METRUM
RESEARCH GROUP

MBMA of combined individual and aggregate data: strategies and issues

Bill Gillespie

Metrum Research Group

26 October 2016

MBMA of combined individual and aggregate data: strategies and issues

- Why combine individual and aggregate data?
  - Pros/cons of aggregate data (AD) MA
  - Pros/cons of individual patient data (IPD) MA
- Methods
  - Focus on nonlinear and longitudinal data models
  - Two-stage approach
  - Hierarchical/multilevel modeling approaches
  - Analytic approximation of aggregate data likelihood
  - Imputation of aggregate data likelihood by simulation
- Closing discussion

Pros/cons of aggregate data (AD) MA

- Pros
  - Relatively easy access to data from public sources

Pros/cons of aggregate data (AD) MA

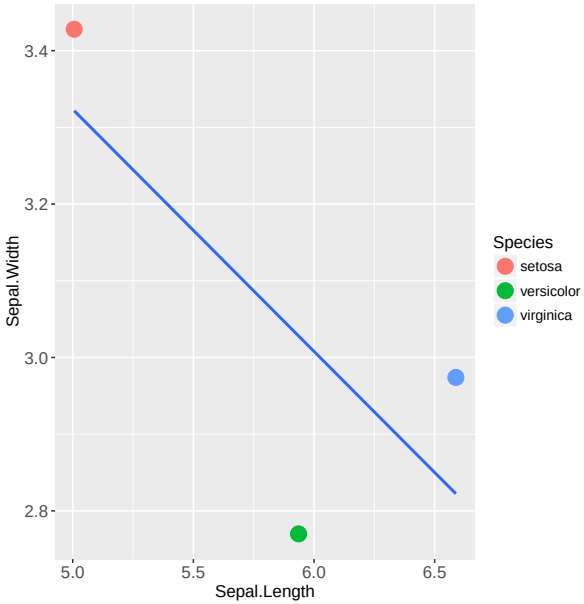
- Pros
  - Relatively easy access to data from public sources
- Cons
  - Not well-suited for inferences about patient-level covariates.
    - Ecological bias/fallacy
    - Aggregate covariate data describes a narrower range of values than individual covariate data
  - For nonlinear models the relationship between the dependent variable and the covariates, e.g., dose or time, is not described by the same function for AD and IPD.
  - Usually no info about correlations among multiple outcomes
  - Model usually not suitable for prediction/simulation of individual outcomes

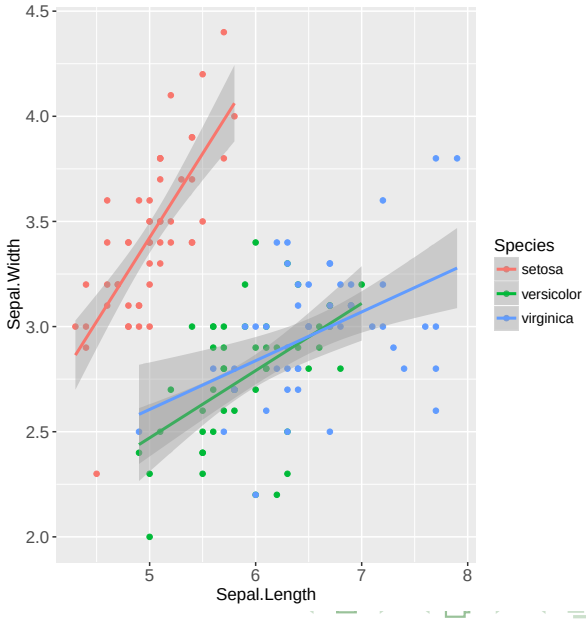
Ecological bias/fallacy

- Errors resulting from attempting to infer individual properties based on aggregate data

Ecological bias/fallacy

- Errors resulting from attempting to infer individual properties based on aggregate data
- Simpson's paradox
  - May happen when trial outcomes differ for reasons not captured in the model or even identifiable with AD, e.g., when the model does not include influential covariates (confounding).





For nonlinear models IPD and AD may not be described by the same function

- 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.

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:

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

Pros/cons of individual patient data (IPD) MA

- Pros
  - “Gold standard”, particularly for longitudinal data
  - Can support inferences about:
    - Patient-level covariates
    - Correlations among outcomes
  - Model is suitable for prediction/simulation of individual outcomes.

