The Perinatal Periods of Risk Approach

Vital Records Data – Access and Preparation
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)
Clearly Define the Study Population

- Agree on geo-political boundaries, racial/ethnic or other subgroup, timeframe

- Include infants and fetal deaths whose mothers were residents of the study area (City, e.g.) at the time of the birth
What vital records data are needed for the PPOR MAP?

1. Live births
2. Fetal deaths
3. Infant deaths, linked to birth records

ALL are produced by every state, but they are sometimes difficult for local health departments to obtain.

Note: Spontaneous and induced abortions are NOT included!
Live Births

- Birth certificate required by state law for every birth
- Hospital or birth attendant is legally responsible for completing and filing with the state
- NCHS compiles data for all births in the U.S.
What PPOR Needs from Birth Certificate

• **Phase 1**
  – Maternal residence
  – Birth year of infant
  – Birth weight
  – (Gestational age for imputation of BW)

• **Reference Group**
  – Maternal age, education, race, ethnicity

• **Phase 2**
  – Everything
## Schematic of Birth Data File

<table>
<thead>
<tr>
<th>ID</th>
<th>Maternal Residence</th>
<th>Mother’s Race / Origin</th>
<th>Maternal Age</th>
<th>Estimated Gestational Age</th>
<th>Detail Birth Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC001</td>
<td>Peoria</td>
<td>8</td>
<td>17</td>
<td>29</td>
<td>798</td>
</tr>
<tr>
<td>BC002</td>
<td>Peoria</td>
<td>6</td>
<td>34</td>
<td>41</td>
<td>2,537</td>
</tr>
<tr>
<td>BC003</td>
<td>Peoria</td>
<td>8</td>
<td>22</td>
<td>38</td>
<td>3,511</td>
</tr>
<tr>
<td>BC004</td>
<td>Peoria</td>
<td>8</td>
<td>25</td>
<td>32</td>
<td>2,314</td>
</tr>
<tr>
<td>BC005</td>
<td>Peoria</td>
<td>8</td>
<td>21</td>
<td>32</td>
<td>1,757</td>
</tr>
<tr>
<td>BC006</td>
<td>Peoria</td>
<td>2</td>
<td>20</td>
<td>40</td>
<td>3,459</td>
</tr>
<tr>
<td>BC007</td>
<td>Peoria</td>
<td>1</td>
<td>26</td>
<td>24</td>
<td>680</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
Infant Deaths

• A certificate of death is required by state law for every death in the United States

• Physicians or medical examiners / coroners are required to pronounce death and to complete the “cause of death” portion of the death certificate

• NCHS compiles data for all deaths in the U.S.
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).
What PPOR Needs from Death Certificate

• Phase 1:
  – Date of death
  – Linkage to birth certificate to obtain age at death

• Phase 2:
  – Cause of death
  – Anything that could be a risk factor
Creating the linked death file

List of infant deaths with all information
Each death is “linked” to the corresponding birth certificate so that birth information is known.
Fetal Death

- Death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy
- Reporting requirements vary from state to state
What PPOR needs from the fetal death certificate

• Phase 1
  – Maternal place of residence
  – Year of delivery
  – Birth weight and gestational age

• Reference Group
  – Maternal age, race, ethnicity, education

• Phase 2
  – Cause of death
  – Everything
Fetal deaths are additional “cases” (data elements related to birth and death)
Preparation of Data-- Assess Data Quality

What does “data quality” mean?

- All cases (babies) are included
- All data items are in each baby’s record
- The data items were accurately known and correctly recorded

Missing information causes biased results
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.

- CityMatCH and the CDC examined the distribution of fetal mortality rates in cities across the country.
For term stillbirths, the 64 largest cities formed a bell-shaped curve.
At 26-28 weeks gestation, the cities still formed a bell-shaped curve.
Many cities seem to have no fetal deaths at 18-19 weeks gestation, while others have extremely high rates.
Fetal mortality "rate" distribution
for three gestational ages (each point is a city)
Fetal Mortality Rate* Distribution across US Cities, by Gestational Age

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

• The cities reporting low rates are the cities that were NOT reporting all deaths

• Some cities, hospitals, and physicians in states requiring reporting had suspiciously low rates

• We cannot reliably study disparities among the smallest babies when there is inconsistent reporting
Solution: Restricting Birthweight and Gestational Age

- FETAL DEATHS $\geq 24$ WEEKS AND $\geq 500$ GRAMS
- LIVE BIRTHS $\geq 500$ 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”
  – 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
Implausible birth weight and gestational age combinations are blacked out

