Transforming Raw Data into Clinical Inferences: 

*MPOG Phenotypes*

Michael Burns, PhD, MD
Assistant Professor
University of Michigan
Outline

• What are phenotypes?
• Data types
• Apples to apples: developing uniform data
• Phenotype browser
• Examples of where phenotypes make an impact
Phenotype

Shareable, reproducible algorithm (derived from EHR data) precisely defining a patient characteristic or clinical event \(^1\)

Data types

- **Ratio Data**: Differences between measurements, true zero exists. Examples: Height, Age, Weekly Food Spending.
- **Interval Data**: Differences between measurements but no true zero. Examples: Temperature in Fahrenheit, Standardized exam score.
- **Ordinal Data**: Ordered Categories (rankings, order, or scaling). Examples: Service quality rating, Student letter grades.
- **Nominal Data**: Categories (no ordering or direction). Examples: Marital status, Type of car owned.
Uniform Data – “Semantic Interoperability”

Electronic health record data can be:

- messy...
- incomplete...
- redundant...
- conflicting...
- inaccurate...
- site-specific...
Example 1: Describe intraoperative ventilation

Descriptive project detailing ventilation during cardiac surgery:

Key component is **lung protective ventilation**
- Defined by intraoperative median values for ventilator settings:
  - Tidal volume: <8 mL/kg PBW
  - PEEP: >0 cm H$_2$O
  - PIP: <30 cm H$_2$O
Example 1: Describe intraoperative ventilation

Easy, right?
What about this period?

Tidal volume = PEEP = PIP = 0 …. but not always
Example 1: Describe intraoperative ventilation

Analysis **now requires** identification of periods of cardiopulmonary bypass

<table>
<thead>
<tr>
<th>Data</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraoperative Event Times</td>
<td>(Usually) accurate when entered</td>
<td>Multiple ways to document Can have multiple CPB runs</td>
</tr>
<tr>
<td>Ventilator / Physiologic Data</td>
<td>Automated capture</td>
<td>“Transition” period from CPB Subject to artifact Ventilator not always off</td>
</tr>
</tbody>
</table>
Concept Mapping
Automated Captured Data

Rule-based approach to improve robustness of phenotype:

Cardiopulmonary bypass “rules”:

(SBP – DBP) < 20 mmHg \textbf{or} \ HR \leq 5 \text{ bpm}

*AND*

RR \leq 2 \text{ breaths/min} \textbf{or} \ ETCO_2 \leq 5 \text{ cm H}_2\text{O}
Phenotypes to the Rescue

Multiple observations from multiple parameters:

**Standard Concepts**

- Fluids
- Labs
- Vitals
- Meds
- Times
- Diagnoses
- Outputs
- Events
- Outcomes

**Standard Phenotypes**

On Cardiopulmonary Bypass
Patient under General Anesthesia
Low tidal volume ventilation achieved
Total opioid analgesia, morphine equivalents
Phenotypes – why are they made?

- Research projects
- Quality projects
- Informatics / analytics projects
- Technical necessity
Example 2: Defining a general anesthetic

<table>
<thead>
<tr>
<th>Data</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuromuscular Blocker Used</td>
<td>Not always accurate charting</td>
</tr>
<tr>
<td>Intubation / Extubation note</td>
<td>Occasionally mislabeled</td>
</tr>
<tr>
<td>Inhaled anesthetic</td>
<td>Detection of trace gases</td>
</tr>
<tr>
<td>Free Text Notes</td>
<td>Natural language processing required to “unlock” the data</td>
</tr>
</tbody>
</table>
Example 3: Defining Height and weight

Not as easy as it sounds:
- Often multiple sources with differing values
- No one reliable source as “ground truth”
- Single values without units (convert or trust?)
- Cases without a value?
Phenotype Browser – continuously updated

http://phenotypes.mpog.org/
<table>
<thead>
<tr>
<th>Phenotype Name</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbidity - Aids \ HIV</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Alcohol Abuse</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Blood Loss Anemia</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Cardiac Arrhythmias</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Chronic Pulmonary Disease</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Coagulopathy</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Congestive Heart Failure</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Deficiency Anemia</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Depression</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Diabetes (complicated)</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Diabetes (uncomplicated)</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Drug Abuse</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Fluid/Electrolyte Disorders</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Hypertension (complicated)</td>
<td>ElixhauserComorbidity</td>
</tr>
<tr>
<td>Comorbidity - Hypertension (uncomplicated)</td>
<td>ElixhauserComorbidity</td>
</tr>
</tbody>
</table>
Comorbidity - Alcohol Abuse

