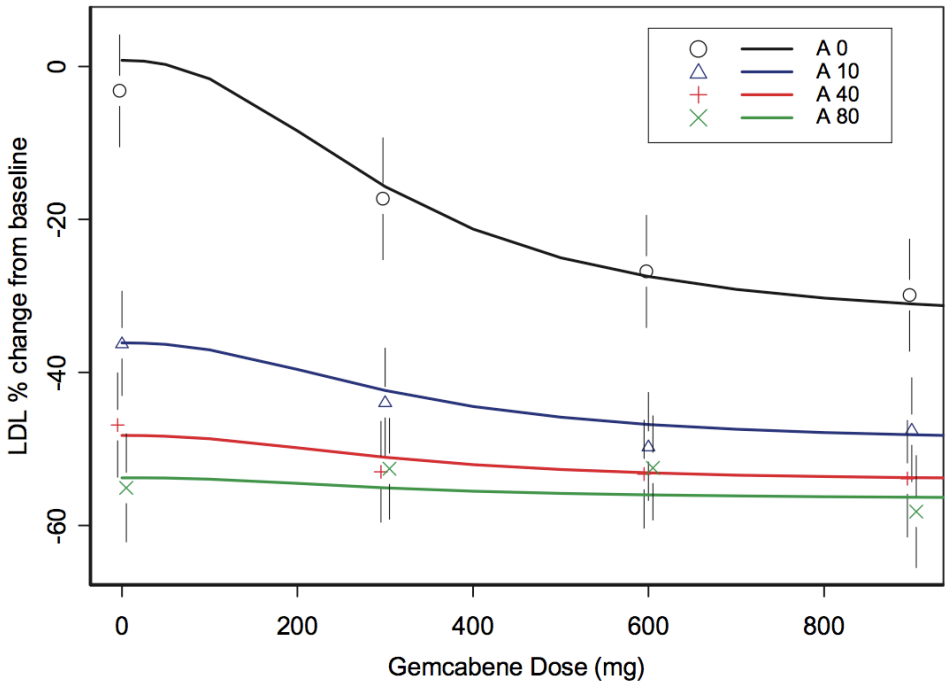
Jaap W. Mandema,[1,3] David Hermann,[2] Wenping Wang,[1,4] Tim Sheiner,[1] Mark Milad,[2] Rebecca Bakker-Arkema,[2] and Daniel Hartman[2]

- Meta-analysis of mean LDL-C % change from baseline
- 25 clinical trials
  - 13 trials of statin monotherapy
  - 4 trials of ezetimibe monotherapy
  - 4 trials of statin/ezetimibe combinations
  - 3 trials of gemcabene monotherapy
  - 1 trial of statin/gemcabene combinations
- The original work also included modeling of CRP, headache, ALT elevation, myalgia, HDL-C and CHD relative risk

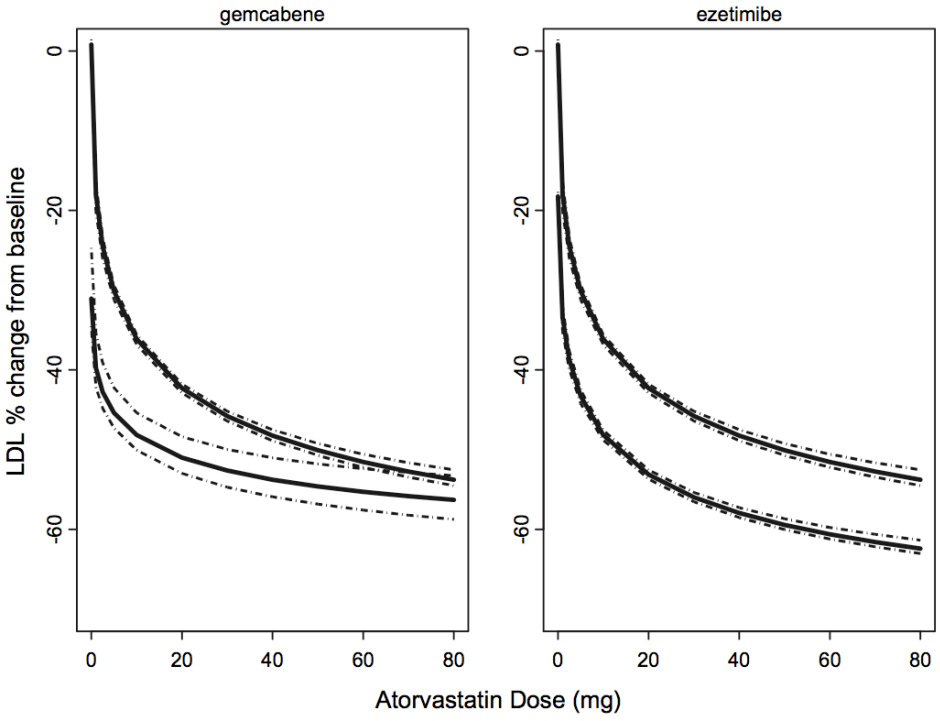
Statin dose-response: monotherapy and combination with ezetimibe 10 mg/d



Incremental effect of gemcabene decreases with increasing statin doses



Statin/gemcabene is inferior to statin/ezetimibe at higher statin doses



Statin/gemcabene is inferior to statin/ezetimibe at higher statin doses

Table 3. Predicted Additional Mean Change in LDL-C Reduction Between Combination Therapy of Atrovastatin With 900 mg of Gemcabene or 10 mg of Ezetimibe and Monotherapy With Atorvastatin*

	900 mg Gemcabene			10 mg Ezetimibe		
Atorvastatin (mg)	**Mean**	**5%**	**95%**	**Mean**	**5%**	**95%**
0 mg	−31.9	−35.4	−25.5	−19.1	−19.8	−18.3
10 mg	−12.0	−13.9	−9.1	−12.0	−12.4	−11.5
20 mg	−8.7	−10.7	−6.0	−10.8	−11.2	−10.4
40 mg	−5.5	−7.8	−2.7	−9.7	−10.0	−9.3
80 mg	−2.5	−5.3	0.4	−8.7	−9.0	−8.3

*The expectation and 90% probability interval are shown.

Conclusion: Terminate the development program because gemcabene cannot compete favorably with ezetimibe as combination therapy with statins

"In conclusion, the availability of the integrated model combined with the model visualization tool (DMX) led to a quick decision to stop the development of gemcabene. The model contributed significantly to this decision, because it provided a quantitative comparison between gemcabene and ezetimibe when administered alone or in combination with a statin. These treatment options were not directly compared in the phase II trial. The integrated model also increased the certainty of the decision to stop development."

Longitudinal Model-Based Meta-Analysis in Rheumatoid Arthritis: An Application Toward Model-Based Drug Development

I Demin[1], B Hamrén[1], O Luttringer[1], G Pillai[2] and T Jung[3]

Clinical Pharmacology & Therapeutics 92:352–359 (2012)

MBMA was used during the canakinumab development program to assess its efficacy relative to several marketed drugs.

- A model was developed to describe ACR20 response over time. for abatacept, adalimumab, anakinra, certolizumab, etanercept, golimumab, infliximab, rituximab, tocilizumab, methotrexate, placebo and canakinumab.
- The results were used to assess whether canakinumab was sufficiently efficacious to warrant further development.

Longitudinal MBMA of RA treatments

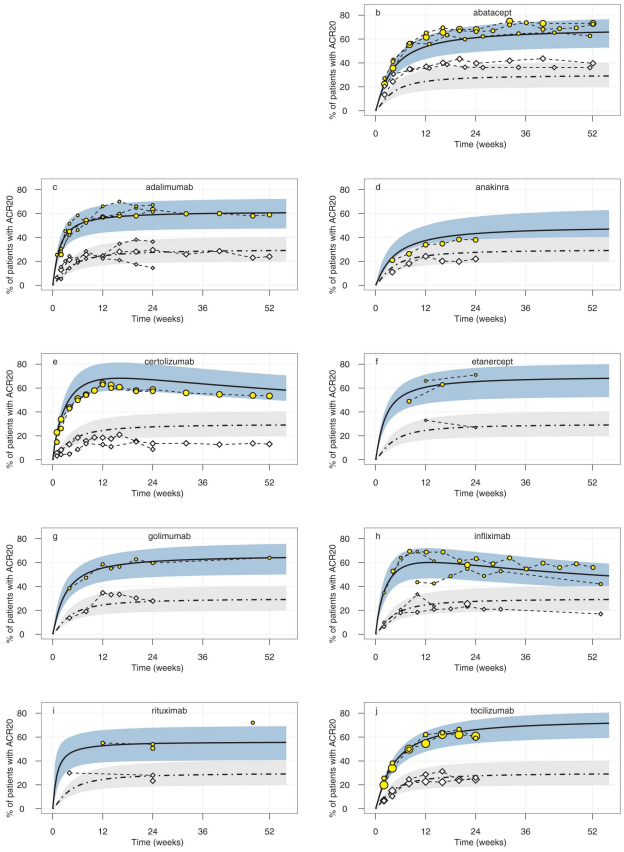
Table 1 Summary-level information about the studies included in the analysis

Drug name	No. of studies[a]	No. of patients[b]
Abatacept	4	951
Adalimumab	7	1,403
Anakinra	1	250
Certolizumab	2	639
Etanercept	7	1,021
Golimumab	3	401
Infliximab	5	994
Rituximab	4	683
Tocilizumab	5	1,525
Methotrexate	29	5,287
Placebo	6	321

[a]Number of studies with the corresponding drug. [b]Number of patients receiving an approved dose of the corresponding treatment.

Model developed by analysis of published data (except canakinumab)

- Fraction of patients achieving ACR20
- 37 double-blind controlled clinical trials
- Longitudinal data up to 54 weeks



Longitudinal ACR20 data and model predictions

The model was used to support a key strategic decision

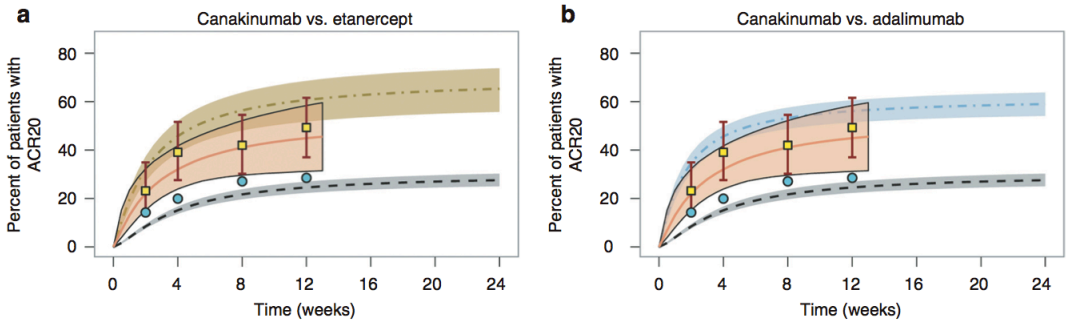


Figure 3 Model-based predictions of the time course of ACR20 responder rates for etanercept (panel **a**, brown dash-and-dot line), adalimumab (panel **b**, blue dash-and-dot line), and placebo (both panels, gray broken lines) vs. canakinumab (both panels, red solid lines) given in combination with methotrexate (MTX) to patients with an inadequate response to MTX. Green circles represent placebo-plus-MTX data from the canakinumab study. Yellow squares represent the observed ACR20 values for canakinumab. The shaded areas are the respective 90% Bayesian confidence intervals for model-based predictions of ACR20 responder rates. The red vertical bars are 95% confidence intervals for canakinumab data.

"...a decision was made not to pursue the RA indication. The integrated assessment of canakinumab data with the quantitative knowledge about current standard-of-care treatments supported this decision, given that the likelihood was low that canakinumab at current doses and regimen could provide incremental benefit to patients as compared with existing therapeutic options."

Therapeutic benefit of eletriptan compared to sumatriptan for the acute relief of migraine pain – results of a model-based meta-analysis that accounts for encapsulation

JW Mandema[1,*], E Cox[1] & J Alderman[2]

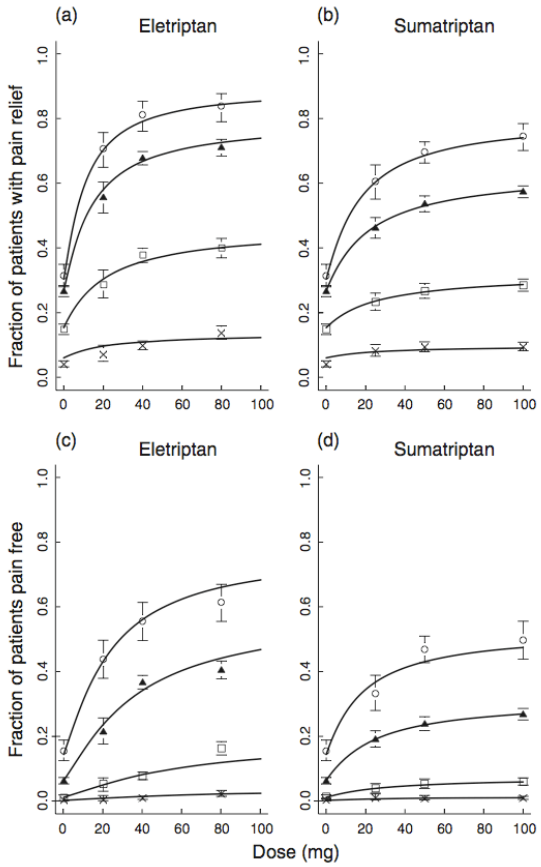
Cephalalgia 2005; 25:715725

MBMA was used during the eletriptan development program to compare the dose- and time-response of eletriptan to sumatriptan. The primary comparisons addressed in the paper were:

- The degree of migraine pain relief resulting from eletriptan 20 and 40 mg compared to sumatriptan 50 and 100 mg.
- The effect of encapsulation (for blinding purposes) on pain relief efficacy of sumatriptan.

MBMA to compare migraine pain relief response of two triptans

- 19 clinical trials
- 2 summary statistics for migraine pain relief
  - Fraction of patients with pain relief. Pain relief = improvement in headache pain score from a baseline of moderate or severe to mild or no pain
  - Fraction of patients pain free. Pain free = improvement in headache pain score from a baseline of moderate or severe to no pain
- Longitudinal data up to 4 hours



Eletriptan & sumatriptan dose-response at 0.5, 1, 2 & 4 h

Figure 1 Fit of the dose-response model for (a,c) eletriptan and (b,d) sumatriptan to the combined data from all trials, stratified by time since dosing ○ 4 h; ▲ 2 h; □ 1 h; × 0.5 h. (a,b) show pain relief, whereas (c,d) shows the results for pain free. The solid lines represent the model-predicted dose response relationships at specific time points after dosing in a typical trial. The symbols reflect the mean (across all trials) observed fraction of patients with pain relief or pain free at each evaluated dose and time point after adjusting for the random trial-to-trial differences. The error bar around each of the symbols reflects a 95% confidence interval on the observed fraction of patients.

Effect of encapsulation on sumatriptan dose-response

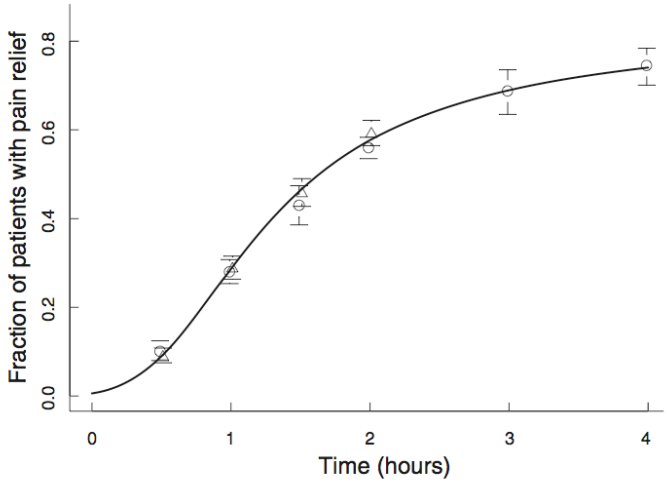


Figure 3 Lack of impact of encapsulation on the time course of response to 100 mg Sumatriptan. The solid line represents the model-predicted time course of response to 100 mg Sumatriptan in a typical trial. The symbols reflect the observed mean response to 100 mg encapsulated sumatriptan (△) and commercial (unencapsulated) sumatriptan (○) after accounting for random trial-to-trial differences. The error bar reflects a 95% confidence interval on the observed fraction of patients with pain relief.

Comparison of eletriptan and sumatriptan efficacy

pain relief

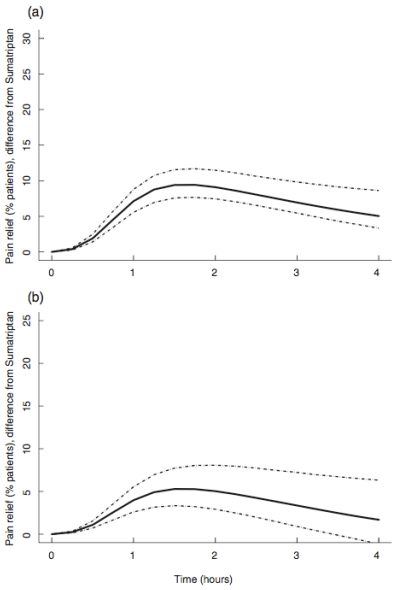


Figure 4 Time course of the absolute difference in the per cent of patients with pain relief (a) between eletriptan 40 mg and sumatriptan 100 mg, and (b) between eletriptan 20 mg and sumatriptan 50 mg. The predictions are shown for a typical patient population in a typical trial. The solid lines represent the model predictions; the dashed lines span a 90% probability interval for those predictions.

pain free

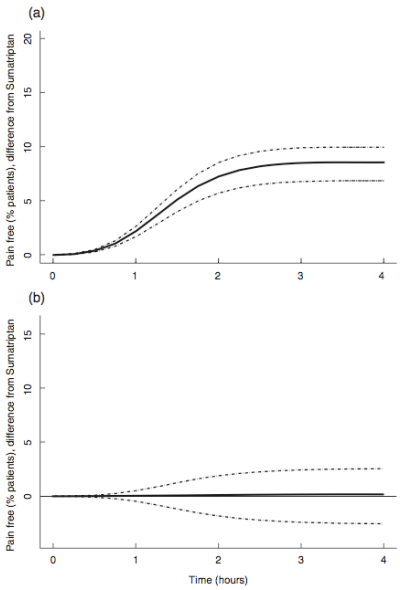


Figure 5 Time course of the absolute difference in the per cent of patients that are pain free (a) between eletriptan 40 mg and sumatriptan 100 mg, and (b) between eletriptan 20 mg and sumatriptan 50 mg. The predictions are shown for a typical patient population. The solid lines represent the model predictions; the dashed lines span a 90% probability interval for those predictions.

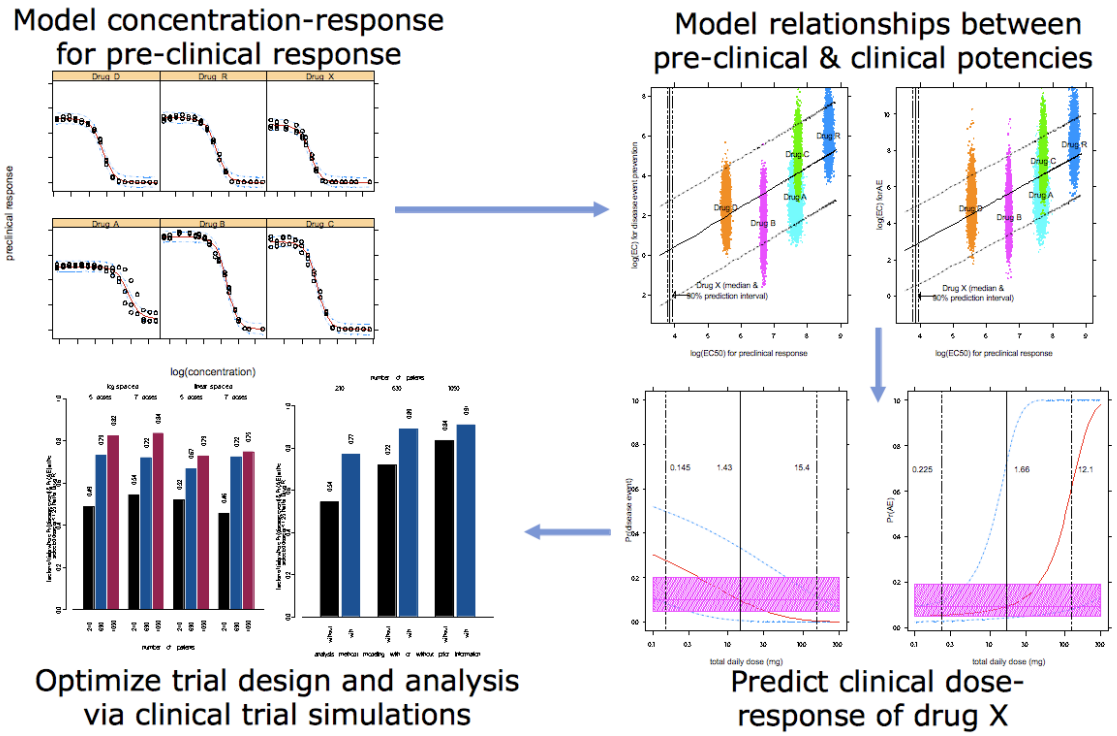
Conclusion: MBMA confirmed superior efficacy of eletriptan and lack of encapsulation effect

The authors concluded:
"In conclusion, the meta-analysis confirms the superior efficacy for pain relief and pain free of eletriptan 40 mg vs. sumatriptan 50 mg and 100 mg up to 4 h after treatment." ... "Encapsulation did not impact the efficacy of sumatriptan during this time frame."

Trial simulation to design a Phase II dose finding strategy

- Phase II objective: Efficiently find a dose of drug X that:
  - Is at least non-inferior to the standard of care (drug R 10 mg/d) with respect to both efficacy and safety.
  - With sufficient certainty that we can risk a Phase III program with only one dose level.
- Efficacy
  - Decrease in fraction of patients with a disease-related event
- Probable dose-limiting AE
  - Same biological mechanism as efficacy
- Available information:
  - Drug X:
    - Clinical pharmacokinetics from Phase I.
    - Pre-clinical biomarker thought to be predictive of clinical outcomes related to mechanism of action (both efficacy and dose-limiting AE)
  - Public-source data on related drugs:
    - Efficacy-related events
    - Dose-limiting AEs
    - Pharmacokinetics

Modeling strategy: Simultaneously model pre-clinical biomarker and frequency of clinical efficacy and AE events



Simulations led to more efficient designs that can find a non-inferior dose with high probability

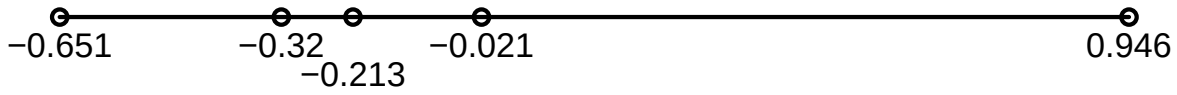
- Results using conventional statistical analysis without prior information
  - Best performance is seen with 7 log-spaced doses
  - Only trials with $>$ 1050 patients/arm offer sufficient certainty to consider risking a Phase III program with only one dose level.
- Bayesian modeling using prior information:
  - Increases the probability of selecting a non-inferior dose, particularly in trials with smaller sample sizes
  - Also improves the probability of making correct pruning decisions

Simulations led to more efficient designs that can find a non-inferior dose with high probability

- Leveraging prior information permits more efficient design and analysis of a Phase II trial to select a dose for Phase III:
  - Optimizes range, number and spacing of doses
  - Adaptive pruning assigns patients to most relevant doses
  - Enhances characterization of dose-response & therefore dose selection
    - Shorter, more informative Phase II program (reduction by $\sim$400-600 patients or $\sim$4-6 months)
    - Shorter time to market (at least 4 months and potentially much greater by avoiding incorrect dose selection and a failed Phase III program)

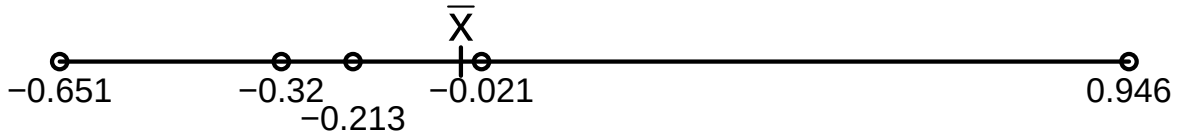
Traditional meta-analysis

How might we get an estimate of the mean of these data?



Values	-0.651	-0.320	-0.213	-0.021	0.946

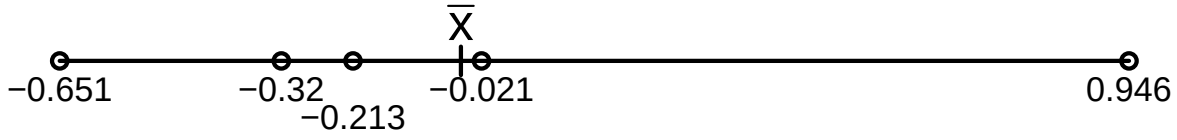
How might we get an estimate of the mean of these data?



Values	-0.651	-0.320	-0.213	-0.021	0.946

A simple average of the values: $\overline{x} = \sum_{i=1}^{5} x_i$

How might we get an estimate of the mean of these data?

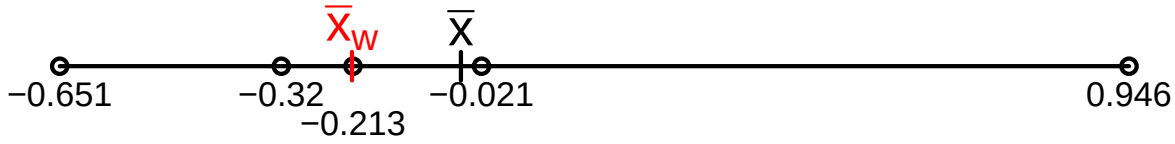


Values	-0.651	-0.320	-0.213	-0.021	0.946

Variances	0.932	0.060	0.125	0.375	0.683

A simple average of the values: $\overline{x} = \sum_{i=1}^{5} x_i$

How might we get an estimate of the mean of these data?



Values	-0.651	-0.320	-0.213	-0.021	0.946
Variances	0.932	0.060	0.125	0.375	0.683
Weights	1.07	16.7	8.00	2.67	1.46

A simple average of the values: $\overline{x} = \sum_{i=1}^{5} x_i$

A weighted average of the values: $\overline{x}_w = \frac{\sum_{i=1}^{5} w_i x_i}{\sum_{i=1}^{5} w_i}$ where $w_i = \frac{1}{\text{Var}(x_i)}$

At its simplest, a meta-analysis estimate is a weighted average

- A fixed effects meta-analysis estimate is essentially the weighted average on the previous slide.
- A random effects meta-analysis estimate lies between the simple and weighted average.
  - It is closer to the simple average when there is large between-study variability.
  - It is closer to the weighted average when there is very little between-study variability.

The End

More formally, meta-analysis is...

- A quantitative review and synthesis of results from related but independent studies (Normand, 1999)
- Aggregate data (AD) meta-analysis
  - Based on summary statistics from each trial
  - Mean response, proportion of responders
  - Mean difference between groups, odds ratio, hazard ratio
- Individual patient data (IPD) meta-analysis
  - Observed response, time-to-event
- If data are collected longitudinally, either approach can be applied at a single time point or as a longitudinal model
  - IPD is better suited for longitudinal models than AD

In contrast to a systematic review

- A (qualitative) summary of literature related to a specific set of research objectives
- More recently, systematic reviews include a quantitative summary of the literature - i.e., a meta-analysis
- Cochrane Collaboration Handbook lists guidelines for performing a systematic review and meta-analysis
- The PRISMA statement provides standards for reporting systematic review and meta-analyses

If you're looking for some starting places, you might consider...

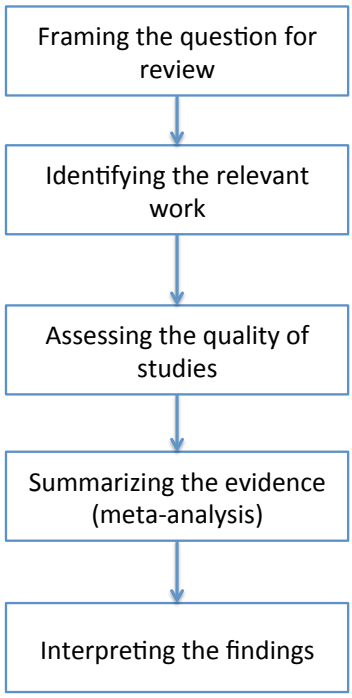
- Borenstein, Hedges, Higgins and Rothstein. Introduction to Meta-Analysis. John Wiley & Sons. 2009.
- Anne Whitehead, Meta-Analysis of Controlled Clinical Trials, John Wiley & Sons: West Sussex. 2002.
- Sharon-Lise T. Normand. Meta-analysis: Formulating, evaluating, combining and reporting. Statistics in Medicine 18, 321-359. 1999.

The importance of writing a meta-analysis plan

Conducting a meta-analysis is conceptually no different than conducting primary research

- This suggests writing a protocol / analysis plan to
  - Describe the objectives of the research
  - Define inclusion / exclusion criteria for the studies in your literature review
  - Describe how you intend to analyze the data
- And building a multi-disciplinary team
  - Clinical subject matter expert
  - Statistician
  - Clinical pharmacologist / pharmacometrician
  - Information scientist

Five steps in performing a meta-analysis



Adapted from Khan KS, Kunz R, Kleijnen J, and Antes G. Five steps to conducting a systematic review. *J R Soc Med.* 2003 96(3): 118-121.

Key components of the analysis plan

- A well formulated research question
- Inclusion/exclusion criteria will rely on
  - Patient population
  - Study designs (e.g., randomized, controlled)
  - Treatments or interventions of interest (and comparators of interest)
  - Definition of the outcome of interest (e.g., overall survival; objective response rate; WOMAC pain score)
- Description of how data will be extracted (e.g., from text, tables, figures etc.)
- Description of the analysis methods

Comparative Effectiveness of Treatment for HCV Infection in Adults (Chou *et al.*)

Objectives

In patients with HCV infection . . .

1. What is the comparative effectiveness of antiviral treatment in improving health outcomes?
2. What is the comparative effectiveness on the rate of sustained virology response (SVR)?
3. What are the comparative harms associated with antiviral treatment?
4. Have improvements in SVR been shown to reduce the rates of adverse health outcomes (e..g, mortality)?

Chou R, Hartung D, Rahman B, Wasson N, Cottrell EB, and Fu R. (2013). Comparative Effectiveness of Antiviral Treatment for Hepatitis C Virus Infection in Adults: A Systematic Review. *Ann Intern Med.* 158(2): 114-123.

Comparative Effectiveness of Treatment for HCV Infection in Adults (Chou *et al.*)

Populations

Adults with HCV infection who have not had previous antiviral drug treatment

Interventions

Pegylated interferon alfa-2a with ribavirin, Pegylated interferon alfa-2b with ribavirin, Protease inhibitors (e.g., telaprevir, boceprevir)

Outcomes

SVR rates, mortality from HCV, withdrawals due to AEs, SAEs

Studies Observational studies, systematic reviews, and clinical trials. (Case studies and small case series excluded.)

Chou R, Hartung D, Rahman B, Wasson N, Cottrell EB, and Fu R. (2013). Comparative Effectiveness of Antiviral Treatment for Hepatitis C Virus Infection in Adults: A Systematic Review. *Ann Intern Med.* 158(2): 114-123.

Chemotherapy and radiotherapy for inoperable advanced pancreatic cancer (Yip et al.)

Interventions

Outcome

Objectives

To assess the effects of chemotherapy and/or radiotherapy [on overall survival] in people with inoperable advanced (including locally advanced and metastatic) or relapsed disease.

Patient population

Types of studies

Randomised controlled trials with a single-blind or double blind design, in which one of the intervention types (chemotherapy or radiotherapy) was contrasted with either placebo or another type of intervention. Both published and unpublished studies were identified and assessed for inclusion.

Yip D, Karapetis C, Strickland A, Steer CB, Goldstein D. Chemotherapy and radiotherapy for inoperable advanced pancreatic cancer. Cochrane Database of Systematic Reviews 2006, Issue 3.

Performing the literature search

- Rely on the people who are trained to do these searches
- Define which bibliographic and abstract databases to search
  - Include SBAs, ClinicalStudyResults.org, etc. in addition to the usual sources
- Don't forget to review the references in the first set of papers you get

The ideal is to include in your analysis all trials (published and unpublished) that fit your inclusion/exclusion criteria

Critical appraisal of the studies is important

- You get to understand the studies that are going into your meta-analysis
- For formal, 'official' meta-analyses, studies are often scored using the Jadad score
  - Total of 5 points
  - Assesses randomization, blinding, description of withdrawal/dropout, and the method to generate randomization
- Important to assess potential for bias (due to study exclusion and biases within the studies)

Jadad AR, et al. Assessing the Quality of Reports of Randomized Clinical Trials: Is Blinding Necessary? Controlled Clin Trials 1996; 17: 1-12.

Extracting the data

- Ideally establish standardized database formats to use for data retrieval and associated processes that can make things easier for your team
- Before you extract any data, think about what you will be modeling
  - This comes back to your analysis plan / protocol

Extract the relevant outcome data

Consider that you may model the response, but present the results in terms of treatment effects.

- Binary data
  - Response: e.g., response rates
  - Measures of effect: risk difference or ratio; odds ratio
- Continuous data
  - Response: e.g., Observed means
  - Measure of effect: e.g., mean difference between groups, effect size
- Time to event data
  - Response: e.g., median time to event, estimated survival rates at specific time points
  - Measure of effect: e.g., hazard ratio

Summarizing the evidence and interpreting the results

This is the focus of the remainder of the course . . .

Simple meta-analysis models

There are at least 3 sources of variation to consider

- Sampling error
  - Large studies typically provide more precise estimates than small studies
- Study-level characteristics
  - Will consider these as ways to explain differences in treatment effects across studies through covariates (meta-regression)
- Inter-study variation
  - The remaining, unexplained, variability in treatment effects across studies

(Normand, Statistics in Medicine 18, 321-359. 1999)

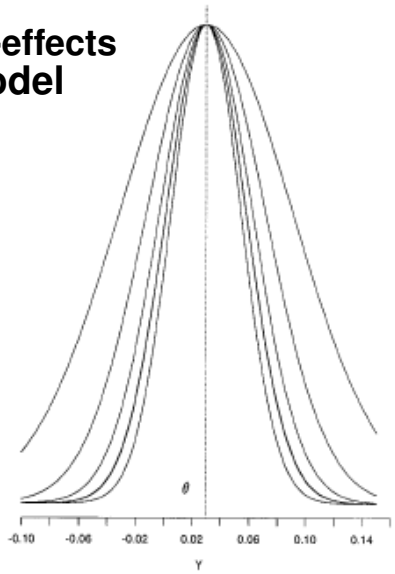
There are two general types of statistical model for meta-analysis

- Fixed effects models
  - Each study is measuring the same underlying parameter
  - There is no inter-study variation in treatment effect (after possibly accounting for covariate effects)
  - After accounting for covariates, the only source of variation is sampling error
- Random effects models
  - Each study is associated with a different, but related, underlying parameter
  - After accounting for covariates, there is still some unexplained inter-study variability in addition to sampling error

(Normand, Statistics in Medicine 18, 321-359. 1999)

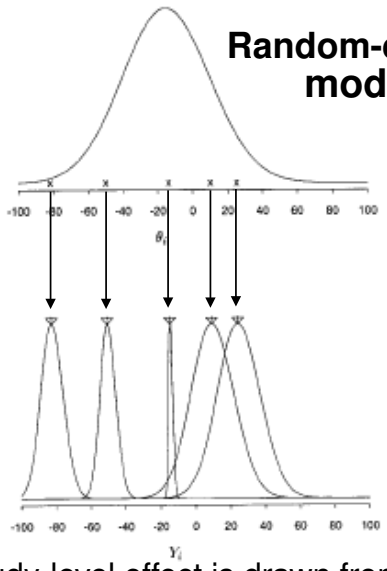
Distribution of 5 hypothetical study statistics under fixed- and random-effects models

Fixed-effects model



Each study provides an estimate of the common mean effect (θ). They differ only in how well each study sample estimates θ.

Random-effects model



Each study-level effect is drawn from a "superpopulation" with mean θ. Estimates from individual studies are centered around these.

Normand, 1999

Fixed effects meta-analysis: the traditional approach

Notation

Let Y_{ij} be the observed summary data in the j^{th} arm of the i^{th} study.

- E.g., Y_{i1} = response rate in the experimental group and
  Y_{i2} = response rate in the control group
- E.g., Y_{i1} = mean Δ baseline response in the experimental group
  Y_{i2} = mean Δ baseline response in the control group

Let $\widehat{\theta}_i$ be the observed treatment effect in the i^{th} study.

- E.g., $\widehat{\theta}_i$ = log odds ratio comparing experimental and control groups
- E.g., $\widehat{\theta}_i$ = mean difference in Δ baseline response comparing experimental and control groups

The simple fixed effects model

Suppose there are r independent studies, each comparing the treatment group with the control group.

θ denotes the measure of treatment difference
$\widehat{\theta}_i$ its estimate from the i^{th} study

Then the general fixed effects model is

$$\widehat{\theta}_i = \theta + \epsilon_i$$

with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \xi_i^2$.

Assume we know $\text{Var}(\epsilon_i)$

Usually we treat ξ_i^2 as known and equal to the estimated variance of $\widehat{\theta}_i$.

Thus,

$$\theta_i \sim N\left(\theta, s_i^2\right)$$

where s_i^2 is the estimated variance of $\widehat{\theta}_i$ and is assumed known.

I've snuck in an additional (and technically unnecessary) assumption here... what is it?

MLE for fixed effects model

When ξ_i^2 is assumed known*, the maximum likelihood estimate for θ is

$$\widehat{\theta}_{FE} = \frac{\sum_{i=1}^{r} W_i \widehat{\theta}_i}{\sum_{i=1}^{r} W_i} \text{ with } W_i = \frac{1}{s_i^2}$$

and Var $\left(\widehat{\theta}_{FE}\right) = \left(\sum_{i=1}^{r} W_i\right)^{-1}$

We can use this to construct confidence intervals for θ based on the normal distribution, since the s_i^2 are assumed known.

The assumption has little impact on the results (Hardy and Thompson, 1996).

* (and we assume the errors follow a normal distribution)

Just for fun: let's show that this is the MLE!

