Outline of Talk

• Sibling Recurrence Risk (SRR) and SRR ratio
• Counting rules used in national surveys
• Diabetes data from 1976 National Health Interview Survey
• Estimators of SRR & SRR ratio
• SRR and SRR ratio of diabetes from 1976 NHIS
• Conclusions

Sibling Recurrence Risk

• **Sibling Recurrence Risk** (SRR) is a measure of aggregation of disease within families, i.e., aggregation among siblings. Indicates genetic cause for a disease.

• SRR is the probability that a sibling of a diseased individual also is diseased.

\[ K = \frac{\text{total # of affected sibs of diseased individuals}}{\text{total # of sibs of diseased individuals}}. \]

OR the proportion of all pairs of siblings both diseased.

• Let there be \( T \) sibships for a population

\[
K = \frac{\sum_{t=1}^{T} a_t (a_t - 1)}{\sum_{t=1}^{T} a_t (s_t - 1)}
\]

\( a_t = \# \text{ affected sibs in sibship } t. \)

\( s_t = \# \text{ sibs in sibship } t. \)
Sibling Recurrence Risk Ratio

- A related measure of family aggregation is sibling recurrence risk ratio (SRR ratio) is $K/p$, $p$ is the population prevalence of the disease.

- Genetic studies obtain $p$ externally.
Estimating Sibling Recurrence Risk in Genetic Studies

• In genetic studies two basic sampling schemes used to estimate $K$ -- *complete ascertainment* and *single ascertainment*.

• Complete ascertainment a sample of sibships is available from which disease status of sibs is obtained.

• Single ascertainment a sample of “probands” of affecteds (diseased) provide disease status of sibs.

• Genetic study samples are convenience samples based on volunteers or ad hoc collection of families – could lead to ascertainment bias.
Survey Sampling & Counting Rules

- Our approach is to apply single ascertainment idea to sample surveys.

- We define a rule for linking sampled individuals in selected households to the population elements, which are the sibships. This rule is called a counting rule.

- Consider two counting rules:
  1. Any sampled individual reports about disease status of their sibship (AllSCR)
  2. Only affected (diseased) sampled individuals report about disease status of their sibship (AffSCR)
Survey Sampling & Counting Rules

• For each counting rule, report disease status of only living siblings. We also consider reporting disease status of both living and deceased siblings.

• Under each counting rule a sibship can be reported multiple times, e.g., AllSCR a sibship can be reported by each sampled living sibling.

• Need to down weight each reported sibship by the # of individuals who can report sibship, called network weight. For AllSCR network weight is the number of living sibs, for AffSCR network weight is number of living affected sibs.

• Note for reporting living & deceased siblings there may be sibships where all the affecteds are deceased -- estimated SRR under AffCR will be biased. We used a hybrid estimator that uses AllCR for these sibships and AffCR for the rest of the sibships.
Living Sibships

Household Population

IA IU
IA
IU IU IA
IA IU

Sibships

SA SA SU
SA SU
SA SU SU
All Sibling Counting Rule (Living Sibships)

Household Population

Sibships
All Sibling Counting Rule with Household Sample (Living Sibships)

Household Population

- $S_A S_A S_U$
- $I_A$
- $S_A S_U$
- $S_A S_U S_U$
- $I_A I_U$

Sibships
Affected Sibling Counting Rule
(Living Sibships)

Household Population

Sibships
Affected Sibling Counting Rule with Household Sample (Living Sibships)

Household Population

- S_A S_A S_U
- I_A
- S_A S_U
- I_A I_U

Sibships
1976 National Health Interview Survey (NHIS)

- NHIS a cross-sectional household survey of all ages of US population. Design is a stratified multistage cluster household sample with sample weights. 41,559 sampled households, n = 113,178 sampled individuals.

- In 1976 a diabetes supplement collected interview data about diabetes from the respondent about his/herself and his/her family:
  
  Q 1. Do you have diabetes?
  
  Q 2. How many living brothers and sisters do you have? *(Not including adopted, step or half siblings)*
  
  Q 3. How many of these brothers and sisters have diabetes?
  
  Q 4. How many of your brothers and sisters are no longer living?
  
  Q 5. How many of these brothers and sisters had diabetes?
Estimators of Sibling Recurrence Risk

- Population recurrence risk

\[ K = \frac{\sum_{t=1}^{T} a_t (a_t - 1)}{\sum_{t=1}^{T} a_t (s_t - 1)} = \frac{\sum_{i=1}^{N} a_i (a_i - 1)}{\sum_{i=1}^{N} a_i (s_i - 1)} \]

- Recurrence risk estimator under AllSCR

\[ \hat{K}_{AllSCR} = \frac{\sum_{i=1}^{n} w_i a_i (a_i - 1)}{\sum_{i=1}^{n} w_i a_i (s_i - 1)} \]
Estimators of Sibling Recurrence Risk

- Recurrence risk estimator under AffSCR

\[ \hat{K}_{\text{AffSCR}} = \sum_{i \in A} \frac{w_i a_i(a_i - 1)}{a_i} = \frac{\sum w_i (a_i - 1)}{\sum w_i (s_i - 1)} \]

where \( A \) is the set affecteds in the sample.
Network Estimators of Sibling Recurrence Risk Ratio

- Population recurrence risk ratio: the recurrence risk divided by the population prevalence of the disease \( \lambda = K/p \)

- Recurrence risk ratio estimator under AllSCR
  \[
  \hat{\lambda}_{AllSCR} = \frac{\hat{K}_{AllSCR}}{\hat{p}_{AllSCR}}
  \]
  where
  \[
  \hat{p}_{AllSCR} = \frac{n}{\sum_{i=1}^{n} \frac{w_i}{s_i} a_i} \div \sum_{i=1}^{n} w_i
  \]

