I like big data and I cannot lie

Or: how multivariate regression can help us make sense of overwhelming amounts of data, improve practice, and save the world

Shannon M. Campbell, MPP
Senior Research & Evaluation Analyst
Mental Health & Addiction Services, Multnomah County Health Department
Portland, Oregon

Contact information: shannon.campbell@multco.us
People who drowned after falling out of a fishing boat correlates with Marriage rate in Kentucky

Correlation: 95.24% (r=0.952407)

Data sources: Centers for Disease Control & Prevention and National Vital Statistics Reports
Why multivariate regression?

- Because correlation *isn’t* causation
- Because it helps us bring our “sliced and diced” data together into one place (here’s B broken down by A, here’s B broken down by C - but what if A explains C?)
Constructing a model: an art & a science

- Many regression models available; many different purposes
- 95% of the battle is getting the right data, in the right form
  - Thinking through all the what ifs
- No such thing as a perfect model
  - Aim: minimize problems while staying true to theory and your data
    - “See what a fraud I am”
Examples of use

- Types of projects informing policy, practice, or both
  - Explanatory models:
    - Identifying if an intervention actually impacted a desired outcome
    - Does connection to treatment after jail reduce the risk of rearrest?
    - Does behavioral healthcare reduce physical healthcare costs?
    - Do longer lengths of hospital stays reduce the readmission rate?
  - Predictive models:
    - Can we know who among our clients is at most risk of hospital admission before it occurs?
  - Identifying disparities:
    - Do different demographics have equal access to treatment?
    - Is there a significant difference in demographics between houseless and housed behavioral health clients?
Real-time stats, real-time problems: informing practice via predictive analytics

- The topic: acute care
  - Inpatient psychiatric hospitalizations, behavioral health-driven ER visits, psychiatric emergency services (Unity)
  - Want to reduce acute care utilization by engaging clients in different levels of care that sustainably address their needs
- We follow up on hospitalizations and ED visits, we craft models showing what factors may impact readmission, but:
  - Question: can we predict which of our members are most at risk of impending psychiatric crisis before they occur, so that we can intervene in a more timely manner?
Preparation

- Our dependent variable:
  - Acute care event
    - Inpatient psychiatric hospitalizations
    - Psychiatric emergency services (PES)
    - Emergency department visits attributable to mental health and/or substance use diagnoses

- Our sample:
  - HSO members with 1+ year coverage & SPMI

- Our time period:
  - January 1, 2015 to June 30, 2017 (2.5 years)

Result: 13,158 clients; 11,222 acute care events
Preparation

- Determining variables to explore
  - What data is available? Will it be available in the future?
  - Information gathering with staff; review of literature
    - What traits do high utilizers seem to have in common? What signs precede hospitalizations?

- Indicators vs. causes
  - What am I trying to find out?
  - Correlation can be sufficient without a clear causal link
    - E.g.: psychiatric hospitalizations occur because people are in crisis, people call the crisis line when they are in crisis and not in the hospital, perhaps calls to the crisis line can be a proxy for crisis
Models: two options

- **Logistic regression**
  - **Positives**: simple model to construct; straightforward assumptions; easy to compute predictive fit
  - **Negatives**: doesn’t assume multiple entries by one person or take time into account; doesn’t explicitly factor in that an event not happening during observation doesn’t mean it never happens

- **Cox multi-failure proportional hazards model (survival analysis)**
  - **Positives**: accounts for variation in time to event, for multiple sequential events by same person (Andersen-Gill method), for possibility of events happening outside observation
  - **Negatives**: requires very specific, complex data structure; not as easy to compute predictive fit

- *Results very similar; decided to use logistic, with clustered S.E.*
Results

- Significant non-demographic variables (odds ratio):
  - No recent mental health outpatient history (4.5)
  - Multiple SPMI-level diagnoses (4.3)
  - History of substance use (2.9)
  - Week with 2+ crisis line calls (2.9)
  - History of homelessness/housing instability (1.7)
  - Receiving SSI for disability (1.7)
  - Healthcare encounters with respiratory (1.6) or pain issues (1.5) as primary diagnosis

- Original model:
  - Area under the ROC curve (AUROC): 0.85
    - 0.9 to 1 considered excellent; 0.8 to 0.89 → very good
Validating results

- **Revalidation #1: Equity models**
  - Avoid systematically under- or over-predicting for any population
    - Ran model *without* demographics included, on each individual race, age, sex, language, as well as random combinations of traits
    - Intent: ensure it works well for different populations
    - Use AUROC to test predictive power of model for each group
    - **Short answer:** yes, does basically predict equally well for different demographics
Validating results

- Revalidation #2: Model with different (but similar) sample population
  - Run the exact same models with all SPMI members with less than 1 year of coverage (pop. of 3,380)
    - Odds ratios virtually the same, AUROC of 0.84; important because we often work with incomplete data → realistic scenario
What now?