Under this model, the log-likelihood for θ is

$$\ell(\theta) = \sum_{i=1}^{r} -\frac{1}{2}\log\left(s_i^2\right) - \frac{1}{2}\frac{\left(\widehat{\theta}_i - \theta\right)^2}{s_i^2}$$

Differentiating the log-likelihood with respect to θ gives

$$\frac{\partial \ell(\theta)}{\partial \theta} = \sum_{i=1}^{r} \frac{\left(\widehat{\theta}_i - \theta\right)}{s_i^2}$$

Setting the derivative equal to 0 and solving for θ gives

$$\sum_{i=1}^{r} \frac{\widehat{\theta}_i}{s_i^2} = \widehat{\theta}_{FE} \sum_{i=1}^{r} \frac{1}{s_i^2} \Longrightarrow \widehat{\theta}_{FE} = \frac{\sum_{i=1}^{r} \frac{\widehat{\theta}_i}{s_i^2}}{\sum_{i=1}^{r} \frac{1}{s_i^2}}$$

$$\text{or } \widehat{\theta}_{FE} = \frac{\sum_{i=1}^{r} W_i \widehat{\theta}_i}{\sum_{i=1}^{r} W_i} \text{ with } W_i = \frac{1}{s_i^2}$$

Conjugate Bayesian analysis

If we use a normal prior distribution for θ, then we can show that the posterior distribution for θ is also a normal distribution.

$$\theta_i \mid \theta \sim N\left(\theta, s_i^2\right)$$
$$\theta \sim N\left(\mu, \sigma^2\right)$$
$$\downarrow$$
$$\theta \mid \text{data} \sim N\left(\mu_{post}, \sigma_{post}^2\right)$$

where

$$\mu_{post} = \frac{\sum_{i=1}^{r} W_i \widehat{\theta}_i + \mu/\sigma^2}{\sum_{i=1}^{r} W_i + \sigma^{-2}}$$
$$\sigma_{post}^2 = \left(\sum_{i=1}^{r} W_i + \sigma^{-2}\right)^{-1}$$

Conjugate Bayesian analysis

If the prior distribution is non-informative (e.g., σ^2 is large relative to μ and $\sum W_i$), then $\mu/\sigma^2 \approx 0$ and $\sigma^{-2} \approx 0$. Thus,

$$\mu_{post} \approx \widehat{\theta}_{FE}$$

and

$$\sigma_{post}^2 \approx \text{Var}\left(\widehat{\theta}_{FE}\right)$$

We can estimate a non-conjugate Bayesian fixed-effects meta-analysis model using a tool like WinBUGS.

What if I want to use the mean response, not the treatment effect?

The same approach holds – nothing changes.

If we assume that $Y_i \sim N\left(\theta, s_i^2\right)$, then

$$\widehat{\theta}_{FE} = \frac{\sum_{i=1}^{r} W_i Y_i}{\sum_{i=1}^{r} W_i} \text{ with } W_i = \frac{1}{s_i^2}$$

and Var $\left(\widehat{\theta}_{FE}\right) = \left(\sum_{i=1}^{r} W_i\right)^{-1}$

If Y_i is a sample mean and we assume that $s_i^2 = \sigma^2/n_i$, (i.e., the variance is the same across all studies, but the sample size could vary), then $W_i = n_i/\sigma^2$ and

$$\widehat{\theta}_{FE} = \frac{\sum_{i=1}^{r} W_i Y_i}{\sum_{i=1}^{r} W_i} = \frac{\sum_{i=1}^{r} n_i Y_i}{\sum_{i=1}^{r} n_i}$$

which only depends on the sample size.

If the standard error isn't reported directly, there are ways to derive it: continuous data

If the treatment effect measure is a difference or a continuous endpoint, you can use the reported

- Confidence interval for the difference
- Standard deviations for the treatment groups
- T-statistic comparing the two groups or p-value from the test
- Imputation based on the studies from which you have, or can derive, the standard error
- See Follmann et al. (1992) for more details

Follmann, Elliott, Suh and Cutler. Variance imputation for overviews of clinical trials with continuous response. J Clin Epi. 45: 769-773. (1992)

Example: continuous data

Suppose we have the following information from an RCT in $n = 212$ patients: "At 6 months, relative to placebo, ramipril was associated with a 75-second (95% CI, 60-89 seconds) increase in mean pain-free walking time ($P < .001$)"

From the CI, we can estimate the standard error for the difference as

$$SE_{diff} = \frac{UL - LL}{2 \times t_{0.975, df=210}} = \frac{89 - 60}{2 \times 1.971} = 7.36$$

From the p-value (assuming $p = 0.001$), we could estimate

$$SE_{diff} = \frac{diff}{t_{1-p/2, df=210}} = \frac{75}{3.34} = 22.5$$

where $t_{\alpha, df}$ is the quantile function from a t-distribution with df degrees of freedom.

Ahimastos *et al.* Effect of Ramipril on Walking Times and Quality of Life Among Patients With Peripheral Artery Disease and Intermittent Claudication. *JAMA* 2013: 309(5): 453-60.

If the standard error isn't reported directly, there are ways to derive it: binomial data

If the treatment effect measure is a log-odds ratio

- The asymptotic variance of the log-odds ratio is a function of the number of subjects with and without events in each group
  - Can calculate this from reported number of subjects and the percent of subjects with events
  - $Var(logOR) = \frac{1}{n_a} + \frac{1}{N_a - n_a} + \frac{1}{n_b} + \frac{1}{N_b - n_b}$
- Can also derive it if a CI or p-value for the odds ratio is reported, using similar methods as for continuous data

Similar approaches can be used for other endpoints (e.g., risk differences, simple proportions, etc.)

If the standard error isn't reported directly, there are ways to derive it: survival data

If the treatment effect measure is a hazard ratio, you can use

- A reported CI
- A p-value for the log-rank test
- Extracted data from the survival curve
- See Parmar et al. (1998) for more details

Parmar MKB, Torri V and Stewart L. Extracting summary statistics to perform meta-analysis of the published literature for survival endpoints. Statistics in Medicine 1998; 17: 2815-2834

What about just using the sample sizes rather than variances?

- Suppose we have a continuous response and don't have any estimate of variability, can we consider using the sample size instead?
  - E.g., pooling mean or median survival times across studies
  - E.g., pooling mean reduction in tumor size
- What assumptions are we making when we do this?
- Is this a reasonable thing to do even if we do have the standard errors?

If we only use sample sizes, the model becomes

If we use sample sizes, the general fixed effects model becomes

$$Y_i = \theta + \epsilon_i$$

with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \sigma^2/n_i$.

A major assumption is that the population variance is the same across all studies.

- Whether or not this is reasonable depends on the situation.

It's generally better to use the reported standard errors or standard deviations if they are reported or can be derived.

Fixed effects meta-analysis of difference from placebo in percent change from baseline LDL for 10 mg Atorvastatin

Study	Mean Difference	SE	Weight ($1/SE^2$)	Relative Weight
981-4	−48.60	3.890	0.07	0.039
P00692	−41.00	2.828	0.12	0.074
4522IL/0008	−40.60	3.023	0.11	0.065
981-8	−37.00	1.118	0.80	0.475
981-96	−37.00	4.884	0.04	0.025
4522IL/0024	−35.00	1.769	0.32	0.190
981-25	−34.00	2.546	0.15	0.092
3	−33.80	3.833	0.07	0.040

Fixed effects estimate for difference from placebo in percent change from baseline LDL

For the difference from placebo in percent change from baseline LDL

$$\widehat{\theta}_{FE} = \frac{\sum_{i=1}^{8} W_i \widehat{\theta}_i}{\sum_{i=1}^{8} W_i} = -37.20$$

$$\text{Std.Err}(\widehat{\theta}_{FE}) = \sqrt{\left(\sum_{i=1}^{8} W_i\right)^{-1}} = 0.771$$

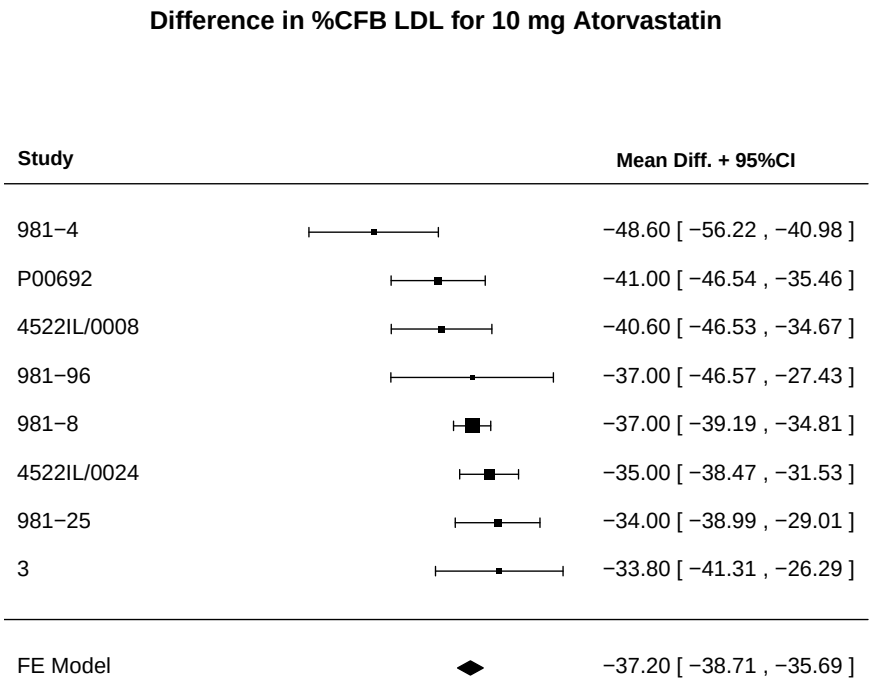
An approximate 95% confidence interval is

$$-37.20 \pm 1.96 \times 0.771$$

or

$$(-38.71, -35.69)$$

Forest plot showing mean difference from placebo and summary estimate



Chemotherapy and radiotherapy for inoperable advanced pancreatic cancer (Review)

Yip D, Karapetis C, Strickland A, Steer CB, Goldstein D

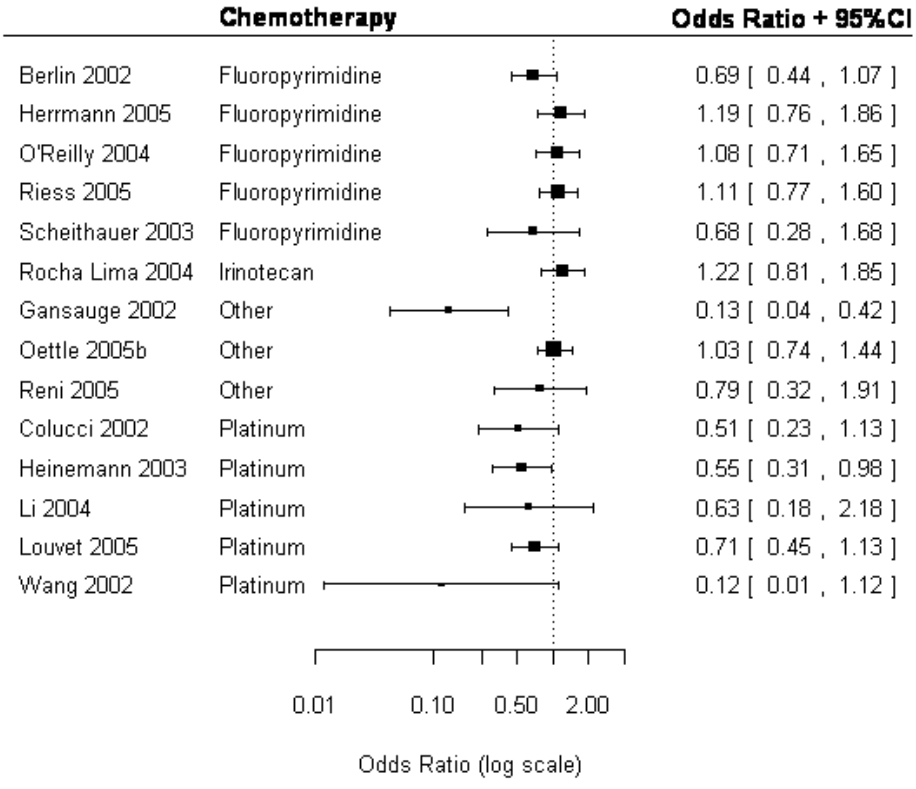
THE COCHRANE COLLABORATION®

This is a meta-analysis published in 2006 under the auspices of the Cochrane Collaboration.

The main objective was to assess the effects of chemotherapy and/or radiotherapy in the management of pancreatic adenocarcinoma in patients with inoperable advanced pancreatic cancer.

We'll fit meta-analysis and meta-regression models to the 6 month mortality incidence from 14 trials comparing gemcitabine to gemcitabine + chemotherapy combination

Forest plot of Month 6 Mortality Incidence



FE meta-analysis for Cochrane Month 6 mortality incidence data

Study	Odds Ratio	log(OR)	Std. error log(OR)	Weight ($1/SE^2$)	Relative Weight
Wang 2002	0.12	-2.16	1.16	0.74	< 0.01
Gansauge 2002	0.13	-2.02	0.58	2.93	0.01
Colucci 2002	0.51	-0.66	0.40	6.19	0.03
Heinemann 2003	0.55	-0.59	0.29	11.67	0.06
Li 2004	0.63	-0.46	0.63	2.50	0.01
Scheithauer 2003	0.68	-0.38	0.46	4.73	0.02
Berlin 2002	0.69	-0.37	0.22	19.93	0.10
Louvet 2005	0.71	-0.34	0.24	17.82	0.09
Reni 2005	0.79	-0.24	0.45	4.89	0.02
Oettle 2005	1.03	0.03	0.17	34.12	0.17
0'Reilly 2004	1.08	0.08	0.21	21.72	0.11
Riess 2005	1.11	0.10	0.19	29.08	0.15
Herrmann 2005	1.19	0.17	0.23	19.17	0.10
Rocha Lima 2004	1.22	0.20	0.21	22.39	0.11

Odds ratios compare gemcitabine+chemo to gemcitabine alone.

Values < 1 indicate benefit of combination.

Fixed effects odds ratio estimate for 6 Month mortality incidence comparing G+C to G alone

For the log odds-ratio:

$$\widehat{\theta}_{FE} = \frac{\sum_{i=1}^{14} W_i \widehat{\theta}_i}{\sum_{i=1}^{14} W_i} = -0.115$$

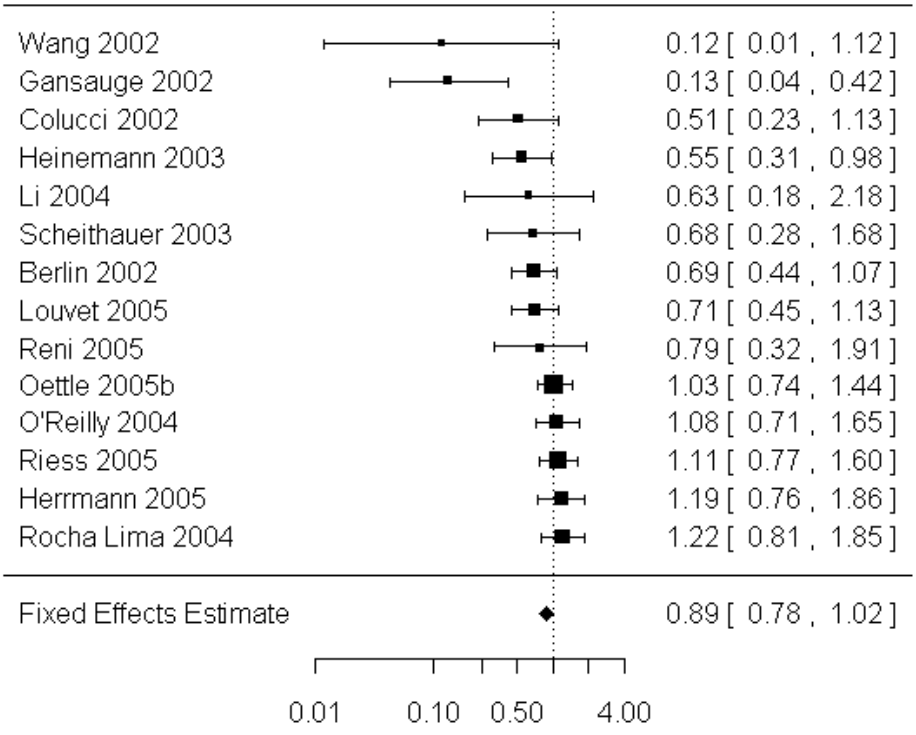
$$\text{Std.Err}(\widehat{\theta}_{FE}) = \sqrt{\left(\sum_{i=1}^{14} W_i\right)^{-1}} = 0.071$$

For the odds ratio:

$$\widehat{OR}_{FE} = \exp(-0.115) = 0.89$$

$$95\%\ \text{CI: } \exp(-0.115 \pm 1.96 \times 0.071) = (0.78, 1.02)$$

Forest plot of Month 6 Mortality Incidence



The Mantel-Haenszel estimator is an alternative for binary data

- For binary data, we typically estimate the treatment effect using the odds ratio.
- The fixed-effects MA estimator of the odds ratio is a weighted average of the log-odds ratios from the individual studies
  - Then exponentiating this estimate
- The M-H estimator of the odds ratio can be viewed as a weighted average of the odds ratios themselves.

How to get the M-H estimate

The estimated odds ratio from study i is

$$\widehat{\psi}_i = \frac{A_i D_i}{B_i C_i}$$

Using the Mantel-Haenszel weights

$$w_i = \frac{B_i C_i}{n_i}$$

the M-H estimator of the odds ratio is

$$\widehat{\psi}_{MH} = \frac{\sum A_i D_i / n_i}{\sum B_i C_i / n_i}$$

	dead	alive
Group 1	A_i	B_i
Group 2	C_i	D_i

$$n_i = A_i + B_i + C_i + D_i$$

Robbins et al (1986) provides formulas for several estimators of $\widehat{\text{Var}}\left(\log\left(\widehat{\psi}_{MH}\right)\right)$.

The MH estimator is preferable when there are no events in one of the treatment arms

- The standard FE estimator excludes that study
  - Because it gets a weight of 0.
  - An ad-hoc solution is to add 0.5 to each cell of the 2x2 tables
- If no events in a *study*, both the M-H and the standard FE estimator exclude that study from the analysis
  - The weight for that study is 0 in either case
- In practice, with no 0's, the difference between the FE and M-H estimators is usually small
- Not clear how a random-effects or meta-regression model might be accommodated in the M-H framework (Whitehead, 2002).
- Cai et al (2010, [CPR10]) propose an alternative approach using Poisson regression

Equivalence of stating model in terms of Y and θ

For many models, we can get the same estimates if we formulate the model in terms of the observed data instead of the treatment effects.

For continuous data:

$$Y_{ij} = \mu_i + \theta \cdot I(j = 1) + \delta_{ij}$$

for $i = 1, \ldots, r$ and $j = 1$ (experimental), 2 (control) where Y_{ij} is observed mean value and Var (δ_{ij}) is known.

For binary data:

$$\text{logit}\left(p_{ij}\right) = \alpha_i + \theta \cdot I(j = 1)$$

where p_{ij} is the observed response rate.

Introduction to `metafor`

At this point, we'll look at how to make a forest plot and fit a fixed-effects meta-analysis using the `metafor` package in R Viechtbauer, 2010 [Vie10]).

Measures of heterogeneity of effect between studies

We should expect some heterogeneity in effect

- Because meta-analyses typically pool studies that are diverse clinically and methodologically
- Could be due to differences in
  - Patient inclusion criteria
  - Background treatment
  - Dosing regimens
  - Study quality

Measures of heterogeneity

- Cochran's Q statistic
- I^2
  - Percentage of total variance that is due to heterogeneity rather than chance
- H
- τ^2
  - Between study variance estimate from a random effects model

Cochran's Q statistic

- Q is a weighted sum of squares of the deviations of individual study estimates from the overall estimate

$$Q = \sum_{i=1}^{r} W_i \left(\widehat{\theta}_i - \widehat{\theta}_{FE}\right)^2 \text{ where } W_i = 1/s_i^2$$

- If there is no heterogeneity, then Q follows a chi-squared distribution with r-1 degrees of freedom ($r = \#$ of studies)
  - Frequently used to test for heterogeneity

Tests based on Cochran's Q statistic aren't very useful

- Problems with power can give misleading results
  - Low power when few studies
  - Excessive power when many studies
- Values of Q cannot be compared across meta-analyses
  - Because its magnitude depends on the number of studies
- Despite these problems, Q is reported more than any other measure of heterogeneity.

Higgins and Thompson. Quantifying heterogeneity in a meta-analysis. Stat. Med 2002; 21:1539-58

I^2 tries to address these issues

- A measure of degree of inconsistency across studies
  - Percentage of total variation that is due to heterogeneity rather than chance

$$I^2 = 100\% \times \frac{\tau^2}{\sigma^2 + \tau^2}$$

  - If using the Dersimonian and Laird estimator of τ^2, then

$$I^2 = 100\% \times (Q - (r - 1)) / Q$$

- No tests based on I^2
- A tentative rule of thumb might be
  - $<$ 25% = low
  - 25 –75% = moderate
  - $>$ 75% = high

Higgins and Thompson. Quantifying heterogeneity in a meta-analysis. Stat. Med 2002; 21:1539-58

Higgins et al. Measuring inconsistency in meta-analyses. BMJ 2003; 327: 557-560

I^2 is preferred to Q

- It has an intuitive interpretation
  - Also doesnt depend on the type of outcome data or effect measure
- Confidence interval easily calculated in closed-form
- Can be directly compared across meta-analyses
  - Does not depend on the number of studies

Higgins and Thompson. Quantifying heterogeneity in a meta-analysis. Stat. Med 2002; 21:1539-58

What about H and τ^2?

- H has nice statistical relationships, but not a clear intuitive interpretation
  - $H^2 = Q/(r-1)$, the relative excess in Q compared to its expectation
  - H is the residual standard deviation from a radial plot (coming later)
- τ^2 provides a measure of extent of heterogeneity, but not a measure of impact
  - Is specific to a particular treatment measure
  - Can't be compared across meta-analyses

Using the LDL data, let's calculate Q

For the mean difference from placebo in %CFB LDL at 10 mg Atorvastatin, Cochran's Q statistic is

$$Q = \sum W_i (Y_i - \hat{\mu}_{FE})^2 = \sum W_i (Y_i - (-37.20))^2 = 15.61$$

Under H_0: No Heterogeneity, $Q \sim \chi^2_7$ since there are 8 studies. The p-value for this test is 0.029.

This indicates that there is statistically significant heterogeneity in the effect. Why do you think this might be?

. . . and also H and I^2

Following Higgins and Thompson,

$$H = \sqrt{\frac{Q}{df}} = \sqrt{\frac{15.61}{7}} = 1.493$$

$$se(\log(H)) = \frac{1}{2} \frac{\log(Q) - \log(r-1)}{\sqrt{2Q} - \sqrt{2r-3}} = 0.202$$

From this we could construct a 95% CI for H.

And for I^2

$$I^2 = 100 \times \frac{Q - (r-1)}{Q} = \frac{H^2 - 1}{H^2} = 55.1$$

There are a number of ways you can get a CI for I^2 . . .

Random effects meta-analysis

The simple random effects model

As before, θ denotes the measure of treatment difference and $\widehat{\theta}_i$, its estimate from the i^{th} study.

Then the general random effects model is

$$\widehat{\theta}_i = \theta_i + \epsilon_i \text{ where } \theta_i \sim N\left(\theta, \tau^2\right)$$

As before, we treat Var (ϵ_i) as known and equal to the estimated Var $\left(\widehat{\theta}_i\right)$.

MLE for random effects model

If τ^2 were known, the maximum likelihood estimate for θ would be

$$\widehat{\theta}_{RE} = \frac{\sum_{i=1}^{r} W_i^* \widehat{\theta}_i}{\sum_{i=1}^{r} W_i^*} \text{ with } W_i^* = \frac{1}{s_i^2 + \tau^2}$$

and Var $\left(\widehat{\theta}_{RE}\right) = \left(\sum_{i=1}^{r} W_i^*\right)^{-1}$

Typically, though, we estimate τ^2 using either a Method of Moments estimator or the Restricted Maximum Likelihood (REML) estimator and plug in to the formula above.

Method of Moments estimator for τ^2 (DerSimonian and Laird)

Probably the most commonly used estimator (for simple random effects meta-analysis) because it is easy to calculate.

Derived by equating the Q statistic with its expected value.

$$\widehat{\tau}^2_{DL} = \max\left\{0, \frac{Q-(r-1)}{\sum W_i - \frac{\sum W_i^2}{\sum W_i}}\right\}, \text{ where } W_i \text{ is the FE weight } \left(1/s_i^2\right)$$

If Q is less than its expected value $(r-1)$ then we estimate the between study variance to be 0.

Confidence intervals for $\widehat{\tau}^2_{DL}$

Biggerstaff and Tweedie (1997) proposed two methods of obtaining a confidence interval for τ^2_{DL}:

- A *symmetric* interval based on the large sample approximation to Var $\left(\widehat{\tau}^2_{DL}\right)$.
- An *asymmetric* interval based on a gamma distribution.

(RE)ML estimator of τ^2

- The ML and REML estimators aren't available in closed-form
  - But can be easily obtained using most standard software for fitting mixed-effects models (S-PLUS, SAS, NONMEM, etc.)
  - In R we'll the `metafor` package (Viechtbauer, 2010)
- The DL and REML estimates will typically be very similar, though there are times when they are not.
  - E.g., when $Q < df$, DL estimate is 0, but REML estimate may be > 0.

The Bayesian random effects meta-analysis model

For the Bayesian model, we add one more level to the hierarchical model – prior distributions for θ and τ^2.

The general Bayesian random effects model is

$$\widehat{\theta}_i \mid \theta_i, s_i^2 \sim N\left(\theta_i, s_i^2\right) \qquad \text{data model}$$

$$\theta_i \mid \theta, \tau^2 \sim N\left(\theta, \tau^2\right) \qquad \text{study-level parameter model}$$

$$\theta \sim f \quad \tau^2 \sim g \qquad \text{prior distributions}$$

For example, we might use $\theta \sim N(\mu, \sigma^2)$ and $\tau \sim U(0, a)$.

There are 5 relevant aspects of a RE meta-analysis

1. Quantification of heterogeneity in results
2. Estimation of the mean effect
3. Estimation of study-specific effects
4. Prediction of the effect in a new study
5. Testing
   - Whether an effect exists in any study,
   - The consistency of the direction of effect
   - Whether an effect is predicted to exist in a new study

Objectives for a specific MA will vary

Higgins, Thompson, and Spiegelhalter. A re-evaluation of random effects meta-analysis. JRSS-A. 2009. 172: 137-159.

Recommendations from HTS

a) **Use visual inspection of a plot,** such as a forest plot, as a preliminary inspection of heterogeneity

b) **Allow for heterogeneity** in a meta-analysis using covariates and/or random effects

c) Interpret random-effects meta-analyses with due consideration of the whole distribution of effects, ideally by **presenting a prediction interval**

d) Focus statistical tests on the important questions of whether an effect exists anywhere and whether it has a consistent direction across studies.

Recommend starting with a random effects meta-analysis

- We should always expect some heterogeneity in the treatment effect or response
  - (Slightly) different treatment regimens, patient populations, summary statistics, etc.
- RE models naturally allow prediction of future studies, accounting for variability above and beyond sampling error
- If the estimate of τ^2 is small (or 0), then the RE and FE estimates will be similar (or identical).
  - So there is no loss in fitting the RE model
- The CIs from a RE model will be (correctly) wider than for the FE model
  - In a review by Schmidt et al., they found that the nominal 95% FE CI's were, on average, 56% CI's
- Caveat: If the number of studies is small, no good choice!

Schmidt, Oh, and Hayes. Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. Br. J. Math. Stat. Psych. 2009. 62: 97-128.

If the number of studies is small, you need to choose among a few options

- A random effects model is problematic because the estimate of τ^2 will have poor precision
- Could perform a fixed-effects analysis
  - Doesn't allow generalization to wider population
- Could take a Bayesian approach
  - Putting a moderately informative prior distribution on τ^2
  - Probably the best option
- Could report separate effects (skipping the MA all together)

Borenstein, Hedges, Higgins and Rothstein. Introduction to Meta-Analysis. 2009. John Wiley & Sons Ltd.

Let's fit a random effects model to the LDL data comparing 10 mg Atorvastatin to placebo

For this data, Y_i is the mean difference from placebo in percent change from baseline LDL in study i, $i = 1, \ldots, 8$.

The random effects model is

$$Y_i \sim N(\mu_i, s_i^2) \text{ and } \mu_i \sim N(\theta, \tau^2)$$

The forest plot of the data was shown in the FE analysis. The plot and the estimates of Q and I^2 demonstrate significant between-study variability.

Recall, from the FE analysis $Q = 15.61$. With $\sum W_i = 1.684$ and $\sum W_i^2 = 0.804$ we have, $\widehat{\tau_{DL}}^2 = \frac{Q-(r-1)}{\sum W_i - \frac{\sum w_i^2}{\sum w_i}} = 7.13$

Random effects model for comparing 10 mg Atorvastatin to placebo

Given $\widehat{\tau_{DL}}^2 = 7.13$, we can calculate $\widehat{\theta}_{RE}$ and $\widehat{Var}\left(\widehat{\theta}_{RE}\right)$.

Specifically, letting $W_i^* = \left(s_i^2 + \widehat{\tau_{DL}}^2\right)^{-1}$, we get

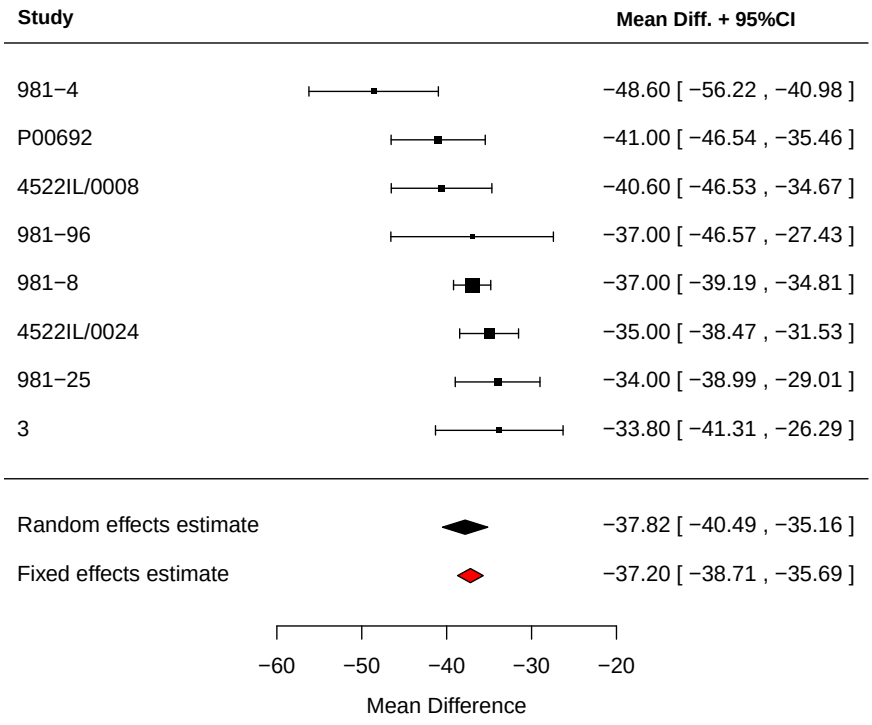
$$\widehat{\theta}_{RE} = -37.82 \text{ and } \widehat{Var}\left(\widehat{\theta}_{RE}\right) = 1.36^2$$

Thus, an approximate 95% CI for the mean difference from placebo is given by

$$-37.92 \pm 1.96 \times 1.36 = (-40.49, -35.16).$$

How do these compare to the fixed effects estimates and 95% CI?
How would we calculate a prediction interval for a new study?

Comparison of Fixed- and Random-effects estimates and confidence intervals



Using `metafor` and WinBUGS for random-effects models

Now, we'll look at using the `metafor` package and WinBUGS for fitting random effects model.

Understanding heterogeneity through meta-regression

The concept of meta-regression is a simple one

- Meta-regression is simply a regression of the outcome (e.g., treatment effect) on study-level covariates
  - Goal is to describe study-to-study variability in treatment effect with between study covariates (e.g., to describe different study populations)
- Covariates could also be on treatment arm-level covariates,
  - In a randomized study this shouldn't be necessary
  - May have some benefit in an observational study
- As we'll see later, MBMA is an extension of meta-regression

A simple extension of the fixed effects model

Suppose there is one covariate, x_i, for study i.

Then the general fixed effects meta-regression model is

$$\widehat{\theta}_i = \alpha + \beta x_i + \epsilon_i$$

with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \xi_i^2$ as before.

We get estimates for α and β using generalized least squares.

Similar approach to the random effects meta-regression

The general random effects meta-regression model is

$$\widehat{\theta}_i = \alpha + \beta x_i + \eta_i + \epsilon_i \text{ where } \eta_i \sim N\left(0, \tau^2\right)$$

with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \xi_i^2$ as before.

We will typically estimate α, β and τ^2 using likelihood-based methods (ML or REML).

A Strategy for dealing with heterogeneity (Whitehead, 2002)

Things to consider before starting your meta-analysis (e.g., to include in your analysis plan)

- Will you test for heterogeneity or assume its there (i.e., fixed vs. random effects models)
- If testing, how will you test for heterogeneity?
- What study-level covariates will be investigated?

Don't interpret meta-regression effects as subject-level effects: ecological bias

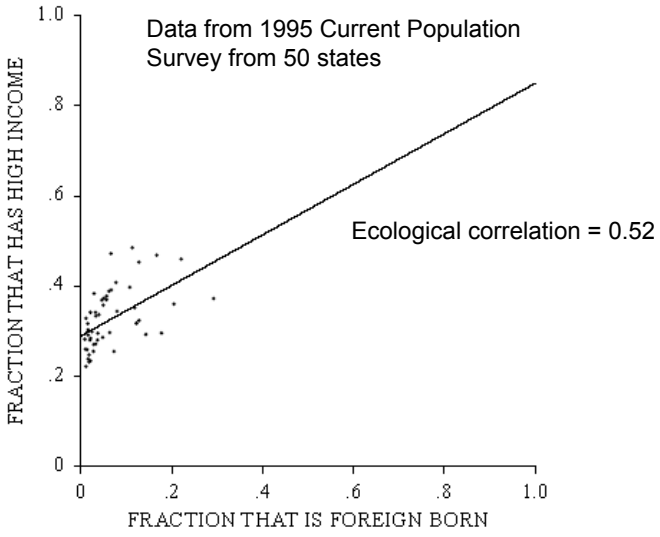
- Typically, group-level analyses must explore covariate effects using characterizations of the studies at the group level
- This is analogous to ecological analyses in which all individuals (e.g., defined geographically) are assigned an average value for a covariate (e.g., proportion of males)
- Ecological bias is the difference between the association at the individual and group levels
  - If we interpret the coefficients as effects at the individual level, we run the risk of making an incorrect interpretation

Two examples of ecological bias

Suicide Rates in 19th C Europe

- In 19th century Europe, suicide rates were higher in countries that were more heavily Protestant.
- Therefore, the social conditions of Protestantism promote suicide (Durkheim 1897)
- So, we might infer that if you were Protestant in 19th century Europe, you were more likely to commit suicide that if you werent (!?!)
- The problem is that the Protestant countries were different from the Catholic countries in many ways besides religion (Confounding)

Income and Nativity



So, we might infer that individuals who are foreign born are more likely to have a high income

But individual level correlation is -0.05

Freedman, Ecological Inference and Fallacy. Encyclopedia of the Social and Behavioral Sciences

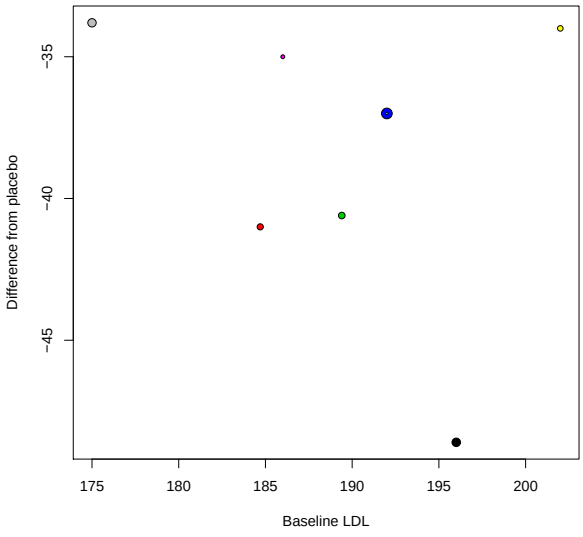
Causes of ecological bias

- Ecological bias arises from the inability of aggregate data to characterize within-group variability in covariates.
- Confounding can lead to ecological bias
  - Confounding between the covariate and response
  - Completely controlling for confounders with AD is generally not possible (Wakefield, 2008)
- Mis-specification of the correct group-level model (aka Pure Specification Bias or Aggregation Bias)
  - AD meta-regression models are rarely formulated as the average of the IPD models
- Wakefield (2008) gives a nice description of this problem and several examples.

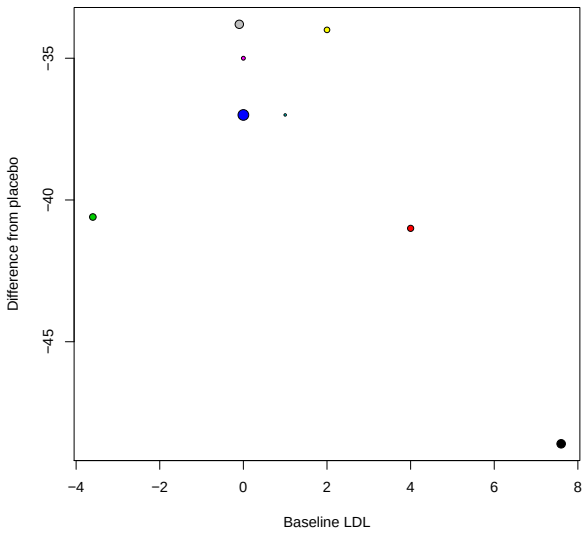
J. Wakefield. Ecological studies revisited. Annu.Rev.Public Health.2008. 29:75-90.

Let's look at the atorvastatin data more closely...