Pros/cons of individual patient data (IPD) MA

- Pros
  - “Gold standard”, particularly for longitudinal data
  - Can support inferences about:
    - Patient-level covariates
    - Correlations among outcomes
  - Model is suitable for prediction/simulation of individual outcomes.
- Cons
  - Access issues
    - IPD may not be obtainable for all trials of interest.
    - May introduce a form of selection bias
  - Potentially much more time consuming
    - Mainly due to delay in obtaining data from external sources
    - May not be feasible for time critical decision making
  - More computationally demanding

Why combine individual and aggregate data?

Addition of AD to enhance/extend inferences from IPD analysis

- Good reasons
  - Indirect comparisons of treatment effects
    - Particularly when comparators are only available in AD
  - Quantifying effects of other group-level covariates (when AD is available for the relevant groups)
  - Quantifying inter-trial variability
  - Improving precision of some model parameter estimates
- Not-so-good reasons
  - Quantifying effects of patient-level covariates

Why combine individual and aggregate data?

Addition of IPD to enhance/extend inferences from AD analysis

- IPD required to inform correlations among individual-level outcomes and covariates
- IPD required to quantify effects of patient-level covariates

Methods

- Focus on nonlinear and longitudinal data models
- Methods for linear models are well-covered in the statistics literature [1, 2, 3, 4, 5, 6].
- Two-stage approach
- Hierarchical/multilevel modeling approaches, e.g., hierarchical related regression
- Approximating or imputing the AD likelihood based on the IPD model
  - Analytic approximation of AD likelihood
  - Imputation of AD likelihood by simulation

Two-stage approach

1. For IPD calculate AD statistics
2. Apply suitable AD meta-analysis method [3]

- Same limitations as AD meta-analysis
- Probably sufficient if the primary objective is treatment comparison.

Hierarchical/multilevel modeling approaches

- Model with a nested hierarchy: observation within patient within study [7, 8, 9, 10, 11]
  - IPD = function of individual and observation level parameters
  - AD = function of study level parameters
- Hierarchical related regression [8, 9, 10, 11]
  - Variation in which related but somewhat different models are used for IPD and AD.
  - IPD model includes patient level covariates.
  - AD model does not.
  - Both models share same treatment effect parameter(s).

Approximating or imputing the AD likelihood based on the IPD model

- A conceptually attractive approach to combined analysis of IPD and AD is use a common IPD model to analyze all of the data.
  - Consider the individual measurements contributing to the AD as missing data.
  - Perform a Bayesian analysis in which all unknowns including those missing individual measurements are parameters of the posterior distribution.

Approximating or imputing the AD likelihood based on the IPD model

- A conceptually attractive approach to combined analysis of IPD and AD is use a common IPD model to analyze all of the data.
  - Consider the individual measurements contributing to the AD as missing data.
  - Perform a Bayesian analysis in which all unknowns including those missing individual measurements are parameters of the posterior distribution.
- Very challenging computational problem—usually impractical due to:
  - Massive increase in dimensionality of the joint posterior distribution.
  - Large number of poorly identifiable parameters.
- So we look to approximations of such an approach.

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.
- The AD model (likelihood) is derived from the IPD model [12, 13, 14].
  - Likelihoods for both sample means and standard deviations

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.
- The AD model (likelihood) is derived from the IPD model [12, 13, 14].
  - Likelihoods for both sample means and standard deviations
  - Both the inter-arm and residual variances should be adjusted for sample size.

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.
- The AD model (likelihood) is derived from the IPD model [12, 13, 14].
  - Likelihoods for both sample means and standard deviations
  - Both the inter-arm and residual variances should be adjusted for sample size.
  - In the case where the individual data model is linear with respect to normally-distributed random effects, the derivation is exact.

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.
- The AD model (likelihood) is derived from the IPD model [12, 13, 14].
  - Likelihoods for both sample means and standard deviations
  - Both the inter-arm and residual variances should be adjusted for sample size.
  - In the case where the individual data model is linear with respect to normally-distributed random effects, the derivation is exact.
  - For the general nonlinear case it is an approximation in 3 senses.
    - The sampling distributions are approximated as normal for the mean and gamma for the variance.
    - The AD model is approximated using the IPD 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.

Analytic approximation of aggregate data likelihood

- We begin with a hierarchical model for IPD with with 3 levels of variation: inter-trial, inter-arm, and residual.
- The AD model (likelihood) is derived from the IPD model [12, 13, 14].
  - Likelihoods for both sample means and standard deviations
  - Both the inter-arm and residual variances should be adjusted for sample size.
  - In the case where the individual data model is linear with respect to normally-distributed random effects, the derivation is exact.
  - For the general nonlinear case it is an approximation in 3 senses.
    - The sampling distributions are approximated as normal for the mean and gamma for the variance.
    - The AD model is approximated using the IPD 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.
- May be implemented in the usual tools, e.g., NONMEM, BUGS, Stan, etc.

