Big Data in OB and OB Anesthesia Research

Brian T. Bateman, MD, MSc
Associate Professor of Anesthesia
Department of Anesthesia, Critical Care, and Pain Medicine
Massachusetts General Hospital
Division of Pharmacoepidemiology and Pharmacoeconomics
Department of Medicine
Brigham and Women’s Hospital
Harvard Medical School
Disclosures

- **Research funding:**
  - National Institutes of Health
  - Icelandic Centre for Research
  - Pfizer
  - Lilly
  - Baxalta
  - GSK
  - Pacira
Outline

• Overview of research based on healthcare utilization data
  ▫ Types of data sources
  ▫ Types of research

• Methodological considerations in comparative effectiveness/safety research
  ▫ Example of statins and congenital malformations

• Novel methods to control for unmeasured confounders
  ▫ Example of induction and the risk of autism
Healthcare utilization data

- Routinely collected for the payment and administration of health services
  - Claims for services/procedures/medications
  - Diagnoses used to justify services
- Not collected primarily for research purposes
Publications in obstetrics based on healthcare utilization data

Source: Pubmed
Advantages of healthcare utilization data

• **Large**
  ▫ Rare events
  ▫ Power to analyze subgroups

• **Representative**
  ▫ Real-world effectiveness and utilization patterns

• **Informative regarding health systems**
  ▫ Information on volume, hospital/physician characteristics, costs

• **Accessible**
  ▫ Data are available at low cost
  ▫ Near real-time
Available data sources

• Types: Longitudinal and episodic

• Longitudinal databases
  ▫ Claims for outpatient, inpatient, laboratory, imaging services, medication dispensing

• Examples:
  • Medicaid Analytic eXtract (MAX)
  • United Healthcare database
  • Kaiser Permanente
Available data sources

- Episodic databases
- Information derived from single, inpatient encounter
- Examples:
  - Nationwide Inpatient Sample
    - 20% stratified-sample of all US hospitalizations
  - State Inpatient Databases
    - Complete claims for admissions from selected states
  - Premier Perspective Database
    - ~1/6 of all US hospitalizations; granular data on hospital charges
Healthcare utilization data

• Types of studies:

  ▫ Descriptive studies
    • Disease/complication epidemiology
    • Healthcare utilization

  ▫ Disease/complication risk prediction

  ▫ Comparative effectiveness/safety research
Healthcare utilization data

- Types of studies:
  - Descriptive studies
    - Disease/complication epidemiology
    - Healthcare utilization
  - Disease/complication risk prediction
  - Comparative effectiveness/safety research
PERIOPERATIVE MEDICINE

Cardiac Arrest during Hospitalization for Delivery in the United States, 1998–2011

Jill M. Mhyre, M.D., Lawrence C. Tsen, M.D., Sharon Einav, M.D., Elena V. Kuklina, M.D., Ph.D., Lisa R. Leffert, M.D., Brian T. Bateman, M.D., M.Sc.
Cardiac Arrest during Hospitalization for Delivery in the United States, 1998–2011

Jill M. Mhyre, M.D., Lawrence C. Tsen, M.D., Sharon Einav, M.D., Elena V. Kuklina, M.D., Ph.D., Lisa R. Leffert, M.D., Brian T. Bateman, M.D., M.Sc.

- Nationwide Inpatient Sample
- 56 million deliveries
- Frequency: 1 in 12,000
- Leading etiology: Hemorrhage (40%)
- Post-arrest survival: 58.9%

*Anesthesiology. 2014 Apr;120(4):810-8.*
Healthcare utilization data

- Types of studies:
  - Descriptive studies
    - Disease/complication epidemiology
    - Healthcare utilization
  - Disease/complication risk prediction
  - Comparative effectiveness/safety research
Patterns of Opioid Utilization in Pregnancy in a Large Cohort of Commercial Insurance Beneficiaries in the United States

Brian T. Bateman, M.D., M.Sc., Sonia Hernandez-Diaz, M.D., Dr.P.H., James P. Rathmell, M.D., John D. Seeger, Pharm.D., Dr.P.H., Michael Doherty, M.S., Michael A. Fischer, M.D., M.S., Krista F. Huybrechts, M.S., Ph.D.
Patterns of Opioid Utilization in Pregnancy in a Large Cohort of Commercial Insurance Beneficiaries in the United States

Brian T. Bateman, M.D., M.Sc., Sonia Hernandez-Diaz, M.D., Dr.P.H., James P. Rathmell, M.D., John D. Seeger, Pharm.D., Dr.P.H., Michael Doherty, M.S., Michael A. Fischer, M.D., M.S., Krista F. Huybrechts, M.S., Ph.D.