Difference from placebo vs. baseline LDL



Difference from placebo vs. placebo response



Placebo response looks more promising as a covariate

The simple fixed effects meta-regression model is

$$\hat{\theta}_i = \beta_1 + \beta_2 \times (\text{Placebo CFB}_i) + \epsilon_i$$

The corresponding random effects model is

$$\hat{\theta}_i = \beta_1 + \beta_2 \times (\text{Placebo CFB}_i) + \eta_i + \epsilon_i$$

$$\eta_i \sim N\left(0, \tau^2\right)$$

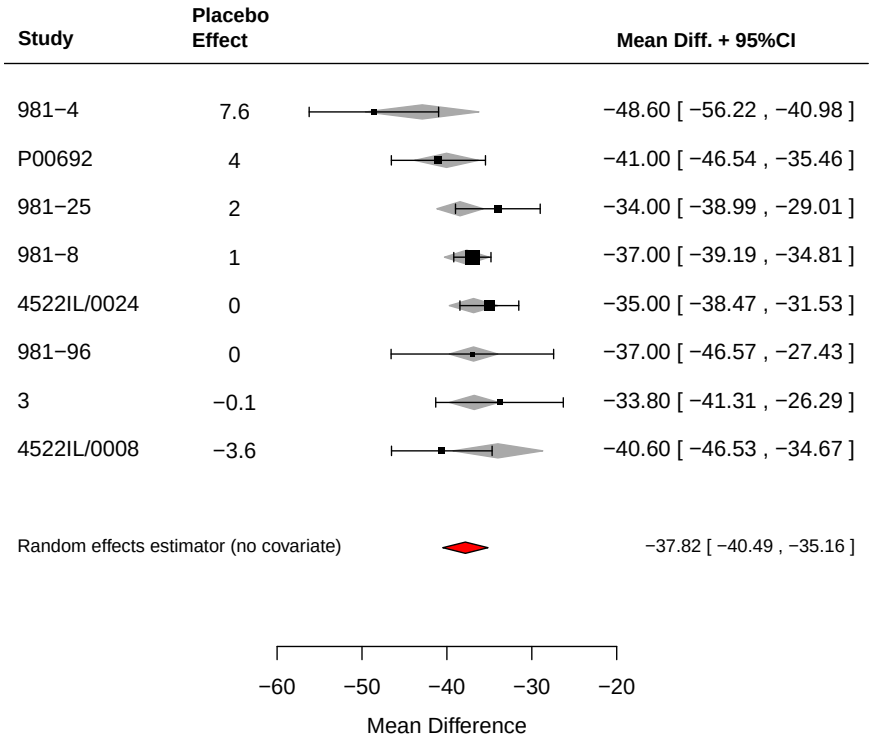
For a categorical variable, the fixed effect model would be coded similarly.