J Pharmacokinet Pharmacodyn
DOI 10.1007/s10928-012-9263-3

ORIGINAL PAPER

Combining patient-level and summary-level data for Alzheimer's disease modeling and simulation: a beta regression meta-analysis

James A. Rogers · Daniel Polhamus · William R. Gillespie · Kaori Ito · Klaus Romero · Ruolun Qiu · Diane Stephenson · Marc R. Gastonguay · Brian Corrigan

Objective:

Develop a model to describe the longitudinal progression of ADAS-cog in Alzheimer's disease patients in both natural history and randomized clinical trial settings, utilizing both IPD and AD.

MBMA of longitudinal ADAS-cog IPD and AD

- IPD
  - CAMD database: 3,223 patients
  - ADNI database: 186 patients
- AD
  - Extracted from 73 literature references: 17,235 patients

MBMA of longitudinal ADAS-cog IPD and AD

- IPD
  - CAMD database: 3,223 patients
  - ADNI database: 186 patients
- AD
  - Extracted from 73 literature references: 17,235 patients
- IPD was most informative about disease progression and patient-level covariates.
- AD contributed information required for inferences about the effects of galantamine, donepezil and rivastigmine.

MBMA of longitudinal ADAS-cog IPD and AD

- IPD
  - CAMD database: 3,223 patients
  - ADNI database: 186 patients
- AD
  - Extracted from 73 literature references: 17,235 patients
- IPD was most informative about disease progression and patient-level covariates.
- AD contributed information required for inferences about the effects of galantamine, donepezil and rivastigmine.
- Resulted in a model suitable for:
  - Simulating individual patient outcomes, e.g., clinical trial simulations,
  - Making inferences about the comparative efficacy of galantamine, donepezil and rivastigmine.