- United Healthcare
- 534,500 pregnancies from 2005 to 2011
- 14.4% opioid exposed during pregnancy
- Significant regional variation
Patterns of Opioid Utilization in Pregnancy in a Large Cohort of Commercial Insurance Beneficiaries in the United States

Brian T. Bateman, M.D., M.Sc., Sonia Hernandez-Diaz, M.D., Dr.P.H., James P. Rathmell, M.D., John D. Seeger, Pharm.D., Dr.P.H., Michael Doherty, M.S., Michael A. Fischer, M.D., M.S., Krista F. Huybrechts, M.S., Ph.D.
Healthcare utilization data

• Types of studies:
  ▫ Descriptive studies
    • Disease/complication epidemiology
    • Healthcare utilization
  ▫ Disease/complication risk prediction
  ▫ Comparative effectiveness/safety research
Development of a Comorbidity Index for Use in Obstetric Patients

Brian T. Bateman, MD, MSc, Jill M. Mhyre, MD, Sonia Hernandez-Diaz, MD, DrPH, Krista F. Huybrechts, MS, PhD, Michael A. Fischer, MD, MS, Andreea A. Creanga, MD, PhD, William M. Callaghan, MD, MPH, and Joshua J. Gagne, PharmD, ScD
Development of a Comorbidity Index for Use in Obstetric Patients

Brian T. Bateman, MD, MSc, Jill M. Mhyre, MD, Sonia Hernandez-Diaz, MD, DrPH, Krista F. Huybrechts, MS, PhD, Michael A. Fischer, MD, MS, Andreea A. Creanga, MD, PhD, William M. Callaghan, MD, MPH, and Joshua J. Gagne, PharmD, ScD

- Medicaid Analytic eXtract
- 854,823; 1.2% severe morbidity
- Includes 20 maternal conditions
- Risk=0.68% for score 0; 10.9% if >10
Healthcare utilization data

• Types of studies:
  ▫ Descriptive studies  
    • Disease/complication epidemiology  
    • Healthcare utilization
  ▫ Disease/complication risk prediction
  ▫ Comparative effectiveness/safety research
Comparative effectiveness and safety research

- Gold standard \( \rightarrow \) Randomized controlled clinical trials
  - Lack problem of confounding
  - May lack generalizability
  - Less useful in establishing safety
  - Limited follow-up time
  - Not all relevant questions will be answered with RCTs
Comparative effectiveness and safety research

• Gold standard $\rightarrow$ Randomized controlled clinical trials
  ▫ Lack problem of confounding
  ▫ May lack generalizability
  ▫ Less useful in establishing safety
  ▫ Limited follow-up time
  ▫ Not all relevant questions will be answered with RCTs
Healthcare utilization data

• **Challenges**

  • Those inherent to all observational research
    ▫ Selection bias
    ▫ Information bias
    ▫ Confounding bias

  • Those particular to healthcare utilization data
    ▫ Strong potential for misclassification/underascertainment of outcomes, exposures, and confounders
Study question

Exposure → Outcome
Study question

Exposure

Confounder

Outcome
Challenges: Outcome misclassification

- Not as problematic as one might think...
- Important to define the outcome with specificity
  - If the outcome is defined with 100% specificity, then relative risk estimates will be unbiased assuming that misclassification is non-differential

The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>
The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>
### The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

**RR = 2**
### The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

**RR = 2**

**Outcome measured with 50% sensitivity**
### The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

RR = 2

### Outcome measured with 50% sensitivity

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>
### The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

### Outcome measured with 50% sensitivity

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>

\[RR = 2\]
### The truth

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

RR = 2

### Outcome measured with 50% sensitivity

<table>
<thead>
<tr>
<th></th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Unexposed</td>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>