We can still use the rma() function to fit the meta-regression model

rma(yi=diff, sei=seDiff, data=ator10, mods=ldlPcfb.pbo, method="REML", slab=trial)

```
Mixed-Effects Model (k = 8; tau2 estimator:  REML)

tau2 (estimate of residual amount of heterogeneity):  6.7521 (SE = 7.7285)
tau (sqrt of the estimate of residual heterogeneity):  2.5985
Test for Residual Heterogeneity:
QE(df = 6) = 11.7588, p-val = 0.0676

Test of Moderators (coefficient(s) 2):
QM(df = 1) = 2.6447, p-val = 0.1039
Model Results:

estimate se zval pval ci.lb ci.ub
intrcpt -36.8594 1.4605 -25.2382 <.0001 -39.7219 -33.9970 ***
mods -0.7956 0.4893 -1.6263 0.1039 -1.7546 0.1633

---

Signif.  codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

From that object we can derive the estimates for each baseline value



Grey diamonds are centered at model predictions and extend to 95% CI.

Be careful with meta-regression

- Occasionally, we will pre-specify a meta-regression
  - E.g., based on prior work or publications
- If it's not pre-specified, we should probably view any significant relationships cautiously
  - These are exploratory analyses
  - Not unlike other kinds of subgroup analyses

Homework for meta-regression

- We will revisit the LDL example to examine if there are any covariates that might explain the heterogeneity of effect.
- Starting with plots of the data (scatter and forest plots)
- Then use the `metafor` package and WinBUGS to fit some meta-regression models

Miscellaneous (but important) topics

Topics

- Selection bias
- Missing data
- Combining different summary statistics

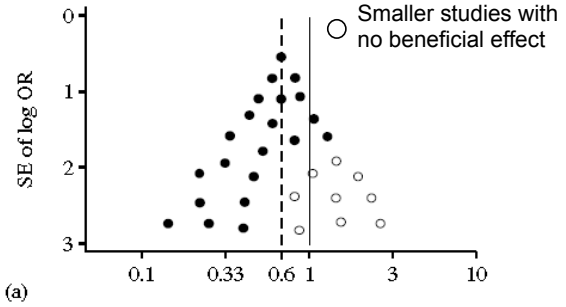
Assessing selection bias

- We can introduce bias if were not careful in how we select studies for inclusion in a meta-analysis
- Bias can result from
  - Including studies which themselves have biased estimates of treatment effects
  - The “file drawer” effect (publication bias)
- Sensitivity analyses can be used in the first case
- Publication bias has received much more attention
  - Publication bias occurs when the research that is readily available (or that you’re using for your analysis) differs in its results from the results of all the research that has been done in an area
- **Funnel plots** are the primary visual tool for the investigation of publication and other bias in meta-analysis.

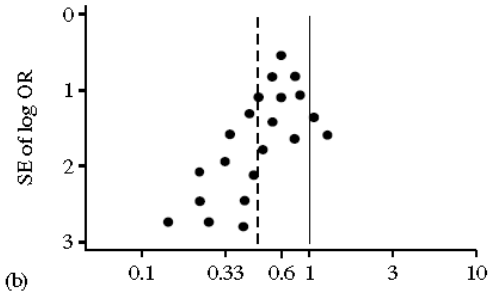
Rothstein, Sutton, and Borenstein. Publication Bias in Meta-Analysis. 2005. John Wiley & Sons

A scatter plot of the treatment effects (x) against a measure of study size (y)

Symmetrical plot in the absence of bias



Asymmetrical plot in the presence of publication bias (smaller studies showing no beneficial effects are missing)



Rothstein et al. Publication Bias in Meta-Analysis. Chapter 5.

In the absence of bias, the plot should look like a (symmetrical) funnel

- Results from small studies scatter at the bottom of the graph, with the spread narrowing among larger studies.
- If there is bias, then the funnel plot will appear asymmetrical
- Publication bias is only one possible cause of funnel plot asymmetry (Egger et al. 1997). Other causes include
  - Heterogeneity in results (e.g., differences in populations)
  - Poor methodological design/analysis of smaller studies
  - Random chance (e.g., more difficult to interpret with few studies)
- Funnel plots should be seen as a generic means of examining small-study effects
  - Rather than as a tool to diagnose specific types of bias.

Rothstein et al. Publication Bias in Meta-Analysis. Chapter 5.

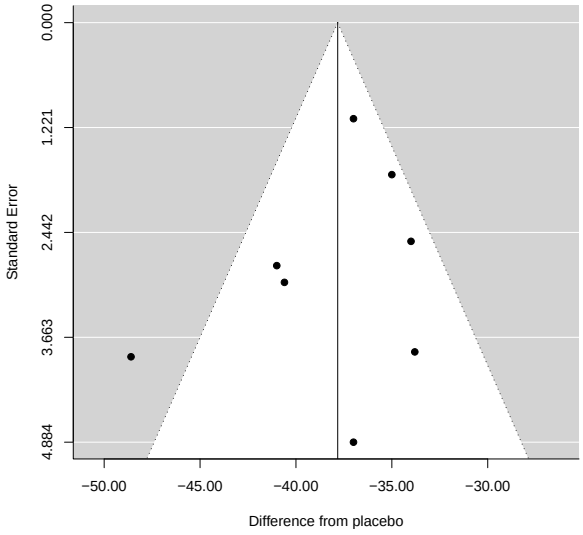
Recommendations for choice of axes for funnel plots

- For studies with binary outcomes
  - Standard error is the best measure of study size
  - Log odds ratios should generally be used as the measure of treatment effect
- For studies with numerical/continuous outcomes
  - Standard error of treatment difference or sample size as a measure of study size
  - Treatment difference as a measure of treatment effect

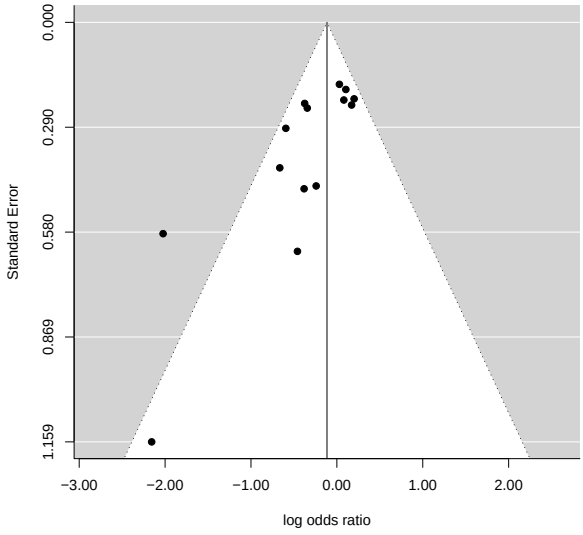
- There are more formal means for testing for selection/publication bias and attempting to correct for it.
  - Rothstein et al, Publication Bias in Meta-analysis, 2005
  - Whitehead, Meta-analysis of controlled clinical trials, 2002.

Do you think these show any indication of small-study effects / publication bias?

Atorvastatin 10 mg LDL data



Pancreatic cancer 6 mo. mortality



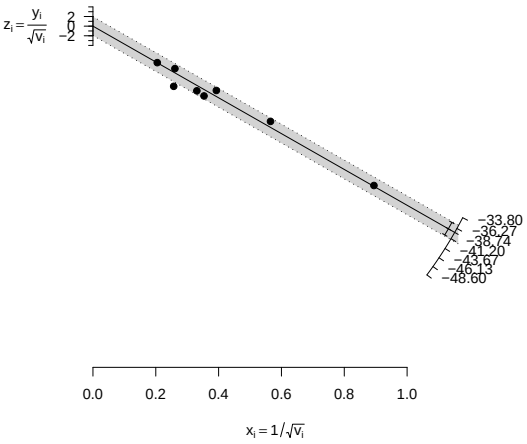
To meke these plots, use the `funnel` function.
For example, `funnel(refit, xlab="Difference from placebo")`

Another useful plot is the radial plot

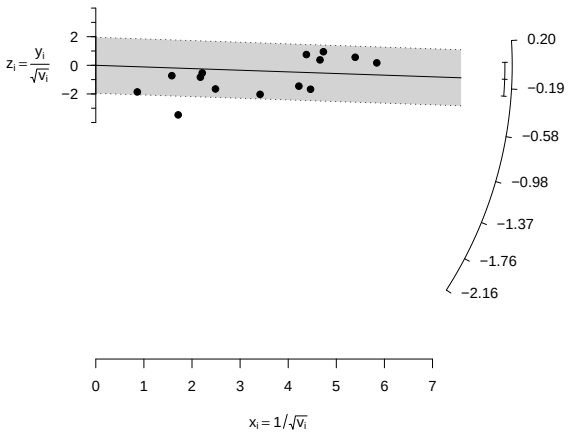
- Also called a Galbraith plot
- A plot of the standardized treatment effect vs. precision
- Circular axis is the treatment effect axis (e.g., log hazard ratio or log odds ratio)
- Regression line through origin points to overall estimate
- Approximate 95% confidence band should include ∼95% of points (if there is a common effect)
- Useful for assessing consistency amongst the studies

Radial plots for the atorvastatin and pancreatic cancer data sets

Atorvastatin 10 mg LDL data



Pancreatic cancer 6 mo. mortality



Complexities due to missing data in a meta-analysis

Missing data can impact a meta-analysis in (at least) two ways

- Missing papers (e.g., due to publication bias or an incomplete search of the literature)
- Values from individual studies are based on missing data in the individual studies

Table I. DSH meta-analysis data.

			Experimental arm				Control arm				
Trial	Author	Year	Successes	Failures	Missing	Total	Successes	Failures	Missing	Total	Per cent missing
1	Chowdhury	1973	54	17	0	71	65	19	0	84	0
2	Welu	1977	59	3	1	63	48	9	0	57	1
3	Hawton	1981	43	5	0	48	41	7	0	48	0
4	Allard	1992	41	22	13	76	44	19	11	74	16
5	Van Heerington	1995	175	21	62	258	161	34	63	258	24
6	Van der Sande	1997	116	24	0	140	114	20	0	134	0

From White, Higgins & Wood, 2008

Is missing data in an individual study really an issue?

- If there is informative missingness in individual studies, then we get biased estimates of effect in that study
- If there is non-informative missingness in individual studies
  - We can get biased estimates of effects if missingness depends on the “true” study-level effect
  - E.g., if there is higher drop-out in studies with small effects, the studies with large effects get too much weight
- This can be a particular issue when doing longitudinal meta-analysis

Methods for evaluating and assessing sensitivity to missing data

- How to handle this issue?
  - Funnel plots, etc. for publication bias (see previous caveats)
  - Frequently, issues with missing data within individual studies is ignored!
- Some principled ways for assessing sensitivity to missing data mechanisms have been proposed
  - White, Higgins and Wood. Allowing for uncertainty due to missing data in meta-analysis Part 1: Two-stage methods. Stat Med. 2008; 27:711-727
  - White, Welton, Wood et al. Allowing for uncertainty due to missing data in meta-analysis Part 2: Hierarchical models. Stat Med. 2008; 27: 728-745
- Some principled methods for handling bias when drop-out rate is related to underlying size of treatment effect
  - Yuan and Little. Meta-analysis of studies with missing data. Biometrics. 2009; 65: 487-496.

Combining trials that report different summary statistics

- What to do if the studies report different summary statistics?
  - E.g., some studies report the hazard ratio (and se) and others report only the % of subjects surviving at one year
  - E.g., Some studies report a responder rate (based on an underlying continuous variable) and others report the mean and sd of the continuous variable
  - E.g., some studies report the mean and others the median response
- What if they report a similar endpoint (e.g., log OR), but from different regression models?
  - E.g., stratifying or adjusting for different covariates
- What if they report results using a variety of different outcome scales/instruments?
  - In Alzheimer's Disease some studies measure cognitive effects using the MMSE and others using ADAS-cog
  - In oncology, various quality of life scales are used

There are ways to handle this, but...

- They rely heavily on making untestable assumptions
- These assumptions can (will) affect the validity and interpretability of the results
- Combining different outcome scales/instruments
  - Frequently, use the standardized difference or response
  - Assumes that clinical important of x units is the same throughout the scales and approximately normally distributed
- Combining different summary statistics
  - Try to convert to a common or shared measure (see following examples)
  - Often relies on completely untestable assumptions
- It is always preferable to have the IPD to do the meta-analysis
  - This prevents these sorts of problems.
  - However, we rarely are in the position of having IPD outside of our own studies.

Examples of combining different endpoints

- Combination of hazard ratios and % surviving at a time point
  - Treat % surviving as interval censored, then the OR approximates the hazard ratio
  - Makes assumptions about censoring mechanism
- Combination of responder rates and continuous summary statistics
  - Approximate responder rate assuming normal (or some other) distribution
  - Makes un-assessable distributional assumptions

More examples of combining different endpoints

- Combination of medians and means
  - Assume constant multiplier for relationship between standard errors
  - Depends on assumed underlying distribution (e.g., if normal, multiplier is $\sim$1.25)
- Combination of ordinal data when only some categories are reported in each study
  - Consider modeling outcomes using ordered categorical model (e.g., PASI analysis; Reich et al. Efficacy of biologics in the treatment of moderate to sever psoriasis: a network meta-analysis of controlled trials. *Br J Dermatol* 2012; 166(1): 179-88)
  - Makes relatively few assumptions

Network meta-analysis

(aka Mixed treatment comparisons)
(aka Multiple treatment meta-analysis)

More to come . . .

Model-based meta-analysis

Model-based meta-analysis (MBMA): What is it?

- The distinction between traditional MA and MBMA is not a sharp one. We're really talking about a continuum from MA to meta-regression to MBMA.
- For our purposes let's use the term MBMA to refer to meta-regression using models based on pharmacometric knowledge and principles.
- These will often employ models that describe some efficacy- or safety-related response as a nonlinear function of drug exposure and possibly time. They may also consider other covariates such as demographics and other baseline measurements.

Role of MBMA: Objectives not adequately addressed by traditional meta-analysis

- Indirect statistical comparisons, i.e., comparisons of treatments that were never administered within the same study. A more "statistical" approach to this problem is known as network MA.
- Predictions and inferences about responses that would occur at values of independent variables that were never directly studied, e.g., responses to doses that were never administered.
- Integration of heterogeneous data and knowledge.

Why Bayesian?

- Natural extension of the types of hierarchical models we commonly use for population modeling
- Model parameter estimation and model-based inference do not require a 2 step process
- Inferences have a more natural and direct interpretation than do those from a frequentist perspective.
- *Prior information about model parameters is easily integrated into the modeling process.*

Why BUGS?

- Flexibility!! The BUGS model specification language is very flexible w.r.t. stochastic structure.
  - As many levels of variability as you want.
  - Easy to simultaneously fit multiple models even if the stochastic structures are very different
  - Large selection of built-in probability distributions
- On the other hand BUGS is not as flexible wrt models where the deterministic components are complex. This is due to the lack of control structures like if-then-else and true loops.
- However the user can program custom functions in Component Pascal, the language in which WinBUGS is written.

Modeling sample statistics I: Sample means

Choosing a sampling distribution and adjusting for sample size

Let's start with the simplest case: one sample mean per treatment arm. The primary independent variables will usually describe the treatments, e.g., the drug and the dose. Frequently the distribution of the residual variation in the sample means may be adequately described by a normal distribution with a variance adjusted for the sample size.

$$\overline{y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Possible justifications for the use of the normal distribution include:

- The underlying individual data is normally distributed, or
- An argument that the mean is asymptotically normal (central limit theorem).

Modeling sample statistics I: Sample means

Choosing a sampling distribution and adjusting for sample size

One variation of this approach is to use the reported standard error of the mean instead of estimating it:

$$\overline{y} \sim N\left(\mu, se^2\right)$$

This has the potential advantage of allowing for inter-trial (and inter-arm) differences in variability.

Modeling sample statistics I: Sample means

Choosing a sampling distribution and adjusting for sample size

Alternative distributions for the sample means may be considered, e.g., log-normal or t, but rigorous treatment of sample size adjustment and justification based on the distribution of individual data are elusive. For example the mean of log-normally distributed individual data is not log-normally distributed. Also if you log-transform the sample means, it is not clear what the appropriate sample size adjustment in the variance should be.

Modeling sample statistics I: Sample means

Original scale vs CFB or PCFB

Clinical measurements are often reported in terms of change from baseline (CFB) or percent change from baseline (PCFB). When modeling endpoint data (as opposed to longitudinal data) the choice of whether to use the measurement on the original scale or baseline adjusted depends on what summary statistics are available and the intended use of the model. If CFB (or PCFB) is commonly used for the primary statistical analysis of clinical trials for the indication of interest, then modeling of CPB (or PCFB) is a logical choice and will probably appeal to the relevant decision-makers. In some cases predictions on both the original and baseline adjusted scales may be desired, so it may be necessary to construct models for both. Either or both models may include the mean baseline response as a covariate.

Modeling sample statistics I: Sample means

Original scale vs CFB or PCFB (cont.)

With longitudinal data that includes a baseline measurement a reasonable approach is to model the entire time course on the original scale. CFB can then be predicted by calculating the difference between the predicted values for a given time and that for baseline.

Modeling sample statistics I: Sample means

Other covariates

In addition to describing the effects of treatment we may want to account for other differences in results among trials and treatment arms. There may be summary statistics for covariates describing the patients in a treatment arm, e.g., mean age, mean weight, fraction of females, fraction taking a particular concomitant medication, etc. In addition there may be covariates general to a whole trial, e.g., geographic location, phase of development, etc.

Modeling sample statistics I: Sample means

Unexplained inter-trial variation

Finally we may want to quantify unexplained variation among trials via random effects, i.e., inter-trial variation. Since there is only one observation per treatment arm, inter-arm variation and other sources of residual variation are not separately estimable.

Modeling sample statistics I: Sample means

Interpretation of covariates and their effects

Because patient-specific covariates are in the form of summary statistics, their values cover a narrower range than the individual values. Consequently they are less informative about the their effects unless the data have been stratified based on them, e.g., separate means for males and females, or for young and elderly.

Example: Dose response model based on sample means

MBMA to assess dose-response of gemcabene-statin combinations relative to marketed treatments

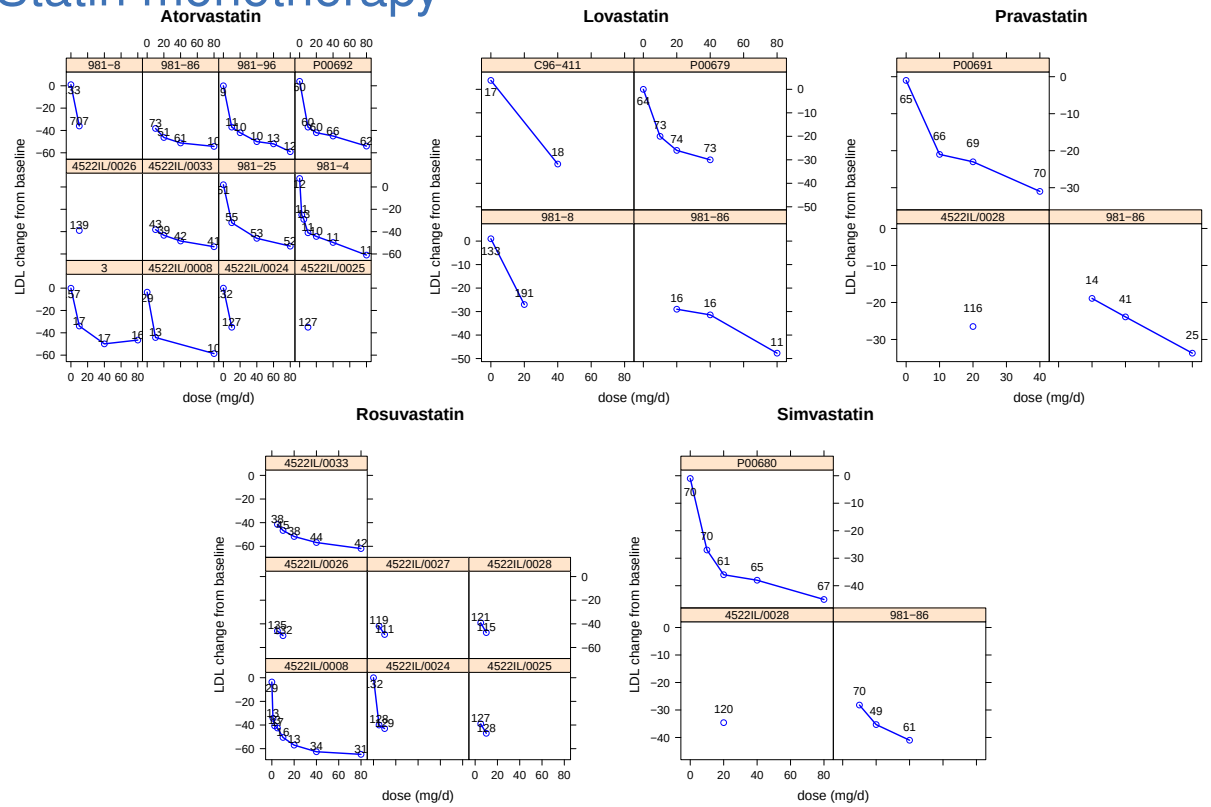
The scenario

This example revisits the MBMA of LDL concentrations described by Mandema et al [MHW+05]. Differences from the original analysis include use of a Bayesian method, and the gemcabene data is simulated. The remaining data were extracted from public sources. The primary objective of the analysis is to assess the benefit of gemcabene add-on therapy relative to marketed treatments including ezetimibe-statin combinations.

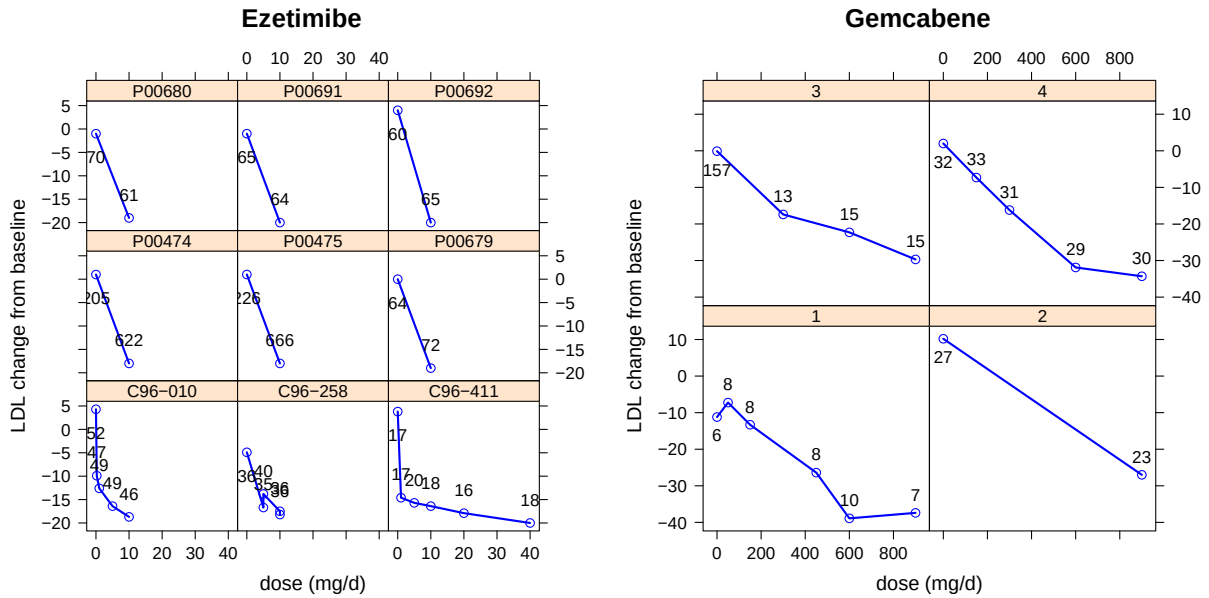
The data

- The data
  - Sample means for LDL % change from baseline at study endpoint from published sources
    - 25 studies
    - 154 treatment arms
    - Data reflects results from 9751 patients
    - 4–16 week treatment duration
  - Available covariates: baseline LDL, treatment duration.
- Objective
  - Construct a model for LDL % change from baseline as a function of drug, daily dose and possibly other covariates.
- Data file: `statinGemcabene/ldlData.csv`

Statin monotherapy

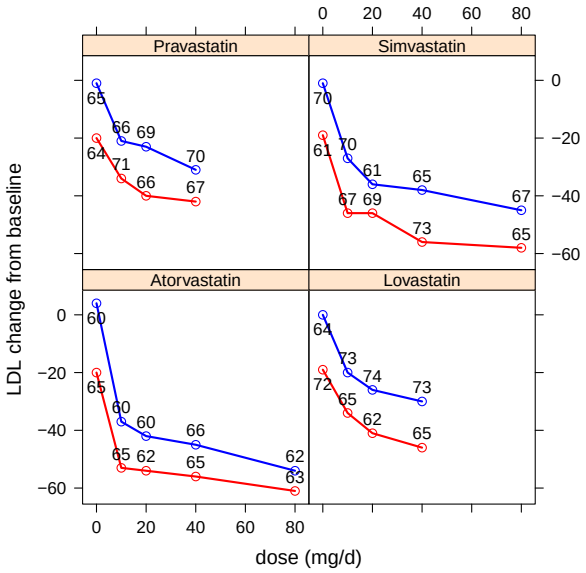


Non-statin monotherapy

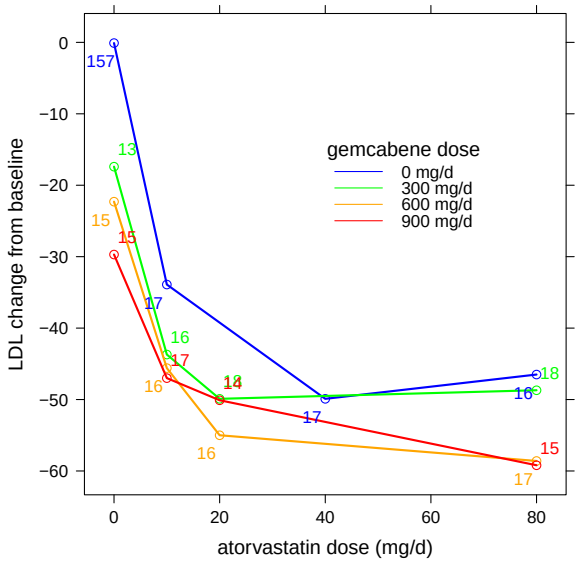


Non-statin / statin combinations

statin dose response with and without ezetimibe 10 mg/d



atorvastatin dose response when combined with a range of gemcabene doses



Proposed model

Emax models with interaction for mean LDL % change from baseline in the i^{th} treatment arm of the j^{th} trial:

$$
\begin{aligned}
E_{ij} &\sim N\left(\overline{E}_{ij}, \frac{\sigma^2}{n_{ij}}\right) \\
\overline{E}_{ij} &= E_{0,j} - E_{\text{statin},ij} - E_{\text{non-statin},ij} + 0.01\gamma_{ij} E_{\text{statin},ij} E_{\text{non-statin},ij} \\
E_{\text{statin},ij} &= \frac{E_{\max,\text{statin}} D_{\text{statin},ij}^{n_{\text{statin}}}}{ED_{50,\text{statin},ij}^{n_{\text{statin}}} + D_{\text{statin},ij}^{n_{\text{statin}}}} \\
E_{\text{non-statin},ij} &= \frac{E_{\max,\text{non-statin},ij} D_{\text{non-statin},ij}^{n_{\text{non-statin},ij}}}{ED_{50,\text{non-statin},ij}^{n_{\text{non-statin},ij}} + D_{\text{non-statin},ij}^{n_{\text{non-statin},ij}}} \\
E_{0,j} &\sim N\left(\overline{E_0}, \omega_{E_0}^2\right)
\end{aligned}
$$

Proposed model

Emax models with interaction for mean LDL % change from baseline in the i^{th} treatment arm of the j^{th} trial:

$$
\begin{aligned}
ED_{50,\text{statin},ij} &= \begin{cases} ED_{50,\text{atorvastatin}}, & \text{statin}_{ij} = \text{atorvastatin} \\ ED_{50,\text{rosuvastatin}}, & \text{statin}_{ij} = \text{rosuvastatin} \\ ED_{50,\text{simvastatin}}, & \text{statin}_{ij} = \text{simvastatin} \\ ED_{50,\text{lovastatin}}, & \text{statin}_{ij} = \text{lovastatin} \\ ED_{50,\text{pravastatin}}, & \text{statin}_{ij} = \text{pravastatin} \end{cases} \\
E_{\text{max,non-statin},ij} &= \begin{cases} E_{\text{max,ezetimibe}}, & \text{non-statin}_{ij} = \text{ezetimibe} \\ E_{\text{max,gemcabene}}, & \text{non-statin}_{ij} = \text{gemcabene} \end{cases} \\
ED_{50,\text{non-statin},ij} &= \begin{cases} ED_{50,\text{ezetimibe}}, & \text{non-statin}_{ij} = \text{ezetimibe} \\ ED_{50,\text{gemcabene}}, & \text{non-statin}_{ij} = \text{gemcabene} \end{cases} \\
n_{\text{non-statin},ij} &= \begin{cases} n_{\text{ezetimibe}}, & \text{non-statin}_{ij} = \text{ezetimibe} \\ n_{\text{gemcabene}}, & \text{non-statin}_{ij} = \text{gemcabene} \end{cases} \\
\gamma_{ij} &= \begin{cases} \gamma_{\text{ezetimibe}}, & \text{non-statin}_{ij} = \text{ezetimibe} \\ \gamma_{\text{gemcabene}}, & \text{non-statin}_{ij} = \text{gemcabene} \end{cases}
\end{aligned}
$$

Proposed model

Prior distributions

$$
\begin{aligned}
\overline{E_0} &\sim N\left(0, 10^6\right) \\
\log\left(ED_{50,\text{statin}}\right) &\sim N\left(0, 10^6\right) \\
E_{\text{max,non-statin}} &\sim U\left(0, 100\right) \\
n_{\text{ezetimibe}} &= 1 \\
\log\left(\gamma_{\text{non-statin}}\right) &\sim N\left(0, 10^6\right) \\
\frac{1}{\omega^2_{E_0}} &\sim \text{gamma}\left(0.01, 0.01\right)
\end{aligned}
\qquad
\begin{aligned}
E_{\text{max,statin}} &\sim U\left(0, 100\right) \\
n_{\text{statin}} &\sim U\left(0.1, 10\right) \\
\log\left(ED_{50,\text{non-statin}}\right) &\sim N\left(0, 10^6\right) \\
n_{\text{gemcabene}} &\sim U\left(0.1, 10\right) \\
\frac{1}{\sigma^2} &\sim \text{gamma}\left(0.01, 0.01\right)
\end{aligned}
$$

Re-parametrization to improve MCMC convergence

PII-108

"TRUNCATED SIGMOID E_{max} MODELS": A REPARAMETERIZATION OF THE SIGMOID E_{max} MODEL FOR USE WITH TRUNCATED PK/PD DATA. WJ Bachman PhD and WR Gillespie PhD, GloboMax LLC, Hanover, MD.

The parameters of the sigmoid E_{max} model are poorly estimated when the range of PK/PD data available is limited to $<0.95E_{max}$ [Dutta et al. J Pharm Sci 85:232 (1996)]. The following reparameterized form of the sigmoid E_{max} model has improved parameter estimation properties:

$$E = E_0 + \frac{(\beta^\gamma + 1)(E^* - E_0)C^\gamma}{C^{*\gamma} + \beta^\gamma C^\gamma}$$

where E is the effect measure and C is a measure of drug exposure (e.g., concentration or dose). The parameter E^* is the estimated effect resulting from C^*, γ is the usual "sigmoidicity" parameter, and E_0 is the baseline effect. β is a measure of the degree to which the function deviates from linearity in C^γ. One approach to applying this parameterization is to fix C^* (or E^*) at a value and estimate the remaining parameters E_0, E^* (or C^*), β, and γ by nonlinear regression. The properties of this approach are evaluated by application to simulated PK/PD data that is truncated at various fractions of E_{max}. When C^* (or E^*) is chosen within the range of the observed data, then the parameters E^* (or C^*) and β are more precisely and accurately estimated than EC_{50} and E_{max} of the standard parameterization.

Model parameterization implemented in statinGemcabeneTruncEmax.txt

CP&T 63:199 (1998) [BG98]

Proposed model

Advanced topic: Re-parametrization to improve MCMC convergence

Alternative prior distributions based on the use of truncated Emax parametrization [BG98]:

$$\begin{aligned}
E_{\text{max,statin}} &= \left(\frac{1}{b_{\text{atorvastatin}}^{n_{\text{statin}}}} + 1\right) E^*_{statin} \\
ED_{50,\text{statin},ij} &= \frac{D^*_{\text{statin}}}{b_{\text{statin},ij}} \qquad b_{\text{statin},ij} = \pi_{\text{statin},ij} b_{atorvastatin} \\
\pi_{\text{statin},ij} &= \begin{cases} 1, & \text{statin}_{ij} = \text{atorvastatin} \\ \pi_{\text{rosuvastatin}}, & \text{statin}_{ij} = \text{rosuvastatin} \\ \pi_{\text{simvastatin}}, & \text{statin}_{ij} = \text{simvastatin} \\ \pi_{\text{lovastatin}}, & \text{statin}_{ij} = \text{lovastatin} \\ \pi_{\text{pravastatin}}, & \text{statin}_{ij} = \text{pravastatin} \end{cases} \\
D^*_{\text{statin}} &= 80 \qquad E^*_{statin} \sim U(0, 100) \\
b_{atorvastatin} &\sim U(0, 1000) \qquad \pi_{\text{statin}} \sim N\left(0, 10^6\right)
\end{aligned}$$

Proposed model

Advanced topic: Re-parametrization to improve MCMC convergence (cont.)

Alternative prior distributions based on the use of truncated Emax parametrization [BG98]:

$$
\begin{aligned}
E_{\text{max,ezetimibe}} &= \left(\frac{1}{b_{\text{ezetimibe}}} + 1\right) E^*_{ezetimibe} \\
ED_{50,\text{ezetimibe}} &= \frac{D^*_{\text{ezetimibe}}}{b_{\text{ezetimibe}}} \\
D^*_{\text{ezetimibe}} &= 40 \quad E^*_{ezetimibe} \sim U(0,100) \quad b_{\text{ezetimibe}} \sim U(0,1000) \\
E_{\text{max,gemcabene}} &= \left(\frac{1}{b^{n_{\text{gemcabene}}}_{\text{gemcabene}}} + 1\right) E^*_{gemcabene} \\
ED_{50,\text{gemcabene}} &= \frac{D^*_{\text{gemcabene}}}{b_{\text{gemcabene}}} \\
D^*_{\text{gemcabene}} &= 900 \quad E^*_{gemcabene} \sim U(0,100) \quad b_{\text{gemcabene}} \sim U(0,1000)
\end{aligned}
$$

Hands-on Problem 3: Dose response model based on sample means

MBMA to assess dose-response of atorvastatin and rosuvastatin

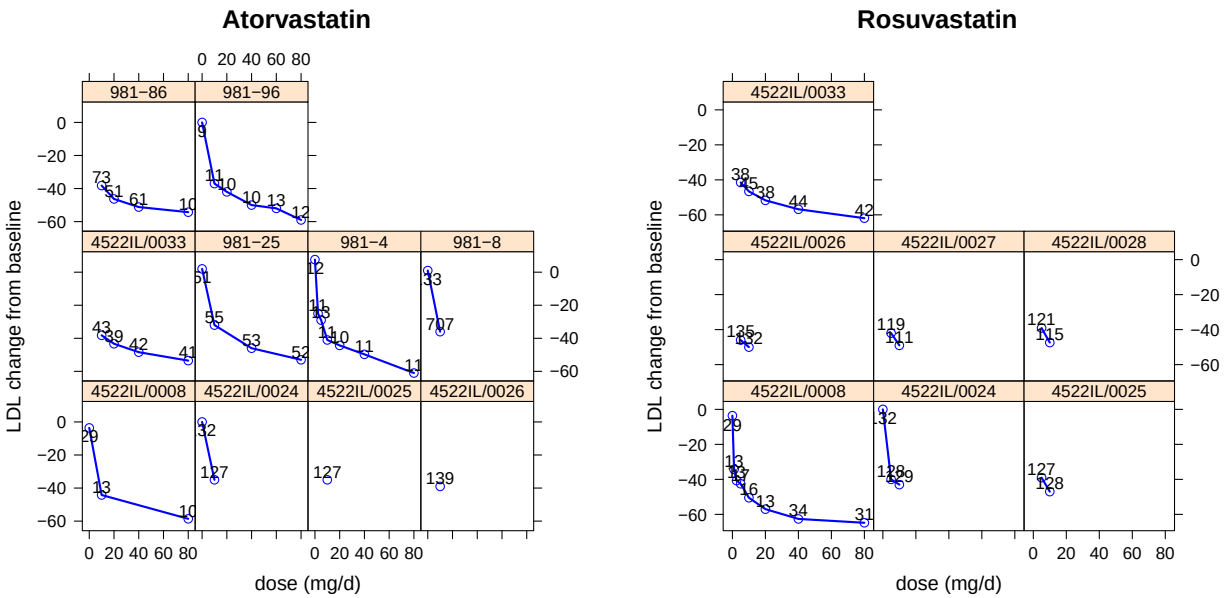
The scenario

This exercise explores the effects of two statins on LDL concentration % change from baseline. The data were extracted from public sources. The primary objective of the analysis is to compare the dose-response of atorvastatin and rosuvastatin.

The data

- The data
  - Sample means for LDL % change from baseline at study endpoint from published sources
    - 12 studies
    - 56 treatment arms
    - Data reflects results from 3721 patients
    - 6–16 week treatment duration
  - Available covariates: baseline LDL, treatment duration.
- Exercise
  - Construct a model for LDL % change from baseline as a function of drug, daily dose and possibly other covariates.
  - Try both standard and truncated Emax parametrizations.
- Data file: `statinHandsOn/ldlData.csv`

LDL % change from baseline following atorvastatin and rosuvastatin



Proposed model

Emax models with interaction for mean LDL % change from baseline in the i^{th} treatment arm of the j^{th} trial:

$$
\begin{aligned}
E_{ij} &\sim N\left(\overline{E}_{ij}, \frac{\sigma^2}{n_{ij}}\right) \\
\overline{E}_{ij} &= E_{0,j} - E_{\text{statin},ij} \\
E_{\text{statin},ij} &= \frac{E_{\max,\text{statin}} D_{\text{statin},ij}^{n_{\text{statin}}}}{ED_{50,\text{statin},ij}^{n_{\text{statin}}} + D_{\text{statin},ij}^{n_{\text{statin}}}} \\
E_{0,j} &\sim N\left(\overline{E_0}, \omega_{E_0}^2\right) \\
ED_{50,\text{statin},ij} &= \begin{cases} ED_{50,\text{atorvastatin}}, & \text{statin}_{ij} = \text{atorvastatin} \\ ED_{50,\text{rosuvastatin}}, & \text{statin}_{ij} = \text{rosuvastatin} \end{cases}
\end{aligned}
$$

Proposed model

Proposed prior distributions

$$
\begin{aligned}
\overline{E_0} &\sim N\left(0, 10^6\right) \\
E_{\max,\text{statin}} &\sim U\left(0, 100\right) \\
\log\left(ED_{50,\text{statin}}\right) &\sim N\left(0, 10^6\right) \\
n_{\text{statin}} &\sim U\left(0.1, 10\right) \\
\frac{1}{\omega_{E_0}^2} &\sim \text{gamma}\left(0.01, 0.01\right) \\
\frac{1}{\sigma^2} &\sim \text{gamma}\left(0.01, 0.01\right)
\end{aligned}
$$

Modeling sample statistics II: Sample standard deviations

Why?

- If the intended application of the model is adequately supported by prediction and comparison of population mean responses for different treatments then analysis of sample mean data is probably sufficient.
  - Including analysis of sample standard deviations will usually provide only marginal benefit—mainly by accounting for inter-trial differences in variability and thereby more appropriately weighting the studies.
- On the other hand some applications, such as trial design and simulation, benefit significantly from better estimates of variability and the degree to which that variability varies from trial to trial. In such cases it is worth the additional effort to model the sample standard deviations along with the means.

Modeling sample statistics II: Sample standard deviations

How?

If the underlying individual data is normally distributed with mean μ and standard deviation σ then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$ where s is the sample standard deviation and n is the sample size. Equivalently,

$$s^2 \sim \text{gamma}\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right)$$

Example: MBMA of gemcabene-statin combinations

The statin-gemcabene example is extended to include modeling of the sample standard deviations. This is demonstrated using the data and code provided in the following folders:

- `statinGemcabeneSDTruncEmax`: Same residual standard deviation for all trials
- `statinGemcabeneSDTruncEmax2`: Inter-trial variation in the residual standard deviation

Hands-on Problem 3b: Dose response model based on sample means and standard deviations

Extend the statin dose-response model you implemented in hands-on problem 3 to use the reported standard error data. Do this in 3 ways:

1. Directly use the observed standard error in the likelihood for the sample mean (ala statinGemcabeneTruncEmax2)
2. Model the sample mean and standard deviation data assuming the same residual standard deviation for all trials (ala statinGemcabeneSDTruncEmax).
3. Model the sample mean and standard deviation data allowing for inter-trial variation in the residual standard deviation (ala statinGemcabeneSDTruncEmax2).

- Compare the results of the 3 models.
- Is it appropriate to compare the model 1 results with those of models 2 and 3 using DIC or mean deviance? Why or why not?

Recent example: Comparative efficacy of DPP-4 inhibitors

A novel model-based meta-analysis to indirectly estimate the comparative efficacy of two medications: an example using DPP-4 inhibitors, sitagliptin and linagliptin, in treatment of type 2 diabetes mellitus

Jorge Luiz Gross,[1] James Rogers,[2] Daniel Polhamus,[2] William Gillespie,[2] Christian Friedrich,[3] Yan Gong,[4] Brigitta Ursula Monz,[4] Sanjay Patel,[5] Alexander Staab,[3] Silke Retlich[3]

BMJ Open 2013;3: e001844 (http://bmjopen.bmj.com/content/3/3/e001844) [GRP+13]

Population simulations

- Here I use the term “population simulation” to refer to simulation of a population estimand, e.g., population mean, population percentile, probability of an event, etc.
- This is in contrast to simulation of clinical trial results or individual observations.
- The idea is to use simulations to characterize the probable range, e.g., 90 or 95% credible interval, for the population estimand for each treatment of interest.

Simulations with models based on analysis of only summary data

- Models based on analysis of summary statistics alone are not ideally suited for simulation of individual patient data.
- They are appropriate for simulation of:
  - The summary statistic on which they are based and
  - A population estimand for which the summary statistic is an estimator.
    - For example a sample mean is an estimator of the population mean, and the fraction of patients that experience a particular outcome is an estimator of the probability of that outcome.

Simulations with models based on analysis of only summary data

- Suppose we have a model of the following form describing the sample mean for a clinical measurement in the i^{th} treatment arm:

$$\overline{y}_i \sim N\left(\mu_i, \frac{\sigma^2}{n_i}\right) \qquad \mu_i = f\left(D_i, \text{drug}_i, \theta\right)$$

  For now let's assume no inter-trial variation in the model parameters (θ).
  - For this case the population mean response to dose D_i of drug drug_i is just μ_i.
  - Given MCMC generated posterior samples of θ we can generate posterior predictions of μ for the dose of any drug in the model.
  - The 50^{th}, 5^{th} and 95^{th} percentiles of those predictions estimates the median and 90% credible interval for the population mean.
- This approach is readily generalized to distributions other than normal, most easily in cases when the expected value for that distribution is an analytic function of the parameters.

Example: Population simulations

Simulations to assess dose-response of gemcabene-statin combinations relative to marketed treatments

- This example uses the MCMC samples generated in statin-gemcabene modeling example to simulate the posterior distribution of population mean LDL % change from baseline predicted to result from various doses of statins, ezetimibe, gemcabene and combinations of statins and non-statins.
- The primary objective of the simulations is to assess the benefit of gemcabene add-on therapy relative marketed treatments including ezetimibe-statin combinations.

Example: Population simulations

Simulation specifications

- Statin doses: c(0:4, seq(5, 80, by = 5)) mg/d
- Ezetimibe doses: c(0, 0.25, 0.5, 1, 2.5, 5, 10, 20, 30, 40) mg/d
- Gemcabene doses: seq(0, 900, by = 50) mg/d
- Number of posterior samples: 1000 per dose combination
- Create plots showing posterior median and 90% credible intervals for the population mean LDL % change from baseline versus dose
  - Statin monotherapy
  - Non-statin monotherapy
  - Atorvastatin dose-response with and without ezetimibe 10 mg/d
  - Atorvastatin dose-response with and without gemcabene 900 mg/d

Simulations with models that include inter-individual variation
Simulations to describe the probable range of a population estimand

- Suppose we want to describe the probable range of a population estimand, but we have a model constructed by analysis of individual data. Usually such a model will include inter-individual variation in some parameters.
- For simplicity let's assume a model of the following form for an observation on the i^{th} occasion in the j^{th} individual:

$$\begin{aligned} y_{ij} &\sim N\left(\widehat{y}_{ij}, \sigma^2\right) \\ \widehat{y}_{ij} &= f\left(t_{ij}, D_j, \text{drug}_j, \theta_j\right) \\ \theta_j &\sim N\left(\widehat{\theta}, \Omega\right) \end{aligned}$$

Again we assume no inter-trial variation.

Simulations with models that include inter-individual variation
Simulations to describe the probable range of a population estimand

- In most cases we do not have a closed-form solution to the function for the population estimand, so we approximate it via simulation:
  - For each MCMC generated posterior sample of the model parameters and set of independent variable values, we simulate many replicates of the predicted response at a desired time t_{ij} and treatment (drug_{ij} and D_{ij}).
  - We can then readily calculate simulated values for the desired population estimand using the appropriate sample statistic, e.g., the sample mean to estimate the population mean, the sample percentile to estimate the corresponding population percentile, etc.

Simulations with models that include inter-individual variation
Simulations to describe the probable range of a population estimand

for $m = 1$ to n_{post} **do**
  sample $\widehat{\theta}^{(m)}$, $\Omega^{(m)}$ and $\sigma_{(m)}$ from the posterior distribution
  for $j = 1$ to n_{pop} **do**
