Bayesian Data Analysis, class 1a

Andrew Gelman

Introductory examples:
Soccer ratings and Population toxicokinetics
Bayesian data analysis: (1) Modeling

- “Generative models”
  - Data are a realization from a (multivariate) probability distribution
  - Data vector \( y \), probability model \( p(y|\theta) \), parameter vector \( \theta \)
- Prior distributions
Bayesian data analysis: (1) Modeling

“Generative models”
- Data are a realization from a (multivariate) probability distribution
- Data vector $y$, probability model $p(y|\theta)$, parameter vector $\theta$
- Prior distributions
Bayesian data analysis: (1) Modeling

- "Generative models"
  - Data are a realization from a (multivariate) probability distribution
  - Data vector $y$, probability model $p(y|\theta)$, parameter vector $\theta$

- Prior distributions
  - In Bayes inference, the parameter vector $\theta$ is a realization from a prior distribution, $\pi(\theta|\phi)$
  - Vector of hyperparameters $\phi$ is specified or itself modeled
Bayesian data analysis: (1) Modeling

“Generative models”
- Data are a realization from a (multivariate) probability distribution
- Data vector $y$, probability model $p(y|\theta)$, parameter vector $\theta$

Prior distributions
- In Bayes inference, the parameter vector $\theta$ is a realization from a prior distribution, $p(\theta|\phi)$
- Vector of hyperparameters $\phi$ is specified or itself modeled
Bayesian data analysis: (1) Modeling

- “Generative models”
  - Data are a realization from a (multivariate) probability distribution
  - Data vector $y$, probability model $p(y|\theta)$, parameter vector $\theta$

- Prior distributions
  - In Bayes inference, the parameter vector $\theta$ is a realization from a prior distribution, $p(\theta|\phi)$
  - Vector of hyperparameters $\phi$ is specified or itself modeled
Bayesian data analysis: (1) Modeling

- “Generative models”
  - Data are a realization from a (multivariate) probability distribution
  - Data vector $y$, probability model $p(y|\theta)$, parameter vector $\theta$

- Prior distributions
  - In Bayes inference, the parameter vector $\theta$ is a realization from a prior distribution, $p(\theta|\phi)$
  - Vector of hyperparameters $\phi$ is specified or itself modeled
“Generative models”
- Data are a realization from a (multivariate) probability distribution
- Data vector \( y \), probability model \( p(y|\theta) \), parameter vector \( \theta \)

Prior distributions
- In Bayes inference, the parameter vector \( \theta \) is a realization from a prior distribution, \( p(\theta|\phi) \)
- Vector of hyperparameters \( \phi \) is specified or itself modeled
Bayesian data analysis: (2) Inference

- Inference is represented by a matrix of posterior simulations
  - 1000 simulations of 90 parameters: a $1000 \times 90$ matrix
- Postprocessing
Inference is represented by a matrix of posterior simulations

- 1000 simulations of 90 parameters: a $1000 \times 90$ matrix

Postprocessing

- Inference for qoI's
- Decision analysis
Inference is represented by a matrix of posterior simulations
- 1000 simulations of 90 parameters: a $1000 \times 90$ matrix

Postprocessing
- Inference for qoi's
- Decision analysis
Inference is represented by a matrix of posterior simulations
- 1000 simulations of 90 parameters: a $1000 \times 90$ matrix

Postprocessing
- Inference for qoi’s
- Decision analysis
Inference is represented by a matrix of posterior simulations
   ▶ 1000 simulations of 90 parameters: a $1000 \times 90$ matrix

Postprocessing
   ▶ Inference for qoi’s
   ▶ Decision analysis
Inference is represented by a matrix of posterior simulations
- 1000 simulations of 90 parameters: a $1000 \times 90$ matrix

Postprocessing
- Inference for qoi’s
- Decision analysis
Do the inferences make sense?
Are the model’s predictions consistent with the data?
Not: Is the model true?
Not: What is Pr (model is true)?
Not: Can we “reject” the model?
Expanding the model
Including more data
Bayesian data analysis: (3) Model checking/improvement

- Do the inferences make sense?
- Are the model’s predictions consistent with the data?
- Not: Is the model true?
- Not: What is Pr (model is true)?
- Not: Can we “reject” the model?
- Expanding the model
- Including more data
Bayesian data analysis: (3) Model checking/improvement

- Do the inferences make sense?
- Are the model’s predictions consistent with the data?
  - Not: Is the model true?
  - Not: What is Pr (model is true)?
  - Not: Can we “reject” the model?
- Expanding the model
- Including more data
Bayesian data analysis: (3) Model checking/improvement

- Do the inferences make sense?
- Are the model’s predictions consistent with the data?
- *Not*: Is the model true?
- *Not*: What is Pr (model is true)?
- *Not*: Can we “reject” the model?
- Expanding the model
- Including more data
Bayesian data analysis: (3) Model checking/improvement

- Do the inferences make sense?
- Are the model’s predictions consistent with the data?
  - *Not*: Is the model true?
  - *Not*: What is $\Pr$ (model is true)?
  - *Not*: Can we “reject” the model?
- Expanding the model
- Including more data
Do the inferences make sense?
Are the model’s predictions consistent with the data?
Not: Is the model true?
Not: What is Pr (model is true)?
Not: Can we “reject” the model?
Expanding the model
Including more data
Do the inferences make sense?