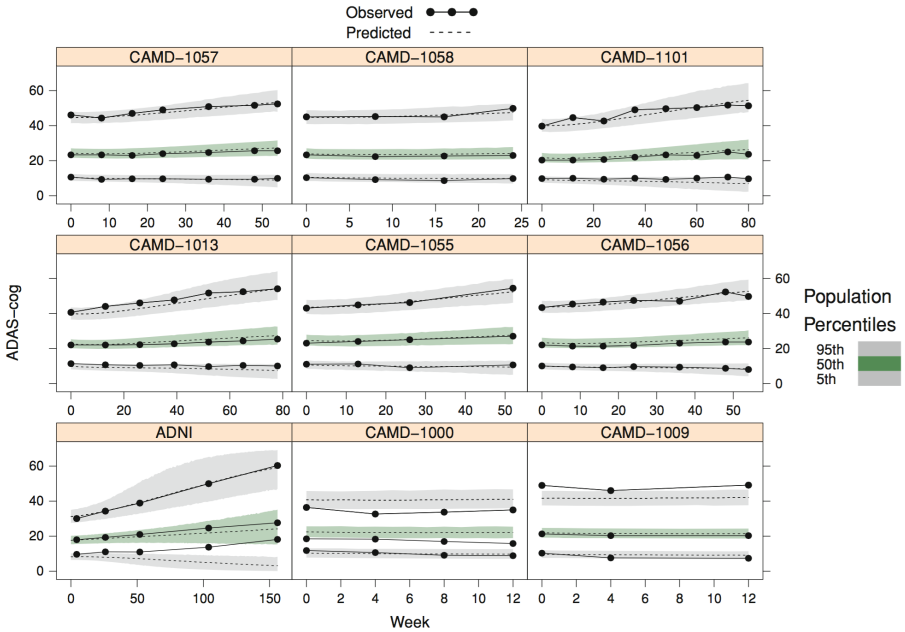
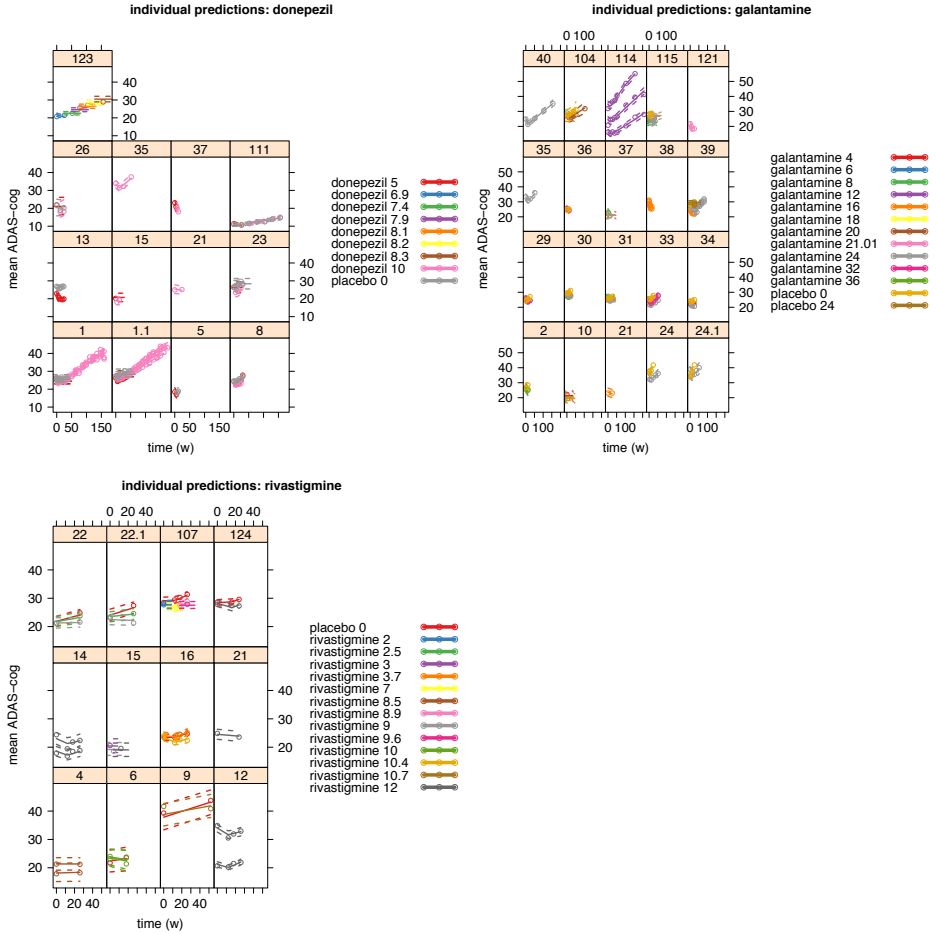


Fig. 7 Plot of unconditional predictive checks for sample population percentiles of ADNI and CAMD studies



$$\log\left(\left(\frac{x^{\zeta}}{1-x^{\zeta}}\right)^{1/\zeta}\right) = \eta_{jk} + \alpha_{jk}\, t_{ijk} + E_{\text{placebo}}\left(t_{ijk}\right) + E_{\text{drug}}\left(t_{ijk}, D_{ijk}\right)$$

$$\frac{\text{ADAS-cog}_{ijk}}{70} \sim \text{Beta}\left(\theta_{ijk}\tau, \left(1-\theta_{ijk}\right)\tau\right)$$

Imputation of aggregate data likelihood by simulation

- An approach that may more closely approximate the AD likelihood is to impute it from simulations of individual data [15, 16, 17, 18].
- For each treatment arm suppose you have a set of means $\overline{y}_i, i = 1, 2, \ldots, n_T$ of longitudinal data for N individuals.
- Impute the joint likelihood of the $\overline{y}_i$'s by:
  - Simulating individual data for a large number of individuals,
  - Calculating the mean vector M_s and covariance matrix Σ_s of the simulated values,
  - Approximating the joint likelihood of $\overline{y}_i, i = 1, 2, \ldots, n_T$ as multivariate normal: $N\left(M_s, \frac{\Sigma_s}{N}\right)$

Hierarchical expectation propagation for Bayesian aggregation of average data*

Sebastian Weber† Andrew Gelman‡ Bob Carpenter ‡ Daniel Lee‡
Michael Betancourt§ Aki Vehtari¶ Amy Racine†

26 Oct 2015

https://arxiv.org/abs/1602.02055

- Details methodology for joint analysis of IPD and AD from one study each [17, 18].
- Readily generalized to multiple IPD and AD studies.
- The AD data likelihood is imputed by simulation.
- That is embedded within an overall Bayesian analysis method involving:
  - Analysis of IPD by HMC (Stan),
  - Analysis of AD data by importance sampling, and
  - Iterative updating of both IPD and AD analyses by expectation propagation.