    sample $\theta_j^{(m)} \sim N\left(\widehat{\theta}^{(m)}, \Omega^{(m)}\right)$
    calculate $\widehat{y}_{ij}^{(m)} = f\left(t_{ij}, D_{ij}, \text{drug}_{ij}, \theta_j^{(m)}\right)$
    sample $y_{ij}^{(m)} \sim N\left(\widehat{y}_{ij}^{(m)}, \sigma_{(m)}^2\right)$
  end for
  calculate the appropriate sample statistic for the desired estimate, e.g., the sample mean for the population mean: $\overline{y}_i^{(m)} = \frac{1}{n_{pop}} \sum_{j=1}^{n_{pop}} \widehat{y}_{ij}^{(m)}$. In the case of the population mean we don't need to sample from the residual variation because $E\left(y_{ij}\right) = \widehat{y}_{ij}$.
end for

Simulations with models that include inter-individual variation
Simulations to describe the probable range of a population estimand

- For such models it has been suggested that another approach to calculating a population mean is to substitute the population mean of the patient-specific parameters into the function for predicting a new response.
- I would argue that estimand is better described as a typical response for an individual patient (or a response for a typical patient).
- The posterior distribution of that estimand might be called the posterior distribution of a typical response.
- However that distribution is not useful for inferences except in cases where it is a good approximation to the distribution of the population mean response, e.g., linear models.

Simulations to describe the distribution of individual observations

- How might we use simulations to describe or make inferences about interindividual variation?
- Let's look at 3 different types of simulations that communicate somewhat different things about the distribution of individual observations.

Simulations to describe the distribution of individual observations
Posterior predictive distribution of individual observations

- Simulations from the posterior predictive distribution of individual observations are appropriate for inferences regarding the probable range of observations in an individual.
- That distribution reflects all sources of variability, e.g., inter-individual and residual, and uncertainty in the model parameters. As a result it is more disperse that one would predict for the distribution of observations in a typical population.

Simulations to describe the distribution of individual observations
Posterior predictive distribution of individual observations

Samples from the posterior predictive distribution may be generated according to:

for $j = 1$ to n_{post} **do**
  sample $\widehat{\theta}^{(j)}$, $\Omega^{(j)}$ and $\sigma_{(j)}$ from the posterior distribution
  sample $\theta_j^{(j)} \sim N\left(\widehat{\theta}^{(j)}, \Omega^{(j)}\right)$
  calculate $\widehat{y}_{ij}^{(j)} = f\left(t_{ij}, D_{ij}, \text{drug}_{ij}, \theta_j^{(j)}\right)$
  sample $y_{ij}^{(j)} \sim N\left(\widehat{y}_{ij}^{(j)}, \sigma_{(j)}^2\right)$
end for

Note that nested loops are not required.

Simulations to describe the distribution of individual observations
Distribution of individual observations for a “typical” population

- The notion here is to simulate a “typical” distribution of observations in order to communicate the scale of variability in a “typical” population.
- This distribution is generally not suitable for formal Bayesian statistical inference but it may be useful as a communication device.

Simulations to describe the distribution of individual observations
Distribution of individual observations for a "typical" population

Samples from the distribution of individual observations for a "typical" population may be generated according to:

calculate posterior means $\overline{\theta} = \frac{1}{n_{\text{post}}} \sum_{m=1}^{n_{\text{post}}} \widehat{\theta}(m)$ and $\overline{\Omega} = \frac{1}{n_{\text{post}}} \sum_{m=1}^{n_{\text{post}}} \Omega^{(m)}$ (or medians) of the posterior samples of the model parameters.
for $j = 1$ to n_{pop} **do**
  sample $\theta_j \sim N\left(\overline{\theta}, \overline{\Omega}\right)$
  calculate $\widehat{y}_{ij} = f\left(t_{ij}, D_{ij}, \text{drug}_{ij}, \theta_j\right)$
  sample $y_{ij} \sim N\left(\widehat{y}_{ij}, \sigma^2\right)$
end for

This is a Bayesian analog to the distribution used in what is often called "visual predictive checks" based on point estimation methods.

Simulations to describe the distribution of individual observations
Posterior distributions of the population median (or mean) and tail percentiles (e.g., 5^{th} and 95^{th})

- Simulations to describe the posterior distribution of tail percentiles provide a more statistically formal approach for communicating the extent of variability.
- While you're at it, you may as well also simulate the population median or mean.
- The medians and credible intervals for these population estimands may be plotted together to depict both central tendency and variability of the observations in the population.

Simulations to describe the distribution of individual observations

Posterior distributions of the population median (or mean) and tail percentiles (e.g., 5^{th} and 95^{th})

The simulation algorithm is the same as that described for population mean simulations except that the final step is calculation of a sample percentile instead of a mean, i.e.,

for $m = 1$ to n_{post} **do**
  sample $\widehat{\theta}^{(m)}$ and $\Omega^{(m)}$ from the posterior distribution
  for $j = 1$ to n_{pop} **do**
    sample $\theta_j^{(m)} \sim N\left(\widehat{\theta}^{(m)}, \Omega^{(m)}\right)$
    calculate $\widehat{y}_{ij}^{(m)} = f\left(t_{ij}, D_{ij}, \text{drug}_{ij}, \theta_j^{(m)}\right)$
    sample $y_{ij}^{(m)} \sim N\left(\widehat{y}_{ij}^{(m)}, \sigma_{(m)}^2\right)$
  end for
  calculate the appropriate sample percentiles of $y_{ij}^{(m)}$
end for

Example: Population mean factor Xa inhibition

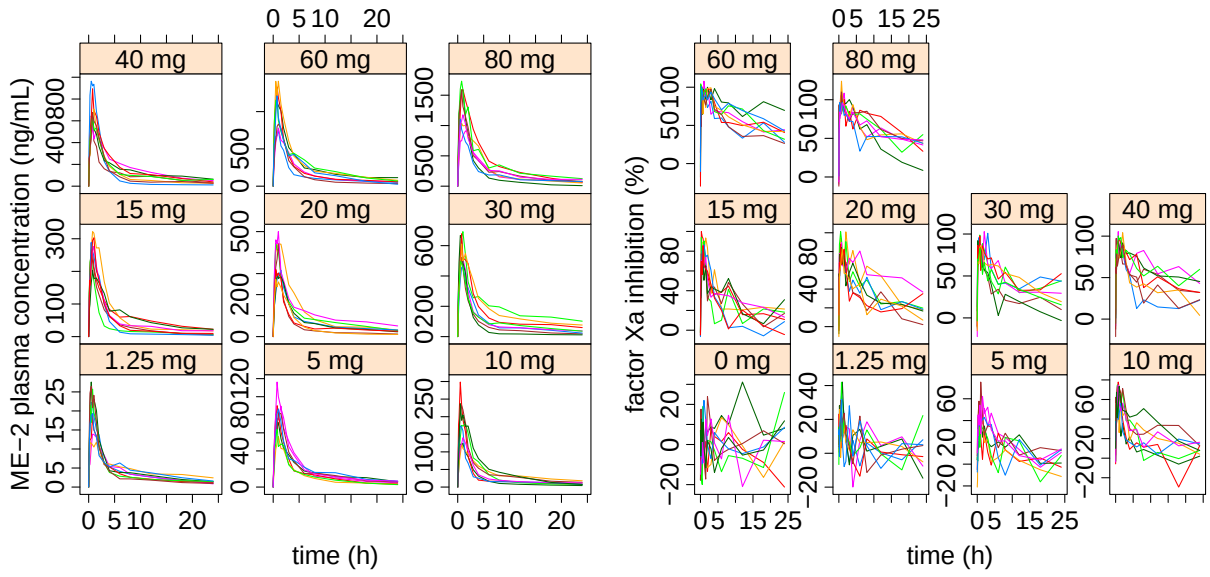
Use the MCMC samples from the popPK and popPD models for a hypothetical factor Xa inhibitor for the following:

- Simulate posterior median and 90% CI's for population mean factor Xa inhibition vs concentration (overall, not stratified by dose)
- Simulate posterior median and 90% CI's for population mean factor Xa inhibition vs time by dose
- Single doses: placebo, 1.25, 5, 10, 15, 20, 30, 40, 60 and 80 mg
- ME-2 concentration range = [0, 1600].
- Time range: [0, 24] hours.
- Summarize as plots showing posterior medians and 90% credible intervals (by time or concentration as appropriate).

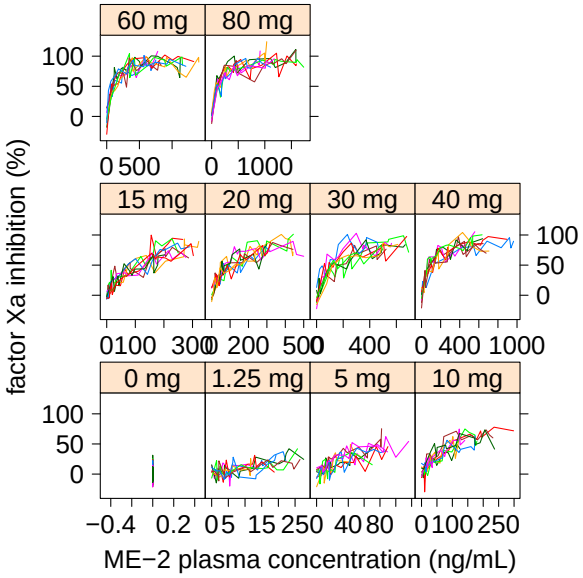
Population PK-PD modeling of time-matched biomarker and PK data

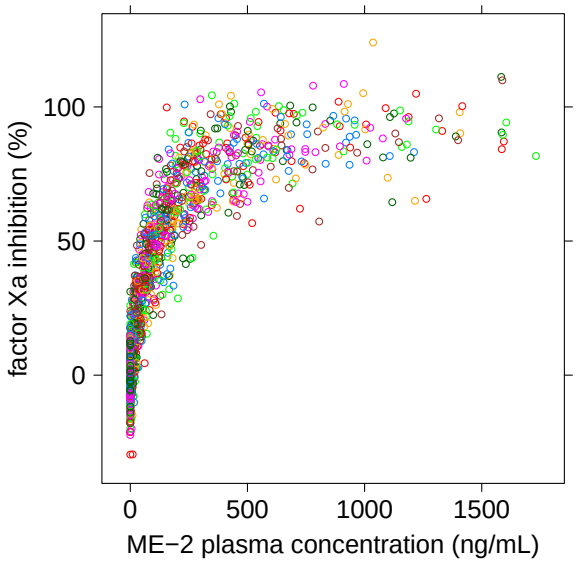
- Phase 1 single dose study in healthy volunteers
  - Parallel dose-escalation design
  - 8 subjects per dose arm
  - Single doses of ME-2
    - Placebo, 1.25, 5, 10, 15, 20, 30, 40, 60 and 80 mg
  - PK: plasma concentrations of parent drug
  - Biomarker: ex vivo inhibition of factor Xa activity in plasma
    - PK and biomarker measured at 0, 0.083, 0.167, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8, 12, 18 and 24 hours after dose.
- Modeling objective:
  - Apply a direct action PK/PD model to the time-matched factor Xa inhibition and ME-2 plasma concentrations.

EDA: PK and biomarker data



EDA: Relationship between biomarker and PK data





PD model description

- Sigmoid Emax model relating % inhibition of factor Xa activity to ME-2 plasma concentration on the i^{th} occasion in the j^{th} subject:

$$\begin{aligned} E_{ij} &\sim N\left(\hat{E}_{ij}, \sigma^2\right) \\ \hat{E}_{ij} &= \frac{E_{max} c_{ij}^{\gamma}}{EC_{50,j}^{\gamma} + c_{ij}^{\gamma}} \\ \log\left(EC_{50,j}\right) &\sim N\left(\log\left(\widehat{EC_{50}}\right), \omega_{EC_{50}}^2\right) \end{aligned}$$

- Some possible weakly informative prior distributions:

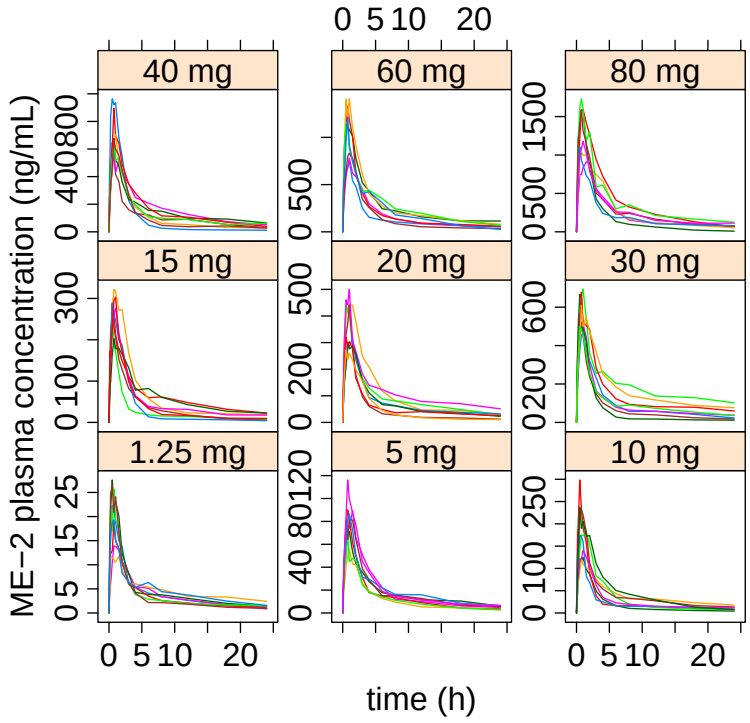
$$\begin{aligned} E_{max} &\sim U(0,100) \quad \log\left(\widehat{EC_{50}}\right) \sim N\left(0, 10^6\right) \\ \gamma &\sim U(0,10) \\ \omega_{EC_{50}} &\sim U\left(0, 10^5\right) \quad \sigma \sim U(0,1000) \end{aligned}$$

Population PK

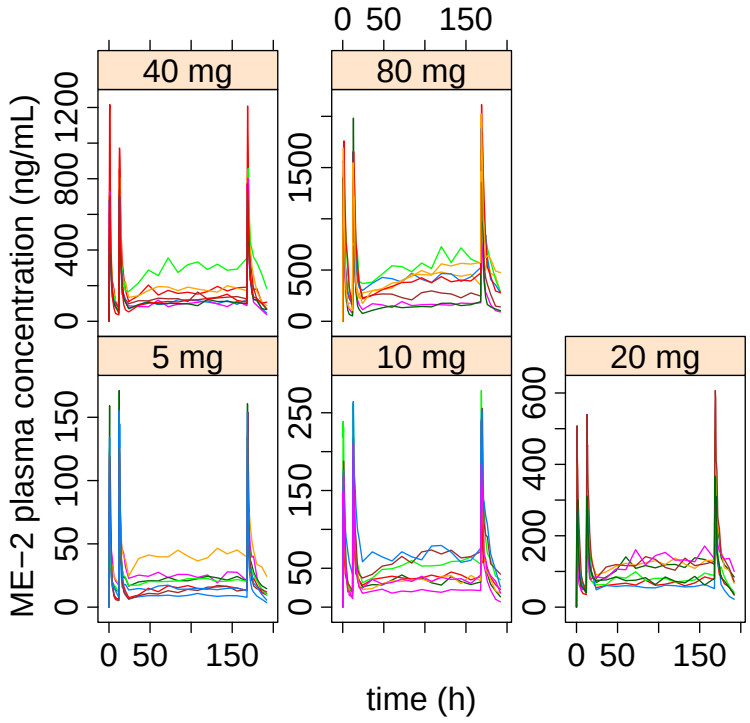
- A phase IIa PoC trial of ME-2 for prevention of post-op VTEs has just been completed.
- One of your tasks is to do a pop PK analysis based on the accumulated ME-2 PK data (Phase I SD (described above), Phase I MD & Phase IIa)
- Phase 1 multiple dose study in healthy volunteers
  - Parallel dose-escalation design
  - 8 subjects per dose arm
  - Placebo or ME-2 5, 10, 20, 40 or 80 mg bid (q12h) x 7 days
  - PK: plasma concentrations of parent drug
    - PK measured at 0, 0.083, 0.167, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8, 12, 12.1, 12.2, 12.2, 12.5, 12.8, 13, 13.5, 14, 15, 16, 18, 20, 24, 36, 48, 60, 72, 84, 96, 108, 120, 132, 144, 156, 168, 168, 168, 168, 168, 169, 169, 170, 170, 171, 172, 174, 176, 180, 186 and 192 hours after the first dose.

Population PK

- Phase IIa trial design:
  - Treatments
    - ME-2 20 mg bid (q12h) x 7 days
    - Enoxaparin 30 mg bid (q12h) x 7 days
  - 100 patients per treatment arm
  - Sparse ME-2 PK data (3-6 samples/patient)
    - LOQ = 10 ng/mL
- Available patient-specific covariates: weight, age, gender

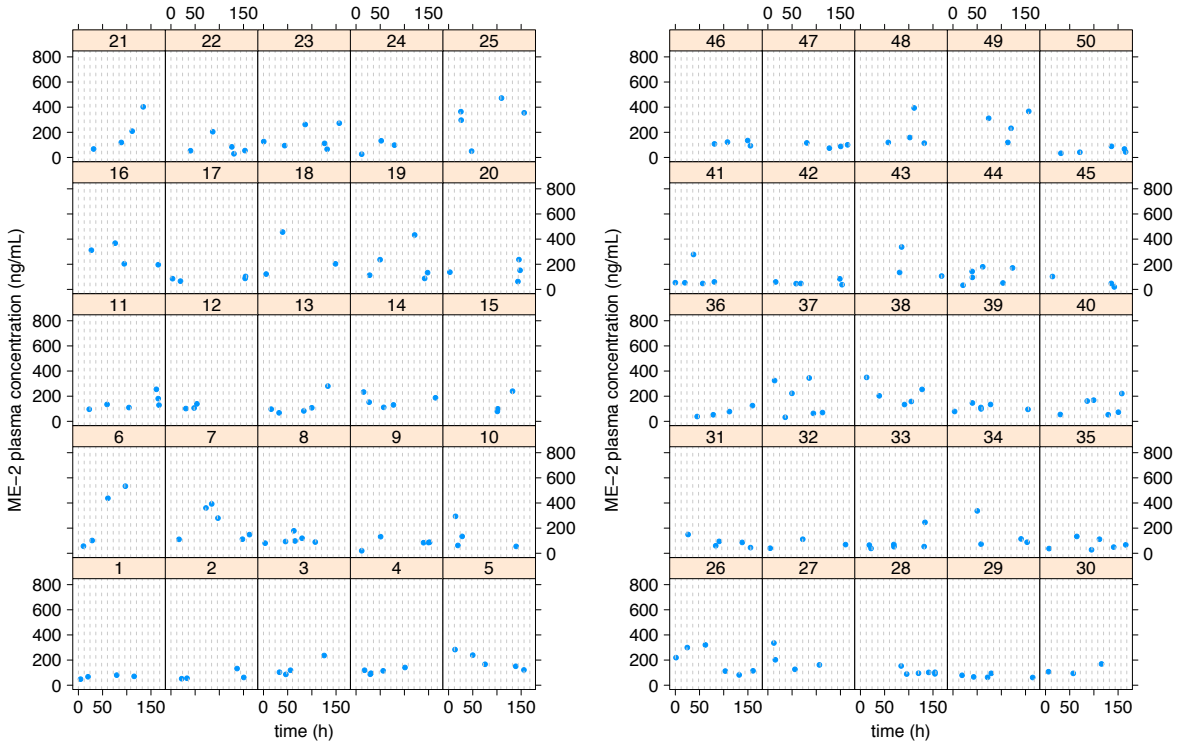


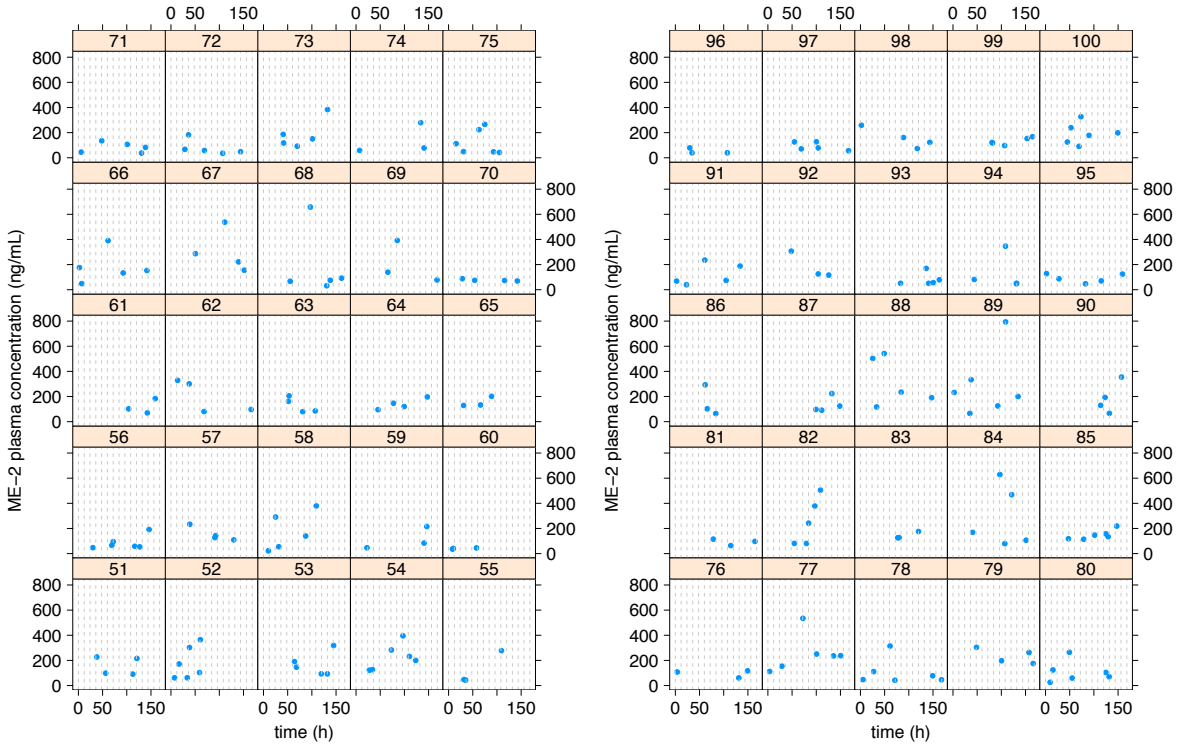
ME-2 PK data from Phase I SD trial



ME-2 PK data from Phase I MD trial

PK data from Phase IIa trial





PK model

- Two compartment model with first order absorption describing ME-2 plasma concentration on the i^{th} occasion in the j^{th} subject as a function of time, dose and body weight:

$$\begin{aligned}
\log\left(c_{ij}\right) &\sim N\left(\log\left(\hat{c}_{ij}\right), \sigma^2\right) \\
\hat{c}_{ij} &= f_{2cpt}\left(t_{ij}, D_j, \tau_j, CL_j, Q_j, V_{1j}, V_{2j}, k_{aj}\right) \\
\log\left(CL_j, Q_j, V_{1j}, V_{2j}, k_{aj}\right) & \\
&\sim N\left(\log\left(\widehat{CL}\left(\frac{bw_j}{70}\right)^{0.75}, \widehat{Q}\left(\frac{bw_j}{70}\right)^{0.75}, \widehat{V}_1\left(\frac{bw_j}{70}\right), \widehat{V}_2\left(\frac{bw_j}{70}\right), \widehat{k}_a\right), \Omega\right)
\end{aligned}$$

- Some possible weakly informative prior distributions:

$$\begin{aligned}
\log\left(\widehat{CL}\right) &\sim N\left(0, 10^6\right) \quad \log\left(\widehat{Q}\right) \sim N\left(0, 10^6\right) \quad \log\left(\widehat{V}_1\right) \sim N\left(0, 10^6\right) \\
\log\left(\widehat{V}_2\right) &\sim N\left(0, 10^6\right) \quad \log\left(\widehat{k}_a - \lambda_1\right) \sim N\left(0, 10^6\right) \quad \sigma \sim U\left(0, 100\right) \\
\Omega^{-1} &\sim \text{Wishart}\left(5\begin{pmatrix} 0.05 & 0 & 0 & 0 & 0 \\ 0 & 0.05 & 0 & 0 & 0 \\ 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0 & 0.05 \end{pmatrix}, 5\right)
\end{aligned}$$

Role of inter-trial variation

- In the previous sections I ignored the issue of inter-trial variation and how it should enter the posterior distributions and the simulations used to approximate them.
- This is, in part, because I still debate with myself about this. It comes down to what I think the inter-trial variation represents.
  - Do all the trials randomly sample from the target population of interest?
  - Or do they sample from a range of populations that may contain the population of interest to us, but is not necessarily restricted to that population?
- If the answer to the first question is yes then I would calculate the population estimand by averaging over the inter-trial variation.
- If the answer to the second question is yes then I would include the inter-trial variation in the posterior distribution of the model parameters. To date I have usually chosen to go with the second approach, in part because it is more conservative in the sense that the estimated uncertainty in the population estimand is greater.

Hands-on Problem 4: Population simulations

Simulations to compare dose-response of atorvastatin and rosuvastatin

- This exercise uses the MCMC samples generated in Hands-on Problem 3 to simulate the posterior distribution of population mean LDL % change from baseline predicted to result from various doses of atorvastatin and rosuvastatin.
- The primary objective of the simulations is to compare the dose-response of the two statins.

Hands-on Problem 4: Population simulations

Suggested simulation specifications

- Statin doses: c(0:4, seq(5, 80, by = 5)) mg/d
- Number of posterior samples: 1000 per dose combination
- Create plots showing posterior median and 90% credible intervals for the population mean LDL % change from baseline versus dose.

Issues arising from analysis of summary data
Applying individual models to summary data

- In MBMA it is common to apply models originally developed to describe responses in individuals to data consisting of summary statistics, particularly sample means.
- However our usual PK and PD models are strictly relevant only for describing responses in individual organisms—not for summary stats for groups.
- Nonlinear individual models do not "collapse" to the same model for sample means except in special cases, e.g., when the model function is linear with respect to individual-specific parameters.

Applying individual models to summary data

Example: One compartment PK model

- Consider a drug with PK following an iv bolus best described by a one compartment model.
- Each patients data is described by monoexponential function, but the mean concentration time-course of n patients data is described by a polyexponential with n exponential terms (unless their elimination rate constants are equal).
  - Concentration-time course in the i^{th} individual following a single dose D (neglecting residual variation to keep things simple):

$$c_i(t) = \frac{D}{V_i} e^{-k_i t}$$

  - Mean concentration-time course in n individuals:

$$\begin{aligned} \overline{c(t)} &= \frac{1}{n}\sum_{i=1}^{n} c_i(t) = \frac{1}{n}\sum_{i=1}^{n} \frac{D}{V_i} e^{-k_i t} \\ &\neq \frac{1}{n}\left(\sum_{i=1}^{n} \frac{D}{V_i}\right) e^{-kt} \\ \text{unless } k_i &= k \text{ for all individuals} \end{aligned}$$

Applying individual models to summary data

Example: Emax model

- Suppose the dose-response in an individual is described by an Emax model.
- The mean dose-response for n patients will not be an Emax model except in the special case where all patients share the same ED50.
  - Dose-response in the i^{th} individual (neglecting residual variation to keep things simple):

$$E_i(D) = \frac{E_{\max,i} D}{ED_{50,i} + D}$$

  - Mean dose-response in n individuals:

$$\begin{aligned} \overline{E(D)} &= \frac{1}{n}\sum_{i=1}^{n} E_i(D) = \frac{1}{n}\sum_{i=1}^{n} \frac{E_{\max,i} D}{ED_{50,i} + D} \\ &\neq \frac{\frac{1}{n}\left(\sum_{i=1}^{n} E_{\max,i}\right) D}{ED_{50} + D} \\ \text{unless } ED_{50,i} &= ED_{50} \text{ for all individuals} \end{aligned}$$

Applying individual models to summary data

- No easy remedy for the discrepancy between models fro individual and summary data.
- One recommendation arising from such discrepancies between model functions for individual and mean data is that mechanistic interpretations of model structure and parameter values should be approached cautiously.

Analysis of longitudinal data

- To some extent treatment arms may be viewed as "super-individuals."
- Like data from individuals within a population analysis, multiple observations within a treatment arm are correlated, and our analysis should account for that.
- This is further complicated by the fact that the treatment arms are not "created equal."
  - Sample sizes differ among the treatment arms, and some treatment arms may be from the same study and some from other studies.
- Relevant publications: [GRIG09, AF10, RPG+12, DF12]

Analysis of longitudinal data

- Sample size affects not only the residual variance but also the inter-arm variance in model parameters—the analog to IIV in population analysis.
- One approach for dealing with these issues is to use a hierarchical model with at least 3 levels of variation: inter-trial, inter-arm, and residual. In addition both the inter-arm and residual variances should be adjusted for sample size.
- The approach can be derived from population models for individual data.
  - In the special case where the individual data model is linear with respect to all random effects and the random effects are normally distributed, the derivation is exact.
  - For the general nonlinear case it is an approximation.

Linear case

Model for individual patient data

- Consider a model that is linear with respect to the inter-patient and residual random effects, and has normally-distributed residual, inter-patient and inter-study variation.
- Dependent variable y_{ijk} on the i^{th} occasion in the j^{th} patient in the k^{th} study:

$$\begin{aligned} y_{ijk} &\sim N\left(\widehat{y}_{ijk}, \sigma_k^2\right) \\ \widehat{y}_{ijk} &= f\left(t_{ijk}, x_{jk}, \theta, \mathbf{H}_{jk}, \mathbf{K}_k\right) \\ &= f_0\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) + \sum_{m=1}^{n_{\mathbf{H}}} f_m\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) \eta_{mjk} \end{aligned}$$

where

$$\begin{aligned} x_{jk} &\equiv \text{independent variables for the } j^{th} \text{ patient and } k^{th} \text{ study, e.g., assigned treatment} \\ \theta &\equiv \text{model parameters} \\ \mathbf{K}_k &= \left\{\kappa_{1k}, \kappa_{2k}, \cdots, \kappa_{n_{\mathbf{K}}k}\right\} = \text{inter-study random effects for } k^{th} \text{ study} \\ &\sim N(0, \Psi) \\ \mathbf{H}_{jk} &= \left\{\eta_{1jk}, \eta_{2jk}, \cdots, \eta_{n_{\mathbf{H}}jk}\right\} = \text{inter-patient random effects for } j^{th} \text{ patient in } k^{th} \text{ study} \\ &\sim N(0, \Omega) \end{aligned}$$

Linear case

Modifications for sample mean and variance

Since $y_{ijk}|\mathbf{H}_{jk}, \mathbf{K}_k \sim N\left(\widehat{y}_{ijk}, \sigma_k^2\right)$, the sample mean $\overline{y}_{ijk}$ on the i^{th} occasion in the j^{th} treatment arm in the k^{th} study is also normally distributed:

$$\overline{y}_{ijk} \sim N\left(\widehat{y}_{ijk}, \frac{\sigma_k^2}{n_{jk}}\right)$$

where

$$\begin{aligned} \widehat{y}_{ijk} &= f\left(t_{ijk}, x_{jk}, \theta, \overline{\mathbf{H}}_{jk}, \mathbf{K}_k\right) \\ &= f_0\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) + \sum_{m=1}^{n_{\mathbf{H}}} f_m\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) \overline{\eta}_{mjk} \\ \overline{\mathbf{H}}_{jk} &= \left\{\overline{\eta}_{1jk}, \overline{\eta}_{2jk}, \cdots, \overline{\eta}_{n_{\mathbf{H}}jk}\right\} = \text{inter-arm random effects for } j^{th} \text{ arm in } k^{th} \text{ study} \\ &\sim N\left(0, \frac{\Omega}{n_{jk}}\right) \end{aligned}$$

Linear case
Modifications for sample mean and variance

In this case $\widehat{y}_{ijk}$ represents the expected value conditioned on treatment arm j and study k. Similarly the distribution of y_{ijk} conditioned only on study k is normally distributed, i.e., $y_{ijk}|\mathbf{K}_k \sim N\left(f_0\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right), \sigma^2_{\text{marginal},ijk}\right)$. It follows that the normalized sample variance $\left(\left(n_{ijk}-1\right)s^2\left(y\right)_{ijk}\right) \Big/ \sigma^2_{\text{marginal},ijk}$ is $\chi^2\left(n_{ijk}-1\right)$ distributed or equivalently:

$$s^2\left(y\right)_{ijk} \sim \text{gamma}\left(\frac{n_{jk}-1}{2}, \frac{n_{jk}-1}{2\sigma^2_{\text{marginal},ijk}}\right)$$

where $\sigma^2_{\text{marginal},ijk}$ is the variance of y_{ijk} conditioned on study k.

Linear case
Modifications for sample mean and variance

An expression in terms of the model parameters is derived below:

$$\begin{aligned}
\sigma^2_{\text{marginal},ijk} &= \text{Var}\left(y_{ijk}|\mathbf{K}_k\right) \\
&= \text{Var}\left(\text{E}\left(y_{ijk}|\mathbf{H}_{jk}, \mathbf{K}_k\right)|\mathbf{K}_k\right) + \text{E}\left(\text{Var}\left(y_{ijk}|\mathbf{H}_{jk}, \mathbf{K}_k\right)|\mathbf{K}_k\right) \\
&= \sum_{\ell=1}^{n_{\mathbf{H}}}\sum_{m=1}^{n_{\mathbf{H}}} f_\ell\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) f_m\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right)\Omega_{\ell m} + \sigma^2_k
\end{aligned}$$

In the case where Ω is a diagonal matrix the equation simplifies to:

$$\sigma^2_{\text{marginal},ijk} = \sum_{m=1}^{n_{\mathbf{H}}} f_m\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right)^2 \omega^2_m + \sigma^2_k$$

where $\omega^2_m = \Omega_{mm} = \text{Var}\left(\eta_m\right)$.

Nonlinear case
Model for individual patient data

- Now consider the more general case where the model may be nonlinear with respect to the inter-patient and residual random effects, and has normally-distributed residual, inter-patient and inter-study variation.
- Dependent variable y_{ijk} on the i^{th} occasion in the j^{th} patient in the k^{th} study:

$$\begin{aligned} y_{ijk} &\sim N\left(\widehat{y}_{ijk}, \sigma_k^2\right) \\ \widehat{y}_{ijk} &= f\left(t_{ijk}, x_{jk}, \theta, \mathbf{H}_{jk}, \mathbf{K}_k\right) \end{aligned}$$

where

$$\begin{aligned} x_{jk} &\equiv \text{independent variables for the } j^{th} \text{ patient and } k^{th} \text{ study, e.g., assigned treatment} \\ \theta &\equiv \text{model parameters} \\ \mathbf{K}_k &= \left\{\kappa_{1k}, \kappa_{2k}, \cdots, \kappa_{n_{\mathbf{K}}k}\right\} = \text{inter-study random effects for } k^{th} \text{ study} \\ &\sim N(0, \Psi) \\ \mathbf{H}_{jk} &= \left\{\eta_{1jk}, \eta_{2jk}, \cdots, \eta_{n_{\mathbf{H}}jk}\right\} = \text{inter-patient random effects for } j^{th} \text{ patient in } k^{th} \text{ study} \\ &\sim N(0, \Omega) \end{aligned}$$

Nonlinear case
Modifications for sample mean and variance

Approximate equations for the sampling distributions of the sample means ($\overline{y}_{ijk}$) and variances ($s^2(y)_{ijk}$) are derived by first approximating the model using a first order Taylor series and then deriving the relationships as described above for the linear case. Begin by approximating $\widehat{y}_{ijk}$ for individual patients with a first order Taylor series where the inter-patient random effects (η's) are expanded about their expected values, i.e., 0:

$$\begin{aligned} \widehat{y}_{ijk} &= f\left(t_{ijk}, x_{jk}, \theta, \mathbf{H}_{jk}, \mathbf{K}_k\right) \\ &\approx \widehat{y}_{\text{approx},ijk} = f\left(t_{ijk}, x_{jk}, \theta, 0, \mathbf{K}_k\right) + \sum_{m=1}^{n_{\mathbf{H}}} f_{\eta_m}\left(t_{ijk}, x_{jk}, \theta, 0, \mathbf{K}_k\right) \eta_{mjk} \end{aligned}$$

where f_{η_m} is the derivative of f with respect to η_m.

Nonlinear case

Modifications for sample mean and variance

Proceeding as before the approximate equations used for fitting the sample means and variances follow.

$$\begin{aligned} \overline{y}_{ijk} &\sim N\left(\widehat{y}_{ijk}, \frac{\sigma_k^2}{n_{jk}}\right) \\ s^2(y)_{ijk} &\sim \text{gamma}\left(\frac{n_{jk}-1}{2}, \frac{n_{jk}-1}{2\sigma^2_{\text{marginal},ijk}}\right) \end{aligned}$$

where

$$\begin{aligned} \widehat{y}_{ijk} &= f\left(t_{ijk}, x_{jk}, \theta, \overline{\mathbf{H}}_{jk}, \mathbf{K}_k\right) \quad \overline{\mathbf{H}}_{jk} \sim N\left(0, \frac{\Omega}{n_{jk}}\right) \\ \sigma^2_{\text{marginal},ijk} &= \sum_{\ell=1}^{n_{\mathbf{H}}} \sum_{m=1}^{n_{\mathbf{H}}} f_{\eta_\ell}\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) f_{\eta_m}\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right) \Omega_{\ell m} + \sigma_k^2 \end{aligned}$$

In the case where Ω is a diagonal matrix the equation simplifies to:

$$\sigma^2_{\text{marginal},ijk} = \sum_{m=1}^{n_{\mathbf{H}}} f_{\eta_m}\left(t_{ijk}, x_{jk}, \theta, \mathbf{K}_k\right)^2 \omega_m^2 + \sigma_k^2$$

where $\omega_m^2 = \Omega_{mm} = \text{Var}(\eta_m)$.

Nonlinear case

Modifications for sample mean and variance

When the individual data model is *not* linear with respect to inter-patient random effects, the sampling distributions for treatment means and variances are approximated in 3 senses:

- The sampling distributions are approximated as normal for the mean and gamma for the variance.
- The conditional expectation for the treatment mean is approximated using the individual data model in which the variances of the inter-arm random effects are sample size adjusted inter-patient variances.
- The marginal variance is approximated via the delta method.

Additional research is required to assess the extent to which these approximations may adversely affect model-based inferences using this approach.

Example: MBMA of the ADAS-cog score time course in patients with Alzheimer's disease

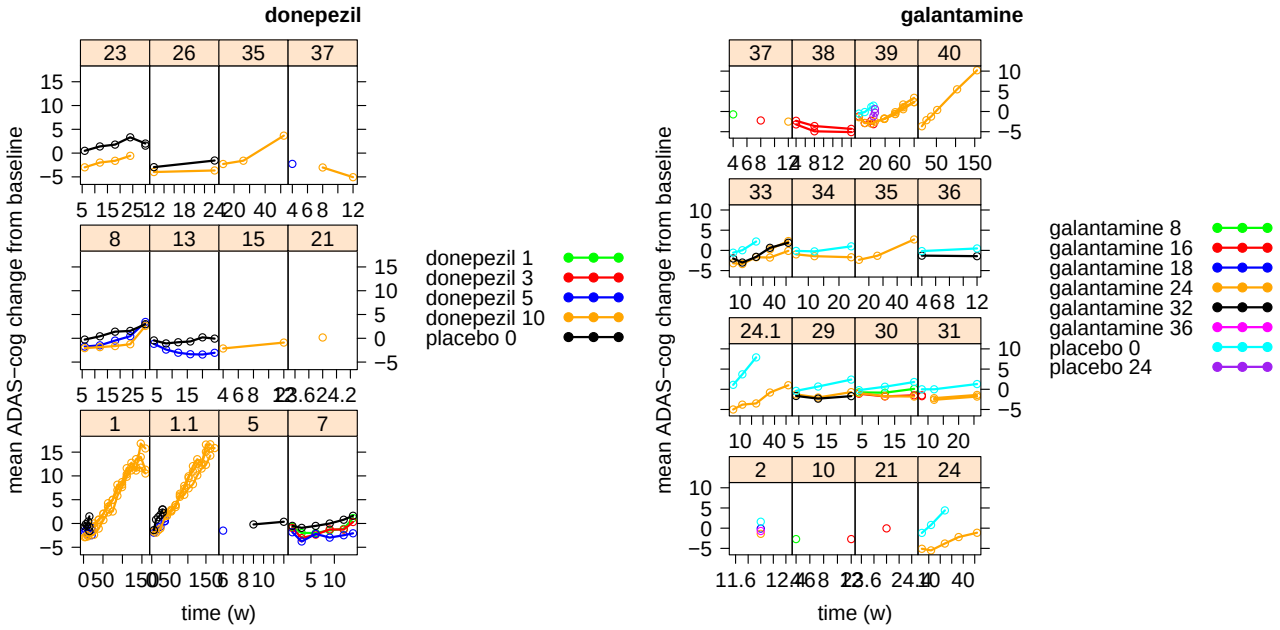
The scenario

- The imagined scenario is that you are involved in the development of one or more new potential drug treatments for Alzheimer's disease (AD).
- The ADAS-cog score is commonly used as a measure of AD symptoms in clinical trials.
- A model of the ADAS-cog time course during treatment with placebo or marketed drugs will be useful to support decision-making regarding trial designs, dose selection and possibly go/no-go in drug development programs.

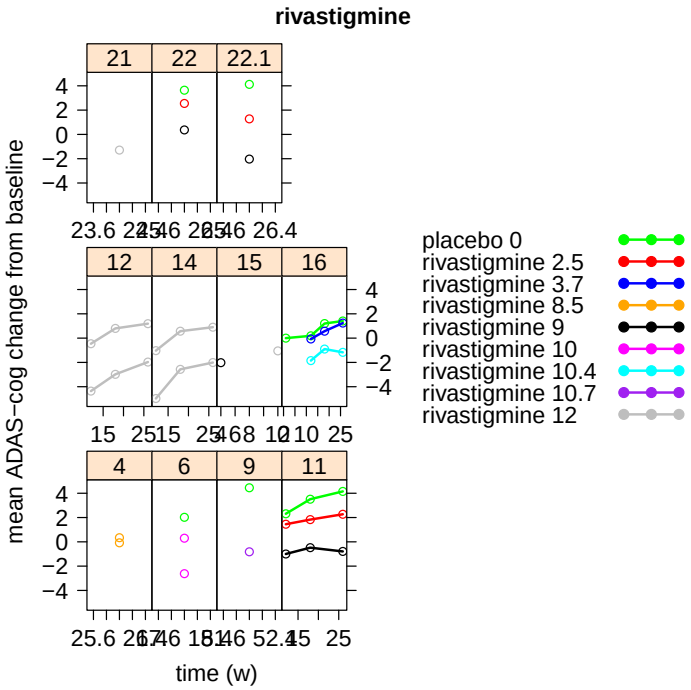
The data

- Post-baseline sample means and sample variances for ADAS-cog change from baseline from published sources
- Data set obtained from `http://OpenDiseaseModels.org`
  - Processed to reduce data management tasks
  - 55 studies
  - 114 treatment arms
    - 465 sample means
    - 263 sample variances
    - Data reflects results from 16543 patients and 68177 observations
- Objective
  - Construct a model for ADAS-cog change from baseline as a function of drug, daily dose, time and possibly other covariates.
- Data file: `adasCog/adasCogData.csv`

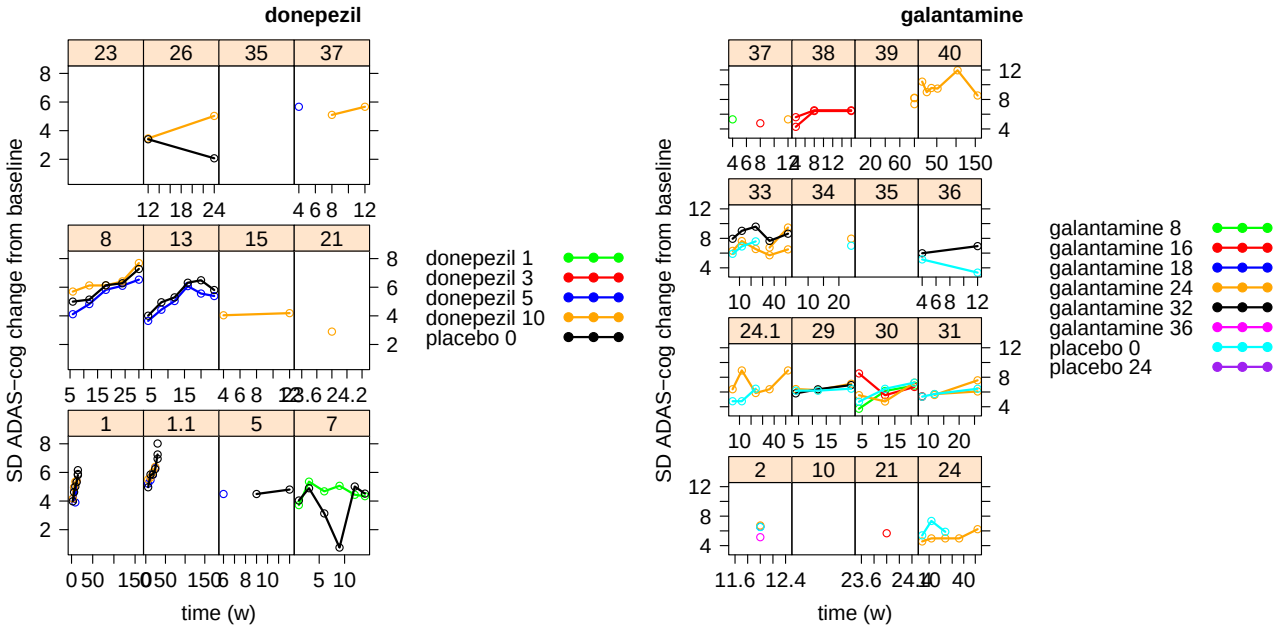
Mean ADAS-cog change from baseline



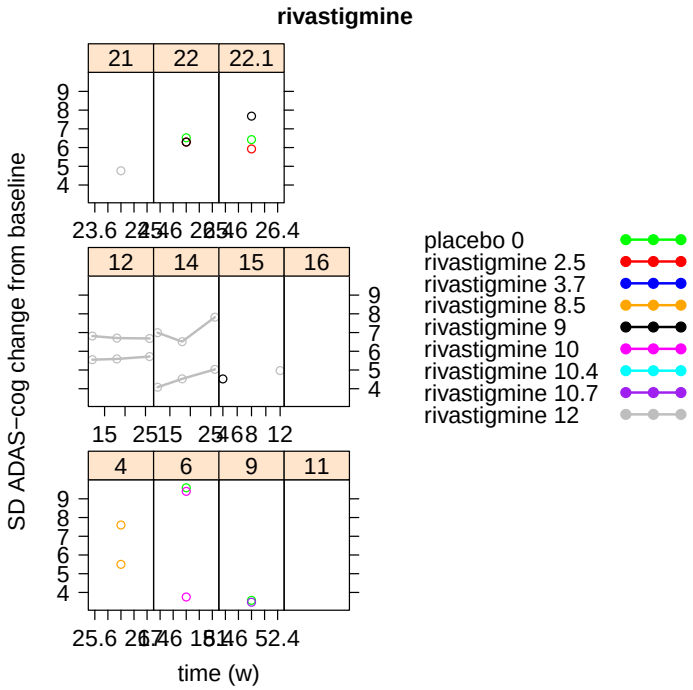
Mean ADAS-cog change from baseline



Standard deviation of ADAS-cog change from baseline



Standard deviation of ADAS-cog change from baseline



Model for individual patient data

We begin by conceptualizing the model in terms of individual patient data. Then we modify as necessary to model the sample statistics.

ADAS-cog change from baseline on the i^{th} occasion in the j^{th} patient in the k^{th} study:

$$\begin{aligned}
\Delta ADAS_{ijk} &\sim N\left(\widehat{\Delta ADAS}_{ijk}, \sigma_k^2\right) \\
\widehat{\Delta ADAS}_{ijk} &= \alpha_{jk} t_{ijk} + E_{\text{placebo,ijk}} + (1 - I_{\text{placebo,jk}}) E_{\text{drug,ijk}} + \eta_{\text{intercept},jk} \\
E_{\text{placebo,ijk}} &= \beta\left(e^{-k_{el} t_{ijk}} - e^{-k_{eq} t_{ijk}}\right) \\
E_{\text{drug,ijk}} &= \left(\frac{D_{jk}}{D_{ref,jk}}\right)^{\gamma_{jk}} \frac{E_{\Delta,jk} e^{\eta_{\text{drug},k}} t_{ijk}}{ET_{50,jk} + t_{ijk}}
\end{aligned}$$

Model for individual patient data

$$\begin{aligned}
D_{ref,jk} &= \begin{cases} 5, \\ 24, \\ 6 \end{cases} \quad E_{\Delta,jk} = \begin{cases} E_{\Delta,\text{donepezil}}, & \text{drug}_{jk} = \text{donepezil} \\ E_{\Delta,\text{galantamine}}, & \text{drug}_{jk} = \text{galantamine} \\ E_{\Delta,\text{rivastigmine}}, & \text{drug}_{jk} = \text{rivastigmine} \end{cases} \\
ET_{50,jk} &= \begin{cases} ET_{50,\text{donepezil}}, \\ ET_{50,\text{galantamine}}, \\ ET_{50,\text{rivastigmine}} \end{cases} \quad \gamma_{jk} = \begin{cases} \gamma_{\text{donepezil}}, & \text{drug}_{jk} = \text{donepezil} \\ \gamma_{\text{galantamine}}, & \text{drug}_{jk} = \text{galantamine} \\ \gamma_{\text{rivastigmine}}, & \text{drug}_{jk} = \text{rivastigmine} \end{cases} \\
\eta_{\text{intercept},jk} &\sim t\left(\eta_{\text{intercept,study},k}, \omega^2_{\text{intercept}}, df_{\text{intercept}}\right) \\
\eta_{\text{intercept,study},k} &\sim N\left(0, \psi^2_{\text{intercept}}\right) \\
\alpha_{jk} &\sim t\left(\alpha_{\text{study},k}, \omega^2_{\alpha}, df_{\alpha}\right) \quad \alpha_{\text{study},k} \sim N\left(\widehat{\alpha}, \psi^2_{\alpha}\right) \\
\eta_{\text{drug},k} &\sim N\left(0, \psi^2_{\text{drug}}\right) \quad \frac{1}{\sigma_k^2} \sim \text{gamma}\left(\alpha_\sigma, \alpha_\sigma \widehat{\sigma}^2\right)
\end{aligned}$$

Model for sample means and variances

Modifications for sample mean and variance of ADAS-cog change from baseline on the i^{th} occasion in the j^{th} treatment arm in the k^{th} study:

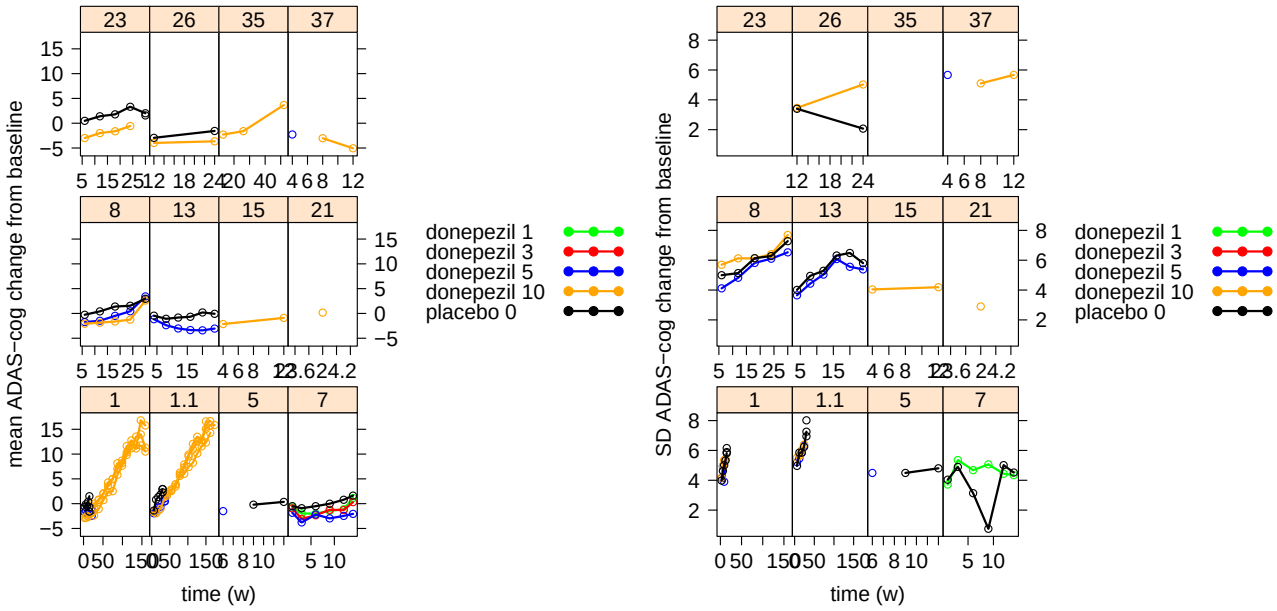
$$
\begin{aligned}
\overline{\Delta ADAS}_{ijk} &\sim N\left(\widehat{\Delta ADAS}_{ijk}, \frac{\sigma_k^2}{n_{jk}}\right) \\
s^2\left(\Delta ADAS\right)_{ijk} &\sim \text{gamma}\left(\frac{n_{jk}-1}{2}, \frac{n_{jk}-1}{2\sigma^2_{\text{marginal},ijk}}\right) \\
\sigma^2_{\text{marginal},ijk} &\approx t_{ijk}^2\omega_\alpha^2 + \omega^2_{\text{intercept}} + \sigma_k^2 \\
\eta_{\text{intercept},jk} &\sim t\left(\eta_{\text{intercept,study},k}, \frac{\omega^2_{\text{intercept}}}{n_{jk}}, df_{\text{intercept}}\right) \\
\alpha_{jk} &\sim t\left(\alpha_{\text{study},k}, \frac{\omega_\alpha^2}{n_{jk}}, df_\alpha\right)
\end{aligned}
$$

Hands-on Problem 5: Longitudinal dose-response model based on longitudinal summary data

MBMA of the ADAS-cog as a function of time and donepezil dose

- Post-baseline sample means and sample variances for ADAS-cog change from baseline following various doses of donepezil from published sources
- Data set obtained from `http://OpenDiseaseModels.org`
  - Processed to reduce data management tasks
  - 12 studies
  - 24 treatment arms
    - 183 sample means
    - 82 sample variances
    - Data reflects results from 2581 patients
- Hands-on exercise
  - Construct a model for ADAS-cog change from baseline as a function of donepezil daily dose and time.
- Data file: `adasCogHandsOn/donepezilAdasData.csv`

Hands-on Problem 5: Mean and SD of ADAS-cog change from baseline



Hands-on Problem 5: Proposed model

Model for individual patient data

ADAS-cog change from baseline on the i^{th} occasion in the j^{th} patient in the k^{th} study:

$$
\begin{aligned}
\Delta ADAS_{ijk} &\sim N\left(\widehat{\Delta ADAS}_{ijk}, \sigma_k^2\right) \\
\widehat{\Delta ADAS}_{ijk} &= \alpha_{jk} t_{ijk} + E_{\text{placebo},jk} + \left(1 - I_{\text{placebo},jk}\right) E_{\text{drug},ijk} \\
E_{\text{drug},ijk} &= \left(\frac{D_{jk}}{D_{ref}}\right)^{\gamma} \frac{E_{\Delta,k}\, t_{ijk}}{ET_{50} + t_{ijk}} \\
D_{ref} &= 5 \\
E_{\Delta,k} &\sim N\left(\widehat{E}_{\Delta}, \psi^2_{E_\Delta}\right) \\
E_{\text{placebo},jk} &\sim N\left(E_{\text{placebo,study},k}, \omega^2_{E_{\text{placebo}}}\right) \\
E_{\text{placebo,study},k} &\sim N\left(\widehat{E}_{\text{placebo}}, \psi^2_{E_{\text{placebo}}}\right) \\
\alpha_{jk} &\sim N\left(\alpha_{\text{study},k}, \omega^2_{\alpha}\right), \quad \alpha_{\text{study},k} \sim N\left(\widehat{\alpha}, \psi^2_{\alpha}\right) \\
\frac{1}{\sigma_k^2} &\sim \text{gamma}\left(\alpha_\sigma, \alpha_\sigma \widehat{\sigma}^2\right)
\end{aligned}
$$

Hands-on Problem 5: Model for sample means and variances

Modifications for sample mean and variance of ADAS-cog change from baseline on the i^{th} occasion in the j^{th} treatment arm in the k^{th} study:

$$\begin{aligned}
\overline{\Delta ADAS}_{ijk} &\sim N\left(\widehat{\Delta ADAS}_{ijk}, \frac{\sigma_k^2}{n_{jk}}\right) \\
s^2\left(\Delta ADAS\right)_{ijk} &\sim \text{gamma}\left(\frac{n_{jk}-1}{2}, \frac{n_{jk}-1}{2\sigma^2_{\text{marginal},ijk}}\right) \\
\sigma^2_{\text{marginal},ijk} &\approx t_{ijk}^2\omega_\alpha^2 + \omega^2_{E_{\text{placebo}}} + \sigma_k^2 \\
E_{\text{placebo},jk} &\sim N\left(E_{\text{placebo,study},k}, \frac{\omega^2_{E_{\text{placebo}}}}{n_{jk}}\right) \\
\alpha_{jk} &\sim N\left(\alpha_{\text{study},k}, \frac{\omega_\alpha^2}{n_{jk}}\right)
\end{aligned}$$

Modeling other types of summary statistics

Number or fraction of patients that experience a particular outcome or event

- Binary outcomes in individual patients are often summarized in terms of the number or fraction of patients with one of the two possible outcomes.
- Such binary outcomes could include the occurrence of some event or an indicator variable for whether or not the patient benefited from treatment. Examples include:
  - Discrete events
    - Death
    - MI
    - DVT
    - AE
    - Dropout
  - Other outcomes
    - Responder or remission status, e.g., achieving a specified degree of improvement

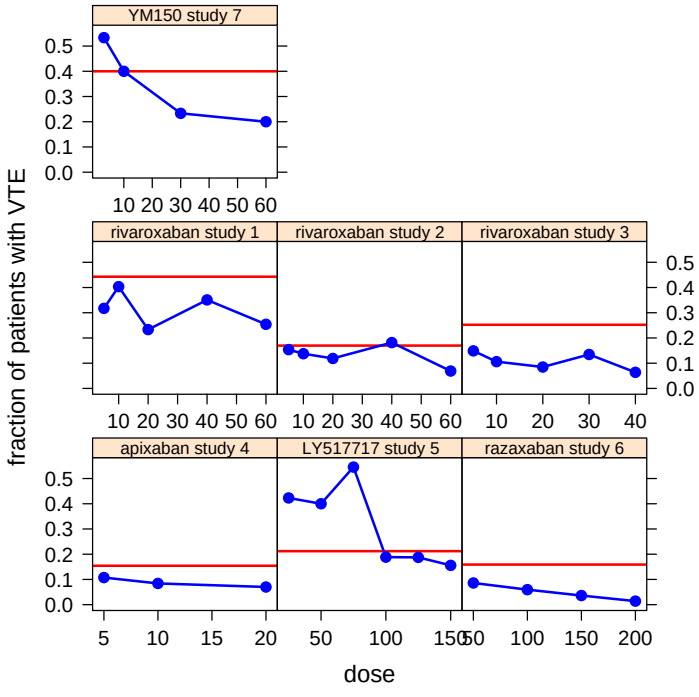
Modeling other types of summary statistics

Number or fraction of patients that experience a particular outcome or event

- The appropriate likelihood function for the number of patients with a particular outcome is the binomial distribution.
- This is parametrized in terms of the probability of an outcome (p) in an individual and the total number of individuals (n).
- The probability p is typically a function of independent variables like daily dose, time and mean demographic variables, and it must be restricted to (0, 1). Beyond that it is difficult to generalize about the specific function for p.
- One approach would be some variation on logistic regression, i.e., describing p using a logit transformation.

Example: Effects of factor Xa inhibitors on incidence of post-op VTE's

- Meta-analysis of factor Xa inhibitor effect on VTE incidence
- Published VTE rates from 7 dose-finding trials comparing new factor Xa inhibitors to enoxaparin. Horizontal lines show observed values for enoxaparin 40-60 mg/d.



Proposed model

- Linear logistic regression model for VTE occurrence in the i^{th} treatment arm of the j^{th} study as a function of dose:

$$\begin{aligned} n_{VTE,ij} &\sim \text{Binomial}\left(p_{VTE,ij}, n_{ij}\right) \\ \text{logit}\left(p_{VTE,ij}\right) &= E_{p,j} + E_{\text{drug},ij} \\ E_{\text{drug},ij} &= \theta_{\text{drug}} \log\left(D_{ij} + 1\right) \\ E_{p,j} &\sim N\left(\theta_p, \sigma^2\right) \end{aligned}$$

  where $n_{VTE,ij}$ is the number of patients experiencing a VTE and n_{ij} is the total number of patients in the in the i^{th} treatment arm of the j^{th} study, respectively.
- Weakly informative priors for drug effects:

$$\theta_{\text{drug}} \sim N\left(0, 10^6\right)$$

- Informative priors for placebo effects (fictional but it illustrates the idea of using historical knowledge about control treatment response):

$$\theta_p \sim N\left(\text{logit}\left(0.45\right), 0.1^2\right) \quad \log\left(\sigma\right) \sim N\left(\log\left(0.4\right), 0.5^2\right)$$

WinBUGS implementation

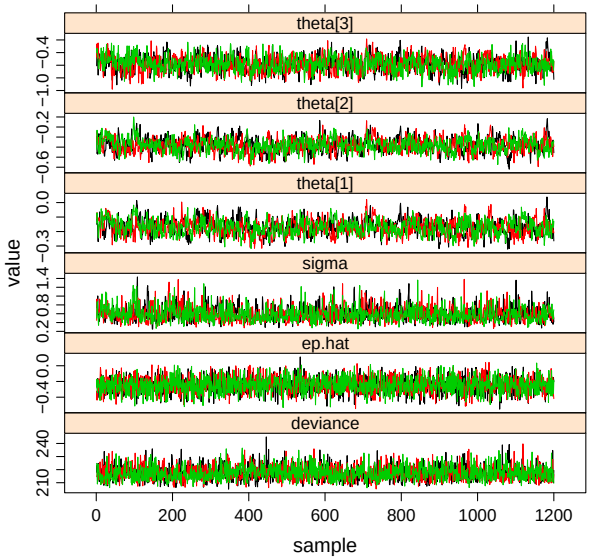
```
model{