Are the model’s predictions consistent with the data?

Not: Is the model true?

Not: What is Pr (model is true)?

Not: Can we “reject” the model?

Expanding the model

Including more data
Bayesian data analysis: (3) Model checking/improvement

- Do the inferences make sense?
- Are the model’s predictions consistent with the data?
  - *Not:* Is the model true?
  - *Not:* What is $\Pr$ (model is true)?
  - *Not:* Can we “reject” the model?
- Expanding the model
- Including more data
Stan goes to the World Cup
The model

- Fit data on signed square roots:
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)
- \( a_i \) and \( a_j \) are “ability parameters”
- \( \sigma_y \) is a scale parameter
- To allow for outliers, use \( t_7 \) instead of normal
- Prior info on abilities:
  \[ a_i \sim N(\mu + b \ast \text{prior.score}_i, \sigma_a^2) \]
- We can set \( \mu = 0 \)
- No further prior info in model
The model

- Fit data on signed square roots:

  \[ y_{ij} = \sqrt{{\text{score differential when team } i \text{ plays team } j}} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)

- \( a_i \) and \( a_j \) are “ability parameters”

- \( \sigma_y \) is a scale parameter

- To allow for outliers, use \( t_7 \) instead of normal

- Prior info on abilities:

  \[ a_i \sim N(\mu + b \times \text{prior.score}_i, \sigma_a^2) \]

- We can set \( \mu = 0 \)

- No further prior info in model
The model

- Fit data on signed square roots:
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)
  - \( a_i \) and \( a_j \) are “ability parameters”
  - \( \sigma_y \) is a scale parameter
  - To allow for outliers, use \( t_7 \) instead of normal
  - Prior info on abilities:
    \[ a_i \sim N(\mu + b \times \text{prior.score}_i, \sigma_a^2) \]
    - We can set \( \mu = 0 \)
    - No further prior info in model
The model

- Fit data on signed square roots:
  
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)

- \( a_i \) and \( a_j \) are “ability parameters”

- \( \sigma_y \) is a scale parameter

- To allow for outliers, use \( t_7 \) instead of normal

- Prior info on abilities:
  
  \[ a_i \sim N(\mu + b \times \text{prior.score}_i, \sigma_a^2) \]

- We can set \( \mu = 0 \)

- No further prior info in model
The model

- Fit data on signed square roots:
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)
- \( a_i \) and \( a_j \) are “ability parameters”
- \( \sigma_y \) is a scale parameter
- To allow for outliers, use \( t_7 \) instead of normal
- Prior info on abilities:
  \[ a_i \sim N(\mu + b \ast \text{prior.score}_i, \sigma_a^2) \]
- We can set \( \mu = 0 \)
- No further prior info in model
The model

- Fit data on signed square roots:
  
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model
  
  \[ y_{ij} \sim N(a_i - a_j, \sigma_y^2) \]

- \( a_i \) and \( a_j \) are “ability parameters”

- \( \sigma_y \) is a scale parameter

- To allow for outliers, use \( t_7 \) instead of normal

- Prior info on abilities:
  
  \[ a_i \sim N(\mu + b \ast \text{prior.score}_i, \sigma_a^2) \]

- We can set \( \mu = 0 \)

- No further prior info in model
The model

- Fit data on signed square roots:
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma^2_y) \)
- \( a_i \) and \( a_j \) are “ability parameters”
- \( \sigma_y \) is a scale parameter
- To allow for outliers, use \( t_7 \) instead of normal
- Prior info on abilities:
  \[ a_i \sim N(\mu + b \times \text{prior.score}_i, \sigma^2_a) \]
- We can set \( \mu = 0 \)
- No further prior info in model
The model

- Fit data on signed square roots:

\[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim \mathcal{N}(a_i - a_j, \sigma_y^2) \)
- \( a_i \) and \( a_j \) are “ability parameters”
- \( \sigma_y \) is a scale parameter
- To allow for outliers, use \( t_7 \) instead of normal
- Prior info on abilities:

\[ a_i \sim \mathcal{N}(\mu + b \ast \text{prior.score}_i, \sigma_a^2) \]

