Studying Pregnancy Outcomes with Electronic Medical Records Data

GroupHealth
RESEARCH INSTITUTE

Sascha Dublin, MD, PhD
Group Health Research Institute
October 24, 2016

Precision Medicine

GroupHealth RESEARCH INSTITUTE

- “An emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person."

https://ghr.nlm.nih.gov/primer/precisionmedicine/definition

Electronic Health Records

GroupHealth
RESEARCH INSTITUTE

- Administrative/claims data or electronic medical records
- Very large sample sizes; representative and often diverse populations; detailed clinical information
- Can identify and study subgroups
- Real-world treatments under real-world conditions

Challenges

GroupHealth RESEARCH INSTITUTE

- May lack key pieces of information
  - Missing data for many environmental or lifestyle factors
- May need to obtain and incorporate supplemental data
- Phenotypes (outcomes?) may be measured inaccurately

Objectives

GroupHealth
RESEARCH INSTITUTE

- Describe a study using electronic health data to address an important clinical question within obstetrics
- Describe challenges that we faced and some solutions
- Review the use of 2-phase study designs to incorporate supplemental data

Clinical Question

GroupHealth RESEARCH INSTITUTE

- Is elective induction of labor at term safe for the mother and baby?
  - At 38 weeks gestation? 39? 40?

EI = initiating labor in a woman with no medical or obstetric reason for immediate delivery

Elective Induction

GroupHealth RESEARCH INSTITUTE

- In the past, up to 10% of US births: 400,000 per year
- May increase Cesarean delivery and neonatal ICU admission, but controversial
- Randomized trials not helpful (few, old, very small)
- Many observational studies exist but had problematic methods and data
- We still do not know whether EI increases risk of cesarean delivery or other outcomes

Precision Medicine

GroupHealth RESEARCH INSTITUTE

- Outcomes of EI may vary greatly based on characteristics of the woman and the pregnancy
  - Parity
  - Gestational age
  - Cervical ripeness
  - Race/ethnicity
  - Obesity
- Would like to be able to describe risks and benefits tailored to a woman's specific characteristics

Specific Aims

GroupHealth RESEARCH INSTITUTE

- To estimate risks of adverse maternal and neonatal outcomes after elective induction compared to expectant management at 38, 39 or 40 weeks' gestation
  - Including Cesarean delivery, neonatal intensive care unit (NICU) stay, others
- To examine how risks vary by maternal race/ethnicity and pre-pregnancy obesity.

NICHD R01 HD071986

Study Overview

GroupHealth RESEARCH INSTITUTE

- Two-phase study of elective induction and pregnancy outcomes in nulliparous women
- Set within 2 integrated healthcare systems, Group Health (GH) and Kaiser Permanente Southern California (KPSC)
- Singleton births, 2007-2013
- Data from automated health plan data, KP pregnancy registry, birth certificates, and detailed review of medical records

NICHD R01 HD071986

Challenges

GroupHealth RESEARCH INSTITUTE

- Mother-infant linkages often not available in healthcare data
- Key variables including gestational age, parity and cervical ripeness are not readily available in these electronic data sources
- Misclassification of exposure status (i.e., elective induction)
- Misclassification of outcome status: postpartum hemorrhage, chorioamnionitis

Use of Birth Certificates

GroupHealth RESEARCH INSTITUTE

- Linkage to state birth certificates can supply crucial data elements needed for pregnancy studies
- Data on gestational age is crucial for ascertaining timing of exposure
  - Gestational age is not routinely available in healthcare data
- Parity
- Maternal race/ethnicity, education, smoking, many others

Misclassification

GroupHealth RESEARCH INSTITUTE

- EI is not well measured in automated health plan or birth certificate data
  - PPV for induction: 60%; for EI: 36%
- Some outcomes also need validation
- Need to review medical records
- Both outcome and exposure are relatively uncommon (5-10%)
- What methods can be used to study uncommon exposure and outcomes with mismeasurement?

Two-Phase Study

GroupHealth RESEARCH INSTITUTE

- Phase I: automated data to identify potential exposure and outcome status
  - N=43,000
- Phase II: medical record review to validate exposure and outcome status and collect supplemental data
  - N=3,125
  - Stratify on Phase I exposure and outcome status
  - Oversample most informative women: those with apparent EI and outcome

Two-Phase Study

GroupHealth RESEARCH INSTITUTE

Hypothetical distribution:

Phase I Induction status	Cesarean Delivery	No Cesarean Delivery
Elective induction	995	5,636
Expectant management	5,527	31,318

- Oversample most informative
- Balanced design is often best
- Simulation used to develop sampling scheme

Methods

GroupHealth RESEARCH INSTITUTE

- Eligible population
  - Singleton, nulliparous, and delivered at 38-42 weeks gestation
  - No contraindications to induction
  - No indication for induction as of 38 weeks
- Identify potential EIs and outcomes using automated data (codes)
- Sample for chart review