  for(i in 1:nobs){
    ## likelihood
    vte[i] ~ dbin(p.vte[i],n[i])
    ## posterior prediction for new observations in same study
    vte.cond[i] ~ dbin(p.vte[i],n[i])
    logit(p.vte[i]) <- ep[study[i]] + edrug[i]
    ## +1 added so that intercept still corresponds to dose = 0
    edrug[i] <- theta[drug[i]]*log(dose[i]+1)

    ## posterior predictions for new observations
    ## in new study of same design
    vte.pred[i] ~ dbin(p.vte.pred[i],n[i])
    logit(p.vte.pred[i]) <- ep.pred[study[i]] + edrug[i]
  }

  for(i in 1:nstudy){
    ## interstudy variation in the intercept
    ep[i] ~ dnorm(ep.hat,tau)
    ep.pred[i] ~ dnorm(ep.hat,tau)
  }
```

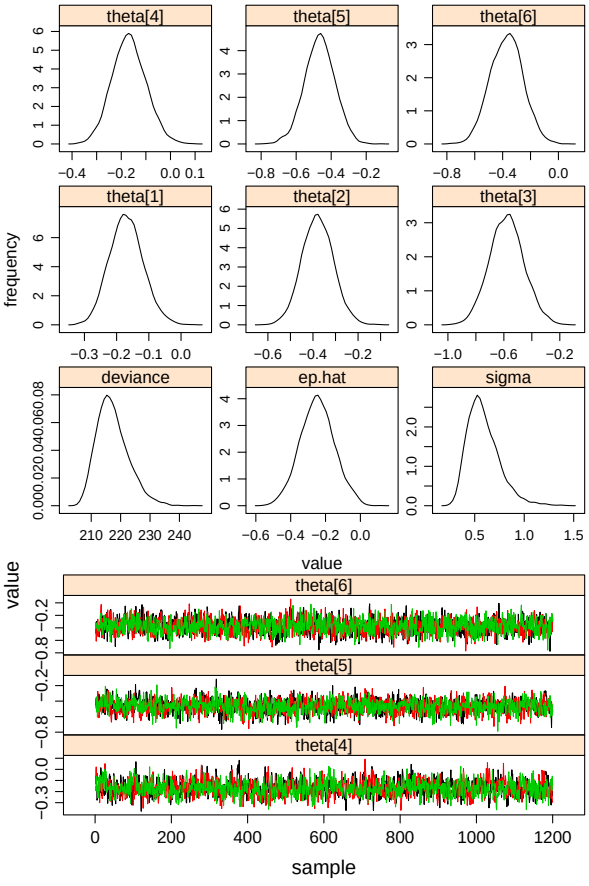
WinBUGS implementation (cont.)