- We can set \( \mu = 0 \)
- No further prior info in model
The model

- Fit data on signed square roots:
  \[ y_{ij} = \sqrt{\text{score differential when team } i \text{ plays team } j} \]

- Model \( y_{ij} \sim N(a_i - a_j, \sigma_y^2) \)
- \( a_i \) and \( a_j \) are “ability parameters”
- \( \sigma_y \) is a scale parameter
- To allow for outliers, use \( t_7 \) instead of normal
- Prior info on abilities:
  \[ a_i \sim N(\mu + b \times \text{prior.score}_i, \sigma_a^2) \]
- We can set \( \mu = 0 \)
- No further prior info in model
data {
    int nteams;
    int ngames;
    vector[nteams] prior_score;
    int team1[ngames];
    int team2[ngames];
    vector[ngames] score1;
    vector[ngames] score2;
    real df;
}
transformed data {
    vector[ngames] dif;
    vector[ngames] sqrt_dif;
    dif <- score1 - score2;
    for (i in 1:ngames)
      sqrt_dif[i] <- (step(dif[i])-.5)*sqrt(fabs(dif[i]));
}
parameters {
    real b;
    real<lower=0> sigma_a;
    real<lower=0> sigma_y;
    vector[nteams] eta_a;
}

transformed parameters {
    vector[nteams] a;
    a <- b*prior_score + sigma_a*eta_a;
}

model {
    eta_a ~ normal(0,1);
    for (i in 1:ngames)
        sqrt_dif[i] ~ student_t(df, a[team1[i]]-a[team2[i]],sigma_y)
}
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

▶ Go into R
▶ Read in the data
▶ Fit the Stan model
▶ Check convergence
▶ Graph the estimated team abilities
▶ Re-fit without prior information
▶ Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Fitting the model

- Go into R
- Read in the data
- Fit the Stan model
- Check convergence
- Graph the estimated team abilities
- Re-fit without prior information
- Compare to model with prior information
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don’t fit the data!!
- Redoing the predictive intervals
- Re-plot, still a problem!
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don’t fit the data!!
- Redoing the predictive intervals
- Re-plot, still a problem!
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don’t fit the data!!
- Redoing the predictive intervals
- Re-plot, still a problem!
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don’t fit the data!!
- Redoing the predictive intervals
- Re-plot, still a problem!
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don't fit the data!!
  - Redoing the predictive intervals
  - Re-plot, still a problem!
Still inside R

For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because \( n \) is large

The predictions don’t fit the data!!

Redoing the predictive intervals

Re-plot, still a problem!
Checking model fit

- Still inside R
- For each game, plot actual score differential and 95% predictive intervals
  - Not cross-validated but no big deal in this case because $n$ is large
- The predictions don't fit the data!!
- Redoing the predictive intervals
- Re-plot, still a problem!
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it's ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
I found the bug!

- Still inside R
  - Re-fit the model on the original scale
  - Display the estimated team abilities
  - Updated plot of data with predictive intervals—now it's ok!
  - Go back and find the bug in the square-root-scale model
  - Re-fit the debugged model
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it’s ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it’s ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it’s ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it’s ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
I found the bug!

- Still inside R
- Re-fit the model on the original scale
- Display the estimated team abilities
- Updated plot of data with predictive intervals—now it’s ok!
- Go back and find the bug in the square-root-scale model
- Re-fit the debugged model
Structure of the course

▶ All other course material is at http://www.stat.columbia.edu/~gelman/bda.course/
▶ In class, we’ll go through the book, section by section:
Structure of the course

▶ Textbook: *Bayesian Data Analysis*, third edition, by Gelman, [Carlin], Stern, Dunson, Vehtari, and [Rubin]

▶ All other course material is at
  http://www.stat.columbia.edu/~gelman/bda.course/

▶ In class, we’ll go through the book, section by section:
  ▶ Stories
  ▶ Tips on what’s tricky
  ▶ Connections and open problems
Structure of the course

- All other course material is at http://www.stat.columbia.edu/~gelman/bda.course/
- In class, we’ll go through the book, section by section:
  - Stories
  - Tips on what’s tricky
  - Connections and open problems
Structure of the course

- All other course material is at [http://www.stat.columbia.edu/~gelman/bda.course/](http://www.stat.columbia.edu/~gelman/bda.course/)
- In class, we’ll go through the book, section by section:
  - Stories
  - Tips on what’s tricky
  - Connections and open problems
Structure of the course

- All other course material is at http://www.stat.columbia.edu/~gelman/bda.course/
- In class, we’ll go through the book, section by section:
  - Stories
  - Tips on what’s tricky
  - Connections and open problems
Structure of the course