Methods

GroupHealth RESEARCH INSTITUTE

- Review records to confirm eligibility and validate exposure and outcome status
- Analyze data using methods for two-phase studies
  - Simplest: reweighting
  - More efficient: semi-parametric maximum likelihood (SPML)
  - Had to extend SPML for our context -- methods development needed

Results (in progress)

- About 43,000 deliveries met inclusion criteria according to Phase I data
- About 6,600 apparent elective inductions (15% of births)
- As of last week, had completed our targeted number of reviews (3125)

Results (in progress)

GroupHealth RESEARCH INSTITUTE

- At KPSC, 1,124 sampled as potential EI
  - 63% confirmed as induced
  - 347 true EI (PPV = 31%)
- At GH, 306 sampled as potential EI
  - 79% confirmed as induced
  - 52 true EI (PPV = 17%)
- Some additional EIs found in the “unexposed” group at both sites
- Total of 510 EI across the sites

Validation of Outcomes

GroupHealth RESEARCH INSTITUTE

Positive predictive values:

Outcome	KPSC	GH
Cesarean delivery	98%	97%
NICU stay	90%	79%
Postpartum hemorrhage	81%	78%
Chorioamnionitis	96%	54%

Next Steps

GroupHealth RESEARCH INSTITUTE

- Data cleaning; create analytic variables
- Conduct primary analyses
- Future directions could include developing a better algorithm to identify induction and EI from automated data

Take-home Messages

GroupHealth RESEARCH INSTITUTE

- Electronic health data offer opportunities to study safety and effectiveness of interventions in pregnancy
- Supplemental data often needed
  - Mother-infant linkages
  - State birth certificates
  - Medical record review

Precision Medicine

GroupHealth RESEARCH INSTITUTE

- Many scenarios may require additional data not readily available from electronic health data:
  - Genetic/genomic information
  - Environmental exposures (smoking, alcohol use, physical activity)
  - Detailed information about the condition, e.g. disease severity, cancer characteristics
- Two-phase studies offer an efficient approach to obtain and incorporate supplemental data

EIPO Team

GroupHealth RESEARCH INSTITUTE

GHRI:

- James Fraser
- Eric Baldwin
- Jennifer Bobb
- Rod Walker
- Mary Shea
- Tammy Dodd

Other institutions:

- Darios Getahun, KPSC, co-PI
- Aaron Caughey, OHSU
- Victoria Holt, UW
- Deborah Wing, UC Irvine

Questions?

GroupHealth
RESEARCH INSTITUTE

Mother-baby linkage

GroupHealth RESEARCH INSTITUTE

- Few plans have birth registries
- Data resources available to link the mothers with infants vary widely between plans
- Need flexible approach, but want to be as standardized as possible
- Created algorithm that provides hierarchy and sets priorities