```
  ## informative prior on placebo response
  ep.hat ~ dnorm(ep.hat.prior.mean, ep.hat.prior.precision)
  ep.hat.prior.mean <- logit(0.45)
  ep.hat.prior.precision <- 1/pow(0.1,2)
  log.sigma ~ dnorm(log.sigma.prior.mean, log.sigma.prior.precision)
  log.sigma.prior.mean <- log(0.4)
  log.sigma.prior.precision <- 1/pow(0.5,2)

  log(sigma) <- log.sigma
  tau <- 1/(sigma*sigma)
  for(i in 1:6){
    theta[i] ~ dnorm(0,1.0E-6)
  }

}
```

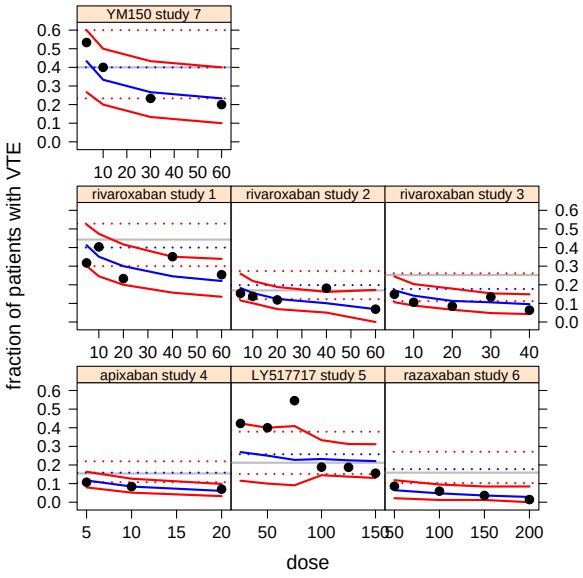

Results

- 10000 × 3 chains
- burn-in = 4000 / chain, thin by 5

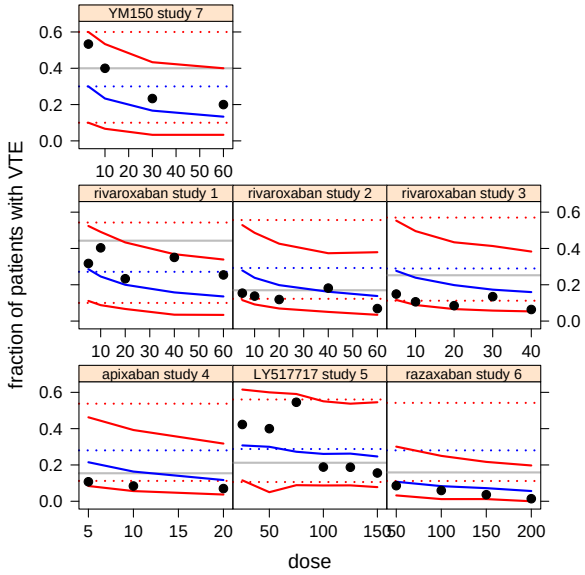




Prediction of new data in same studies



Prediction of new data in new studies of the same design



Horizontal lines show observed (gray) and predicted values for enoxaparin 40-60 mg/d

Model parameter estimates

parameter	mean	sd	2.5%ile	25%ile	median	75%ile	97.5%ile	effective N
deviance	217	5.37	209	214	217	220	230	2420
θ_p	−0.247	0.0976	−0.44	−0.313	−0.248	−0.183	−0.0509	3450
σ	0.59	0.164	0.351	0.472	0.564	0.682	0.984	1710
θ_1	−0.17	0.052	−0.267	−0.205	−0.171	−0.136	−0.0629	708
θ_2	−0.379	0.0693	−0.514	−0.426	−0.38	−0.332	−0.241	809
θ_3	−0.579	0.123	−0.817	−0.662	−0.579	−0.498	−0.331	1130
θ_4	−0.168	0.0689	−0.3	−0.215	−0.17	−0.122	−0.031	1430
θ_5	−0.466	0.0861	−0.64	−0.523	−0.465	−0.409	−0.299	1840
θ_6	−0.371	0.117	−0.597	−0.454	−0.369	−0.293	−0.145	2290

Modeling other types of summary statistics
Number or fraction of patients within each level of an ordinal scale

- Ordinal variables are often used to measure symptom severity or treatment response.
  - Frequently the results of such measurements are reported as means.
  - Occasionally they may be reported in terms of the fraction of patients within each level of the ordinal scale.
- One option for dealing with such data is to lump the levels together to produce only 2 levels as was done in the triptan example. Then you can apply the approach described in the previous section.
- If you prefer to retain all levels of the ordinal scale then the numbers of patients within each level for a particular treatment arm and time may be modeled with a multinomial likelihood.
  - One approach to modeling the required probability vector is to use a cumulative logit model [SBD97, LWRP01].

Modeling other types of summary statistics
Mean number of events per patient

- Another type of outcome is the number of events within a specified period of time.
  - Examples include number of seizures, number of emesis events, number of adverse events, etc.
- Most often such data are summarized as the mean number of events over either the total study period or for the periods between study visits.
- When modeling individual patient count data the Poisson distribution or some over-disperse or zero-inflated extension of the Poisson distribution is often used [God07].

Modeling other types of summary statistics
Mean number of events per patient

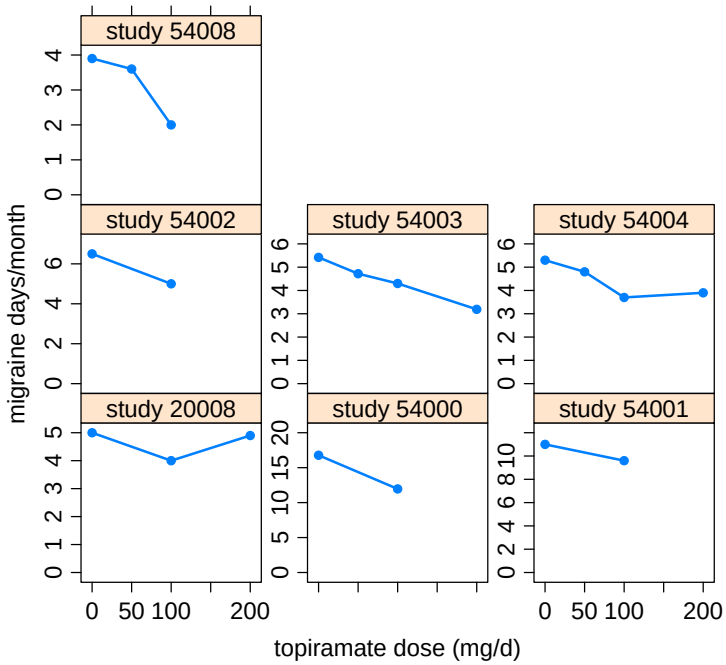
- One approach to modeling of mean count data is to rely on the central limit theorem and use a normal likelihood.
  - If it is reasonable to assume that the hazard for an event is approximately constant over time, then the mean should be proportional to the duration of the observation period.
  - The variance should also be proportional to that duration and inversely proportional to the sample size.
- Alternatively you could take advantage of the addition property of the Poisson distribution, i.e., the sum of Poisson distributed random variables is also Poisson distributed. In particular if $\theta_i \sim \text{Poisson}(\lambda_i)$ for all i, then $\sum_{i=1}^{n} \theta_i \sim \text{Poisson}\left(\sum_{i=1}^{n} \lambda_i\right)$.
  - It follows that if the individual count data are Poisson distributed, then the sum of those counts is also Poisson distributed.
  - That sum can be calculated by multiplying the mean count by the number of patients.

Example: Effect of topiramate on mean number of migraine days per month

Meta-analysis of topiramate's effect on migraine frequency.

Published number of migraine days/month from 7 migraine prophylaxis trials in which one or more doses of topiramate was compared to placebo.

A migraine day is defined as any calendar day during which a patient had a migraine headache of at least 30-min duration.



Proposed initial model

Poisson model for total number of migraine days in the i^{th} treatment arm of the j^{th} study:

$$\begin{aligned} n_{\text{migraine},ij} &\sim \text{Poisson}\left(n_{ij}\lambda_{ij}\Delta t_{ij}\right) \\ \log\left(\lambda_{ij}\right) &= a_{ij} + b_{ij}D_{ij} \\ a_{ij} &= a_{1j} + a_2 \log\left(\frac{\lambda_{0ij}}{6}\right) \\ b_{ij} &= b_{1j} + b_2 \log\left(\frac{\lambda_{0ij}}{6}\right) \\ a_{1j} &\sim N\left(\overline{a}, \omega_a\right) \quad b_{1j} \sim N\left(\overline{b}, \omega_b\right) \end{aligned}$$

where λ_{ij} is the model estimated number of migraine days/month for an individual, λ_{0ij} is the mean baseline migraine days/month, Δt_{ij} is the elapsed time over which the migraine events were counted and D_{ij} is the topiramate daily dose.

Suggested homework

- Improve upon the non-optimal preliminary model while retaining the strategy of modeling event counts as discrete random variables.
- Implement a model using a normal likelihood.
- Compare the results.

Modeling other types of summary statistics
Summary statistics for time-to-event measurements

- The modeling of time-to-event measurements in individual patients typically has to deal with the complication of right censoring, i.e., the case where no event occurs prior to the end of the study.
- Unless an event occurs in all (or at least a large majority) of the patients, a mean event time is not a very good statistic for meta-analysis.

Modeling other types of summary statistics
Summary statistics for time-to-event measurements

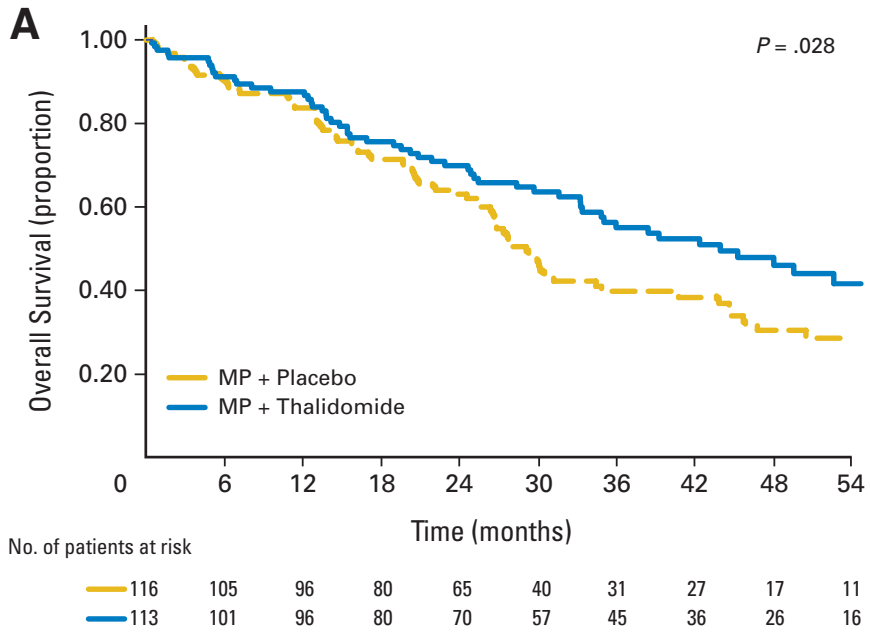
- Some publications may present the results in the form of a Kaplan-Meier plot that depicts the fraction of patients that have not yet experienced an event. You could tabulate values from the plot and use those values to construct a model.
- However, it is likely that most of the publications contributing to your data base will not provide such plots.
- More likely the data will be summarized in the form of the fraction or number of patients that experienced an event during the study period.
  - In that case the approach for binary data described previously is appropriate though the probability of an event must be adjusted for the trial duration given some assumption about the hazard time course.
- For cases when survival curves are available, modeling approaches have been proposed in the statistical literature ([PTS98, AHS08, Dea94] and [SAJ$^+$00, pp 277–283]) and could be adapted for a Bayesian approach.

Meta-analysis of longitudinal time-to-event data

Analysis of time-to-event data

- The gold-standard for meta-analysis of time-to-event data is using IPD
- When only AD is available, analysis of time-to-event data is typically done using
  - (Log) Hazard ratios if using comparative studies
  - Median endpoint values (median OS, median PFS, etc.)
  - Estimated survivor rates at specific times (% surviving at 1 year, at 2 years, etc.)
- The first approach requires all studies to be comparative—eliminating information in the single-arm trials
- The other two approaches consider each outcome separately
  - E.g., separate models for % surviving at 1 year and % surviving at 2 years
- While these approaches are fairly easy to implement, they may be under-utilizing the available data
  - E.g., If not all studies report the same time points
  - This leads us to consider jointly modeling survivor rates

Example Kaplan-Meier curves



C. Hulin et al. J Clin Oncol. 2009. 27(22): 3664-70.

Two related approaches have been described in the literature

- Models for the observed survival rates
  - Dear, KBG. (1994). Iterative Generalized Least Squares for Meta-Analysis of Survival Data at Multiple Times. Biometrics. 50: 989–1002.
  - Arends LR, Hunink MGM and Stijnen T. (2008). Meta-analysis of summary survival curve data. Statistics in Medicine 27: 4381–96.
- Models for the number of events between time points
  - Ouwens MJMN, Philips Z, and Jansen JP. (2010). Network meta-analysis of parametric survival curves. Res Syn Meth. 1: 258–71.
  - Jansen JP. (2011). Network meta-analysis of survival data with fractional polynomials. BMC. Med Res Methodol 11:61.
- For this course, we'll
  - Outline the key elements/assumptions
  - Work through an example of each

Fixed effects model of Dear

Let s_{ijk} be the true survival probability in treatment arm j in study i at the k^{th} time point.

We assume that we have estimates, $\widehat{s}_{ijk}$, along with their standard errors.

Dear fitted a linear regression model relating $\widehat{s}_{ijk}$ to between- and within-study covariates, such as time, study, treatment characteristics and their interactions. Specifically, the model is

$$\widehat{s}_i = X_i\theta + \epsilon_i$$

where $\widehat{s}_i$ is a vector of $\widehat{s}_{ijk}$ and ϵ_i is a vector of residuals with

$$\epsilon_i \sim N(0, V_i).$$

The covariance matrix has a fixed structure

The key to Dear's model is recognizing that V_i will have a fixed structure that depends only on the reported standard errors and the survival probabilities.

Specifically,

- V_i is block diagonal with blocks corresponding to treatment arms within the study.
- The main diagonal is set to the reported squared standard errors.
- The correlations of the proportions at times $t_{ijk} > t_{ijl}$ are

$$\text{corr}\left(s_{ijk}, s_{ijl}\right) = \sqrt{\frac{s_{ijk}\left(1 - s_{ijl}\right)}{\left(1 - s_{ijk}\right)s_{ijl}}}$$

Estimation is done iteratively

1. Start by using the observed survival proportions to approximate V_i.
2. Given V_i, use generalized least squares to estimate θ.
3. Using this estimate of θ, compute the model-based estimates of the s_{ijk}'s and plug into equation for V_i.
4. Iterate until convergence.

Key differences between Arends et al. and Dear models

- Arends et al. use a linear mixed effects model in contrast to a fixed effects model.
- They use a transformation of the s_{ijk} to improve linearity and interpretability.
- They focus on using parametric models for the effects of time (and presumably covariates).

Mixed effects model of Arends

Arends et al. propose the following model:

$$\log\left(-\log\left(\widehat{s}_i\right)\right) = X_i\theta + Z_i\eta_i + \epsilon_i$$

where

$$\begin{aligned}\eta_i &\sim N(0,\Omega)\\ \epsilon_i &\sim N(0, V_i)\end{aligned}$$

and V_i is block diagonal with elements

$$\frac{se_{ijk}}{\widehat{s}_{ijk}\log\left(\widehat{s}_{ijk}\right)}\sqrt{\frac{s_{ijk}\left(1-s_{ijl}\right)}{\left(1-s_{ijk}\right)s_{ijl}}}\frac{se_{ijl}}{\widehat{s}_{ijl}\log\left(\widehat{s}_{ijl}\right)}$$

Estimation is done iteratively as before but using mixed effects software.

Fitting the Arends model using standard software

- We need to have software that can
  - Fix the residual error variance
  - Output the fitted survival estimates for updating the Vi matrix.
- This can be done in SAS PROC MIXED and in S-Plus function lme
  - Not possible in the R function lme (because we can't fix the residual error variance)
  - Could conceivably be done in a combination of R (or some other programming language) and NONMEM, but not NONMEM alone
- We'll fit a related version of this model using WinBUGS

Obtaining the standard error of KM estimates

- Both of these approaches require having the standard error of the survival probabilities
- Sometimes (okay, rarely) they will be reported in the paper
- A bit more often a confidence interval may be included with the KM plot.
  - Can digitize the figure, read the CI at each time point of interest and derive the standard errors.. ugh.
  - E.g., if a 95% CI then the approximate standard error is (upper limit $-$ lower limit) / $(2 \cdot 1.96)$.
- If neither of these are available, youll need to do some additional approximation based on the survivor curve and the number of patients at risk.

The standard error of the estimated survival proportion with no censoring

If there is no censoring, it is simple to show that

$$\text{Var}\left(\widehat{S}(t)\right) = \frac{\widehat{S}(t)\left(1-\widehat{S}(t)\right)}{N}$$

where N is the number of patients in the treatment group at the start of the study.

This could be a reasonable estimate when there is a small amount of censoring.

However, when there is a moderate to large degree of censoring, this will underestimate Var $\left(\widehat{S}(t)\right)$. Thus, we might consider a different estimate in this case.

An approximation to the standard error of the estimated survival proportion

From first principles, we know that $\widehat{S}(t) = \exp\left(\widehat{\Lambda}(t)\right)$. Using the Delta method approximation, we have

$$\text{Var}\left(\widehat{S}(t)\right) \approx \left(\widehat{S}(t)\right)^2 \text{Var}\left(\widehat{\Lambda}(t)\right).$$

where $\widehat{\Lambda}(t)$ is the estimated cumulative hazard.

Also,

$$\text{Var}\left(\widehat{\Lambda}(t)\right) = \sum_{i:t_i<t} \frac{\#\text{ events at time } t_i}{(\#\text{ at risk at time } t_i)^2}$$

where the t_i are the event times.

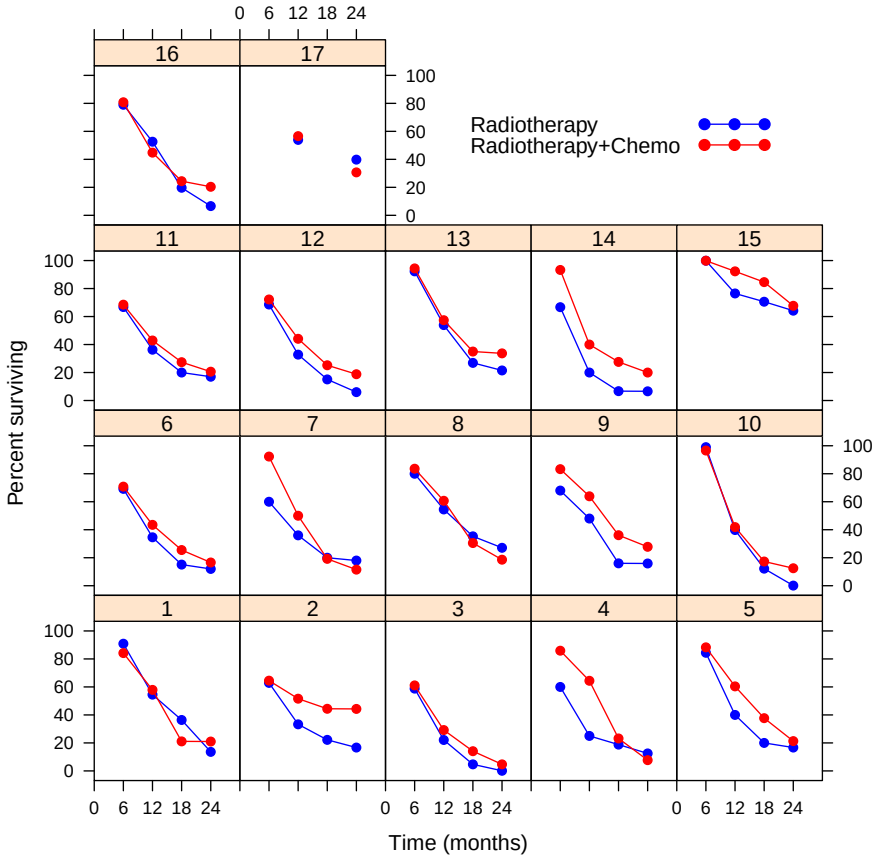
Parmar et al. [PTS98] provide a method for approximating the number of events and number of patients at risk during each time interval.

Example of combining K-M curves

We'll use the data published in Fine et al (1993) and reproduced in Dear [Dea94].

The data come from a meta-analysis of 17 randomized controlled trials that tested the addition of chemotherapy to radiotherapy in postoperative malignant glioma in adults.

They estimate the survival separately at 6, 12, 18 and 24 months after surgery.



Model for malignant glioma survival data

To start, we'll fit a model that assumes the survival follows an exponential distribution with between study variabily in the event rates.

$$S_i^g(t) = \exp\left(-\lambda_i^g t\right)$$
$$\log\left(\lambda_i^g\right) = \mu_i + \delta_i I(g = \text{RC})$$
$$\mu_i \sim N\left(\overline{\mu}, \omega_\mu^2\right)$$
$$\delta_i \sim N\left(\overline{\delta}, \omega_\delta^2\right)$$

where $i = 1, \ldots, 17$ denotes the study and $g \in \{R, RC\}$ denotes the treatment group.

Model for malignant glioma survival data

Using the logit transformation, we model the observed survival estimates as

$$\text{logit}\left(\widehat{S}_i^g\left(t_{ij}\right)\right) = \text{logit}\left(S_i^g\left(t_{ij}\right)\right) + \epsilon_{ij}^g$$

We'll use a covariance model for the ϵ_{ij}^g residuals similar to the one given by Arends but adapted to the logit transformation. Specifically,

$$\epsilon_i^g \sim N\left(0, V_i^g\right)$$

and V_i^g is block diagonal with elements

$$\frac{se_{ijk}}{\sqrt{\widehat{S}_{ijk}\left(1-\widehat{S}_{ijk}\right)}}\sqrt{\frac{\widehat{S}_{ijk}\left(1-\widehat{S}_{ijl}\right)}{\left(1-\widehat{S}_{ijk}\right)\widehat{S}_{ijl}}}\frac{se_{ijl}}{\sqrt{\widehat{S}_{ijl}\left(1-\widehat{S}_{ijl}\right)}}$$

The key difference is that we've replaced $S(t)$ with $\widehat{S}(t)$ in the middle term.

Model for malignant glioma survival data

Finally, we will use relatively non-informative prior distributions for the mean log event rate in the R arms ($\overline{\mu}$), the mean log hazard ratio comparing RC to R ($\overline{\delta}$) and the beween study standard deviations:

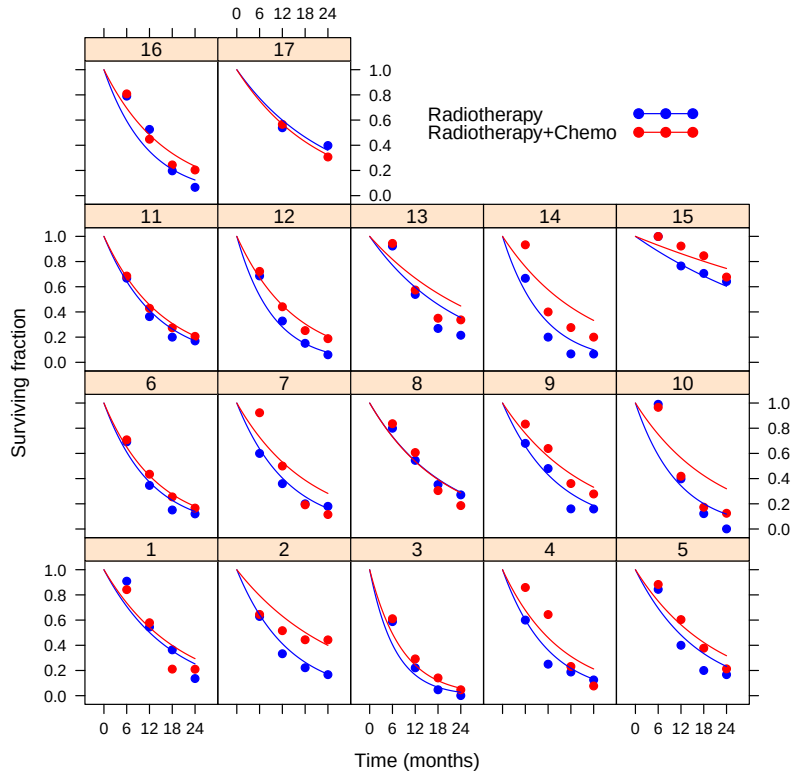
$$\begin{aligned}\overline{\mu} &\sim N(0, 100)\\ \overline{\delta} &\sim N(0, 100)\\ \omega_\mu &\sim U(0, 10)\\ \omega_\delta &\sim U(0, 10)\end{aligned}$$

We will fit the model using WinBUGS.

BUGS code for fitting the model

```
model {
for (i in 1:nArms) {
ostran[i,1:4] ~ dmnorm( mutran[i,1:4] , prec.os[i,1:4,1:4] )
for (j in 1:maxObsPerArm) {
mu[i,j] <- exp( -lam[i] * time[i,j]/12 )
mutran[i,j] <-  logit(mu[i,j])
}
lam[i] <- exp(loglam0[trial[i]] + logdelta[trial[i]]*equals(trt[i],2))
}

# Model for study-level effects
for (j in 1:nTrials) {
loglam0[j] ~ dnorm(mu.lam0, prec.lam0)
logdelta[j] ~ dnorm(mu.delta, prec.delta)
}

# Prior distributions
mu.lam0 ~ dnorm(0, 1.0E-2)
mu.delta ~ dnorm(0, 1.0E-2)
prec.lam0 <- pow(omega.lam0, -2)
omega.lam0 ~ dunif(0,10)
prec.delta <- pow(omega.delta, -2)
omega.delta ~ dunif(0,10)
}
```


Exponential model fit to Fine data



Time-to-event homework #1

The exponential model is a poor fit for some of the studies.

There appears to be a consistent underprediction of survival at the early time points, suggesting that a model with a time-varying hazard might provide a better fit.

For homework, fit an alternative survival model such as a Weibull, Gompertz or log-normal survival function. I would suggest allowing the shape parameter to vary across studies but be shared by the treatment arms within the same study.

You may also want to compare the results to a model that does not account for the correlation within a treatment arm.

An alternative approach to modeling survival data

Ouwens et al. [OPJ11] provide an alternative to the Arends et al. approach.

Instead of modeling the estimated survival curves directly, they model the number of events in non-overlapping time intervals.

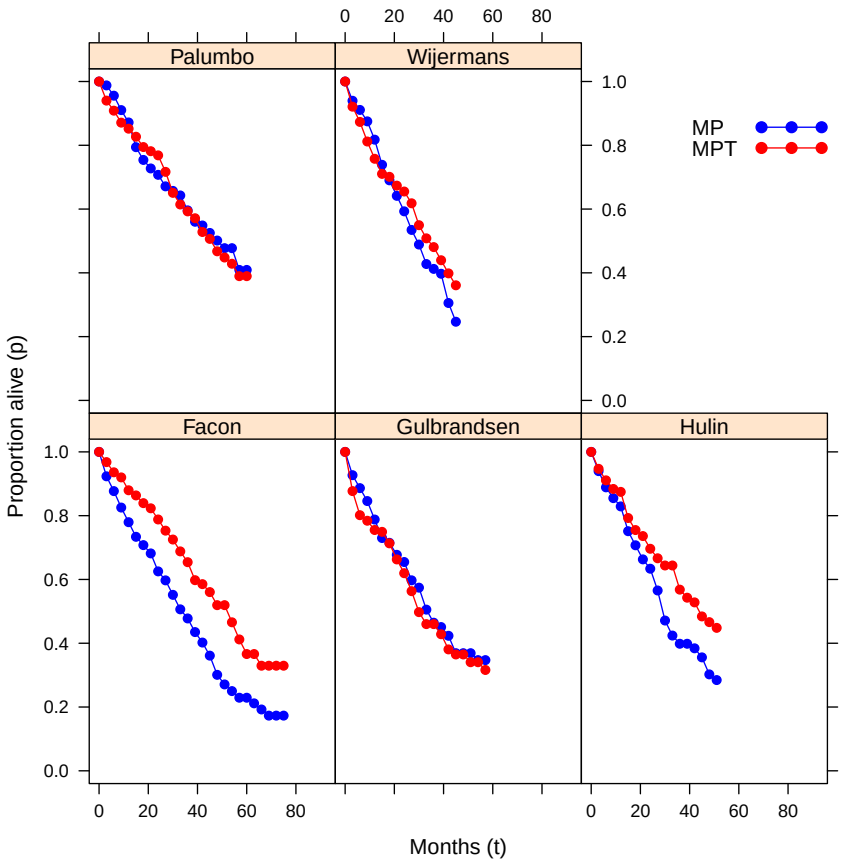
They develop their model in the context of network meta-analysis but we can easily extend this to a model-based meta-analysis framework.

Multiple Myeloma data

Ouwens et al. [OPJ11] present data from 6 clinical trials comparing treatments for multiple myeloma.

We'll focus on the 5 studies comparing melphalan and prednisone (MP) with melphalan-prednisone-thalidomide (MPT) in patients not eligible for bone marrow transplant.

The objective of the analysis was to compare the overall survival across first-line treatments for multiple myeloma in patients non-eligible for transplant based on RCT evidence.



Model for the multiple myeloma data

The proportion of subjects alive at time t_j is given by $S\left(t_j\right)$

The number of subjects alive at time t_j who die between time t_j and time t_{j+1} can be modeled using a binomial distribution. Specifically,

$$r_j \sim \text{Binomial}\left(n_j, p_j\right),$$

where

- r_j is the number of patients dying in the interval $[t_j\ ,\ t_{j+1})$,
- n_j is the number of patients alive (i.e., at risk of dying) at time t_j and
- $p_j = \frac{S\left(t_j\right) - S\left(t_{j+1}\right)}{S\left(t_j\right)}$ is the conditional probability of dying in the time interval given a patient is alive at time t_j

Because of the way they're defined, the r_j are independent across time intervals within a treatment arm.

Model for the multiple myeloma data

To start, we will model $S\left(t\right)$ using an exponential distribution with treatment effects on the rate constant. Specifically,

$$\begin{aligned} S_i^g\left(t\right) &= \exp\left(-\lambda_i^g t\right) \\ \log\left(\lambda_i^g\right) &= \mu_i + \delta_i I\left(g = \text{MPT}\right) \\ \mu_i &\sim N\left(\overline{\mu}, \omega_\mu^2\right) \\ \delta_i &\sim N\left(\overline{\delta}, \omega_\delta^2\right) \end{aligned}$$

where $i = 1, \ldots, 5$ denotes the study and $g \in \{MP, MPT\}$ denotes the treatment group.

Model for the multiple myeloma data

We will use relatively non-informative prior distributions for the mean log event rate in the MP arms ($\overline{\mu}$), the mean log hazard ratio comparing MPT to MP ($\overline{\delta}$) and the beween study standard deviations:

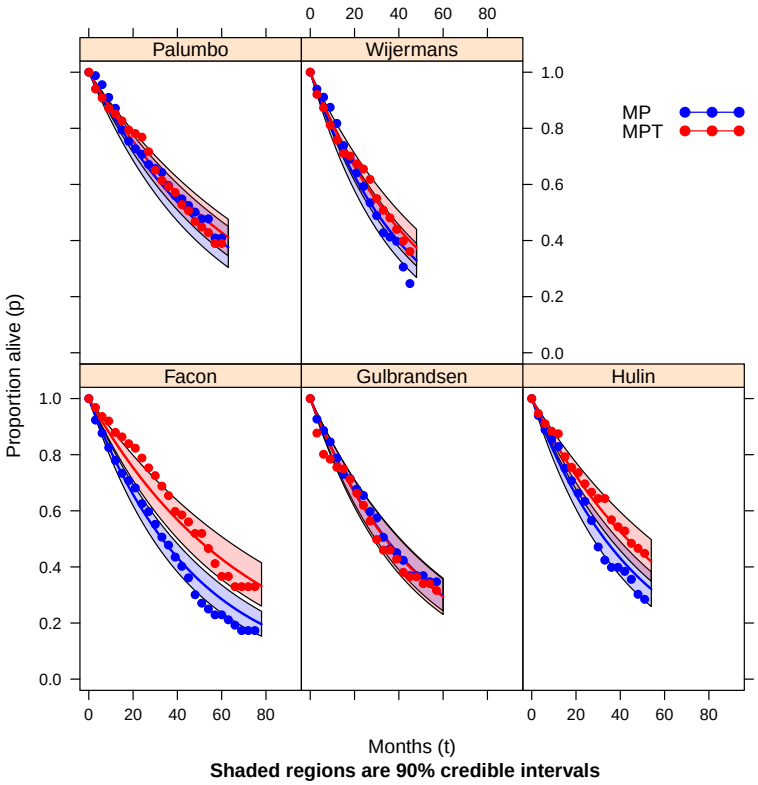
$$\begin{aligned}\overline{\mu} &\sim N(0, 100)\\ \overline{\delta} &\sim N(0, 100)\\ \omega_{\mu} &\sim U(0, 10)\\ \omega_{\delta} &\sim U(0, 10)\end{aligned}$$

We will fit the model using WinBUGS.

BUGS code for fitting the model

```
model {
# Model
for (i in 1:nObs) {
r[i] ~ dbin(p[i], n[i])
Sstart[i] <- exp(-lam[i] * start[i])
Send[i] <- exp(-lam[i] * end[i])
p[i] <- (Sstart[i] - Send[i])/Sstart[i]

log(lam[i]) <- lambda0[study[i]] + delta[study[i]] * equals(group[i],2)
}

for (j in 1:nStudies) {
lambda0[j] ~ dnorm(mu.lambda0, prec.lambda0)
delta[j] ~ dnorm(mu.delta, prec.delta)
}