- We have a set of variables collectively predicting upcoming acute care events with high degree of power that does not appear to create or further existing disparities
- An interesting model...but how do we apply it?
### Quantitative analysis

High risk clients for outreach, 2/28/2019

| Event happened                  | Odds Ratio | Robust Std. Err. | z   | P>|z| | 95% Conf. Interval |
|--------------------------------|------------|------------------|-----|-----|-------------------|
| multipedx_binary               | 4.334528   | .2200021         | 24.96| 0.000| 3.923353         | 4.78855 |
| substance use                  | 2.926588   | .157988          | 18.92| 0.000| 2.626755         | 3.25291 |
| english                        | 1.11377    | .0593887         | 1.21| 0.227| 1.035955         | 1.194641 |
| race1_black                    | .0045963   | .0008099         | -1.94| 0.056| .7762024         | 1.080442 |
| race2_asian                    | .7792272   | .1073113         | -1.81| 0.070| .6985957         | 1.030675 |
| race6_white                    | 1.061151   | .0225686         | 2.49| 0.013| 1.017783         | 1.106329 |
| race6_hispanic                 | .0403447   | .0094123         | -4.49| 0.000| .0200417         | 1.059322 |
| race5_native                   | .9595662   | .1648083         | -0.24| 0.809| .7835758         | 1.153251 |
| race6_pac1                     | .9774699   | .2765273         | -0.45| 0.653| .7698951         | 1.369184 |
| male                           | 1.539767   | .0791646         | 2.50| 0.012| 1.492169         | 1.603534 |
| homeless_or_supported          | 1.198726    | .155224          | 5.84| 0.000| 1.027501         | 1.399905 |
| wek República vs. 2 calls       | 2.892722   | .1485717         | 22.66| 0.000| 2.638547         | 3.170549 |
| pain                           | 1.164671   | .0783958         | 6.60| 0.000| 1.002273         | 1.370792 |
| respiratory                    | 1.060156   | .0936461         | 1.15| 0.247| 1.452786         | 1.803455 |
| no_op_stall                    | 1.819399   | .190065          | 36.05| 0.000| 4.160787         | 9.906747 |
| age_18to20                     | 1.219406   | .060222          | 4.09| 0.000| 1.106012         | 1.362457 |
| age_20to30                     | 1.01554    | .0450383         | 0.35| 0.732| .9509841         | 1.070763 |
| age_30to40                     | .871333    | .0473594         | -2.61| 0.009| .7791078         | .965104 |
| age_40to49                     | .7704137   | .0407774         | -4.06| 0.000| .6980794         | .850203 |
| age_60to69                     | .610254    | .0508332         | -5.93| 0.000| .5283033         | .712943 |
| cons                           | .0792867   | .0083596         | -6.48| 0.000| .0655482         | .0935038 |

Note: _one estimates baseline odds._

Condensing complex information into something easily interpreted and actionable: how do we get from A to B?

- **Jack**: Risk score: **10**
- **Diane**: Risk score: **8**
## Hypothetical client: “Harry Potter”

<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
<th>Odds ratio</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>No recent mental health outpatient history (last 120 days)?</td>
<td>Yes (1)</td>
<td>4.518399</td>
<td>= 4.518399</td>
</tr>
<tr>
<td>Multiple SPMI diagnoses (last 12 months)?</td>
<td>No (0)</td>
<td>4.334528</td>
<td>= 0</td>
</tr>
<tr>
<td>Substance use history (last 12 months)?</td>
<td>Yes (1)</td>
<td>2.928598</td>
<td>= 2.928598</td>
</tr>
<tr>
<td>Week with 2+ crisis line calls (last 3 weeks)?</td>
<td>Yes (1)</td>
<td>2.892232</td>
<td>= 2.892232</td>
</tr>
<tr>
<td>SSI for disability (any time)?</td>
<td>No (0)</td>
<td>1.696737</td>
<td>= 0</td>
</tr>
<tr>
<td>History of housing instability (any time)?</td>
<td>Yes (1)</td>
<td>1.687269</td>
<td>= 1.687269</td>
</tr>
<tr>
<td>Primary respiratory complaint at healthcare visit (last year)?</td>
<td>No (0)</td>
<td>1.606196</td>
<td>= 0</td>
</tr>
<tr>
<td>Primary pain complaint at healthcare visit (last year)?</td>
<td>No (0)</td>
<td>1.546471</td>
<td>= 0</td>
</tr>
<tr>
<td>Constant term</td>
<td>1</td>
<td>0.0372867</td>
<td>= 0.0372867</td>
</tr>
</tbody>
</table>

**Subtotal** = 12.0637847

**Scaling to range of 0 to 10**

\[
\text{Subtotal} / 2.124772 = 6
\]

**Final score**
Percent of cases experiencing acute care events, by risk score.
Back to validation of results

- One more test: how will this work in the “real world”? 
  - If we used that scoring system on our entire population, how accurate would it be?
    - “Freeze” scores on specific date
    - Track actual events for next 30 days
    - Use score as main predictor
Building the tool

