The Perinatal Periods of Risk Approach

Preparation of Data

- Define study population
- Obtain the “raw” data files
- Assess data quality
- Restrict study population by birthweight and gestational age (excluding extremely premature cases)
- Assure sufficient number of deaths (at least 60 deaths in at most 5 years)

Place: occurrence vs. residence

- In the US, births and deaths are reported in the place (city, county, state) in which the event occurs
- Place of residence is also recorded
- If a birth mother resides in a state different from where she gave birth, a copy of the birth certificate is sent to mother’s state of residence
- Death certificates are sent to the decedent’s state of residence, if known

Cause of Death

- Part I -- reporting a chain of events leading directly to death, with the immediate cause of death (the final disease, injury, or complication directly causing death) on line a and the underlying cause of death (the disease or injury that initiated the chain of events that led directly and inevitably to death) on the lowest used line.
- Part II -- reporting all other significant diseases, conditions, or injuries that contributed to death but which did not result in the underlying cause of death given in Part I.

Underlying cause of death

- Underlying cause-of-death is coded by NCHS based on the conditions entered by the physician on the cause of death section of the death certificate
- Determined by the sequence of conditions on the certificate, provisions of the ICD, and associated selection rules and modifications
- Classified in accordance with the International Classification of Disease using the Tenth Revision (ICD-10)

Infant Death Data

- Death certificates in a state are maintained by state vital statistics registrar
  - If a person dies in a state different from the one in which he / she resided, then a copy of certificate is sent to decedent’s state of residence
- All infant deaths (less than 1 year) are linked to the baby’s birth certificate
What does it mean to “link” the death record to the birth record?

- When a baby dies we find the birth information for that same baby and put the two records together into one “row” in a computer data file.
- Most birth certificates are easy to find and match exactly (same mother, same birthday, same address etc.)
- If the mother has moved or changed names or if the baby was adopted this can be more difficult but eventually almost all are found.

Why do we need to link the death record with the birth record?

- The death record tells us the cause of death and the age at death
- The birth record tells us birth weight
- Baby’s health
- Mother’s health history & demographics
- Birth circumstances

Creating the linked death file

- Live Birth Certificate Data File
- Death Certificates
- List of infant deaths with all information

Each death is “linked” to the corresponding birth certificate so that birth information is known.

Data quality problems: Under-reporting

- Under-reporting is probably the largest source of bias, especially for fetal deaths
- Under-reporting can be difficult to detect
- Here is an example --- our investigation of fetal death certificates

The 64 largest cities formed a bell-shaped curve

At 26-28 weeks gestation, the cities still formed a bell-shaped curve
No more bell-shaped curve. Many cities seem to have no fetal deaths at 18-19 weeks gestation, while others have extremely high rates. Why the discrepancy?

The same data plotted by gestational age:

Fetal Mortality Rate* Distribution across US Cities, by Gestational Age

(*fetal deaths per thousand live births <28 weeks gestation)

Solution: Restricting Birthweight and Gestational Age

• FETAL DEATHS >=24 WEEKS AND >= 500 GRAMS
• LIVE BIRTHS >=500 GRAMS

Below these limits, reporting is NOT consistent between hospitals, among cities, and across states
• Comparisons can be invalid

Missing data elements introduce bias

• PPOR needs maternal residence and infant weight at birth
• An infant death that has not been linked to the birth certificate cannot be used in PPOR, artificially decreasing the mortality rate
• If a birth certificate is missing the birth weight data element, it cannot be used in the numerator or denominator
• Often, higher percentages of necessary information are missing among infant deaths than among the births that survived, artificially decreasing the mortality rate

Implausible data elements should not be used (treat as missing)

• Check for very large or very small values, the so-called “outliers”
• Scan and sort data to check for ‘curious’ or potentially mis-coded data
  – E.g., birth weight entered as pounds and ounces instead of grams
• Check for combinations of data elements that are impossible or “implausible”
  – E.g., a baby weighing 2900 grams at only 20 weeks gestation
• Check for out-right errors
Implausible birth weight and gestational age combinations are blacked out. There are other schemes; you don’t have to use these limits.

```
if ((gest_lmp<20 and grams>=500)
   or (gest_lmp>=20 and gest_lmp<24 and grams>=2000)
   or (gest_lmp>=24 and gest_lmp<28 and grams>=3000)
   or (gest_lmp>=28 and gest_lmp<32 and grams>=4000)
   or (gest_lmp>=32 and gest_lmp<47 and grams<1000 and plur=1))
then gest_lmp=99;
```

**Procedures for Assessing Data Quality**

- Count missing data elements and unlinked deaths
- If the fetal and infant mortality rates calculated from your data files do not match published rates, you should find out why
- Test for implausible values (such as very high birth weight with very low gestational age)

**Imputing missing data elements**

- If more than 5-10% of births, deaths, and fetal deaths are missing key data items (such as birth weight, gestational age, maternal residence, age at death), then imputation is recommended
  - e.g. estimating birth weight based on gestational age
How do we measure gestational age?