# Priors
mu.lambda0 ~ dnorm(0, 1.0E-2)
mu.delta ~ dnorm(0, 1.0E-2)
omega.lambda0 ~ dunif(0,10)
omega.delta ~ dunif(0,10)

# Transformations
prec.lambda0 <- pow(omega.lambda0 , -2)
prec.delta <- pow(omega.delta , -2)


}
```

Exponential fit to the multiple myeloma data



Shaded regions are 90% credible intervals

Time-to-event homework #2

Ouwens et al. [OPJ11] include data for a sixth trial that compares MP and cyclophosphomide, thalidomide, and dexamethasone attenuated (CTDa).

- Re-fit the model including data from this study.
- Evaluate whether you think an exponential model provides an adequate fit or if a more flexible model (e.g., Weibull or Gompertz) provides a better fit.

Which approach should you use?

The Ouwens approach may appear to make fewer assumptions than the Arends approach.

- No need for standard errors of $\widehat{S}(t)$
- No assumptions about / approximations for the correlation
- No arbitrary choice of transformation of $S(t)$

However, there are assumptions underlying this approach.

- Censoring is assumed to either be non-existant or to occur within an interval before the events
- Need to derive/approximate the number of events in an interval from reported survival curves
- Most reliable implementation needs number of patients at risk over time (e.g., below K-M plots)

Which approach should you use?

It really depends on what data are available to you.

- If you have numbers at risk over time, then the Ouwens approach seems preferable.
  - They also provide and algorithm for calculating the number of events in a time interval.
  - See also Guyot et al. [GAOW12] for an algorithm for calculating r_j
- Otherwise, the Arends approach can be quite useful
- Because we are modeling the underlying survival function in both approaches, some combination of the two approaches could also be used if you have numbers at risk for some studies and not for others.

Summary statistics when there are dropouts

Summary statistics such as treatment means are typically summaries of:

- LOCF (last observation carried forward) data for all intent to treat patients or
- OC (observed cases) data

Both types of statistics are potentially biased estimates of as treated outcomes

- If assigned treatment affects dropout behavior
- If the measured quantity systematically changes with time

LOCF statistics are most commonly reported and may be the only values available.

Combining summary and individual data

Simultaneous analysis of summary and individual data

The modeling approach for longitudinal data described in section on longitudinal data may also be used to simultaneously model summary and individual data. Recall that the derivation of that method began with a population model for individual data.

Sequential analysis

Sequential analysis

- Another approach is to leverage the Bayesian machinery to construct a model based on summary and individual data sequentially.
- For example a model might initially be developed based only on summary data.
- The posterior samples (MCMC samples) of the parameters resulting from that analysis could be approximated by parametric distributions. Those distributions could be used as prior distributions for subsequent analysis of the individual data.
- This seems a reasonable approach to use if you already have a model “on the shelf” and new data becomes available.

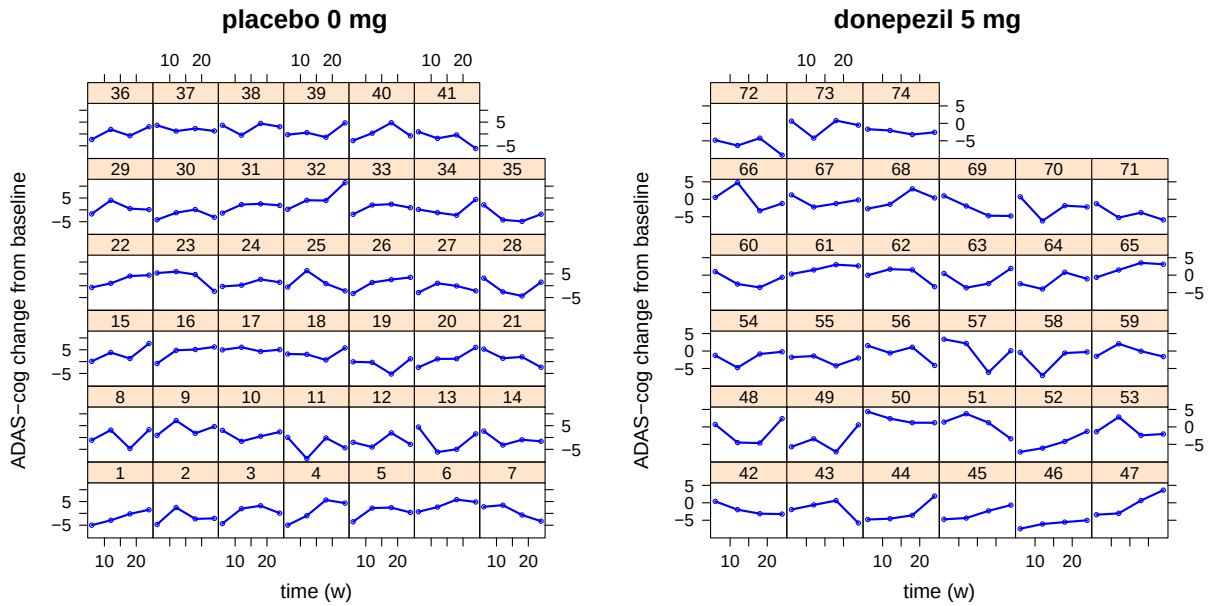
Hands-on Problem 6: Longitudinal dose-response model based on a combination of summary and individual data

- In this exercise we extend Problem 5 by the addition of individual patient data obtained from a (simulated) dose-response trial of the new drug metupezil.
- You will simultaneously analyze the individual patient data from that trial and the same summary data used in Problem 5.
- This approach has the potential to enhance the analysis of metupezil dose-response by (1) borrowing information from prior data and (2) permitting more precise comparative inferences.

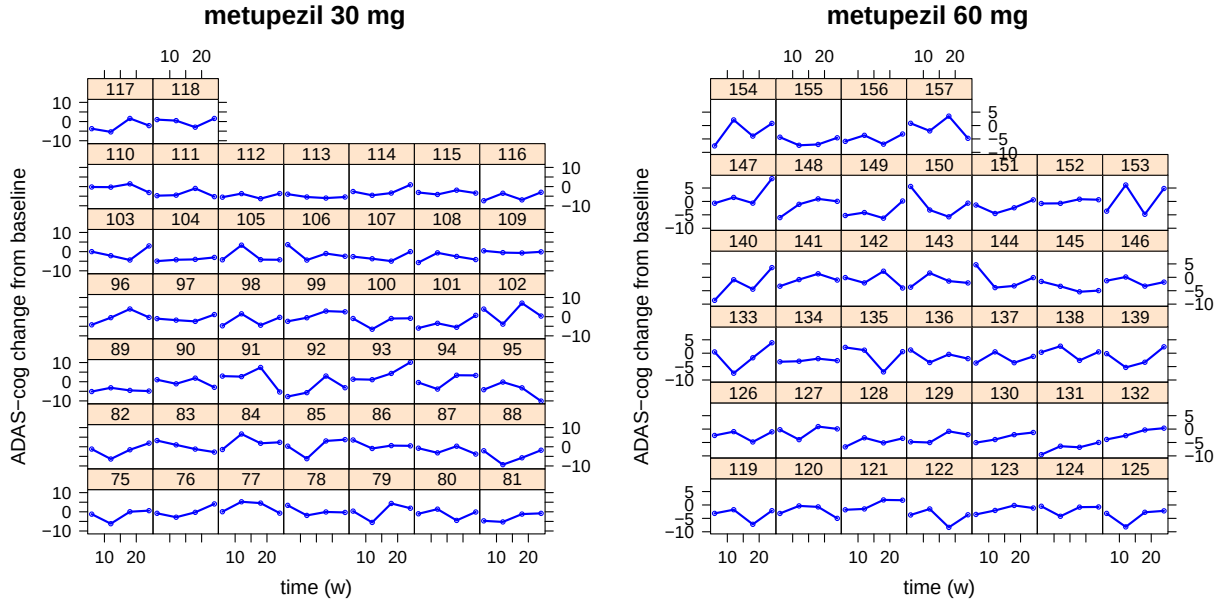
Hands-on Problem 6: The data

- Post-baseline sample means and sample variances for ADAS-cog change from baseline from published sources
  - See Problem 5 for details.
- Post-baseline ADAS-cog change from baseline observed in a metupezil dose-response trial
  - Treatments
    - Placebo: 41 patients
    - Donepezil 5 mg/d: 33 patients
    - Metupezil 30 mg/d: 44 patients
    - Metupezil 60 mg/d: 39 patients
    - Metupezil 120 mg/d: 39 patients
  - Observations at 6, 12, 18 and 24 weeks
- Hands-on exercise
  - Construct a model for ADAS-cog change from baseline as a function of drug, daily dose, time and possibly other covariates.
- Summary data file: `adasCog2HandsOn/donepezilAdasData.csv`
- Individual patient data file: `adasCog2HandsOn/metupezil.csv`

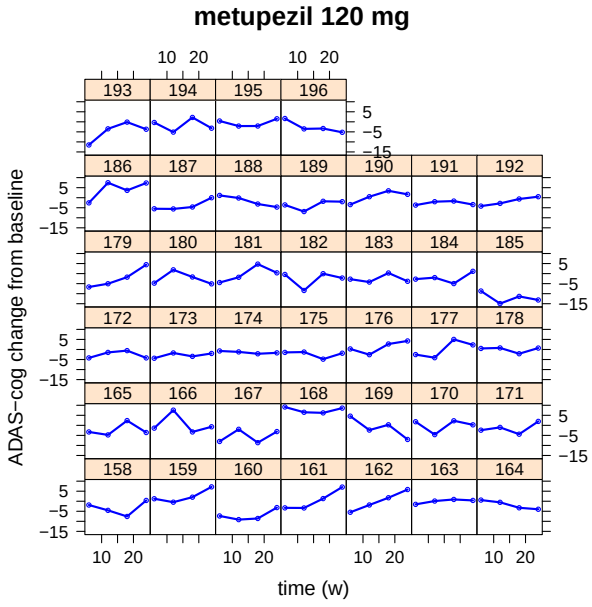
Metupezil individual data



Metupezil individual data



Metupezil individual data



Incorporating a broader range of data and knowledge

So far the course has focused on MBMA as a somewhat isolated process. Logically it should be an integral part of a larger model-based effort to support decision-making. Elements of such an approach include:

- Leveraging the Bayesian framework to incorporate additional quantitative knowledge via informative prior distributions
  - This may (and probably should) include use of more mechanistic models in which the parameters have more direct physiologic interpretation. Prior quantitative knowledge about such parameters might be obtained from a variety of sources.
- Integrating models of the relationships among preclinical, biomarker and clinical outcome data to improve prediction and decision-making in early clinical development

Trial simulation to design a Phase II dose finding strategy

- Example illustrating integration of several Bayesian M&S components including:
  - Model-based meta-analysis
  - Modeling of a preclinical-to-clinical relationship
  - Bayesian trial analysis including use of prior information and adaptive elements

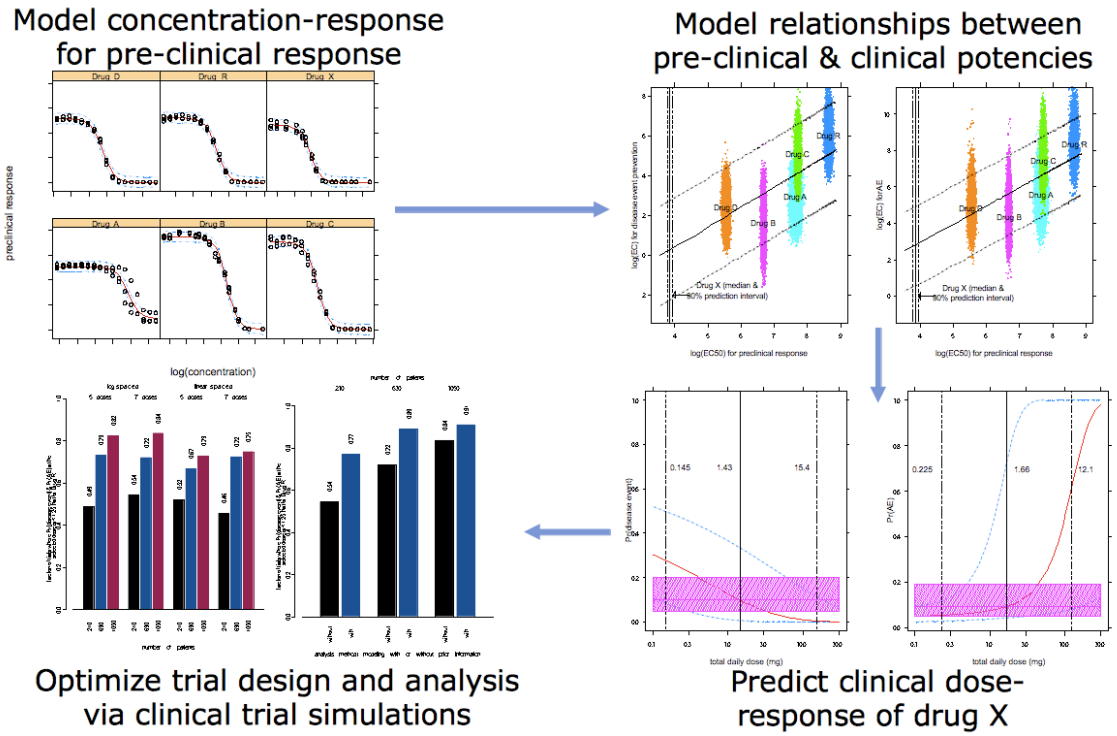
Trial simulation to design a Phase II dose finding strategy

- Phase II objective: Efficiently find a dose of drug X that:
  - Is at least non-inferior to the standard of care (drug R 10 mg/d) with respect to both efficacy and safety.
  - With sufficient certainty that we can risk a Phase III program with only one dose level.
- Efficacy
  - Decrease in fraction of patients with a disease-related event
- Probable dose-limiting AE
  - Same biological mechanism as efficacy

Available information

- Drug X:
  - Clinical pharmacokinetics from Phase I.
  - Pre-clinical response thought to be predictive of clinical outcomes related to mechanism of action (both efficacy and dose-limiting AE)
    - Data for drug X and competitors
- Public-source data:
  - 5 marketed drugs, 27 clinical trials, 77 treatment arms
  - Disease-related events
    - Number of patients with events and total number of patients for each treatment arm
  - Dose-limiting AEs
    - Number of patients with events and total number of patients for each treatment arm
  - Pharmacokinetics
    - Mean clearance
  - Two categories of patients known to have different risks for both disease events and AEs (designated group 1 and group 2)

Modeling strategy: Simultaneously model pre-clinical biomarker and frequency of clinical efficacy and AE events



Implementing the modeling strategy

- Hierarchical model
  - Binomial models for numbers of disease and adverse events
  - Normal model for preclinical responses
  - Inter-trial variation in clinical response
  - Inter-drug variation in the preclinical-to-clinical model to account for model misspecification
- Bayesian data analysis
  - Implemented with WinBUGS
    - Easy to implement complicated probabilistic structure
    - Rigorous approach to quantifying uncertainty in model parameters and predictions
  - Relatively non-informative priors

Integrated preclinical-clinical model: disease event, AE and preclinical response

- Preclinical response sub-model

$$PCR = E_0^{PCR}\left(1 - E_D^{PCR}\right)$$

$$E_D^{PCR} = \frac{c^{\gamma_{PCR}}}{\left(EC_{50}^{PCR}\right)^{\gamma_{PCR}} + c^{\gamma_{PCR}}}$$

Integrated preclinical-clinical model: disease event, AE and preclinical response

- Disease event sub-model

$$n_{DE} \sim \text{binomial}\left(p_{DE}, n\right)$$

$$\text{logit}\left(p_{DE}\right) = E_P^{DE} + I_{group2}\theta_{group2}^{DE} - E_D^{DE} \qquad E_D^{DE} = \frac{\overline{c}^{\gamma_{DE}}}{\left(EC^{DE}\right)^{\gamma_{DE}}}$$

$$\log\left(EC^{DE}\right) \sim N\left(\log\left(\mu_{EC^{DE}}\right), \sigma^2_{EC^{DE}}\right) \qquad \mu_{EC^{DE}} = \theta_{PCR\rightarrow DE} EC_{50}^{PCR}$$

Integrated preclinical-clinical model: disease event, AE and preclinical response

- AE sub-model

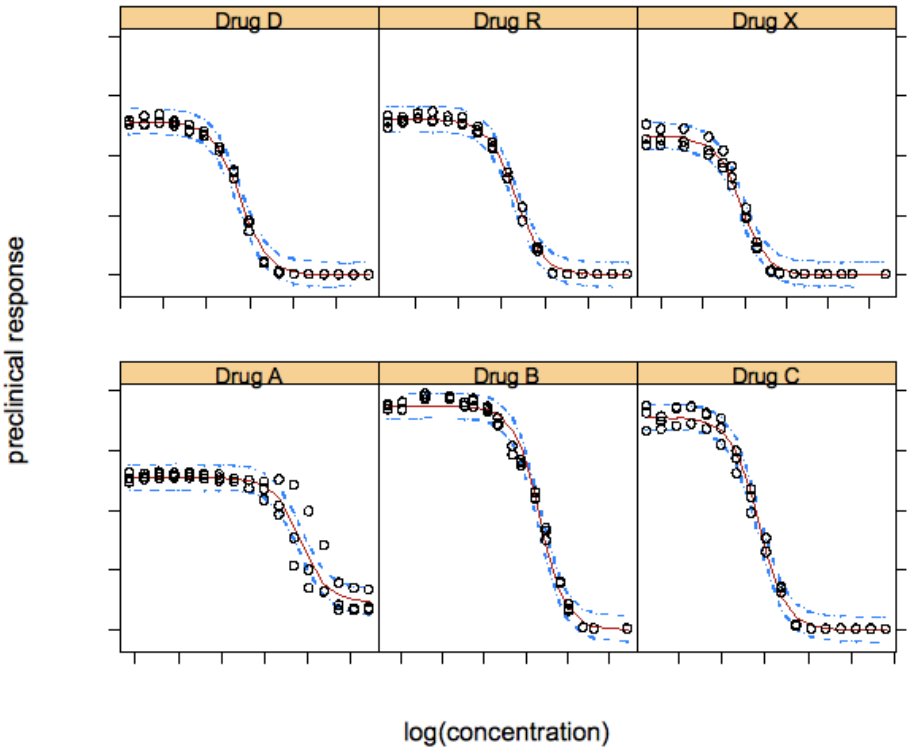
$$n_{AE} \sim \text{binomial}\left(p_{AE}, n\right)$$

$$\text{logit}\left(p_{AE}\right) = E_P^{AE} + I_{group2}\theta_{group2}^{AE} + E_D^{AE} \qquad E_D^{AE} = \frac{\overline{c}^{\gamma_{AE}}}{\left(EC^{AE}\right)^{\gamma_{AE}}}$$

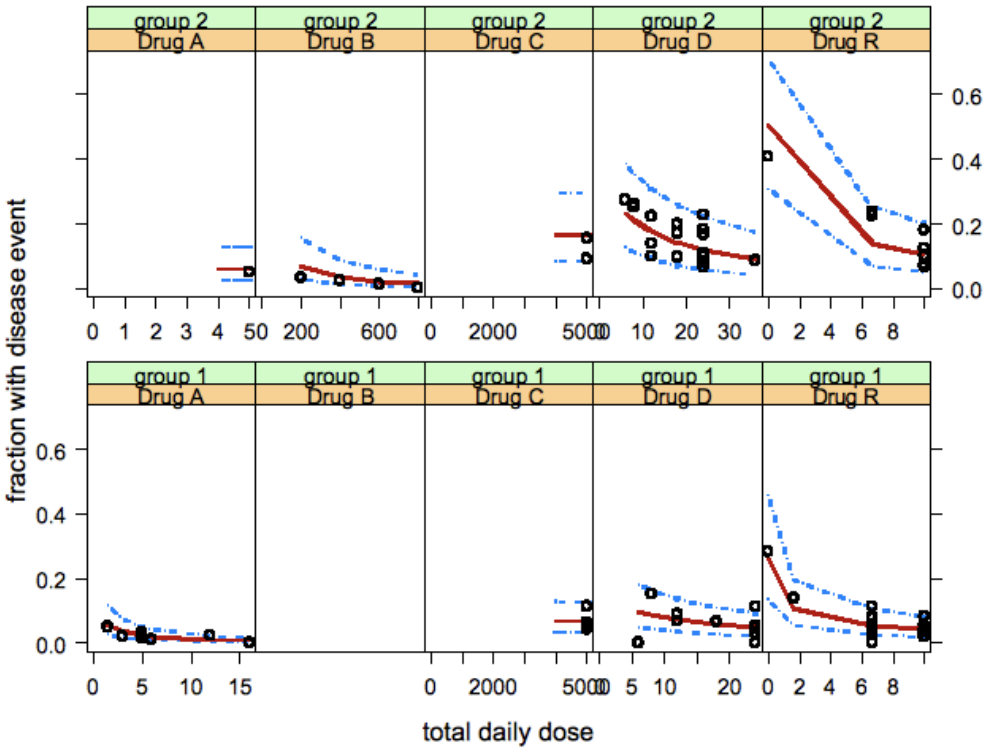
$$\log\left(EC^{AE}\right) \sim N\left(\log\left(\mu_{EC^{AE}}\right), \sigma^2_{EC^{AE}}\right) \qquad \mu_{EC^{AE}} = \theta_{PCR \to AE} EC_{50}^{PCR}$$

This is a simplified description of the model. The gory details are available on request.

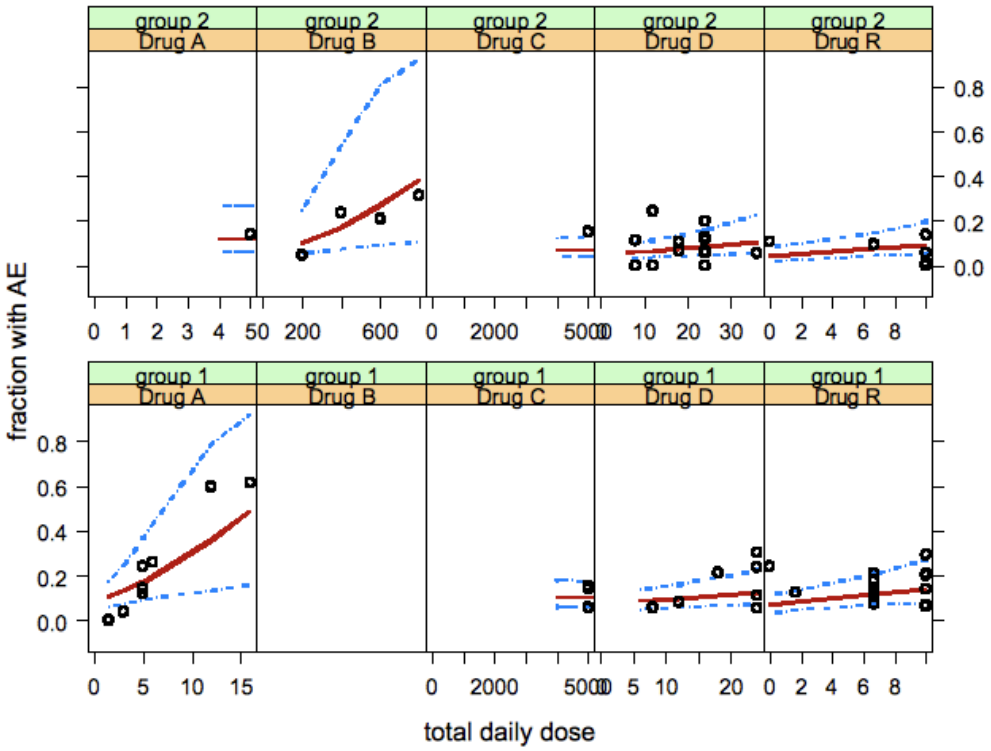
Inhibition of preclinical response



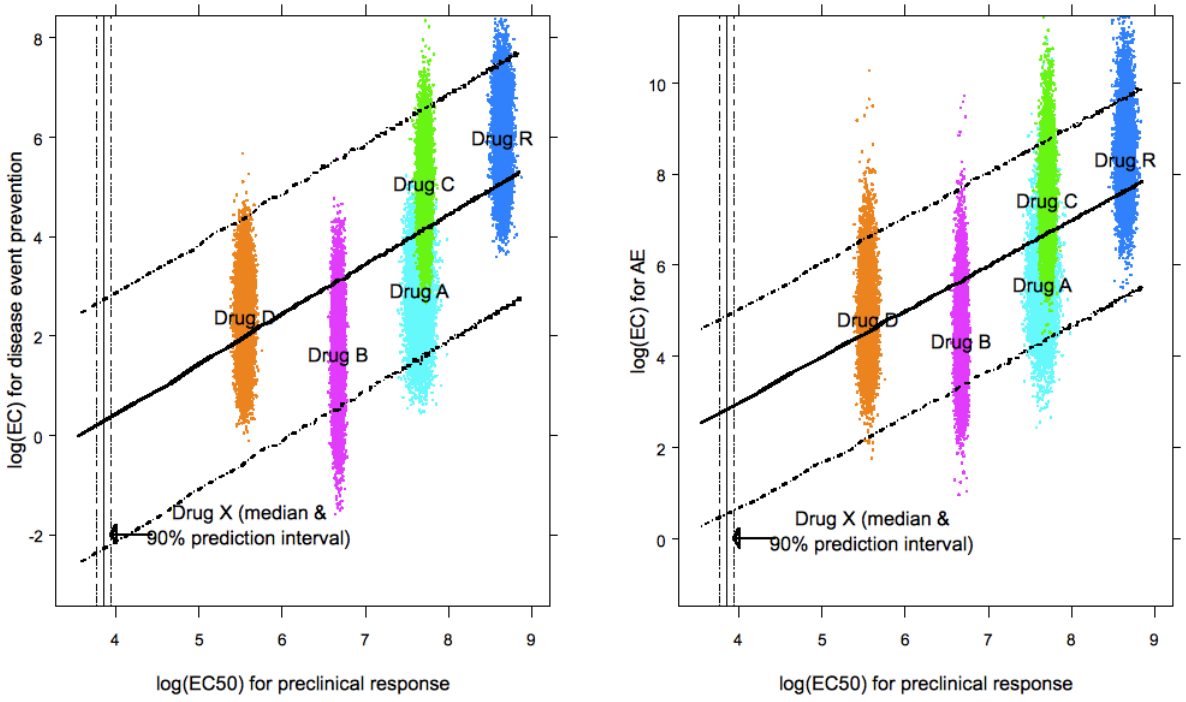
Prevention of disease event: Dose-response by drug



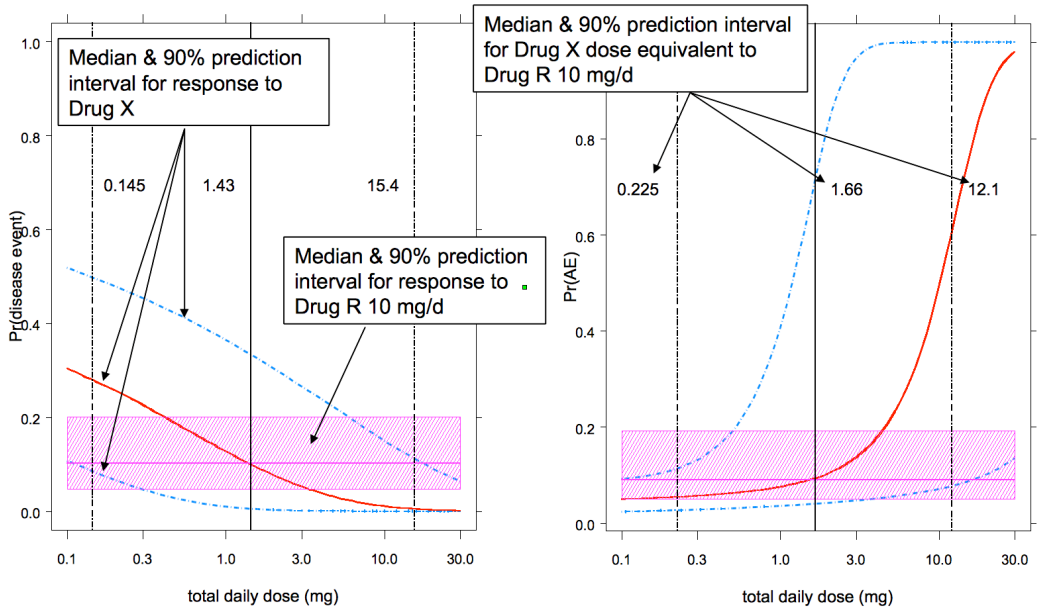
Major AE: Dose-response by drug



Preclinical-to-clinical outcome: Modeled relationships between preclinical and clinical potencies



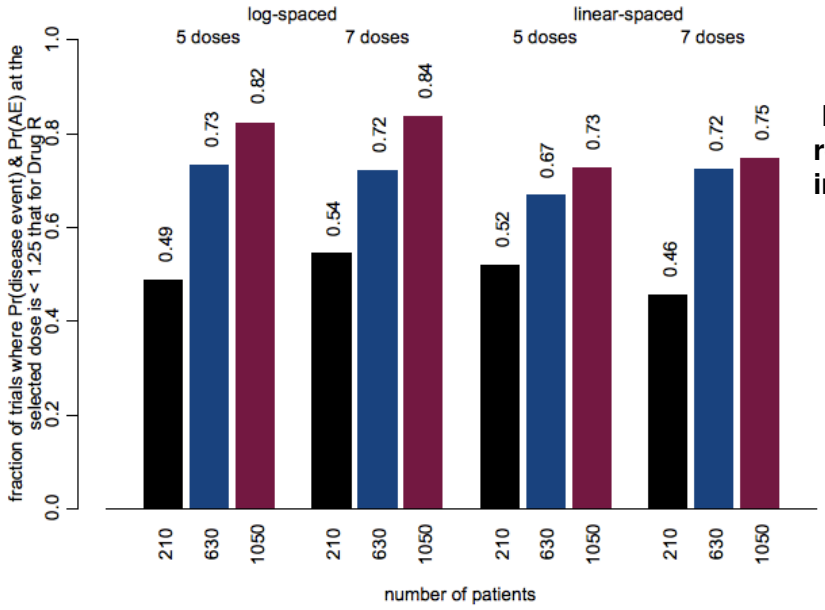
Modeling of disease events and AEs lead to consistent but highly uncertain predictions of the drug X dose equivalent to drug R 10 mg/d



Clinical design and analysis options considered

- Number of patients per treatment arm (210, 630 or 1050)
- Number (5 or 7) and spacing (linear- or log-spaced) of doses
- Trial analysis
  - Dose-response modeling using conventional logistic regression
  - Bayesian modeling using prior information, e.g., the dose-response model described in the previous slides
  - In either case the dose-response model is used to select a dose with efficacy equivalent to drug R 10 mg/d
- Adaptive pruning of treatment arms for lack of efficacy or excess AEs
  - Based on frequentist confidence intervals and observed fraction of events for drug R 10 mg/d
  - Based on posterior probabilities from Bayesian modeling
- Trial performance is assessed by whether the selected dose is non-inferior to drug R 10 mg/d (under the simulated truth)
  - Fraction of simulated trials where Pr(disease event) and Pr(AE) at the selected dose are less than 1.25 times that for drug R 10 mg/d

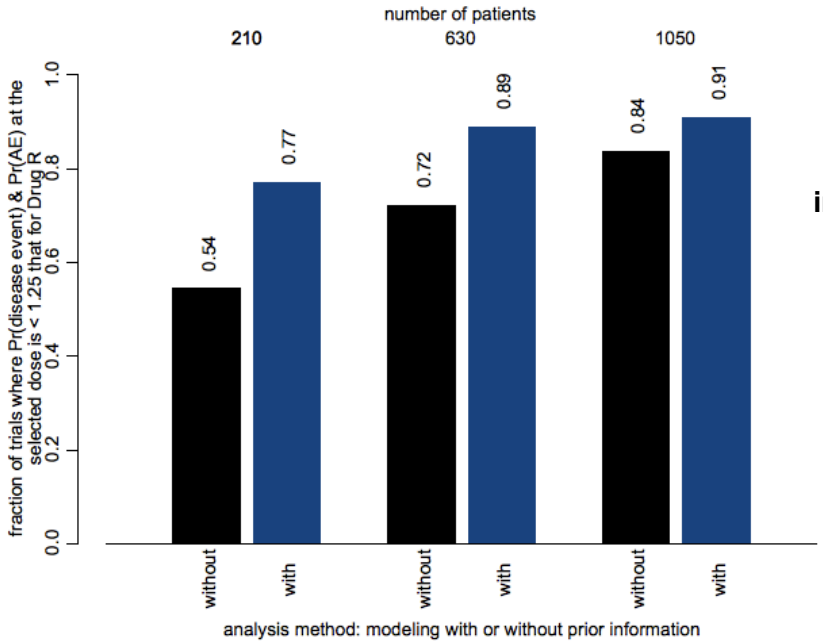
Trial simulation results: Effects of number and spacing of doses and sample size



Results using logistic regression without prior information

- Best performance is seen with 7 log-spaced doses
- Only trials with $\geq$1050 patients offer sufficient certainty to consider risking a Phase III program with only one dose level.

Trial simulation results: Impact of using prior information



Bayesian modeling using prior information

- Increases the probability of selecting a non-inferior dose, particularly in trials with smaller sample sizes
- Also improves the probability of making correct pruning decisions

- **Bottom line: Simulations led to more efficient designs that can find a non-inferior dose with high probability**

Summary of key points from the case study

- Integration of pre-clinical and public-source clinical data permits construction of a model for predicting clinical outcomes for the NCE.
- Leveraging prior information permits more efficient design and analysis of a Phase II trial to select a dose for Phase III:
  - Optimizes range, number and spacing of doses
  - Adaptive pruning assigns patients to most relevant doses
  - Enhances characterization of dose-response and therefore dose selection
  - $\Longrightarrow$ Shorter, more informative Phase II program

Network meta-analysis

What is Network meta-analysis?

- Classical meta-analysis focuses on the direct pairwise comparison of two treatments
- In the MBMA examples weve considered, weve implicitly been making comparisons between treatments using a combination of direct comparisons and indirect comparisons.
- Network meta-analysis (aka mixed treatment comparisons meta analysis or indirect comparisons meta-analysis) is an extension of classical meta-analysis methods to do the same thing

What is Network meta-analysis?

- There has been a tremendous amount of theoretical work done on NMA over the past 10 years.
- This has led to a very rich set of papers. . . Well only cover the basics and point you to some additional interesting and important papers.
- While NMA has a developed theoretical background (unlike MBMA) it is still not without its detractors . . .

A few references to get started

General background

- Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated March 2013; available from http://www.nicedsu.org.uk
- Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health.* 2011 Jun;14(4):417-28.
- David C. Hoaglin, Neil Hawkins, Jeroen P. Jansen, David A. Scott, Robbin Itzler, Joseph C. Cappelleri, Cornelis Boersma, David Thompson, Kay M. Larholt, Mireya Diaz, Annabel Barrett, Conducting Indirect-Treatment-Comparison and Network-Meta-Analysis Studies: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 2, Value in Health, Volume 14, Issue 4, June 2011, Pages 429-437

Classic references

- Bucher HC, Guyatt GH, Griffith LE, Walter SD. The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trials. Journal of Clinical Epidemiology 1997;50:683-691.
- Lumley T. Network meta-analysis for indirect treatment comparisons. Statistics in Medicine 2002;21:2313-2324
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Statistics in Medicine 2004;23:3105-3124.
- Salanti, G., Higgins, J.P.T., Ades, A.E., Ioannidis, J.P.A. Evaluation of networks of randomized trials. Statistical Methods in Medical Research 2008; 17(3):279-301.

A network meta-analysis of treatments for depression

OPEN ACCESS Freely available online

PLoS one

Effectiveness and Cost-Effectiveness of Antidepressants in Primary Care: A Multiple Treatment Comparison Meta-Analysis and Cost-Effectiveness Model

Joakim Ramsberg[1]*, Christian Asseburg[2], Martin Henriksson[3]

1 Department of Learning, Informatics, Management and Ethics, Medical Management Center, Karolinska Institutet, Stockholm, Sweden, 2 ESiOR Oy, Kuopio, Finland, 3 AstraZeneca Nordic, Södertälje, Sweden

Ramsberg J, Asseburg C, and Henriksson M. (2012) Effectiveness and Cost-Effectiveness of Antidepressants in Primary Care: A Multiple Treatment Comparison Meta- Analysis and Cost-Effectiveness Model. PLoS ONE 7(8): e42003.

A network meta-analysis of treatments for depression

Abstract

Objective: To determine effectiveness and cost-effectiveness over a one-year time horizon of pharmacological first line treatment in primary care for patients with moderate to severe depression.

Design: A multiple treatment comparison meta-analysis was employed to determine the relative efficacy in terms of remission of 10 antidepressants (citalopram, duloxetine escitalopram, fluoxetine, fluvoxamine mirtazapine, paroxetine, reboxetine, sertraline and venlafaxine). The estimated remission rates were then applied in a decision-analytic model in order to estimate costs and quality of life with different treatments at one year.

Data Sources: Meta-analyses of remission rates from randomised controlled trials, and cost and quality-of-life data from published sources.

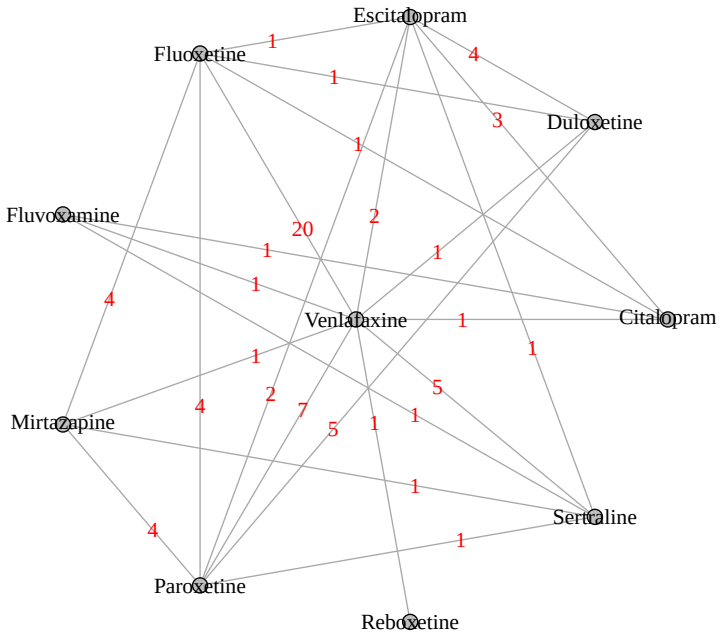
Results: The most favourable pharmacological treatment in terms of remission was escitalopram with an 8- to 12-week probability of remission of 0.47. Despite a high acquisition cost, this clinical effectiveness translated into escitalopram being both more effective and having a lower total cost than all other comparators from a societal perspective. From a healthcare perspective, the cost per QALY of escitalopram was €3732 compared with venlafaxine.

Conclusion: Of the investigated antidepressants, escitalopram has the highest probability of remission and is the most effective and cost-effective pharmacological treatment in a primary care setting, when evaluated over a one year time-horizon. Small differences in remission rates may be important when assessing costs and cost-effectiveness of antidepressants.

There were 73 studies for these 10 treatments.

Ramsberg J, Asseburg C, and Henriksson M. (2012) Effectiveness and Cost-Effectiveness of Antidepressants in Primary Care: A Multiple Treatment Comparison Meta- Analysis and Cost-Effectiveness Model. PLoS ONE 7(8): e42003.

A network meta-analysis of treatments for depression



As an example, we'll focus on the network meta-analysis of remission rates among these 10 treatments for depression:

Citalopram,
Duloxetine,
Escitalopram,
Fluoxetine,
Fluvoxamine,
Mirtazapine,
Paroxetine,
Reboxetine,
Sertraline,
Venlafaxine

Ramsberg data

Here are a few rows of the data from the Ramsberg paper.

Study	Drug.1	Drug.2	N.drug.1	Remission.drug.1	N.drug.2	Remission.drug.2
Ventura et al 2007	Escitalopram	Sertraline	104	51	107	57
Behnke et al 2003	Mirtazapine	Sertraline	171	76	168	73
Rossini et al 2005	Sertraline	Fluvoxamine	48	25	40	28
Haffmans et al 1996	Citalopram	Fluvoxamine	108	15	109	9
Schwartz et al 2002	Reboxetine	Venlafaxine	80	20	87	21
Blier et al 2009	Mirtazapine	Paroxetine	21	4	19	5

They also report covariate values for `duration of study` and `setting`.

Consider two pairwise comparisons

Let's recall the models we used for pairwise comparisons.

We'll focus on the Venlafaxine vs. Sertraline (n=5) and Venlafaxine vs. Paroxetine (n=7) comparisions.

We would model these two differences separately, using models similar to this one for the V vs. S data:

$$Y_{ij} \sim \text{Binomial}(N_{ij}, p_{ij})$$
$$\text{logit}\left(p_{ij}\right) = \mu_{i,V} + \delta_{i,VS} \times I\left(j = \text{S}\right)$$
$$\delta_{i,VS} \sim N\left(\Delta_{VS}, \tau^2_{VS}\right)$$

where i denotes study and j denotes treatment group - e.g., $j \in \{\text{V}, \text{S}\}$.

The pairwise models

The objective would be to estimate

- Δ_{VS} = mean difference in log odds betwen S and V
- Δ_{VP} = mean difference in log odds betwen S and V
- τ^2_{VS} = the between-study variances in log odds in S-V studies
- τ^2_{VP} = the between-study variances in log odds in P-V studies

We would use only the direct, head-to-head data to estimate these parameters and would estimate them separately.

The assumption $\delta_{i,VS} \sim N\left(\Delta_{VS}, \tau^2_{VS}\right)$ means that the $\delta_{i,VS}$'s are exchangeable across the V-S studies.

What if we are interested in the difference between Paroxetine and Sertraline?

Use of indirect evidence

- For the P-S comparision we have one study with direct, head-to-head comparison
  - From that study, we could estimate Δ_{PS} directly from the head-to-head data.
- However, we also have some information from the indirect comparisions through Venlafaxine (and also Mirtazapine).
  - From the Venlafaxine studies, we can get an indirect estimate using $\Delta_{VP} - \Delta_{VS}$
- If we assume the direct and indirect evidence are *consistent*, then we can combine the direct and indirect effects
  - $\Delta_{PS} = \Delta_{VP} - \Delta_{VS}$

The consistency assumption

- The consistency assumption is necessary to make the network meta-analysis inference theoretically justifiable.
- Within a study the consistency assumption holds by definition.
  - That is, in a 3-arm study $\delta_{i,PS} = \delta_{i,VP} - \delta_{i,VS}$
- The assumption is a property of loops of evidence, not of specific head-to-head comparisons
- It can only be evaluated when there is both direct and indirect information

When might consistency not-hold?

- Consistency might not hold when there is a covariate that affects the treatment effects which is also imbalanced across the studies.
- Suppose, we have the following:
  - $\Delta_{VS} = \Delta_{VP} = 0$ in men
  - $\Delta_{VS} = \Delta_{VP} = 2$ in women
  - In a head-to-head trial, the direct effect estimate is $\Delta_{SP}^{\text{direct}} \approx 0$.
- Now, suppose the V-S studies are mostly in men and the V-P studies are mostly in women.
  - The estimate of Δ_{VP} (ignoring the sex effect) is close to 0
  - The estimate of Δ_{VS} (ignoring the sex effect) is close to 2
  - Based on the indirect evidence, $\Delta_{SP}^{\text{indirect}} = \Delta_{VS} - \Delta_{VP} \approx 2$
- In this case, the direct and indirect estimates are not consistent with each other.

The network meta-analysis model

- Let's assume (for now) that all studies are two-arm studies.
- Under the consistency assumption, we have fewer identifiable parameters than actual comparisions.
  - For example, for the P-S-V loop there are three comparisons: $\Delta_{VP}, \Delta_{VS},$ and Δ_{SP}.
- However, we only need to define $(\#\textit{treatments} - 1)$ treatment differences because we can derive all of the others from these.
- These $T - 1$ parameters are called the *fundamental parameters* but are somewhat arbitrary.
  - Conventionally, they are chosen to be the pair-wise differences from a single reference treatment.
  - For the depression analysis, we'll use Venlafaxine because all drugs (but not all studies) have that as a reference point.
  - Note, that we can use any of the treatments as the reference treatment.

The network meta-analysis model

- We'll start by giving numeric labels to all treatments, with the reference treatment getting a value of 1.
  - E.g., Venlafaxine=1, Citalopram=2, Duloxetine=3, . . . , Reboxetine=9, Sertraline=10
- We define a data variable, t_{ij} that defines which treatment the j^{th} arm in the i^{th} study receives.
  - By convention, $j = 1$ for the reference treatment in the i^{th} study.
  - For example, if study i compares Citalopram and Venlafaxine then $t_{i1} = 1$ (Venlafaxine) and $t_{i2} = 2$ (Citalopram)
  - If study i compares Reboxetine and Sertraline, then $t_{i1} = 9$ (Reboxetine) and $t_{i2} = 10$ (Sertraline)
- The fundamental parameters are the pairwise differences from treatment 1: $\Delta_2, \Delta_3, \ldots, \Delta_{10}$
- By defintion $\Delta_1 = 0$

The network meta-analysis model

With that notation, the network meta-analysis model for the depression data is:

$$\begin{aligned} Y_{ij} &\sim \text{Binomial}\left(N_{ij}, p_{ij}\right) \\ \text{logit}\left(p_{ij}\right) &= \mu_i + \delta_{ij} \\ \delta_{i2} &\sim N\left(\Delta_{t_{i2}} - \Delta_{t_{i1}}, \tau^2\right) \\ \delta_{i1} &= 0 \end{aligned}$$

where

$$\begin{aligned} i &= 1, \ldots, \#\text{ studies} \\ j &= 1, 2 \text{ (for now, we assume 2 arms per study)} \\ \mu_i &= \text{log odds of remission in the control treatment in study } i \\ \delta_{i2} &= \text{difference between treatment } t_{i2} \text{ and treatment } t_{i1} \\ \tau^2 &= \text{ between study variance in treatment differences} \end{aligned}$$

Exchangeability as a way to consistency

This seemingly small part of the model is very important:

$$\delta_{ij} \sim N\left(\Delta_{t_{ij}} - \Delta_{t_{i1}}, \tau^2\right)$$

This says that the δ_{ij} values are exchangeable across ***all studies in the network***, not just the studies that involve that particular comparison!

Suppose Study *i* compares Citalopram and Venlafaxine and study *j* compares Sertraline and Venlafaxine. There is a hypothetical (unobserved) comparision of Sertraline and Venalfaxine in study *i* that coms from the same distribution as study *j*.

As a result of this assumption, consistency in the network automatically follows. Consistency is not an additional assumption.

Key assumptions of network meta-analysis model

- Study-specific treatment effects are exchangeable across ***all*** studies
  - Implies consistency in the network
- Equal variance, τ^2, for all between-study effects
  - This can be relaxed.

Evaluating the consistency assumption

This is an active area of research. Commonly used approaches include:

- For simple networks
  - Bucher method for single loops of evidence - a statistical test comparing direct and indirect estimates
- For more general networks
  - Repeated application of Bucher's method
  - Fitting an inconsistency model - assuming no consistency in the network - and compare results (model fit and posterior summary values) to those from the standard model
  - Node splitting (Dias et al. 2010)

Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., Guobing, L. and Ades, A.E. NICE DSU Technical Support Document 4. Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2011.

Dias, S., Welton, N. J., Caldwell, D. M., and Ades, A. E. Checking consistency in mixed treatment comparison meta analysis. Statistics in Medicine 2010; 29 932-944.

Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR: Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Syn Meth 2012, 3(2):98110.

Analyzing the depression network

We'll perform the analysis using all drugs for which there are at least 2 studies, which is all of the original 10 except Reboxetine.

We have $N = 72$ trials in the resulting dataset and $T = 9$ treatments.

We'll use Venlafaxine as our reference treatment and estimate $T - 1 = 8$ parameters for the differences between the other treatments and Venlafaxine.

Analyzing the depression network

Our model for the number of remitters in arm j of study i is

$$\begin{aligned} Y_{ij} &\sim \text{Binomial}\left(N_{ij}, p_{ij}\right) \\ \text{logit}\left(p_{ij}\right) &= \mu_i + \delta_{ij} \\ \delta_{i2} &\sim N\left(\Delta_{t_{i2}} - \Delta_{t_{i1}}, \tau^2\right) \\ \delta_{i1} &= 0 \end{aligned}$$

where

$$\begin{aligned} i &= 1, \ldots, 72 \\ j &= 1, 2 \\ \mu_i &= \text{log odds of remission in the control treatment in study } i \\ \delta_{i2} &= \text{difference between treatment } t_{i2} \text{ and treatment } t_{i1} \\ \tau^2 &= \text{ between study variance in treatment differences} \end{aligned}$$

Analyzing the depression network

We'll fit the model using WinBUGS and use the following prior distributions

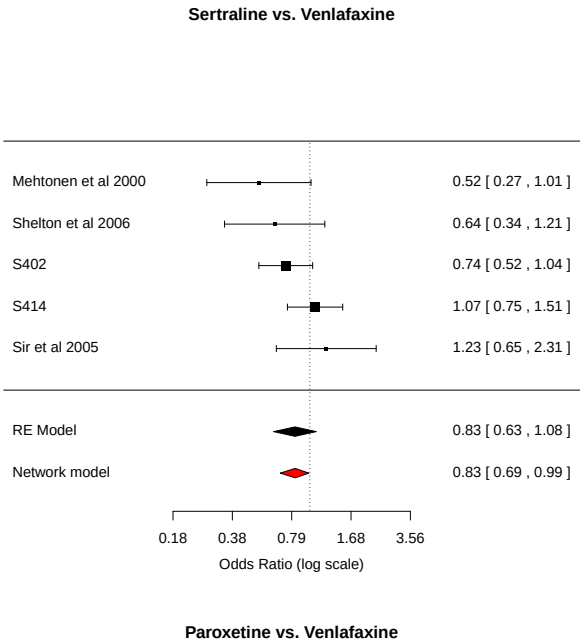
$$\begin{aligned} \mu_i &\sim N\left(0, 1000\right), & i &= 1, \ldots, 72 \\ \Delta_j &\sim N\left(0, 1000\right), & j &= 2, \ldots, 9 \\ \tau &\sim U\left(0, 10\right) \end{aligned}$$

We'll review the WinBUGS code during the next lab session.

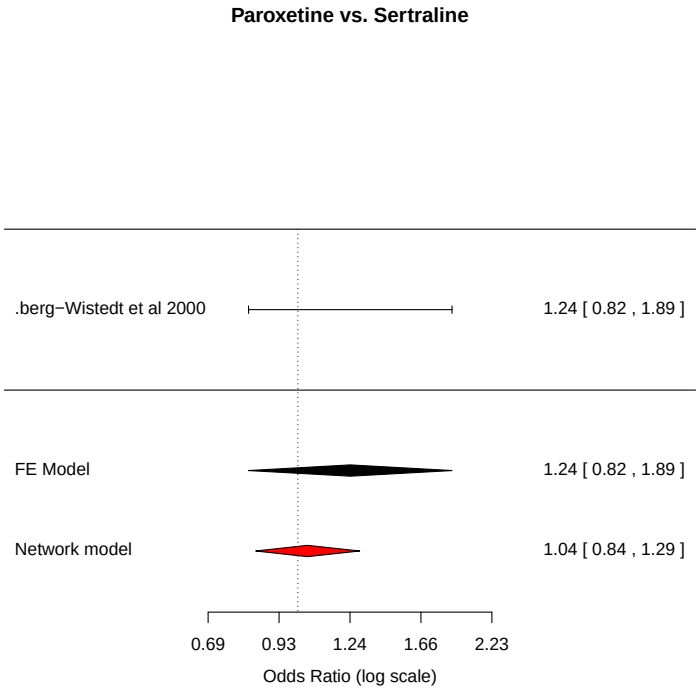
How can we use the network model to compare treatments?

- Pairwise comparisons of effects (i.e., differences in Δ_j values)
- Comparisons of rank probabilities
  - Probability of being the best treatment
  - Probability of being one of the best 2 treatments
  - Median or average rank
  - Surface under the cumulative ranking curve (SUCRA) ([SAI11])
- Probability of being within x% of the best treatment
- Could also perform a more formal analysis using utilities

Pairwise comparisons of effects

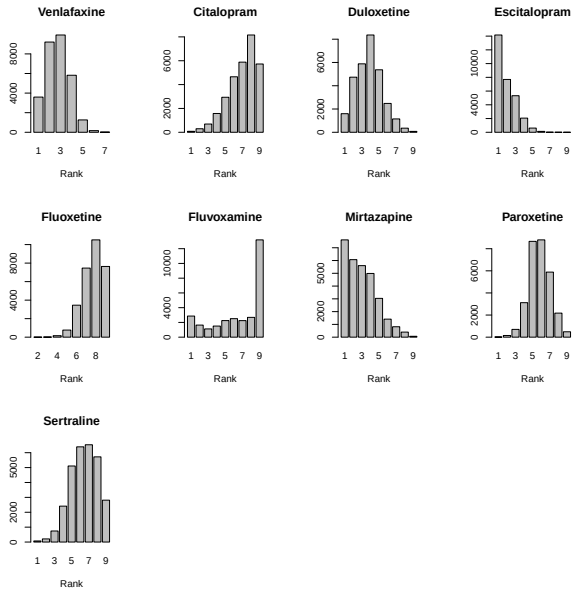


Pairwise comparisons of effects

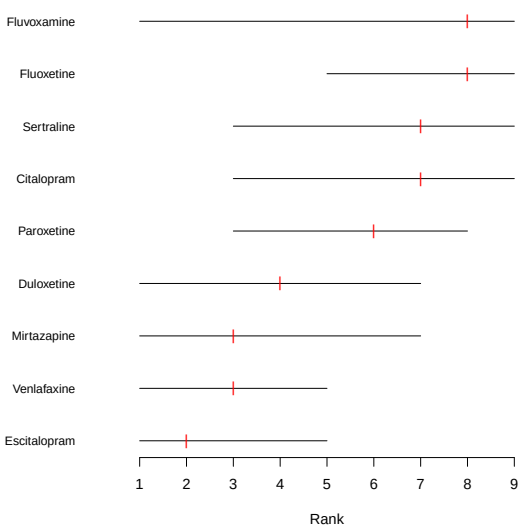


Comparision of rank probabilities

Posterior distributions of ranks

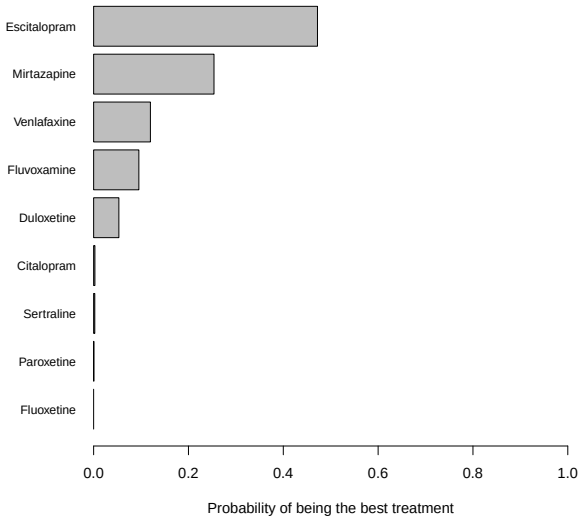


95% credible intervals for ranks

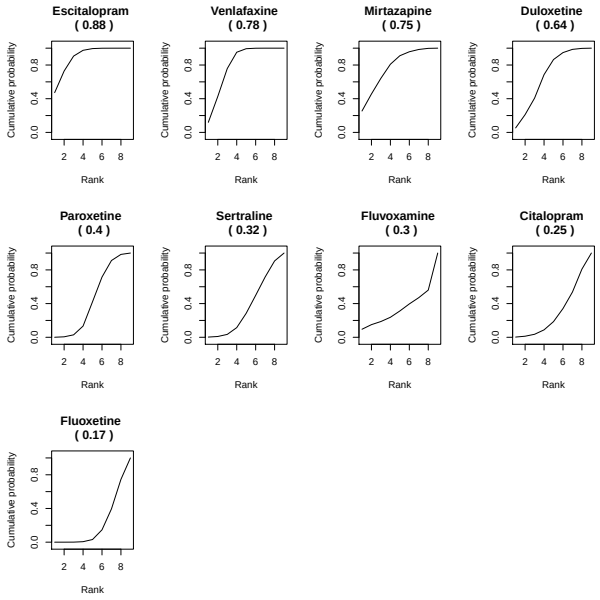


Comparision of rank probabilities

Probability of treatment being the best



Cumulative probability of rank (and SUCRA)



Extension of the model to multi-arm trials

Suppose now that we have trials that include more than two treatment arms.

The framework which we presented previously can easily be extended to include this situation.

Each multi-arm trial will have a vector of random effects, δ_i.

Each trial will have one fewer random effect than treatments (since we're modeling the differences within a study)

A three-arm trial will have two random effects, a four-arm trial will have three random effects, etc.

Extension of the model to multi-arm trials

Under the assumption that all betwen-study effects have the same variance, then

$$\delta_{i2} \sim N\left(\Delta_{t_{i2}} - \Delta_{t_{i1}}, \tau^2\right)$$

becomes, for a 4-arm study

$$\boldsymbol{\delta}_i = \begin{pmatrix} \delta_{i2} \\ \delta_{i3} \\ \delta_{i4} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \Delta_{t_{i2}} - \Delta_{t_{i1}} \\ \Delta_{t_{i3}} - \Delta_{t_{i1}} \\ \Delta_{t_{i4}} - \Delta_{t_{i1}} \end{pmatrix}, \tau^2 \begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix}\right)$$

Why is the correlation 0.5?

The correlation is 0.5 because we have assumed the between-study variance is the same for all of the treatment differences.

This assumption implies that $Var\left(\delta_{i2}\right) = Var\left(\delta_{i3}\right) = Var\left(\delta_{i2} - \delta_{i3}\right)$.

From first principles and this assumption, we have

$$\begin{aligned} Var\left(\delta_{i2} - \delta_{i3}\right) &= Var\left(\delta_{i2}\right) + Var\left(\delta_{i3}\right) - 2Cov\left(\delta_{i2}, \delta_{i3}\right) \\ \tau^2 &= \tau^2 + \tau^2 - 2Cov\left(\delta_{i2}, \delta_{i3}\right) \\ Cov\left(\delta_{i2}, \delta_{i3}\right) &= \tau^2/2 \end{aligned}$$

which means $Corr\left(\delta_{i2}, \delta_{i3}\right) = 1/2$.

Multi-arm network meta-analysis models for treatment differences

When modeling treatment differences, there is an additional correlation induced because the differences within a trial are taken with respect to the same reference group.

This correlation does not arise from model assumptions but from the data themselves.

The correlation can be calculated in closed-form. Specifically, $Cov\left(Y_{AB}, Y_{AC}\right) = Var\left(Y_A\right)$.

Multi-arm network meta-analysis models for treatment differences

Using a 4-arm study as an example, the resulting model is

$$\boldsymbol{Y}_i = \begin{pmatrix} Y_{i,21} \\ Y_{i,31} \\ Y_{i,41} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \theta_{i2} \\ \theta_{i3} \\ \theta_{i4} \end{pmatrix}, \begin{pmatrix} V_{i2} & se_{i1}^2 & se_{i1}^2 \\ se_{i1}^2 & V_{i3} & se_{i1}^2 \\ se_{i1}^2 & se_{i1}^2 & V_{i4} \end{pmatrix}\right)$$

$$\boldsymbol{\theta}_i = \begin{pmatrix} \theta_{i2} \\ \theta_{i3} \\ \theta_{i4} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \Delta_{t_{i2}} - \Delta_{t_{i1}} \\ \Delta_{t_{i3}} - \Delta_{t_{i1}} \\ \Delta_{t_{i4}} - \Delta_{t_{i1}} \end{pmatrix}, \tau^2 \begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix}\right)$$

where V_{ij} is the variance of difference between arms j and 1 and se_{i1} is the standard error of the response in group 1.

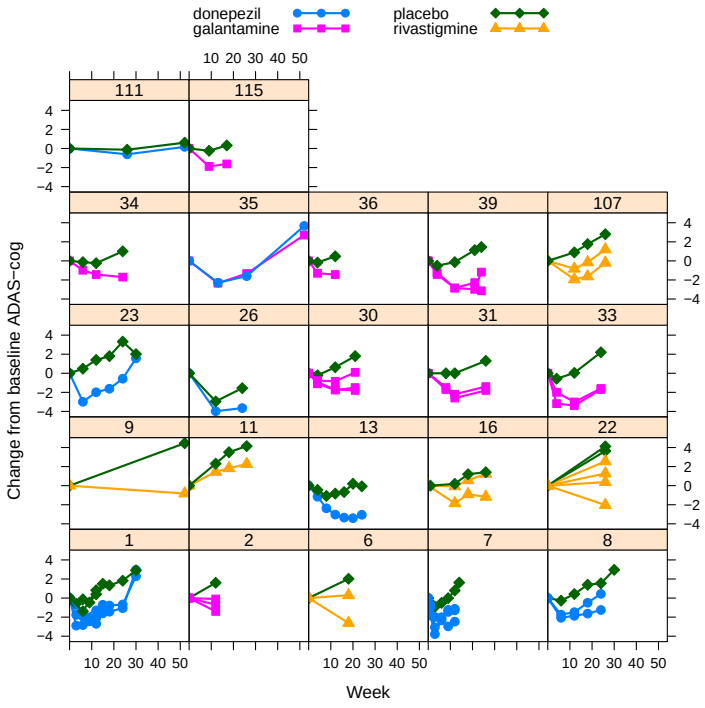
How does Network meta-analysis compare to MBMA?

Both approaches account for multiple treatment comparisions and make use of indirect comparisons to make inferences.

NMA has been developed with regard to the theory and there are many examples in the peer-reviewed literature.

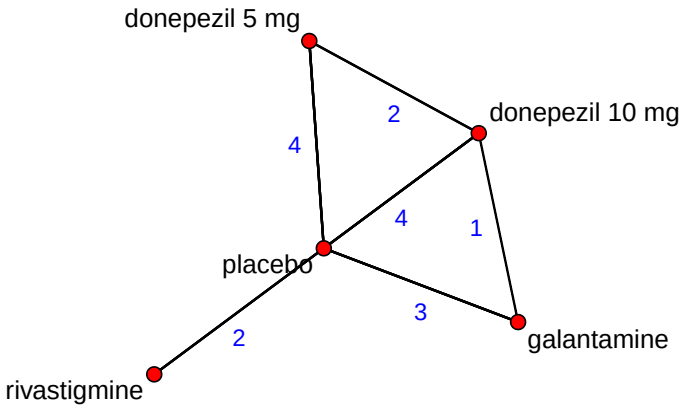
Let's look at an example . . . .

Recall the ADAS-cog data



Recall the ADAS-cog data

We will focus on the 12 studies that evaluated registered doses of donpezil, galantamine and rivastigmine and have Month 6 data.



There are a number of potential objectives of interest

- What are the relative effects of marketed doses at Month 6?
  - Donepezil 10 mg vs. placebo
  - Donepezil 10 mg vs. Galantamine 24 mg
  - Donepezil 10 mg. vs. Rivastigmine
- What are the effects at other time points (e.g., Months 1, 3, 12)?
- What are the effects at other doses?
- What is the rate of change (disease progression) in this patient population?
- What is the effect and/or rate of change in different populations?

For the ADAS-cog data, we will use a model very similar to the longitudinal dose-response model published by Ito et al. (2009):

$$Y_{ijk} = \eta_{1i} + \alpha_i \cdot t_{ijk} + \beta \cdot \left(e^{-Keq \cdot t_{ijk}} - e^{-Kel \cdot t_{ijk}} \right) + \frac{\text{Emax}_j \cdot \left(\frac{d_{ijk}}{RD_j} \right)^{\theta_j} \cdot t_{ijk}}{\text{ET50}_j \cdot e^{\eta_{3i}} + t_{ijk}} + \epsilon_{ijk}$$

where

$$\alpha_i = \alpha + \eta_{2i}$$

i indexes study, *j* indexes treatment arm, *k* indexes visit

Emax_j, ET50_j, and θ_j are drug-specific parameters.

$$\eta_i = (\eta_{1i}, \eta_{2i}, \eta_{3i})^T \sim MVN(\mathbf{0}, \Omega)$$
$$\epsilon_{ijk} \sim N(0, \sigma^2 / N_{ijk})$$

Estimating differences with the MBMA model

The traditional and network meta-analysis models, model the difference between groups directly.

With the MBMA model, we derive these from the model. For example, at Month 6

$$\delta_{i,D10} = \frac{\text{Emax}_D \cdot \left(\frac{10}{5} \right)^{\theta_D} \cdot 6}{\text{ET50}_D \cdot e^{\eta_{3i}} + 6}$$

and

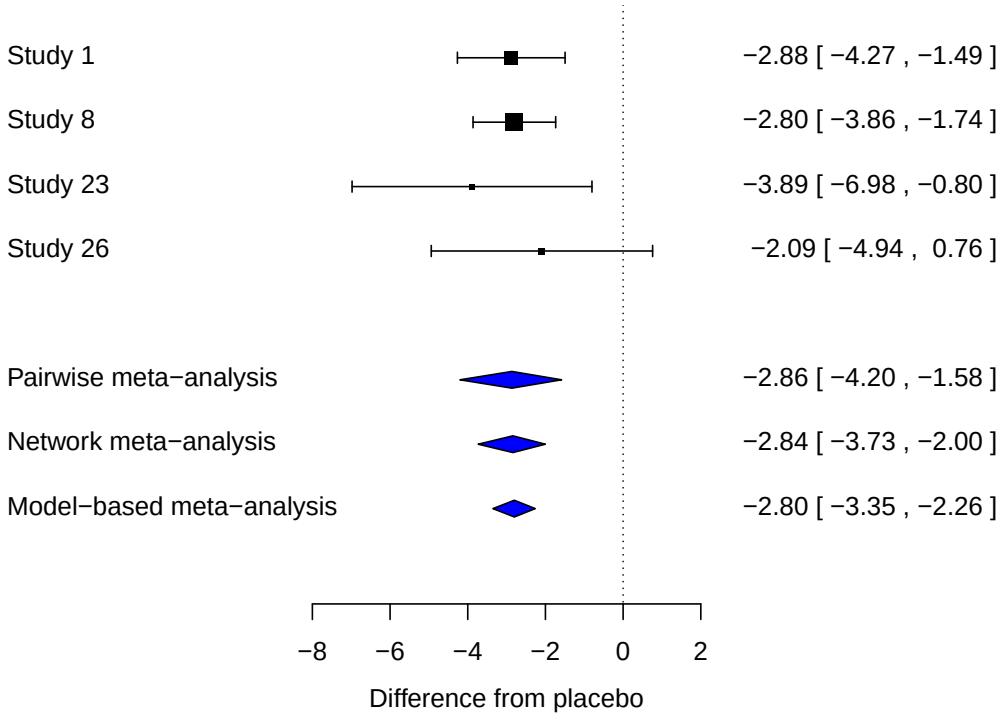
$$\Delta_{D10} = median[\delta_{i,D10}] = \frac{\text{Emax}_D \cdot \left(\frac{10}{5} \right)^{\theta_D} \cdot 6}{\text{ET50}_D + 6}$$

Analysis of ADAS-cog data

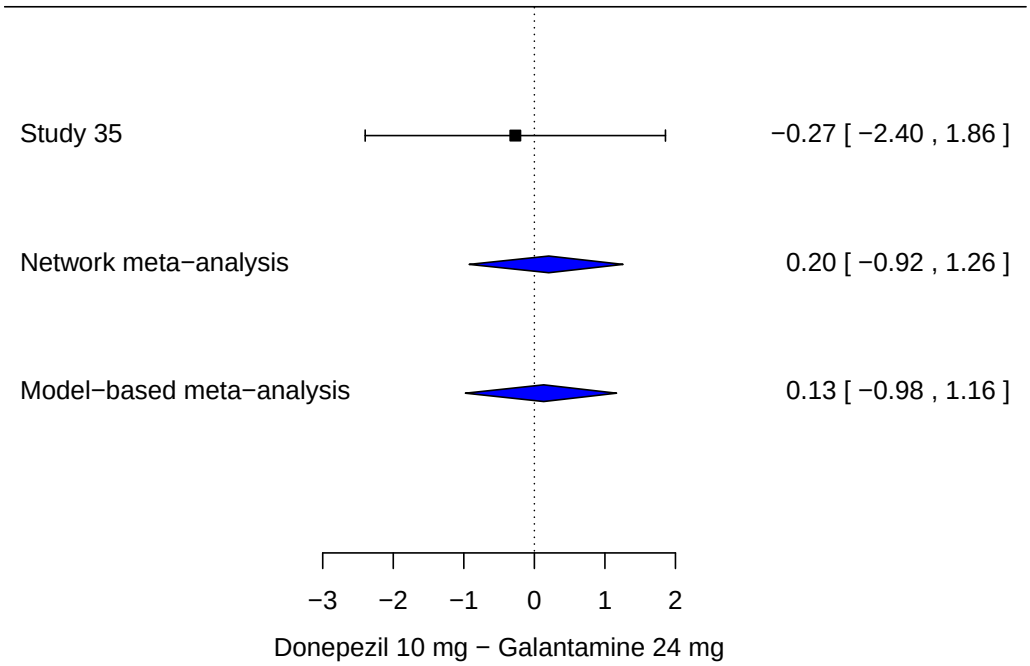
Compare estimates and uncertainty from various models and approaches

- Donepezil 10 mg vs. placebo at Month 6
  - Direct and indirect evidence of effect
- Donepezil 10 mg vs Galantamine 24 mg at Month 6
  - Limited direct evidence + indirect evidence
- Donepezil 10 mg vs. Rivastigmine at Month 6
  - Indirect evidence only
- Effects of donepezil at doses other than 5 and 10 mg
  - Indirect evidence via the model

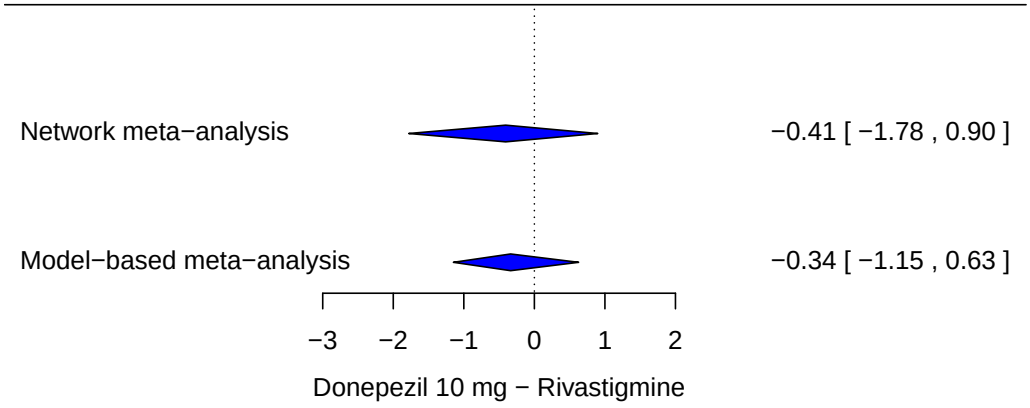
Donepezil 10 mg vs. Placebo at Month 6



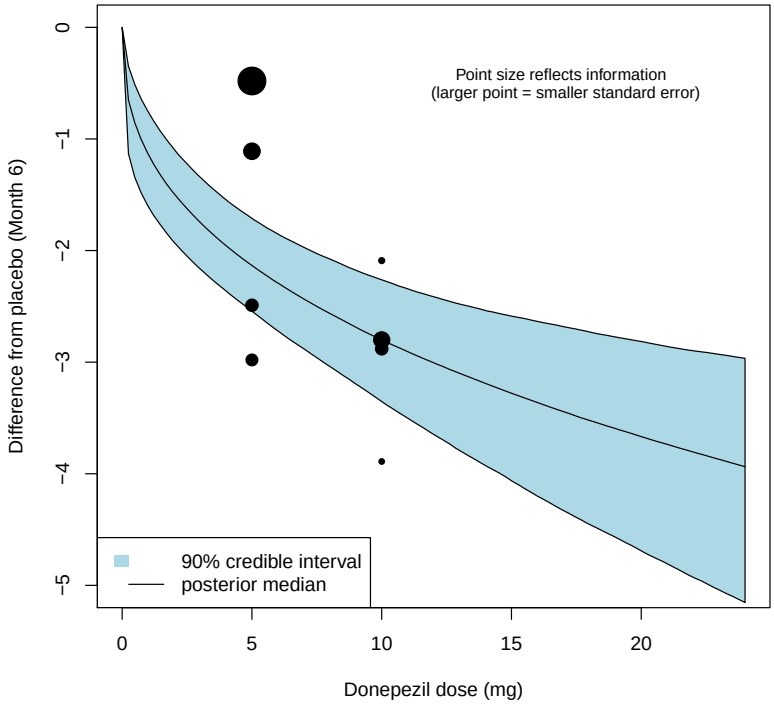
Donepezil 10 mg vs. Galantamine at Month 6



Donepezil 10 mg vs. Rivastigmine at Month 6



Donepezil dose response at Month 6



COMPLETE COURSEWARE LICENSE AGREEMENT I

COMPLETE COURSEWARE LICENSE AGREEMENT II

"You" means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the Work, or who has received express permission from the Licensor to exercise rights under this License despite a previous violation.
"Publicly Perform" means to perform public recitations of the Work and to communicate to the public those public recitations, by any means or process, including by wire or wireless means or public digital performances; to make available to the public Works in such a way that members of the public may access these Works from a place and at a place individually chosen by them; to perform the Work to the public by any means or process and the communication to the public of the performances of the Work, including by public digital performance; to broadcast and rebroadcast the Work by any means including signs, sounds or images.
"Reproduce" means to make copies of the Work by any means including without limitation by sound or visual recordings and the right of fixation and reproducing fixations of the Work, including storage of a protected performance or phonogram in digital form or other electronic medium.
"Commercial" means any copy, application or Adaptation of the Work for the direct or indirect purpose of revenue generation
"Non-Commercial" means any copy, application or Adaptation of the Work for purposes that are not directly or indirectly related to revenue generation.

2. Fair Dealing Rights. Nothing in this License is intended to reduce, limit, or restrict any uses free from copyright or rights arising from limitations or exceptions that are provided for in connection with the copyright protection under copyright law or other applicable laws.

3. License Grant. Subject to the terms and conditions of this License, Licensor hereby grants You fair use of this copyrighted Work for *noncommercial educational* purposes or for commercial and noncommercial *research* purposes. For example, You may adapt example code sets (the Work) and apply them to a commercial research project.

4. Restrictions. The license granted in Section 3 above is expressly made subject to and limited by the following restrictions: Under the terms of this License, You may not Distribute, Reproduce, or Publicly Perform the Work. For example, You may not use courseware (the Work or any adaptations) to create a new public training course, whether commercial or non-commercial.
You may not implicitly or explicitly assert or imply any connection with, sponsorship or endorsement by the Original Author, Licensor and/or Attribution Parties, as appropriate, of You or Your use of the Work, without the separate, express prior written permission of the Original Author, Licensor and/or Attribution Parties.

COMPLETE COURSEWARE LICENSE AGREEMENT III

For the avoidance of doubt:
Non-waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme cannot be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License;
Waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme can be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License if Your exercise of such rights is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(b) and otherwise waives the right to collect royalties through any statutory or compulsory licensing scheme; and,
Voluntary License Schemes. The Licensor reserves the right to collect royalties, whether individually or, in the event that the Licensor is a member of a collecting society that administers voluntary licensing schemes, via that society, from any exercise by You of the rights granted under this License that is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(c).

5. Representations, Warranties and Disclaimer
UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING, LICENSOR OFFERS THE WORK AS-IS AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE WORK, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTIBILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OF ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE. SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO SUCH EXCLUSION MAY NOT APPLY TO YOU.

6. Limitation on Liability. EXCEPT TO THE EXTENT REQUIRED BY APPLICABLE LAW, IN NO EVENT WILL LICENSOR BE LIABLE TO YOU ON ANY LEGAL THEORY FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR EXEMPLARY DAMAGES ARISING OUT OF THIS LICENSE OR THE USE OF THE WORK, EVEN IF LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

7. Termination This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License. Subject to the above terms and conditions, the license granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different license terms

COMPLETE COURSEWARE LICENSE AGREEMENT IV

or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this License (or any other license that has been, or is required to be, granted under the terms of this License), and this License will continue in full force and effect unless terminated as stated above.

8. Miscellaneous If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable. No term or provision of this License shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent. This License constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This License may not be modified without the mutual written agreement of the Licensor and You. The rights granted under, and the subject matter referenced, in this License were drafted utilizing the terminology of the Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979), the Rome Convention of 1961, the WIPO Copyright Treaty of 1996, the WIPO Performances and Phonograms Treaty of 1996 and the Universal Copyright Convention (as revised on July 24, 1971). These rights and subject matter take effect in the relevant jurisdiction in which the License terms are sought to be enforced according to the corresponding provisions of the implementation of those treaty provisions in the applicable national law. If the standard suite of rights granted under applicable copyright law includes additional rights not granted under this License, such additional rights are deemed to be included in the License; this License is not intended to restrict the license of any rights under applicable law.

References I

[AF10] Jae Eun Ahn and Jonathan L French.
Longitudinal aggregate data model-based meta-analysis with nonmem: approaches to handling within treatment arm correlation.
J Pharmacokinet Pharmacodyn, 37(2):179–201, Apr 2010.

[AHS08] Lidia R Arends, M G Myriam Hunink, and Theo Stijnen.
Meta-analysis of summary survival curve data.
Stat Med, 27(22):4381–4396, Sep 2008.

[BG98] W. J. Bachman and W. R. Gillespie.
"Truncated sigmoid emax models": A reparameteriation of the sigmoid emax model for use with truncated pk/pd data.
Clinical Pharmacology and Therapeutics, 63(2):199, 1998.

[CPR10] Tianxi Cai, Layla Parast, and Louise Ryan.
Meta-analysis for rare events.
Statistics in Medicine, 29(20):2078–2089, September 2010.

[Dea94] K B Dear.
Iterative generalized least squares for meta-analysis of survival data at multiple times.
Biometrics, 50(4):989–1002, Dec 1994.

References II

[DF12] Ying Ding and Haoda Fu.
Bayesian indirect and mixed treatment comparisons across longitudinal time points.
Stat Med, Dec 2012.

[DHL+12] I Demin, B Hamrén, O Luttringer, G Pillai, and T Jung.
Longitudinal model-based meta-analysis in rheumatoid arthritis: an application toward model-based drug development.
Clin Pharmacol Ther, 92(3):352–9, Sep 2012.

[DL86] R DerSimonian and N Laird.
Meta-analysis in clinical trials.
Control Clin Trials, 7(3):177–188, Sep 1986.

[GAOW12] Patricia Guyot, AE Ades, Mario JNM Ouwens, and Nicky J Welton.
Enhanced secondary analysis of survival data: reconstructing the data from published kaplan-meier survival curves.
BMC Medical Research Methodology, 12(9), 2012.

[God07] C. J. Godfrey.
Mixed effects modeling analysis of count data.
In E. I. Ette and P. J. Williams, editors, *Pharmacometrics: the science of quantitative pharmacology*, pages 699–721. Wiley, Hoboken, 2007.

References III

[GRIG09] William R. Gillespie, James A Rogers, Kaori Ito, and Marc R Gastonguay.
Population dose-response model for adas-cog scores in patients with alzheimers disease by meta-analysis of a mixture of summary and individual data.
In *American Conference on Pharmacometrics*, Mashantucket, CT, October 2009.

[GRP+13] Jorge Luiz Gross, James Rogers, Daniel Polhamus, William Gillespie, Christian Friedrich, Yan Gong, Brigitta Ursula Monz, Sanjay Patel, Alexander Staab, and Silke Retlich.
A novel model-based meta-analysis to indirectly estimate the comparative efficacy of two medications: an example using DPP-4 inhibitors, sitagliptin and linagliptin, in treatment of type 2 diabetes mellitus.
BMJ Open, 3:e001844 (http://bmjopen.bmj.com/content/3/3/e001844), 2013.

[IAC+09] K Ito, S Ahadieh, B Corrigan, J French, T Fullerton, T Tensfeldt, and AD Disease Working Group.
A disease progression meta-analysis model in Alzheimer's disease.
Alzheimer's & Dementia, Accepted May 13, 2009.

[IRA+08] K. Ito, M Rosario, S Ahadieh, B. W. Corrigan, J French, T Fullerton, R Zhang, P Lockwood, Q Zhao, R Qiu, T Russell, and T Tensfeldt.
A disease progression meta-analysis model for cognitive deterioration with alzheimer's disease.
Clin Pharmacol Ther, 83:S40, 2008.

References IV

[JSS+02] J.A.Berlin, J. Santanna, C.H. Schmid, L.A. Szczech, and H.I. Feldman.
Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head.
Statistics in Medicine, 21:371–387, 2002.

[LKH+07] R. L. Lalonde, K. G. Kowalski, M. M. Hutmacher, W. Ewy, D. J. Nichols, P. A. Milligan, B. W. Corrigan, P. A. Lockwood, S. A. Marshall, L. J. Benincosa, T. G. Tensfeldt, K. Parivar, M. Amantea, P. Glue, H. Koide, and R. Miller.
Model-based drug development.
Clinical Pharmacology and Therapeutics, 82(1):21–32, 2007.

[LRNV10] K T Luu, S R Raber, D J Nickens, and P Vicini.
A model-based meta-analysis of the effect of latanoprost chronotherapy on the circadian intraocular pressure of patients with glaucoma or ocular hypertension.
Clin Pharmacol Ther, 87(4):421–5, Apr 2010.

[LWRP01] D. J. Lunn, J. Wakefield, and A. Racine-Poon.
Cumulative logit models for ordinal data: a case study involving allergic rhinitis severity scores.
Stat Med, 20(15):2261–85, 2001.

[LYC+07] L. Li, M. Yu, R. Chin, A. Lucksiri, D. A. Flockhart, and S. D. Hall.
Drug-drug interaction prediction: a bayesian meta-analysis approach.
Statistics in Medicine, 26(20):3700–3721, 2007.

References V

[MBD11] J W Mandema, R A Boyd, and L A DiCarlo.
Therapeutic index of anticoagulants for prevention of venous thromboembolism following orthopedic surgery: a dose-response meta-analysis.
Clin Pharmacol Ther, 90(6):820–7, Dec 2011.

[MCA05] J. W. Mandema, E. Cox, and J. Alderman.
Therapeutic benefit of eletriptan compared to sumatriptan for the acute relief of migraine pain–results of a model-based meta-analysis that accounts for encapsulation.
Cephalalgia, 25(9):715–725, 2005.

[MHW+05] J. W. Mandema, D. Hermann, W. Wang, T. Sheiner, M. Milad, R. Bakker-Arkema, and D. Hartman.
Model-based development of gemcabene, a new lipid-altering agent.
AAPS J, 7(3):E513–E522, 2005.

[MSBG11] J W Mandema, D H Salinger, S W Baumgartner, and M A Gibbs.
A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis.
Clin Pharmacol Ther, 90(6):828–35, Dec 2011.

[OPJ11] Mario J N M Ouwens, Zoe Philips, and Jeroen P Jansen.
Network meta-analysis of parametric survival curves.
Research Synthesis Methods, 2011.

References VI

[OS98] I. Olkin and A. Sampson.
Comparison of meta-analysis versus analysis of variance of individual patient data.
Biometrics, 54:317–322, 1998.

[PTS98] M K Parmar, V Torri, and L Stewart.
Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints.
Stat Med, 17(24):2815–2834, Dec 1998.

[RGB86] J. Robins, S. Greenland, and N. Breslow.
A general estimator for the variance of the mantel-haenszel odds ratio.
Am. J. Epi, 124:719–723, 1986.

[RLP+] J. Rogers, P. Lockwood, D. Polhamus, Y. Brault, A. Desmet, K. Ito, K. Romero, R. Qiu, B. Gillespie, B. Corrigan, and M. Gastonguay.
Model-based analysis to support strategic decision making: A case study from the development of a 5HT6 antagonist for the treatment of alzheimer's disease
`http://metrumrg.com/images/stories/publications/icad-sam531.pdf`.
Alzheimer's Association International Conference(AAIC) 2012 in Vancouver, British Columbia, Canada, July 14-19, 2012.

References VII

[RPG+] J. Rogers, D. Polhamus, W. Gillespie, C. Friedrich, A. Staab, and S. Retlich.
A model-based meta-analysis comparison of the effects of linagliptin and sitagliptin on HbA1c levels in patients with type 2 diabetes mellitus.
PAGE 2012.

[RPG+12] James A Rogers, Daniel Polhamus, William R Gillespie, Kaori Ito, Klaus Romero, Ruolun Qiu, Diane Stephenson, Marc R Gastonguay, and Brian Corrigan.
Combining patient-level and summary-level data for alzheimer's disease modeling and simulation: a beta regression meta-analysis.
J Pharmacokinet Pharmacodyn, Jul 2012.

[RSB05] H. Rothstein, A.J. Sutton, and M. Borenstein, editors.
Publication Bias in Meta-Analysis.
John Wiley & Sons, Chichester, UK, 2005.

[SAI11] Georgia Salanti, A E Ades, and John P A Ioannidis.
Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.
J Clin Epidemiol, 64(2):163–71, Feb 2011.

[SAJ+00] A.J. Sutton, K. R. Abrams, D.R. Jone, T.A. Sheldon, and F Song.
Methods for Meta-Analysis in Medical Research.
Wiley, Chichester, UK, 2000.

References VIII

[SBD97] L. B. Sheiner, S. L. Beal, and A. Dunne.
Analysis of nonrandomly censored ordered categorical longitudinal data from analgesic trials.
Journal of the American Statistical Association, 92(440):1235–1244, 1997.

[SGT00] D.D. Smith, G.H. Givens, and R.L Tweedie.
Adjustment for publication bias and quality bias using random-effects models.
In D.K. Stangl and D. A. Berry, editors, *Meta-Analysis in Medicine and Health Policy*, pages 277–304. Dekker, 2000.

[vHAS02] Hans C van Houwelingen, Lidia R Arends, and Theo Stijnen.
Advanced methods in meta-analysis: multivariate approach and meta-regression.
Stat Med, 21(4):589–624, Feb 2002.

[Vie10] W. Viechtbauer.
Conducting meta-analysis in r with the metafor package.
Journal of Statistical Software, 36(3):1–48, 2010.

[WHW08a] White, Higgins, and Wood.
Allowing for uncertainty due to missing data in meta-analysis – part 1: Two-stage methods.
Statistics in Medicine, 27:711–727, 2008.

References IX

[WHW08b] White, Higgins, and Wood.
Allowing for uncertainty due to missing data in meta-analysis – part 2: Hierarchical models.
Statistics in Medicine, 27:728–745, 2008.

[YL08] Y. Yuan and R. J. Little.
Meta-analysis of studies with missing data.
Biometrics, 2008.