- Valid model, valid score - how do staff use it?
  - Information available to staff via a Tableau dashboard, automatically updated every 24 hours
    - Look up specific members individually
    - View all members enrolled in a certain type of service
    - View members by risk level (e.g., list of all of today’s high risk members)
    - Explore population averages for different demographics or types of services
Acute Care Risk

Total Clients: 109,391
Avg. Risk Score: 2.31

Clients per Risk Score:
- 0: 1,187
- 1: 3,861
- 2: 68,002
- 3: 22,221
- 4: 10,275
- 5: 599
- 6: 294
- 7: 31
- 8: 10
- 9: 1

Risk Group:
- Low Risk: 75,950
- Medium-Low Risk: 32,496
- Medium-High Risk: 924
- High Risk: 11

Gender:
- Female: 59,375
- Male: 50,015

Race:
- AFRICAN-AMERICAN: 9,452
- ASIAN: 8,054
- CAUCASIAN: 46,847
- HISPANIC: 4,820
- NATIVE AMERICAN: 966
- OTHER: 38,881
- PACIFIC ISLANDER: 391
## COPPER 2019 // Quantitative analysis

### Client Details

<table>
<thead>
<tr>
<th>Policy Id</th>
<th>Full Name</th>
<th>Score</th>
<th>Phone</th>
<th>Gender</th>
<th>Race</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC123</td>
<td>LEIA ORGANA SOLO</td>
<td>5</td>
<td>123-456-7890</td>
<td>Female</td>
<td>WHITE</td>
<td>50+</td>
</tr>
<tr>
<td>DEF456</td>
<td>TYRON LANNISTER</td>
<td>7</td>
<td>234-567-8901</td>
<td>Male</td>
<td>OTHER</td>
<td>30-39</td>
</tr>
<tr>
<td>GHI789</td>
<td>HERMIONE GRANGER</td>
<td>5</td>
<td>345-678-9012</td>
<td>Female</td>
<td>AFRICAN-AMERIC...</td>
<td>18-29</td>
</tr>
<tr>
<td>JKL012</td>
<td>ARAGORN S O ARATHO...</td>
<td>6</td>
<td>456-789-0123</td>
<td>Male</td>
<td>WHITE</td>
<td>50+</td>
</tr>
<tr>
<td>MNO345</td>
<td>KVOTHE ARILIDENSON</td>
<td>7</td>
<td>Null</td>
<td>Male</td>
<td>OTHER</td>
<td>18-29</td>
</tr>
<tr>
<td>PQR678</td>
<td>SHURI</td>
<td>5</td>
<td>567-890-1234</td>
<td>Female</td>
<td>AFRICAN-AMERIC...</td>
<td>Under 18</td>
</tr>
<tr>
<td>STU901</td>
<td>RA S AL GHUL</td>
<td>7</td>
<td>678-901-2345</td>
<td>Male</td>
<td>ASIAN</td>
<td>40-49</td>
</tr>
</tbody>
</table>

### Authorization Type

<table>
<thead>
<tr>
<th>Authorization Type</th>
<th>Referral Id</th>
<th>Vendor Nm</th>
<th>Elig Eff Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLOBAL - Level B Adult Global - Primary A...</td>
<td>RESIST999</td>
<td>Lifeworks Northwest</td>
<td>1/1/2014</td>
</tr>
<tr>
<td>Outpatient SUD 1/1/17 - Medication Assi...</td>
<td>WESTER999</td>
<td>CODA CD</td>
<td>2/15/2014</td>
</tr>
<tr>
<td>Null</td>
<td>Null</td>
<td>Null</td>
<td>3/31/2015</td>
</tr>
<tr>
<td>Null</td>
<td>Null</td>
<td>Null</td>
<td>4/1/2015</td>
</tr>
<tr>
<td>Null</td>
<td>Null</td>
<td>Null</td>
<td>5/15/2016</td>
</tr>
<tr>
<td>Null</td>
<td>Null</td>
<td>Null</td>
<td>6/30/2016</td>
</tr>
<tr>
<td>Exceptional Needs - Supported Employm...</td>
<td>GOTHAM999</td>
<td>Cascadia Behavioral Health</td>
<td>7/1/2017</td>
</tr>
</tbody>
</table>
Ethical considerations

- Responsibility to clearly communicate limits of analysis and principles of use
  - Only proactively offer help/services (never denying)
  - Respect for client autonomy
  - Not overriding clinical judgment
  - Human behavior too nuanced, messy to reduce to single number; intended as an additional data point, not the definitive word on a person or their life
Talk data to me: discussion

- Communicating with decision-makers (other staff, the public, your best friend’s cousin’s dogwalker…)
  - What can we say with certainty?
    - “We failed to reject the null hypothesis” = not nearly as cool as “I proved this works!”
  - How do we communicate complex mathematical principles (or do we)?
    - And how do we ensure the limits of our analyses are understood?
  - What is the evaluator’s role in public service?
    - Is there such a thing as just “objective” facts?
    - What are our ethical obligations?