RISK ASSESSMENT MODELS FOR FAMILIAL BREAST CANCER

Giovanni Parmigiani
gp@jhu.edu
Johns Hopkins University

December 5, 2002
1. BACKGROUND

- BRCA1 & BRCA2 are tumor suppressor genes

- Germline deleterious mutations increase susceptibility to breast & ovarian cancer colon, prostate and other cancers.

- Inheritance of deleterious mutations is autosomal dominant.
1.1. SUSCEPTIBILITY PREDICTION MODELS

Family history can be very informative about the presence of a mutation.

Predicting mutations is possible and useful in two contexts:

HIGH RISK CLINICS

- Counseling about testing decisions
- Interpretation test outcomes *for individuals*
- Predicting who will develop cancer

GENE CHARACTERIZATION RESEARCH

- Selecting high risk subjects
- Building measures of susceptibility
1.2. TOPICS

Approaches to susceptibility prediction

Empirical vs Mendelian Modeling
Principles of Mendelian Modeling

Software

BRCAPRO
CaGene

Validation Studies

“301” Study Results
CGN Study Protocol
2. SUSCEPTIBILITY PREDICTION

OTHER FAMILIES

- BRCA1
- BRCA2
- BRCAX

POLYGENETIC / ENVIRONMENTAL

"HIGH RISK" FAMILIES

CHANCE CLUSTER
2.1. EMPIRICAL MODELING

\[ P \left( \begin{array}{c}
\text{Positive Genetic Test} \\
\text{Pedigree Information}
\end{array} \right) \]

Correlates genetic testing results
to features of family history

Relies on AI/statistics to infer
the **genotype | phenotype** relationship
and the **mode of inheritance**

Generally gives broad classes of families
2.2. MENDELIAN MODELING

\[ P \left( \begin{array}{c}
\text{Deleterious Mutation at Susceptibility Gene} \\
\text{Pedigree Information}
\end{array} \right) \]

Derives carrier probabilities from genetic parameters.

Relies on statistics to infer the phenotype | genotype relationship.

Relies on Mendel’s laws for the mode of inheritance.
2.3. RELATIONSHIP BETWEEN SCALES OF EMPIRICAL AND MENDELIAN PREDICTIONS

\[
\begin{align*}
\beta \times P & \left(\frac{\text{Pedigree Information}}{\text{Deleterious Mutation at Susceptibility Gene}}\right) \\
& = \left(\frac{\text{Positive Genetic Test}}{\text{Pedigree Information}}\right)
\end{align*}
\]

\(\beta\) is the Test Sensitivity

EMPIRICAL

MENDELIAN

Assumes perfect specificity
2.4. EXAMPLES OF EMPIRICAL MODELS FOR BRCA GENES

MYRIAD 1 (Shattuck-Eidens et al, JAMA, 1995)
PENN (Couch et al, NEJM, 1997)
MYRIAD 2 (Frank et al, JCO, 1998)
NCI (Hartge et al, AJHG, 1999)
FHAT (Gilpin CG 2000)
FINLAND (Vahteristo BJC, 2001)
2.5. EXAMPLES OF MENDELIAN MODELS FOR BRCA GENES

CLAUS (Claus et al, Cancer, 1994)

BRCAPRO (Berry et al, JNCI 1997, Parmigiani, AHJG, 1998)

LEIDEN (Houwing-Duistermaat, GE 1997)

ABC (Antoniou et al, BMJ, 2002)
2.6. LOGIC BEHIND MENDELIAN RISK PREDICTION: EXAMPLE (with many minor details omitted)

Computing the carrier probability for the unaffected daughter

**STEP 1.** What is the chance that the mother is a carrier?

**STEP 2.** What is the chance that the daughter is a carrier
- given that the mother is?
- given that she isn’t?
- on average?
STEP 1: mother

\[
\frac{P(\text{mutation} \mid H)}{P(\text{no mutation} \mid H)} = \frac{P(\text{mutation})}{P(\text{no mutation})} \frac{P(H \mid \text{mutation})}{P(H \mid \text{no mutation})}
\]

for example

\[
\frac{1/1000}{999/1000} \times \frac{1/100}{1/10000} = \frac{100}{999} \approx \frac{1}{10}
\]

\[P \approx 0.09\]
STEP 2: daughter

\[
\frac{P(\text{mutation} \mid H)}{P(\text{no mutation} \mid H)} = \frac{P(\text{mutation} \mid \text{mutation in mother})}{P(\text{no mutation} \mid \text{mutation in mother})} \times \frac{P(H \mid \text{mutation})}{P(H \mid \text{no mutation})}
\]

for example

\[
\frac{1/2 \cdot 95/100}{1/2 \cdot 99/100} = \frac{95}{99} \quad P \approx .49
\]

\[
\frac{1/1000 \cdot 95/100}{999/1000 \cdot 99/100} = \frac{95}{98901} \quad P \approx .001
\]

carrier probability for daughter:

\[
\approx .09 \times .49 + .91 \times .001 \approx 0.045
\]
2.7. **LOGIC BEHIND MENDELIAN RISK PREDICTION: GENERAL APPROACH**