Weber et al approach to imputation of aggregate data likelihood by simulation

- Simulation studies indicate good performance on a range of problems including PKPD applications requiring numerical solution of ODEs.
- Avoids most of the approximations used for the previously discussed analytic approximation approach.
- The main remaining approximation is use of the multivariate normal for the joint likelihood of longitudinal means.
- Expectation propagation introduces additional approximations, and
- Requires substantial custom programming to implement the expectation propagation and importance sampling methods.

Weber et al approach to imputation of aggregate data likelihood by simulation

Stay tuned: The authors have since developed and tested a version 2.0 that is a simpler implementation within Stan.

- No expectation propagation.
- No importance sampling.

Summing up: Role of MBMA of combined AD and IPD

- Addition of AD to enhance/extend inferences from IPD analysis is most valuable for:
  - Indirect comparisons of treatment effects when key comparators are not represented in IPD.
  - Quantifying inter-trial variability

Summing up: Role of MBMA of combined AD and IPD

- Addition of AD to enhance/extend inferences from IPD analysis is most valuable for:
  - Indirect comparisons of treatment effects when key comparators are not represented in IPD.
  - Quantifying inter-trial variability
- Addition of IPD to enhance/extend inferences from AD analysis is most valuable for
  - Estimating correlations among individual-level outcomes and covariates.
  - Quantifying effects of patient-level covariates

Summing up: Methods for MBMA of combined AD and IPD

- Analytic approximation of aggregate data likelihood
  - Easiest to implement with familiar tools.
  - Potentially questionable approximation of AD model with IPD model.
  - Could cause unacceptable estimation error/bias, particularly with highly nonlinear models.

Summing up: Methods for MBMA of combined AD and IPD

- Analytic approximation of aggregate data likelihood
  - Easiest to implement with familiar tools.
  - Potentially questionable approximation of AD model with IPD model.
  - Could cause unacceptable estimation error/bias, particularly with highly nonlinear models.
- Imputation of aggregate data likelihood by simulation
  - More plausible approximation of AD likelihood.
  - Harder to implement with standard PMX tools.

Summing up: Methods for MBMA of combined AD and IPD

- Analytic approximation of aggregate data likelihood
  - Easiest to implement with familiar tools.
  - Potentially questionable approximation of AD model with IPD model.
  - Could cause unacceptable estimation error/bias, particularly with highly nonlinear models.
- Imputation of aggregate data likelihood by simulation
  - More plausible approximation of AD likelihood.
  - Harder to implement with standard PMX tools.
- A hierarchical related regression (HRR) approach is applicable to both methods
  - May be advisable to exclude patient-level covariates from AD model to reduce risk of ecological bias.

Summing up: Methods for MBMA of combined AD and IPD

- Analytic approximation of aggregate data likelihood
  - Easiest to implement with familiar tools.
  - Potentially questionable approximation of AD model with IPD model.
  - Could cause unacceptable estimation error/bias, particularly with highly nonlinear models.
- Imputation of aggregate data likelihood by simulation
  - More plausible approximation of AD likelihood.
  - Harder to implement with standard PMX tools.
- A hierarchical related regression (HRR) approach is applicable to both methods
  - May be advisable to exclude patient-level covariates from AD model to reduce risk of ecological bias.
- The absolute and relative performance of these methods remains an open research question.

Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away

Apples and oranges (and pears, oh my!): The search for moderators in meta-analysis

Apples and apples or apples and oranges? A meta-analysis of objective and subjective measures of salesperson performance

Apples, oranges, and placebos: Heterogeneity in a meta-analysis of placebo effects

Meta-analysis: Can we mix apples and oranges?

Multivariate meta-analysis: modelling the heterogeneity mixing apples and oranges; dangerous or delicious?

Can a meta-analysis that mixes apples with oranges be used to demonstrate that levosimendan reduces mortality after coronary revascularization?

Can a Meta-Analysis That Mixes Apples With Oranges Be Used to Demonstrate That Pancreatic Enzymes Do Not Decrease Abdominal Pain in Patients With Chronic Pancreatitis?

Meta-analysis: apples and oranges, or fruitless

Meta-analysis of anatomic resection versus nonanatomic resection for hepatocellular carcinoma: are they comparing apples with oranges?

Multivariate meta-analysis: Modeling the heterogeneity; Mixing apples and oranges; dangerous or delicious?

The most critical question when reading a meta-analysis report: Is it comparing apples with apples or apples with oranges?