▶ All other course material is at http://www.stat.columbia.edu/~gelman/bda.course/
▶ In class, we’ll go through the book, section by section:
  ▶ Stories
  ▶ Tips on what’s tricky
  ▶ Connections and open problems
Structure of the course

- All other course material is at http://www.stat.columbia.edu/~gelman/bda.course/
- In class, we’ll go through the book, section by section:
  - Stories
  - Tips on what’s tricky
  - Connections and open problems
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
  - Just-in-time-teaching assignments due one hour before every class

- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
    - Making sure you do the reading and are prepared for class
- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email

- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class

- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class
- Final exam
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email

- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class

- Final exam
  - Take home
  - Comparable to homework problems
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class
- Final exam
  - Take home
  - Comparable to homework problems
Class assignments

- Homeworks due at beginning of class 2b, 3b, …
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class
- Final exam
  - Take home
  - Comparable to homework problems
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class
- Final exam
  - Take home
  - Comparable to homework problems
Class assignments

- Homeworks due at beginning of class 2b, 3b, ...
  - Assignments and data at http://www.stat.columbia.edu/~gelman/bda.course/
  - Each week: one theory problem, one computing problem, one applied problem
  - Hand in hard copies, no email
- Just-in-time-teaching assignments due one hour before every class
  - Testing background knowledge
  - Making sure you do the reading and are prepared for class
- Final exam
  - Take home
  - Comparable to homework problems
Your first jitt!

- Pull it up on the screen
- Spend 3 minutes on it
- Then we (briefly) discuss
Your first jitt!

- Pull it up on the screen
- Spend 3 minutes on it
- Then we (briefly) discuss
Your first jitt!

- Pull it up on the screen
- Spend 3 minutes on it
- Then we (briefly) discuss
Your first jitt!

- Pull it up on the screen
- Spend 3 minutes on it
- Then we (briefly) discuss
An example from toxicology

- Central story: 4-compartment model of toxicokinetics of perchloroethylene
- Bayesian inference combines prior information and data
- Unresolved questions
- How the model all fits together
An example from toxicology

- Central story: 4-compartment model of toxicokinetics of perchloroethylene
- Bayesian inference combines prior information and data
- Unresolved questions
- How the model all fits together
An example from toxicology

- Central story: 4-compartment model of toxicokinetics of perchloroethylene
- Bayesian inference combines prior information and data
- Unresolved questions
- How the model all fits together
An example from toxicology

- Central story: 4-compartment model of toxicokinetics of perchloroethylene
- Bayesian inference combines prior information and data
- Unresolved questions
  - How the model all fits together
An example from toxicology

- Central story: 4-compartment model of toxicokinetics of perchloroethylene
- Bayesian inference combines prior information and data
- Unresolved questions
- How the model all fits together
Toxicokinetics of perchloroethylene

Goal:
- How much PERC is metabolized at low doses
- Population distribution

Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

4-compartment model, metabolism in liver

Our analysis:
Toxicokinetics of perchloroethylene

Goal:
- How much PERC is metabolized at low doses
- Population distribution

Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks
- 4-compartment model, metabolism in liver

Our analysis:
Toxicokinetics of perchloroethylene

Goal:
- How much PERC is metabolized at low doses
- Population distribution

Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks
- 4-compartment model, metabolism in liver

Our analysis:
Toxicokinetics of perchloroethylene

Goal:
- How much PERC is metabolized at low doses
- Population distribution

Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

4-compartment model, metabolism in liver

Our analysis:
Toxicokinetics of perchloroethylene

- **Goal:**
  - How much PERC is metabolized at low doses
  - Population distribution

- **Experimental data:** Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

- **4-compartment model, metabolism in liver**

- **Our analysis:**
  - Simple data-fitting did not work.
  - Use Bayes to combine data and prior info within model.
Toxicokinetics of perchloroethylene

Goal:
- How much PERC is metabolized at low doses
- Population distribution

Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

4-compartment model, metabolism in liver

Our analysis:
- Simple data-fitting did not work
- Use Bayes to combine data and prior info within model
Toxicokinetics of perchloroethylene

- Goal:
  - How much PERC is metabolized at low doses
  - Population distribution
- Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks
- 4-compartment model, metabolism in liver
- Our analysis:
  - Simple data-fitting did not work
  - Use Bayes to combine data and prior info within model
Toxicokinetics of perchloroethylene

- **Goal:**
  - How much PERC is metabolized at low doses
  - Population distribution

- **Experimental data:** Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

- **4-compartment model, metabolism in liver**

- **Our analysis:**
  - Simple data-fitting did not work
  - Use Bayes to combine data and prior info within model
Toxicokinetics of perchloroethylene

