Biodemography, Health, and Mortality

S.Harper\textsuperscript{1,2}

\textsuperscript{1}Epidemiology, Biostatistics & Occupational Health, McGill University

\textsuperscript{2}Centre on Population Dynamics, McGill University

Population Association of America, San Diego, 2 May 2015
Your manuscript as submitted

... and after peer review and revision

minor revisions

introduction and discussion should be significantly expanded

the latest top-mounted laser cannon. Because.

more replicates needed

front windshield needs to be removed or tinted red

please add necessary circular and triangular windows

(horse hitch "cause that's how we always did it")

*reviewer3 sells odd-shaped windows*

please try this alternate approach

REDPEN/BLACKPEN   http://redpenblackpen.jasonya.com
Belsky: Good foundations

- Interesting question: How are young people aging?
- Very impressive number of biomarkers.
- Well-characterized cohort.
- Excellent follow-up, good retention.
- Some interesting outcomes (independent assessments of “aging”).
Belsky: A few reasons to be skeptical

- Lots of “smoothing”
  - z-scores for all biomarkers
  - sum of 18 slope coefficients
  - then scaled so that 1-year physiological change = 1 chronological year

- Used a 10-biomarker “Biological Age” from NHANES (US)
  - What about generalizability? Expect US correlations = NZ correlations?
  - What about the standard error of prediction?

- Standardized regression coefficients
  - Are not reliable effect measures.
  - How much variation is there in X for young people?

- Correlation coefficients (not useful as measures of effect)
  - is the null hypothesis even interesting?
Goldman et al.

- Excellent background and discussion of discrimination statistics
- Very nice to do multiple imputation!
- Good to add in changes to biomarkers
- Helps to identify potentially more relevant information for disease risk.
- Very good, conservative conclusions.
- Nice discussion of limitations.
Questions for Goldman et al.

- Why not use age as the time axis?
- Would be nice to use an independent dataset to evaluate model performance.
- How much of the improvement occurs at different risk thresholds?
- Challenges with comparing the explanatory power of individual biomarkers because of the temporal ordering of exposure?
- What does it mean to calculate discrimination for death?
- Include a table showing reclassification
Impact of reclassification

- Kerr et al (Epidemiology, 2014):

<table>
<thead>
<tr>
<th>Risk Category</th>
<th>Nonevent (n = 5,669)</th>
<th>Event (n = 209)</th>
<th>Nonevent (n = 5,669), New Model</th>
<th>Event (n = 209)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–3%</td>
<td>67.1</td>
<td>27.3</td>
<td>70.6</td>
<td>24.4</td>
</tr>
<tr>
<td>3–10%</td>
<td>30.6</td>
<td>55.0</td>
<td>22.3</td>
<td>38.8</td>
</tr>
<tr>
<td>&gt;10%</td>
<td>4.4</td>
<td>17.7</td>
<td>7.1</td>
<td>36.8</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*With coronary artery calcium score.
Interesting take on a common questions

Formal description of model parameters very helpful.

One very important assumption:

“Individual environments are assumed to be uncorrelated with the outcome; that is, there cannot be some unobserved factor aside from genetic endowments and common environments that affects both schooling attainment and the health-related behavioral outcome, though the violation of this condition produces predictable bounds on the causal estimates”
Challenges for Sudharsanan et al.

- Selection bias (36% and 65% complete cases for the 2 cohorts).
- Measurement error in self-reported schooling (attenuation bias stronger for within estimate).
- Standardized coefficients.
- Do we really care about “explained variance”?
- What about confounding by other health behaviors that differ between twins?
- Questionable modeling assumptions:
  - Equal environments
  - Non-shared individual factors unrelated to health behaviors?
Sudharsanan: Important assumptions

- “individual environments related to schooling do not affect health behavior.”, i.e., $E_{ij}^x \rightarrow y$ forced to be null.
- Is this plausible? (non-cognitive, non-shared factors like motivation, personality, etc.)
Challenges with “explained variance”

- Find that the majority (~60%) of variance in BMI due to “genetics”
- Recent rise of obesity (and mean BMI) cannot be solely attributable to genetic changes.
- High BMI clearly not separable
- Twins perfectly matched on birth cohort, experience the same changes in environment.
- Drivers of population increases in BMI are ubiquitous in the US environment, so cannot explain individual variation.
  - Recall Geoffrey Rose: if everyone smoked, lung cancer would be a genetic disease
Focused on cross-country descriptions of relations between age, education, gender, and various biological measures related to health.