Meta-analysis of bone marrow transplantation treatment studies: mixing 'apples and oranges'

Comparing Apples to Oranges in Meta-analysis Studies

Mixing apples and oranges and other methodological problems with a meta-analysis of long term psychodynamic psychotherapy

Meta-analysis: Adding apples and oranges?

Meta-analysis of anatomic resection versus nonanatomic resection for hepatocellular carcinoma: are they comparing apples with oranges?

Meta-analysis of survival in mesothelioma: can we mix apples and oranges?

References I

[1] Richard D Riley, Mark C Simmonds, and Maxime P Look.
Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods.
J Clin Epidemiol, 60(5):431–9, May 2007.

[2] Richard D Riley, Susanna R Dodd, Jean V Craig, John R Thompson, and Paula R Williamson.
Meta-analysis of diagnostic test studies using individual patient data and aggregate data.
Statistics in medicine, 27(29):6111–6136, 2008.

[3] Richard D Riley, Paul C Lambert, Jan A Staessen, Jiguang Wang, Francois Gueyffier, Lutgarde Thijs, and Florent Boutitie.
Meta-analysis of continuous outcomes combining individual patient data and aggregate data.
Statistics in medicine, 27(11):1870–1893, 2008.

[4] Richard D Riley, Paul C Lambert, and Ghada Abo-Zaid.
Meta-analysis of individual participant data: rationale, conduct, and reporting.
Bmj, 340:c221, 2010.

[5] Alex J Sutton, Denise Kendrick, and Carol AC Coupland.
Meta-analysis of individual-and aggregate-level data.
Statistics in medicine, 27(5):651–669, 2008.

References II

[6] Howard HZ Thom, Gorana Capkun, Annamaria Cerulli, Richard M Nixon, and Luke S Howard.
Network meta-analysis combining individual patient and aggregate data from a mixture of study designs with an application to pulmonary arterial hypertension.
BMC medical research methodology, 15(1):1, 2015.

[7] Harvey Goldstein, Min Yang, Rumana Omar, Rebecca Turner, and Simon Thompson.
Meta-analysis using multilevel models with an application to the study of class size effects.
Journal of the Royal Statistical Society: Series C (Applied Statistics), 49(3):399–412, 2000.

[8] Christopher Jackson, Nicky Best, and Sylvia Richardson.
Improving ecological inference using individual-level data.
Statistics in medicine, 25(12):2136–2159, 2006.

[9] Christopher Jackson, Sylvia Richardson, et al.
Hierarchical related regression for combining aggregate and individual data in studies of socio-economic disease risk factors.
Journal of the Royal Statistical Society: Series A (Statistics in Society), 171(1):159–178, 2008.

[10] Sebastien J-P A Haneuse and Jonathan C Wakefield.
Hierarchical models for combining ecological and case-control data.
Biometrics, 63(1):128–36, Mar 2007.

References III

[11] Sebastien Haneuse and Scott Bartell.
Designs for the combination of group-and individual-level data.
Epidemiology (Cambridge, Mass.), 22(3):382, 2011.

[12] 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.

[13] 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.
Journal of pharmacokinetics and pharmacodynamics, 39(5):479–498, 2012.

[14] Patanjali Ravva, Mats O Karlsson, and Jonathan L French.
A linearization approach for the model-based analysis of combined aggregate and individual patient data.
Statistics in medicine, 33(9):1460–1476, 2014.

References IV

[15] Yusuke Yamaguchi, Wataru Sakamoto, Shingo Shirahata, and Masashi Goto.
A meta-analysis method based on simulated individual patient data.
In *Bulletin of the ISI 58th World Statistics Congress of the International Statistical Institute*, pages 4702–4707, 2011.

[16] Yusuke Tamaguchi.
Meta-Analysis Methods Based on Simulated Individual Patient Data.
PhD thesis, THE GRADUATE SCHOOL OF ENGINEERING SCIENCE, OSAKA UNIVERSITY, Osaka, Japan, March 2013.

[17] Sebastian Weber, Andrew Gelman, Bob Carpenter, Daniel Lee, Michael Betancourt, Aki Vehtari, and Amy Racine.
Hierarchical expectation propagation for bayesian aggregation of average data.
arXiv preprint arXiv:1602.02055, 2016.

[18] S Weber, A Gelman, B Carpenter, D Lee, F Bois, and A Racine.
Bayesian aggregation of average data in hierarchical models.
Bayes 2015, Basel, Switzerland, May 19–21 2015.