- Goal:
  - How much PERC is metabolized at low doses
  - Population distribution

- Experimental data: Expose 6 healthy volunteers to PERC for four hours, then measure concentrations in blood and air for 2 weeks

- 4-compartment model, metabolism in liver

- Our analysis:
  - Simple data-fitting did not work
  - Use Bayes to combine data and prior info within model
4-compartment model

- **ALV**: Exhaled Air
- **VPR**: Venous Blood
- **Pba**: WELL PERFUSED
  - $V_{wp}$, $P_{wp}$
- **Fwp**: POORLY PERFUSED
  - $V_{pp}$, $P_{pp}$
- **Fpp**: FAT
  - $V_f$, $P_f$
- **Ff**: LIVER
  - $V_l$, $P_l$
- **Fl**: METABOLITES
  - $V_{MI}$, $K_{MI}$
Some data
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, . . .)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, ...)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
  - Discuss an example of prediction
  - Discuss an example of causal inference
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, . . .)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
  - Discuss an example of prediction
  - Discuss an example of causal inference
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, . . .)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
  - Discuss an example of prediction
  - Discuss an example of causal inference
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, ...)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
  - Discuss an example of prediction
  - Discuss an example of causal inference
Connections to general Bayesian principles

- Sometimes the model comes first, based on substantive considerations (toxicology, economics, ...)
- Sometimes the model is chosen based on the data collection (traditional statistics of surveys and experiments)
- Other times the data come first (other statistics examples)
- Usually it’s a mix
  - Discuss an example of prediction
  - Discuss an example of causal inference
Simple statistical ideas did not work

- Fitting 4-compartment model directly to data
- Assisted model fit
- 1 or 2-compartment model
- Simulation from prior distribution
Fitting 4-compartment model directly to data
- Assisted model fit
- 1 or 2-compartment model
- Simulation from prior distribution
Simple statistical ideas did not work

- Fitting 4-compartment model directly to data
- Assisted model fit
  - 1 or 2-compartment model
  - Simulation from prior distribution
Simple statistical ideas did not work

- Fitting 4-compartment model directly to data
- Assisted model fit
- 1 or 2-compartment model
- Simulation from prior distribution
Simple statistical ideas did not work

- Fitting 4-compartment model directly to data
- Assisted model fit
- 1 or 2-compartment model
- Simulation from prior distribution
Simple statistical ideas that did not work:

Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param.
    - "8 kg liver"
- Pooling data and estimating parameters for "the standard man"
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - "8 kg liver"
- Pooling data and estimating parameters for “the standard man"
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - “8 kg liver”
- Pooling data and estimating parameters for “the standard man”
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - “8 kg liver”
- Pooling data and estimating parameters for “the standard man”
  - Not useful for our goal of population inference
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - “8 kg liver”
- Pooling data and estimating parameters for “the standard man”
  - Not useful for our goal of population inference
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - “8 kg liver”
- Pooling data and estimating parameters for “the standard man”
  - Not useful for our goal of population inference
Simple statistical ideas that did not work:
Fitting 4-compartment model directly to data

- Nonlinear least squares
- Fitting to each person separately:
  - Unstable: approx 30 data points, 15 param
  - "8 kg liver"
- Pooling data and estimating parameters for "the standard man"
  - Not useful for our goal of population inference
Simple statistical ideas that did not work:
Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
Simple statistical ideas that did not work:
Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
  - Couldn’t fit the data well
  - Difficult to get fixed values for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
  - Couldn’t fit the data well
  - Difficult to get fixed values for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work: Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
  - Couldn’t fit the data well
  - Difficult to get fixed values for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
  - Couldn’t fit the data well
  - Difficult to get fixed values for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
Assisted model fit

- Set some parameters to fixed values (from the pharmacology literature)
- Estimate the other parameters
- Results:
  - Couldn’t fit the data well
  - Difficult to get fixed values for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
1 or 2-compartment model

- Simpler model can be estimated easily and robustly
- Does not fit the data well
  - Most of the PERC leaves in a few hours, but some stays in the body after a week or more
- Not realistic for low-dose extrapolation
Simple statistical ideas that did not work: 1 or 2-compartment model

- Simpler model can be estimated easily and robustly
- Does not fit the data well
  - Most of the PERC leaves in a few hours, but some stays in the body after a week or more
- Not realistic for low-dose extrapolation
Simple statistical ideas that did not work: 1 or 2-compartment model

- Simpler model can be estimated easily and robustly
- Does not fit the data well
  - Most of the PERC leaves in a few hours, but some stays in the body after a week or more
- Not realistic for low-dose extrapolation
Simple statistical ideas that did not work: 1 or 2-compartment model