- Recurrence risk estimator under AffSCR
  \[
  \hat{\lambda}_{AffSCR} = \frac{\hat{K}_{AffSCR}}{\hat{p}_{AffSCR}}
  \]
  where
  \[
  \hat{p}_{AffSCR} = \frac{\sum_{i \in A} \frac{w_i}{a_i} a_i}{\sum_{i=1}^{n} w_i} \div \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} w_i}
  \]
Variance Estimation

• Since national household surveys such as NHIS are stratified multistage cluster samples, apply survey methods for variance estimation of nonlinear estimators: 

\[
\hat{K}_{AllCR}, \hat{K}_{AffCR}, \hat{\lambda}_{AllCR}, \hat{\lambda}_{AffCR}
\]

• Taylor linearization method of variance estimation, other possibilities are replication methods such leaving-one-out jackknife or balanced repeated replication methods

• Since the magnitude of estimators vary by counting rule we use relative standard error (RSE) = SE(estimator)/estimator.
Estimation of Sibling Recurrence Risk and Recurrence Risk Ratio for Diabetes from the 1976 National Health Interview Survey
### Number of sampled individuals in 1976 NHIS reporting sibships with diabetes

<table>
<thead>
<tr>
<th>Counting rule</th>
<th>Diabetics in the sibship</th>
<th># of Sampled individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≥1 living diabetic</td>
<td>5,831</td>
</tr>
<tr>
<td>All Siblings</td>
<td>≥2 living diabetic</td>
<td>1,112</td>
</tr>
<tr>
<td></td>
<td>≥1 living diabetic</td>
<td>2,005</td>
</tr>
<tr>
<td></td>
<td>≥2 living diabetic</td>
<td>571</td>
</tr>
<tr>
<td>Affected Siblings</td>
<td>≥1 living or deceased diabetic</td>
<td>6,914</td>
</tr>
<tr>
<td></td>
<td>≥2 living or deceased diabetic</td>
<td>1,533</td>
</tr>
<tr>
<td></td>
<td>≥1 living or deceased diabetic</td>
<td>2,197</td>
</tr>
<tr>
<td></td>
<td>≥2 living or deceased diabetic</td>
<td>712</td>
</tr>
</tbody>
</table>
## NHIS Estimates for Family Aggregation of Diabetes

<table>
<thead>
<tr>
<th>Counting Rule</th>
<th>Prevalence % (RSE%)</th>
<th>Sibling Recurrence Risk % (RSE%)</th>
<th>Sibling Recurrence Risk Ratio (RSE %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reporting Living Siblings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected Siblings</td>
<td>2.16 (2.61)</td>
<td>12.51 (4.56)</td>
<td>5.79 (5.12)</td>
</tr>
<tr>
<td>All Siblings</td>
<td>1.97 (1.79)</td>
<td>15.10 (4.03)</td>
<td>7.66 (3.76)</td>
</tr>
<tr>
<td><strong>Reporting Living or Deceased Siblings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected Siblings</td>
<td>2.34 (2.43)</td>
<td>13.75 (4.43)</td>
<td>5.87 (4.38)</td>
</tr>
<tr>
<td>All Siblings</td>
<td>2.16 (1.72)</td>
<td>15.37 (3.89)</td>
<td>7.10 (3.46)</td>
</tr>
</tbody>
</table>

**RSE** relative standard error
# NHIS Estimates for Family Aggregation of Diabetes

<table>
<thead>
<tr>
<th>Counting Rule</th>
<th>Prevalence % (RSE%)</th>
<th>Sibling Recurrence Risk % (RSE%)</th>
<th>Sibling Recurrence Risk Ratio (RSE %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reporting Living Siblings with a Diabetic Parent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected Siblings</td>
<td>6.13 (4.19)</td>
<td>18.08 (6.26)</td>
<td>8.37* (6.74)</td>
</tr>
<tr>
<td>Any Siblings</td>
<td>6.05 (2.74)</td>
<td>20.93 (5.20)</td>
<td>10.62* (5.04)</td>
</tr>
<tr>
<td><strong>Reporting Living Siblings w/o Parental Diabetes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected Siblings</td>
<td>1.68 (2.89)</td>
<td>9.50 (5.99)</td>
<td>4.40* (6.37)</td>
</tr>
<tr>
<td>Any Siblings</td>
<td>1.47 (2.05)</td>
<td>11.45 (5.27)</td>
<td>5.81* (5.62)</td>
</tr>
</tbody>
</table>

*RSE* relative standard error

* Prevalences 2.16 and 1.97 used.
Summary of Results from NHIS Example

- The prevalence of diabetes is slightly larger for AffSCR than AllSCR – indicating that the reporting may be more accurate from affected probands (assuming miss-reporting results in under estimation).

- The prevalence of diabetes increased more than three fold when parents had diabetes.

- Estimates of SRR & SRR ratio were smaller for AffSCR than AllSCR.

- The RSEs were smaller for the AllSCR than AffSCR since the networks were larger under AllSCR.

- SRR similar studies at about the same time period. SRR decrease in more recent time periods – indicating changes in environmental factors.
Conclusions

• National household surveys can be used to estimate of SRR and SRR ratios for diseases not too rare.

• Work needed on accuracy of reporting disease status under different counting rules, accuracy of reporting network sizes used for network weights.
  -- NHIS restricted to noninstitutionalized US population – no restriction for reporting siblings so network sizes would be too large and network weights too small and estimates biased low.

• Other counting rules that could be considered, e.g., combining counting rules or hybrid counting rules, to reduce sampling variance and to reduce reporting bias.
References