If plurality is greater than 1, the combinations in BLUE become plausible.
SAS code
for implausible combinations
of gestational age and birthweight

```sas
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
Imputation Algorithm for Fetal Deaths

GA ≥ 32

BW Unknown

Y

N

GA ≥ 24

BW ≥ 1500

<500

<24

500 ≤ BW < 1500

GA ≥ 24

BW ≥ 500

Vital Records Access / Preparation - 31
Imputation Algorithm for Live Births

BW Unknown

- **GA >= 31**
  - Y: **BW >= 1500**
  - N: **GA >= 22**
    - Y: **500 <= BW < 1500**
    - N: <500

- **GA Unknown**
  - N / A
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
After processing, how many deaths?

- Does your study sample have at least 60 fetal and infant deaths that meet PPOR criteria?
- If necessary, re-define the study population (geo-political boundaries, racial/ethnic group, timeframe)
Why 60 deaths minimum?

Rates based on a small number of events fluctuate AT RANDOM, even when there is no real underlying change in conditions (they are “unstable”).

![Graph showing death rates in different cities with 95% confidence intervals.](image)
This recommendation is based on a PPOR analysis of US cities with at least 250,000 or more population which showed that most US cities with 60 feto-infant deaths would generally have no fewer than 10 feto-infant deaths in any one risk period.

If the expected number of deaths is 10
   – the 95% confidence interval is about 5 to 16 (+/- 50%)

Online Poisson Confidence Interval Calculators
http://statpages.org/confint.html
http://anesi.com/poisson.htm
Minimum Number of Deaths

• At least 60 fetal and infant deaths that meet PPOR criteria, for each population being studied (minimum of 10 per cell)

• May combine up to 5 years (no more, due to changes in medical practice and public health systems)

• Phase 2 analyses require even more deaths.
What if your numbers are too small to create stable rates (<20? <10?)

1. Combine more years of data
   **ASSUME ALL YEARS ARE ALIKE**

2. Combine groups of people (e.g. all race/ethnicities instead of only black)
   **ASSUME ALL GROUPS ARE ALIKE**

3. Combine geographical areas (e.g. county instead of city, three neighboring counties)
   **ASSUME ALL GEO AREAS ARE ALIKE**

4. Use state data (also allows use of state datasets such as PRAMS)
   **ASSUME YOUR CITY IS LIKE THE STATE**

Each of these requires an assumption
1. Green Box Infant Health period: do not use 5-6 categories for cause of death analysis. Instead, collapse all categories with small numbers into one category (other). Many cities end up with SUID/Other.

2. Blue Box Maternal Health/Prematurity Period: We have proposed alternate birth weight stratification limits for Kitagawa that combine strata with the smallest numbers:

<table>
<thead>
<tr>
<th>Birth Weight Range</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-499</td>
<td></td>
</tr>
<tr>
<td>500-699</td>
<td></td>
</tr>
<tr>
<td>700-899</td>
<td></td>
</tr>
<tr>
<td>900-1,199</td>
<td></td>
</tr>
<tr>
<td>1,200-1,499</td>
<td></td>
</tr>
<tr>
<td>1,500-1,999</td>
<td></td>
</tr>
<tr>
<td>2,000-2,499</td>
<td></td>
</tr>
<tr>
<td>2,500+</td>
<td></td>
</tr>
</tbody>
</table>

(note that only rows 2 and 3 are used in estimating the reasons for excess mortality in the MH/P period of risk)
If you don’t think any of those assumptions is valid,

- You can assume your own opinion is correct (or your stakeholders’ opinions)

- You can use several of those methods and see how the estimates vary, or do a more formal sensitivity analysis

- Or you can use a synthetic estimate, which means you partly believe in the state data, but adjust based on known differences between your population and theirs.
Example: State of MIND

- State of MIND has a 44.84% rate of unintended pregnancy based on PRAMS statewide data.
  - The 12% of mothers who are teens have a 95% rate.
  - The older mothers have a 38% rate of unintended.

<table>
<thead>
<tr>
<th></th>
<th>Number of mothers</th>
<th>Number of unintended</th>
<th>% with unintended pregnancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teens</td>
<td>12,000</td>
<td>11,400</td>
<td>95%</td>
</tr>
<tr>
<td>Adults</td>
<td>88,000</td>
<td>33,440</td>
<td>38%</td>
</tr>
<tr>
<td>Total</td>
<td>100,000</td>
<td>44,840</td>
<td>44.84%</td>
</tr>
</tbody>
</table>
You happen to know there is an important population difference between YOUNGCOUNTY and STATE OF MIND.

YOUNGCOUNTY has more teen births.

So you would expect YOUNGCOUNTY to have a higher unintended rate than STATE OF MIND.
In synthetic estimation

- We assume the unintended rates for teens and older mothers are the same in YOUNGCOUNTY as in STATE OF MIND (95% and 38% respectively), i.e. use the STATE OF MIND unintended rates.

- We use the actual YOUNGCOUNTY numbers of mothers

<table>
<thead>
<tr>
<th>YOUNGCOUNTY</th>
<th>State % with unintended pregnancy</th>
<th>YOUNG COUNTY Number of mothers</th>
<th>Estimated Number of YOUNGCOUNTY unintended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teens</td>
<td>95%</td>
<td>3,000</td>
<td>2,850</td>
</tr>
<tr>
<td>Adults</td>
<td>38%</td>
<td>7,000</td>
<td>2,660</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>10,000</td>
<td>5,510</td>
</tr>
</tbody>
</table>

YOUNGCOUNTY Estimate is 55.1 % unintended
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