- Simpler model can be estimated easily and robustly
- Does not fit the data well
  - Most of the PERC leaves in a few hours, but some stays in the body after a week or more
- Not realistic for low-dose extrapolation
Simple statistical ideas that did not work: 1 or 2-compartment model

- Simpler model can be estimated easily and robustly
- Does not fit the data well
  - Most of the PERC leaves in a few hours, but some stays in the body after a week or more
- Not realistic for low-dose extrapolation
Simple statistical ideas that did not work: Simulation from prior distribution

- Get prior information on parameters from pharmacology literature
- Try to fit data within these prior constraints
- Does not fit the data well
- Difficult to get good prior information for PERC-specific parameters such as equilibrium concentration ratios
Simulation from prior distribution

- Get prior information on parameters from pharmacology literature
- Try to fit data within these prior constraints
- Does not fit the data well
- Difficult to get good prior information for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work: Simulation from prior distribution

- Get prior information on parameters from pharmacology literature
- Try to fit data within these prior constraints
  - Does not fit the data well
  - Difficult to get good prior information for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
Simulation from prior distribution

- Get prior information on parameters from pharmacology literature
- Try to fit data within these prior constraints
- Does not fit the data well
- Difficult to get good prior information for PERC-specific parameters such as equilibrium concentration ratios
Simple statistical ideas that did not work:
Simulation from prior distribution

- Get prior information on parameters from pharmacology literature
- Try to fit data within these prior constraints
- Does not fit the data well
- Difficult to get good prior information for PERC-specific parameters such as equilibrium concentration ratios
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information

- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Bayesian inference

- 4-compartment model
- 15 parameters for each person
- Prior information
  - Strong for some parameters (e.g., volume of liver)
  - Weak for others (e.g., Michaelis-Menten coef)
  - Model includes uncertainty and variation
- Posterior simulation: random walk through parameter space
- Inference for parameters and predictions
- Model checking
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$
  - $\log \theta_j \sim N(\mu, \tau^2)$
  - $\mu \sim N(\log 16, (\log 10)^2)$
  - $\tau \approx \log 2$

- Large uncertainty, small variation

- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, \( \theta_j \) for person \( j \)
  - \( \log \theta_j \sim N(\mu, \tau^2) \)
  - \( \mu \sim N(\log 16, (\log 10)^2) \)
  - \( \tau \approx \log 2 \)
- Large uncertainty, small variation
- Can learn about \( \mu \) using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$

- $\log \theta_j \sim N(\mu, \tau^2)$
- $\mu \sim N(\log 16, (\log 10)^2)$
- $\tau \approx \log 2$

- Large uncertainty, small variation
- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$
  - $\log \theta_j \sim N(\mu, \tau^2)$
  - $\mu \sim N(\log 16, (\log 10)^2)$
  - $\tau \approx \log 2$

- Large uncertainty, small variation
- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$

- $\log \theta_j \sim N(\mu, \tau^2)$
- $\mu \sim N(\log 16, (\log 10)^2)$
- $\tau \approx \log 2$

- Large uncertainty, small variation
- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$
  - $\log \theta_j \sim N(\mu, \tau^2)$
  - $\mu \sim N(\log 16, (\log 10)^2)$
  - $\tau \approx \log 2$

- Large uncertainty, small variation
  - Can learn about $\mu$ using data from several people
  - Can’t do this without a hierarchical model
  - Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$
  - $\log \theta_j \sim N(\mu, \tau^2)$
  - $\mu \sim N(\log 16, (\log 10)^2)$
  - $\tau \approx \log 2$

- Large uncertainty, small variation
- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, \( \theta_j \) for person \( j \)
  - \( \log \theta_j \sim N(\mu, \tau^2) \)
  - \( \mu \sim N(\log 16, (\log 10)^2) \)
  - \( \tau \approx \log 2 \)

- Large uncertainty, small variation

- Can learn about \( \mu \) using data from several people

- Can’t do this without a hierarchical model

- Transformations and prior correlations (why transformations are particularly important for Bayesians)
Hierarchical prior distributions

- Prior distribution for a rate parameter in the metabolism, $\theta_j$ for person $j$
  - $\log \theta_j \sim N(\mu, \tau^2)$
  - $\mu \sim N(\log 16, (\log 10)^2)$
  - $\tau \approx \log 2$