RR = 2
<table>
<thead>
<tr>
<th>Condition</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstetric hemorrhage</td>
<td>75</td>
<td>99.8</td>
</tr>
<tr>
<td>Episiotomy</td>
<td>70</td>
<td>99.9</td>
</tr>
<tr>
<td>Obstetric wound infection</td>
<td>68</td>
<td>98</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>87</td>
<td>98</td>
</tr>
<tr>
<td>Renal failure</td>
<td>88</td>
<td>99.4</td>
</tr>
<tr>
<td>Endometritis</td>
<td>46</td>
<td>98</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>87</td>
<td>98</td>
</tr>
</tbody>
</table>

- Romano and Mark, Med Care, 1994; 32: 81-90
- Kiyota et al, Am Heart J, 2004; 148: 99-104
- Shaklee et al, Infect Dis J 2011;30:e38-40
Challenges: Outcome misclassification

• Specificity further improved with treatment codes or LOS
  ▫ Infection: Appropriate antibiotics
  ▫ PPH: Transfusion
  ▫ Severe preeclampsia: Magnesium
  ▫ MI: LOS > 3 days

• Kiyota et al, Am Heart J, 2004; 148: 99-104
• Shaklee et al, Infect Dis J 2011;30:e38-40
Challenges: Exposure misclassification

- Can often be measured with sensitivity and specificity
  - Charges for drugs, procedures

- When the exposure is a diagnosis, misclassification may create bias of effect estimates
Challenges: Exposure misclassification

• Particularly problematic in episodic databases

  ▫ **Overcoding:** Patient with peridelivery MI may have more codes for known risk factors than comparable patient without the complication

  ▫ **Undercoding:** Patients that have a long and complicated hospital course may have fewer codes for chronic conditions
    • DM paradoxically appears to decrease the risk peridelivery sepsis in administrative datasets
Challenges: Confounder underascertainment/misclassification

- Misclassification of or inability to measure confounders is the greatest threat to validity of studies using administrative data
Challenges: Confounder underascertainment/ misclassification

- Minimize the degree of confounding at the design phase
  - Use of active comparators

- Use proxies to identify relevant comorbidities
  - Insulin $\rightarrow$ DM
  - Labetalol $\rightarrow$ HTN
  - Number of outpatient medications as a marker for general health status
• Draws on range of data dimension including diagnoses, procedures, laboratory tests, and medications

• Empirically identifies candidate covariates from thousands of codes

• Prioritizes covariates, and integrates them into a propensity-score

• Improves confounding control in some circumstances
Statins and congenital malformations
Statins and congenital malformations

- Food and Drug Administration Category X
  - Animal data showing the potential for toxicity
  - Role of cholesterol biosynthesis for prenatal development

- Few human data
  - Mixed results
  - Registries, small cohort studies, and spontaneous reports

- Important:
  - Use in women of reproductive age
  - Potential use in preventing preeclampsia
Medicaid Analytic eXtract 2000-2007
- N = 886,996
- 1,152 (0.13%) statin exposed in 1st trimester

Outcome → Congenital malformations
- Defined by codes of 2 separate dates
- One code + corrective surgery (high specificity)

50 covariates drawn from claims
- Demographics, diagnoses, medications, healthcare utilization (robust confounder control incld many proxies)
## Statins and congenital malformations

<table>
<thead>
<tr>
<th></th>
<th>Risk</th>
<th>Unadjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unexposed</td>
<td>3.55%</td>
<td>Referent</td>
</tr>
<tr>
<td>Exposed</td>
<td>6.34%</td>
<td>1.79 (1.43 to 2.23)</td>
</tr>
</tbody>
</table>
## Statins and congenital malformations

<table>
<thead>
<tr>
<th></th>
<th>Risk</th>
<th>Unadjusted</th>
<th>Stratified on diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unexposed</strong></td>
<td>3.55%</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td><strong>Exposed</strong></td>
<td>6.34%</td>
<td>1.79 (1.43 to 2.23)</td>
<td>1.34 (1.07 to 1.68)</td>
</tr>
</tbody>
</table>
## Statins and congenital malformations

<table>
<thead>
<tr>
<th></th>
<th>Risk</th>
<th>Unadjusted</th>
<th>Stratified on diabetes</th>
<th>Propensity-score matched (1:3)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unexposed</strong></td>
<td>3.55%</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td><strong>Exposed</strong></td>
<td>6.34%</td>
<td>1.79 (1.43 to 2.23)</td>
<td>1.34 (1.07 to 1.68)</td>
<td>1.04 (0.79 to 1.37)</td>
</tr>
</tbody>
</table>
Novel design approaches to overcome residual confounding