**INFORMATION SOURCES**

**MENDEL** \( P(\text{GR} | \text{G}) \)

**EPI: penetrance** \( P(\text{H} | \text{GR}, \text{G}) = \prod_m P(H_m | G_m) \)

**EPI: prevalence** \( P(\text{G}) \)

**RISK PREDICTION**

**MARGINAL** \( P(\text{H} | \text{G}) = \sum_{GR} P(H | \text{GR,G}) \ P(\text{GR} | \text{G}) \)

**BAYES** \( P(\text{G} | \text{H}) \propto P(\text{G}) \ P(\text{H} | \text{G}) \)
2.8. MENDELIAN VS EMPIRICAL

Mendelian models are closely tailored to individual families; and are more efficient at incorporating information about patterns of inheritance and pedigree structure.

Empirical models are easier to develop and implement and potentially able to capture familial susceptibility that don’t follow simple modes of inheritance.
3. BRCAPRO

GENOTYPE

BRCA1 & BRCA2

FAMILY HISTORY

I-st and II-nd degree relatives of counseland
Breast and ovarian cancer history (m & f)
Age of onset, age of death or current age

PENETRANCES

Meta-analysis

PREVALENCES

Meta-analysis

Berry, Parmigiani, Sanchez, Schildkraut & Winer *JNCI*, 1997
Parmigiani, Berry & Aguilar *AJHG*, 1998
BRCA1  BRCA2  NEITHER

<table>
<thead>
<tr>
<th></th>
<th>nonAJ</th>
<th>nonAJ pen/2</th>
<th>AJ</th>
<th>AJ pen/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>nonAJ</td>
<td>.025</td>
<td>.005</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>nonAJ pen/2</td>
<td>.05</td>
<td>.01</td>
<td>.94</td>
<td></td>
</tr>
<tr>
<td>AJ</td>
<td>.44</td>
<td>.26</td>
<td>.30</td>
<td></td>
</tr>
<tr>
<td>AJ pen/2</td>
<td>.38</td>
<td>.19</td>
<td>.43</td>
<td></td>
</tr>
</tbody>
</table>

**AJ:** \( \pi_1 = .012, \pi_2 = .013 \)

**OTHERS:** \( \pi_1 = .0012, \pi_2 = .00044 \)
Do we know what the right answer is, for a specific proband? for a specific model?

Risk versus imperfect knowledge
3.1. Software

CaGene

- User-friendly interface
- BRCAPRO & other models
- Educational materials
- Cancer Risk Predictions
- HNPCC Dutch model

CYRILLIC

PROGENY

OTHER LINKAGE ANALYSIS PACKAGES

- e.g. MLINK version of BRCAPRO (Pisa model)
- e.g. MENDEL (Antoniou model)
3.2. Software in progress

R library for familial risk prediction

All–purpose infrastructure for carrier probability calculations in autosomal dominant genes

BRCAPRO

CRCAPRO

Menu of penetrances and prevalences

Long term goal: object oriented environment for family data analysis
4. VALIDATION STUDIES

DESIGN

Sample of tested pedigrees
High-risk clinics versus population-based

ANALYSIS

Calibration & Refinement
Accounting for sensitivity

DATASETS

“301” study (Berry et al, JCO 2002)
Cancer Genetics Network dataset (in progress)
4.1. Overall calibration

BRCAPRO

MYRIAD
NA: 157/301
NA+: 77/157

301 families from Berry et al. JCO (2002)
### 4.2. Summaries of prediction performance

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>LOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCAPRO (all)</td>
<td>0.18</td>
<td>-.55</td>
</tr>
<tr>
<td>BRCAPRO (myr)</td>
<td>0.18</td>
<td>-.53</td>
</tr>
<tr>
<td>MYRIAD</td>
<td>0.24</td>
<td>-.67</td>
</tr>
<tr>
<td>% GAIN</td>
<td>0.28</td>
<td>.22</td>
</tr>
</tbody>
</table>
5. Conclusions

Family history is a powerful predictor of genotype for major cancer genes.

Subjective assessment of family history is useful but modeling enables accurate and economical identification of candidates for testing.

Software is available and user-friendly. Use requires at least basic statistical training.

Mendelian models appear to provide a substantial gain in prediction accuracy over empirical models. Whether this gain justifies the additional data collection depends on the clinical context.
CREDITS

BRCAPRO

Duke Breast SPORE
D. Berry
O. Aguilar, S. Chen, D. Gudbjartsson, E. Iversen, X. Zhou

JCO Validation
Creighton, Dana-Farber, Georgetown, Lahey, U Pittsburgh, UT San Antonio

http://astor.som.jhmi.edu/brcapro