- Large uncertainty, small variation
- Can learn about $\mu$ using data from several people
- Can’t do this without a hierarchical model
- Transformations and prior correlations (why transformations are particularly important for Bayesians)
### Hierarchical prior distributions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Population prior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventilation/perfusion ratio (VPR)</td>
<td>$1.6(x \div 1.3)$</td>
</tr>
<tr>
<td>Blood flow, well-perfused tissues (Fwp)</td>
<td>$.47(x \div 1.17)$</td>
</tr>
<tr>
<td>Blood flow, poorly perfused tissues (Fpp)</td>
<td>$.20(x \div 1.22)$</td>
</tr>
<tr>
<td>Blood flow, fat (Ff)</td>
<td>$.07(x \div 1.27)$</td>
</tr>
<tr>
<td>Blood flow, liver (Fl)</td>
<td>$.25(x \div 1.15)$</td>
</tr>
<tr>
<td>Volume, well-perfused tissues (Vwp)</td>
<td>$.27(x \div 1.36)$</td>
</tr>
<tr>
<td>Volume, poorly perfused tissues (Vpp)</td>
<td>$.55(x \div 1.17)$</td>
</tr>
<tr>
<td>Volume, liver (VI)</td>
<td>$.033(x \div 1.1)$</td>
</tr>
<tr>
<td>Partition coeff, blood/air (Pba)</td>
<td>$12(x \div 1.5)$</td>
</tr>
<tr>
<td>Partition coeff, well-perfused (Pwp)</td>
<td>$4.8(x \div 1.5)$</td>
</tr>
<tr>
<td>Partition coeff, poorly perfused (Ppp)</td>
<td>$1.6(x \div 1.5)$</td>
</tr>
<tr>
<td>Partition coeff, fat (Pf)</td>
<td>$125(x \div 1.5)$</td>
</tr>
<tr>
<td>Partition coeff, liver (Pl)</td>
<td>$4.8(x \div 1.5)$</td>
</tr>
<tr>
<td>Max metabolic rate in liver (VMI)</td>
<td>$.042(x \div 10)$</td>
</tr>
<tr>
<td>$K_m$</td>
<td>$16(x \div 10)$</td>
</tr>
<tr>
<td>in liver (KMI)</td>
<td>$x \div 1.5$</td>
</tr>
</tbody>
</table>
What we did

- Set up a hierarchical prior distribution with uncertainty and population variation for a 4-compartment model
- Fit the model to data (much computation)
- Checked inferences about parameters to see that they made sense
- Re-ran model under hypothetical low-dose, high-dose exposures
What we did

- Set up a hierarchical prior distribution with uncertainty and population variation for a 4-compartment model
- Fit the model to data (much computation)
- Checked inferences about parameters to see that they made sense
- Re-ran model under hypothetical low-dose, high-dose exposures
What we did

- Set up a hierarchical prior distribution with uncertainty and population variation for a 4-compartment model
- Fit the model to data (much computation)
  - Checked inferences about parameters to see that they made sense
  - Re-ran model under hypothetical low-dose, high-dose exposures
What we did

- Set up a hierarchical prior distribution with uncertainty and population variation for a 4-compartment model
- Fit the model to data (much computation)
- Checked inferences about parameters to see that they made sense
- Re-ran model under hypothetical low-dose, high-dose exposures
What we did

- Set up a hierarchical prior distribution with uncertainty and population variation for a 4-compartment model
- Fit the model to data (much computation)
- Checked inferences about parameters to see that they made sense
- Re-ran model under hypothetical low-dose, high-dose exposures
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
- Re-run the model several times to simulate what would happen under different conditions
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
- Re-run the model several times to simulate what would happen under different conditions
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
- Re-run the model several times to simulate what would happen under different conditions
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
- Re-run the model several times to simulate what would happen under different conditions
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
  - Re-run the model several times to simulate what would happen under different conditions
Fitting and using the model

- Use Gibbs sampler and Metropolis algorithm to take a random walk through parameter space
- Computationally intensive
  - Each step requires evaluation of the numerical differential equation solver
- Check inferences: Do they make sense?
- Re-run the model several times to simulate what would happen under different conditions
Inference for 6 individuals
Prediction of data from a new study

![Graph showing data over time](image-url)
Sensitivity to priors

(a) Plot of Percent Metabolized vs. Fat/Blood Partition Coefficient.
(b) Plot of Percent Metabolized vs. Blood/Air Partition Coefficient.

Graphs show data distribution and relationship between variables.

Andrew Gelman
Bayesian Data Analysis, class 1a
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

➤ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model  
(b) Hierarchical population model  
(c) Prior information  
(d) Experimental data  
(e) Bayesian inference  
(f) Computation  
(g) Model checking  

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Putting it all together

(a) Physiological pharmacokinetic model
(b) Hierarchical population model
(c) Prior information
(d) Experimental data
(e) Bayesian inference
(f) Computation
(g) Model checking

▶ We need all of these!
Without a physiological model, there is no good way to get prior information on the parameters.