Consistent evidence of wide variability in the relations between education and biological measures of health.

Make a compelling case for the utility of standardized biological measurements across diverse populations.
Questions for Frankenberg et al.

- How were education-specific estimates “adjusted for age”?
  - “estimates by education are based on linear splines with knots at 6 and 12 years of education and flexible controls for age.”

- Adjusted predictions require specifying values at which age is “held constant”
  - Standardized to a common age distribution?
  - “Stratify before you standardize” (Choi et al. Am J Epid, 1997)
“Stratify before you standardize”

- Adjusted associations may reflect differences in the distribution of education, age, and cholesterol.
- Could be different across countries
- More description of distributions (and methods of adjustment) would enhance the paper.
Questions for Frankenberg et al.

- Meaning of education across countries
  - how to compare 12 years in US with 12 years in Mexico or 12 years in Ghana?

- Measurement error in education (how reliable in lower income environments?)

- More work on explanation of country variations.
  - Slope of BMI on risk markers somewhat consistent across US, Indonesia, China, Mexico
  - Intercepts vary widely (US much higher on cholesterol, Mexico on HbA1c).
Broader issues

▷ What is the goal of evaluating biomarkers of aging?

▷ Goldman et al: “from the point of view of predicting mortality – a well-measured outcome highly correlated with myriad health measures – these findings provide strong support for biomarker collection within household surveys and moderate support for longitudinal collection of such markers – particularly with respect to inflammatory markers.”

▷ Belsky et al.: “There are also potential clinical applications. Early identification of accelerated aging before chronic disease becomes established may offer opportunities for prevention.”
Broader issues

▶ What is the goal of evaluating biomarkers of aging?

▶ Goldman et al: “from the point of view of predicting mortality – a well-measured outcome highly correlated with myriad health measures – these findings provide strong support for biomarker collection within household surveys and moderate support for longitudinal collection of such markers – particularly with respect to inflammatory markers.”

▶ Belsky et al.: “There are also potential clinical applications. Early identification of accelerated aging before chronic disease becomes established may offer opportunities for prevention.”

▶ Neils Bohr: Prediction is very difficult, especially if it’s about the future.
How accurately could we screen for individual risk? Using summary data to examine discriminatory accuracy of a risk marker

Beverly J. Levine *, Douglas W. Levine

<table>
<thead>
<tr>
<th>Absolute risk of disease in unexposed</th>
<th>Relative risk</th>
<th>Prevalence of exposure</th>
<th>Concordance statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>0.001</td>
<td>1.5</td>
<td>0.502</td>
<td>0.521</td>
</tr>
<tr>
<td>0.001</td>
<td>3.0</td>
<td>0.510</td>
<td>0.575</td>
</tr>
<tr>
<td>0.001</td>
<td>10.0</td>
<td>0.541</td>
<td>0.713</td>
</tr>
<tr>
<td>0.010</td>
<td>1.5</td>
<td>0.502</td>
<td>0.522</td>
</tr>
<tr>
<td>0.010</td>
<td>3.0</td>
<td>0.510</td>
<td>0.576</td>
</tr>
<tr>
<td>0.010</td>
<td>10.0</td>
<td>0.541</td>
<td>0.717</td>
</tr>
<tr>
<td>0.050</td>
<td>1.5</td>
<td>0.503</td>
<td>0.523</td>
</tr>
<tr>
<td>0.050</td>
<td>3.0</td>
<td>0.510</td>
<td>0.580</td>
</tr>
<tr>
<td>0.050</td>
<td>10.0</td>
<td>0.543</td>
<td>0.736</td>
</tr>
</tbody>
</table>
Considerations for prevention

- It is uncommon in epidemiology to find risk factors for disease that are both associated with high relative risks for disease, and that are also prevalent in 10–40% of the population. [Levine 2007]