Algorithm

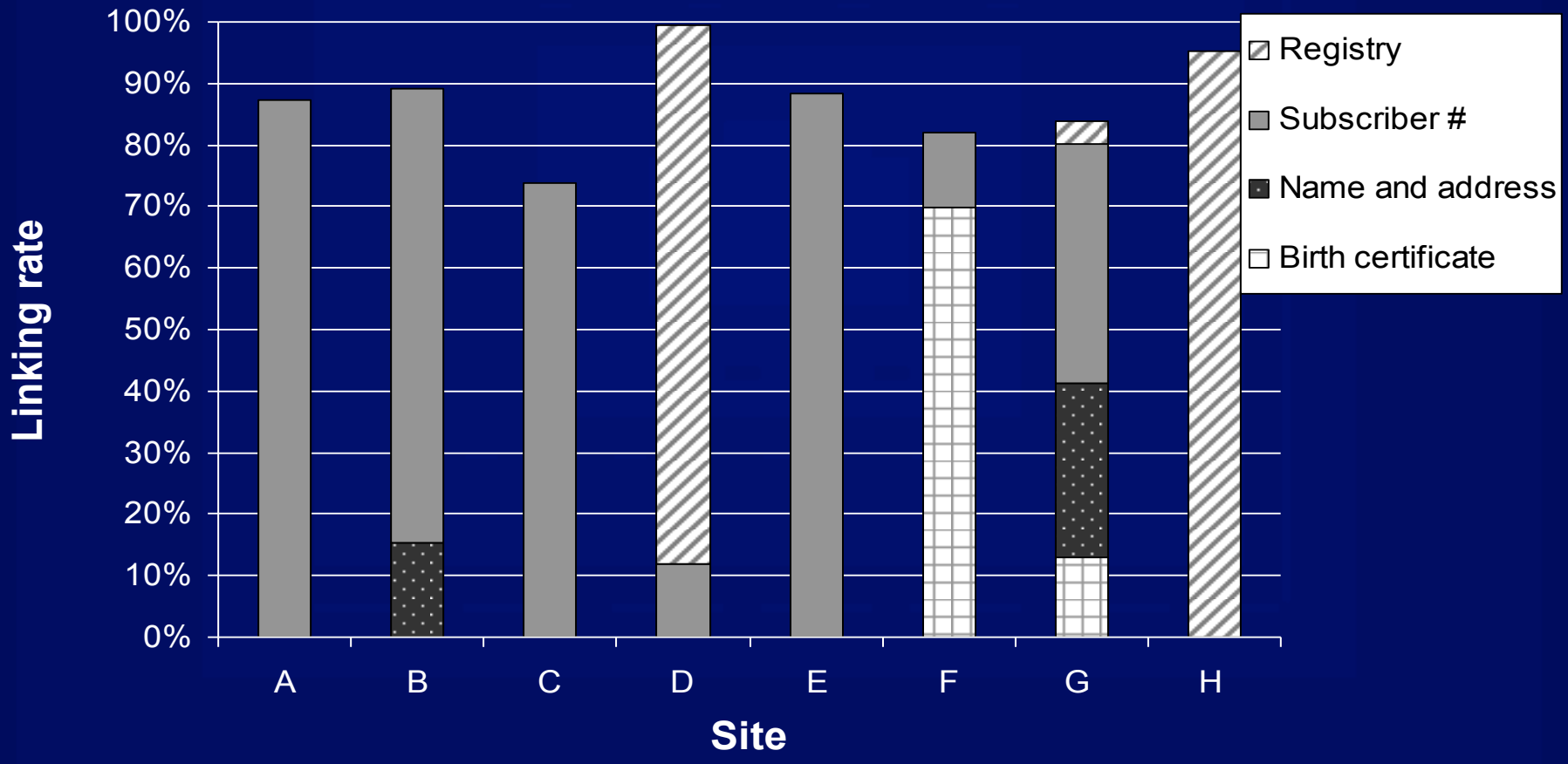
GroupHealth RESEARCH INSTITUTE

- Birth registry if available
- Subscriber number or insurance contract number
- Name and address matching (last names, address)
- Other methods
  - Birth certificate linkage

MEPREP Linkage Methods

GroupHealth RESEARCH INSTITUTE

There is great variability across sites.



Johnson et al., Pharmacoepidemiology & Drug Safety, 2013

Context: Policy Changes

GroupHealth RESEARCH INSTITUTE

- National initiatives to reduce or eliminate early elective delivery (< 39 weeks)
- In some states, Medicaid will not pay for elective deliveries at < 39 weeks
- Washington state initiatives in 2013
  - Eliminate early elective delivery < 39 weeks
  - Reduce elective induction at 39-40 weeks
    - The justification is that EI increases the risk of cesarean delivery
    - Lacks sound evidence base

Evidence Base

GroupHealth RESEARCH INSTITUTE

- Evidence to support these policies is weak
  - Studies showing infants born at < 39 weeks did worse
    - Indicated preterm deliveries for maternal illness, e.g. severe hypertension
    - Spontaneous preterm births – infection?
  - In observational studies, best outcomes seen with delivery at 39 weeks – and worse at later gestational ages
  - Induction recommended at 41 weeks to prevent stillbirth

Methodologic Issues

GroupHealth RESEARCH INSTITUTE

- Who is the right comparison group?
  - Most studies compared EI to spontaneous labor at same gestational age
  - Not clinically relevant
  - Alternative to induction is waiting, with delivery at later gestational age
  - With waiting, complications can develop, and woman may ultimately be induced later or have urgent cesarean delivery

Elective Induction

GroupHealth RESEARCH INSTITUTE

- In obstetrics, there are major evidence gaps with respect to many clinical decisions
- Electronic health data offer opportunities to learn from real-world care
- Challenge: many existing datasets lack key data elements needed
  - Supplemental data may be needed

Group Health

GroupHealth
RESEARCH INSTITUTE

- Integrated healthcare delivery system in Northwest US
  - Provides health care and insurance coverage
  - About 600,000 members and 6500 deliveries per year
  - 2/3 receive care within delivery system (richer data available)

Research Resources

GroupHealth
RESEARCH INSTITUTE

- Defined, accessible population of enrolled members
- High quality, clinically relevant automated data (current and historical) that are organized for research use

Resources and Capabilities

GroupHealth RESEARCH INSTITUTE

Automated data files

Ambulatory Care with Dx
Pharmacy
Inpatient
Radiology
Laboratory
Pathology
Costs of care
Disease registries
Immunization registry
Breast cancer screening registry
Electronic Medical Record

Research capabilities

Survey Research Program
Research Clinic
Medical Records Abstraction
Data Management
Records Linkage Studies
“Real-world” intervention trials
Multi-Center Studies

GroupHealth
RESEARCH INSTITUTE

MEPREP Members