We need physiological parameters (not just curve-fitting of the data) to efficiently combine information across different people.
Without a physiological model, there is no good way to get prior information on the parameters.

We need physiological parameters (not just curve-fitting of the data) to efficiently combine information across different people.
Without a physiological model, there is no good way to get prior information on the parameters.

We need physiological parameters (not just curve-fitting of the data) to efficiently combine information across different people.
Without a population model, there generally are not enough data to estimate the parameters separately for each individual.

And there is too much variation among bodies (even among healthy young male volunteers) to pool all the data together and estimate common parameters.
Without a population model, there generally are not enough data to estimate the parameters separately for each individual. And there is too much variation among bodies (even among healthy young male volunteers) to pool all the data together and estimate common parameters.
Without a population model, there generally are not enough data to estimate the parameters separately for each individual. And there is too much variation among bodies (even among healthy young male volunteers) to pool all the data together and estimate common parameters.
We need prior information. Otherwise, our estimates don’t make sense (the 8 kg liver).

We need experimental data to learn about perchloroethylene in particular.
We need prior information. Otherwise, our estimates don’t make sense (the 8 kg liver).

We need experimental data to learn about perchloroethylene in particular.
We need prior information. Otherwise, our estimates don’t make sense (the 8 kg liver).

We need experimental data to learn about perchloroethylene in particular.
Using Bayesian inference, we can find parameters that are consistent with both prior information and data, if such agreement is possible.

Automatically includes uncertainty and variability, so inferences can be plugged in directly to risk assessment and decision analysis.
Using Bayesian inference, we can find parameters that are consistent with both prior information and data, if such agreement is possible. Automatically includes uncertainty and variability, so inferences can be plugged in directly to risk assessment and decision analysis.
Using Bayesian inference, we can find parameters that are consistent with both prior information and data, if such agreement is possible.

Automatically includes uncertainty and variability, so inferences can be plugged in directly to risk assessment and decision analysis.
Our models are big. Least squares, maximum likelihood, etc., are not enough.

Old-fashioned differential-equation solver is still sitting inside the model.

Our computers are never fast enough. We want more, more, more!
Our models are big. Least squares, maximum likelihood, etc., are not enough

Old-fashioned differential-equation solver is still sitting inside the model

Our computers are never fast enough. We want more, more, more!
Our models are big. Least squares, maximum likelihood, etc., are not enough.

Old-fashioned differential-equation solver is still sitting inside the model.

Our computers are never fast enough. We want more, more, more!
Our models are big. Least squares, maximum likelihood, etc., are not enough.

Old-fashioned differential-equation solver is still sitting inside the model.

Our computers are never fast enough. We want more, more, more!
(g) Model checking

- Check inferences about parameters
  - Do they make sense?
  - Are they consistent with prior distributions?
- Check fit to data
- Check predictions on new data
(g) Model checking

- Check inferences about parameters
  - Do they make sense?
  - Are they consistent with prior distributions
- Check fit to data
- Check predictions on new data
(g) Model checking

- Check inferences about parameters
  - Do they make sense?
  - Are they consistent with prior distributions
- Check fit to data
- Check predictions on new data
(g) Model checking

- Check inferences about parameters
  - Do they make sense?
  - Are they consistent with prior distributions
- Check fit to data
  - Check predictions on new data
(g) Model checking

- Check inferences about parameters
  - Do they make sense?
  - Are they consistent with prior distributions
- Check fit to data
- Check predictions on new data
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
- Using models to combine different sources of information (partial pooling)
- Better dialogue with subject-matter experts (more windows into the model and data)
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
- Using models to combine different sources of information (partial pooling)
- Better dialogue with subject-matter experts (more windows into the model and data)
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
- Using models to combine different sources of information (partial pooling)
- Better dialogue with subject-matter experts (more windows into the model and data)
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
  - Using models to combine different sources of information (partial pooling)
  - Better dialogue with subject-matter experts (more windows into the model and data)
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
- Using models to combine different sources of information (partial pooling)
- Better dialogue with subject-matter experts (more windows into the model and data)
Using Bayesian ideas to improve existing analyses

- Regularization (for example, avoiding estimates on the boundary of parameter space)
- Accounting for uncertainty (especially for decisions)
- Checking model fit
- Using models to combine different sources of information (partial pooling)
- Better dialogue with subject-matter experts (more windows into the model and data)
Summary of example

- Population pharmacokinetic models have many moving parts.
- Complexity in one place can make it easier, not harder, to add information in other places.
Population pharmacokinetic models have many moving parts
Complexity in one place can make it easier, not harder, to add information in other places
Population pharmacokinetic models have many moving parts.

Complexity in one place can make it *easier*, not harder, to add information in other places.