- Physician estimate of gestational age takes into account LMP, ultrasound, physical exams
- Last menstrual period estimate of gestational age requires error checking
  ***use whichever is best***
- Imputing gestational age from birth weight (and vice versa) is reasonable

Strengths and Limitations of Vital Records

- Strengths
  - Electronic availability
  - Completeness for births, deaths
  - Linkage possible
  - Expansion of surveillance systems – e.g., PRAMS, BRFSS, YRBS
- Limitations
  - Timeliness of vital records
  - Access to local level data
  - Completeness, adequate response rate
  - Changes in vital records systems
  - Small numbers

Limitations of Mortality Data

- Underlying cause of death may exclude information pertaining to the immediate cause of death, contributory causes, and intervening causes
- Quality of data
  - Validity studies of death certificates compared with hospital and autopsy records generally find higher validity for certain diseases
- Differences in ICD coding categories and regulations

Special Considerations for Phase 2 Analysis

- Vital records and other data sources should be used
- How to set up a dataset for regression modeling and other examination of the relationship between risk factors and outcomes
- Data quality concerns

To examine prevalence of risk factors

- To examine prevalence of risk factors, the data source should be representative of the population at risk (e.g., all live births plus fetal deaths etc. are at risk for premature birth)
- The data source should include information needed to define the study and reference populations (i.e. mother’s race, ethnicity, age, education, and place of residence at time of the baby’s birth)

To examine the relationship between risk factors and outcomes

- To examine the relationship between risk factors and outcomes (death, age at death, cause of death) the data source needs to be linked to the mortality record so that the outcome of each individual pregnancy is known.
  - The analyst must create an analysis file that includes all live births (and maybe fetal deaths), not just those that resulted in death.
  - Since there are thousands of births, this is done with computer software such as SAS.
Creating a birth cohort file
for births and fetal deaths occurring in year x

- Fetal file is “stacked” on birth file

SAS code for merging births
with linked deaths
by matching birth certificate numbers

- (both datasets previously sorted by birth certificate number)
  - data merged;
  - merge births linked;
  - By bcert;
  - run;

SAS code for stacking the merged file
on top of the fetal death file

data all; set merged fetal; run;

- Assure that variables named and formatted consistently, e.g.
  - For birth weight, “grams” will not stack on top of “bwt”
  - Don’t stack a character variable on top of a numeric

Creating a birth cohort from linked death files

<table>
<thead>
<tr>
<th>Linked Death File</th>
<th>Year Born</th>
<th>Year Died</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>2008</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>2010</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>2010</td>
</tr>
</tbody>
</table>

Phase 2 Data Preparation:
Schematic of Data File

<table>
<thead>
<tr>
<th>ID</th>
<th>Outcome</th>
<th>Birth Certificate</th>
<th>Death Cert.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Birth-weight</td>
<td>Maternal Age</td>
</tr>
<tr>
<td>Fet01</td>
<td>Fetal</td>
<td>798</td>
<td>17</td>
</tr>
<tr>
<td>Fet02</td>
<td>Fetal</td>
<td>2537</td>
<td>34</td>
</tr>
<tr>
<td>LB01</td>
<td>Survive</td>
<td>3511</td>
<td>22</td>
</tr>
<tr>
<td>LB02</td>
<td>Death</td>
<td>2314</td>
<td>25</td>
</tr>
<tr>
<td>LB03</td>
<td>Survive</td>
<td>1293</td>
<td>21</td>
</tr>
<tr>
<td>LB04</td>
<td>Death</td>
<td>631</td>
<td>26</td>
</tr>
</tbody>
</table>

PPOR is not intended to be a research method, results may vary.
(May common sense prevail.)

- Observational data
- Reporting inconsistencies
- Dependence among cases
- Models often do not fit the data well. The limitations of models should be carefully assessed.
Vital Statistics Resources

- National Center for Health Statistics – VitalStats
  - Collection of vital statistics products including tables, data files, and reports
  - US, state, county, MSA, and city level data
  - Allows users to access and examine vital statistics and population data interactively
  - Use prebuilt tables and reports for quick access to statistics
  - Use the data files to create your own tables—choosing from over 100 variables

- CDC -- WONDER
  - An easy-to-use, menu-driven system that makes the information resources of the CDC available to public health professionals and the public at large
  - http://wonder.cdc.gov/