- Circumstances where traditional designs are likely to result in residual confounding

- In this circumstance novel epidemiological designs are needed

- Example: Induction of labor and autism
Background

Autism Spectrum Disorders (ASD):
• Permanent developmental disabilities
  ▫ Impairment in social interaction
  ▫ Language development
  ▫ Stereotyped or repetitive behaviors
  ▫ Incidence: 1 in 90

Genetics and early environmental exposures implicated in pathophysiology
○ Investigation of association with perinatal exposures

Background

**Oxytocin:**
- Key role in social function and cognition
- Used to induce/augment labor

**Hypothesis:**
- *In-utero* exposure causes down-regulation of oxytocin receptors predisposing to ASD

Glasson et al. Arch Gen Psychiatry 2004;61:618-27
Wahl. Medical hypotheses 2004;63:456-60
• Linked information on 625,042 births to educational records
  o **Exposure:** Induction
  o **Outcome:** Autism exceptionality designation
  o After adjustment for confounders:
    • Relative risk: 1.27 (95% CI 1.05-1.52)
Autism Might Be Linked With Inducing And Speeding Up Labor (STUDY)

Induced Labor Linked to Raised Risk of Autism, Study Suggests

Brain changes of autism may begin in the womb

Induced labor 'linked to autism'

Study Links Inducing/Augmenting Labor with Modestly Higher Autism Risk

Can inducing labor make your child autistic?
ACOG issued a Committee Opinion:

- “against a change in current guidance regarding counseling and indications for and methods of labor induction...”

Avoiding induction may lead to more stillbirth, cesareans, and adverse effects on maternal health
A significant concern regarding prior study is potential for residual confounding.

Use Swedish nationwide population register data to identify differentially exposed siblings.
Methods

Study Population:
- Births in Sweden 1992 to 2001
- N=978,98

Exposure:
- Induction indicator on standardized delivery chart

Outcome:
- Diagnoses of Autism Spectrum Disorder
- 1.1%
Methods

Analysis:
• Association between induction and ASD modelled using Cox proportional hazard regression

• Models:
  ▫ Model 1: Baseline
  ▫ Model 2: Stable maternal characteristics:
    Demographics
  ▫ Model 3: Model 2+ individual birth characteristics:
    Risk factors for induction like postdates, preeclampsia, etc
Methods

Model 4:
• Fixed effects model in which the underlying hazard is allowed to vary between mothers (=fixed for all births to same woman)
• Contrast is made within siblings
  ▫ Only siblings discordant with respect to induction status contribute to the estimation of effect
• Maintain adjustment for factors unique to each birth
• Controls for all factors shared by siblings
## Results

<table>
<thead>
<tr>
<th>Autism spectrum disorders</th>
<th>Baseline HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.27 (1.19 - 1.37)</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Stable maternal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autism spectrum disorders</strong></td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td></td>
<td>1.27 (1.19 - 1.37)</td>
<td>1.27 (1.18 - 1.36)</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Baseline HR (95% CI)</th>
<th>Stable maternal HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autism spectrum disorders</strong></td>
<td>1.27 (1.19 - 1.37)</td>
<td>1.27 (1.18 - 1.36)</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Baseline HR (95% CI)</th>
<th>Stable maternal HR (95% CI)</th>
<th>Individual birth HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism spectrum disorders</td>
<td>1.27 (1.19 - 1.37)</td>
<td>1.27 (1.18 - 1.36)</td>
<td>1.18 (1.09 - 1.27)</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Stable maternal</th>
<th>Individual birth</th>
<th>Within siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autism spectrum disorders</strong></td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td></td>
<td>1.27 (1.19 - 1.37)</td>
<td>1.27 (1.18 - 1.36)</td>
<td>1.18 (1.09 - 1.27)</td>
<td>1.05 (0.87 - 1.27)</td>
</tr>
</tbody>
</table>
Conclusions

- Big data are a powerful tool for research in OB and OB anesthesia

- Attention to study design and limitations of data needed

- Novel approaches can help minimize the risk for residual confounding and yield valid estimates