Description

ICD-9, ICD-10 diagnoses from hospital discharge administrative codes, anesthesia history and physical diagnoses, or problem summary list diagnoses. Does not include professional billing diagnoses. ICD-9 and ICD-10-based comorbidity measures. Component 1 (reference 1) ICDs are the only ones pulled, the component 2 (reference 2) ICDs were identified, but this phenotype only focuses on the Quan referenced ICDs.

Data source:
aims_billingdiagnoses
aims_preop

Prior time frame - one year prior (before the operation date)
Post time frame - extends to date of service end

Reference:
(2) Additional codes determined by MPOG to fit this Elchhauser comorbidity through ICD-9 and ICD-10 manual searches

Limitation

ICD 9/10 codes may exist on hospital admission, after, or both. This collation does not differentiate.

Value Type

Categorical
Enumeration

<table>
<thead>
<tr>
<th>Value</th>
<th>Value Code</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
<td>-999</td>
<td>Patient does not have any ICD-9/10 codes for the given dependencies</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td>Patient does have ICD-9/10 codes, but does not have any for this specific Elixhauser comorbidity spec</td>
</tr>
<tr>
<td>Yes</td>
<td>1</td>
<td>Patient has ICD-10 or Enhanced ICD-9-CM codes for this specific Elixhauser comorbidity</td>
</tr>
</tbody>
</table>

Return Columns

<table>
<thead>
<tr>
<th>Column Name</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPOG_Case_ID</td>
<td>uniqueidentifier</td>
</tr>
<tr>
<td>Value</td>
<td>int</td>
</tr>
</tbody>
</table>

Granularity

One value per case
1 – Quan, et. al. “Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data”, Medical Care, 43 (11), Nov 2005
2 – Additional codes determined by MPOG to fit this Elixhauser comorbidity through ICD-9 and ICD-10 manual searches

<table>
<thead>
<tr>
<th>ICD-9 Codes</th>
<th>Description</th>
<th>ICD-9 Code conversion (derived, non-standardized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G80</td>
<td>Toxic effect of (an alcohol)</td>
<td>T51 T51.9</td>
</tr>
<tr>
<td>265.2</td>
<td>Pellagra (Niacin deficiency)</td>
<td>E62</td>
</tr>
<tr>
<td>291.1-35-39</td>
<td>Alcohol abuse (withdrawal, mental changes), Acute alcohol intoxication, Non-dependent Alcohol abuse</td>
<td>F10.9</td>
</tr>
<tr>
<td>303 [log]</td>
<td>Alcoholic polyneuropathy</td>
<td>G62.1</td>
</tr>
<tr>
<td>305.0</td>
<td>Alcoholic cardiomyopathy</td>
<td>I42.6</td>
</tr>
<tr>
<td>535.3</td>
<td>Alcoholic Gastritis</td>
<td>K29.2</td>
</tr>
<tr>
<td>571.0-33</td>
<td>Alcoholic fatty liver, hepatitis, cirrhosis</td>
<td>K70.039</td>
</tr>
<tr>
<td>V113</td>
<td>Problem w/ alcohol use</td>
<td>Z72.1</td>
</tr>
<tr>
<td></td>
<td>Alcohol Abuse Counseling / Rehab</td>
<td>Z50.2 Z71.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Components2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alcohol abuse, dependence</td>
<td>F10.0</td>
</tr>
<tr>
<td></td>
<td>Alcoholic hepatitis</td>
<td>K70.1</td>
</tr>
</tbody>
</table>

Dependencies
No dependency available.
Anesthesia Technique: General

Description

This phenotype attempts to classify each case in terms of general anesthesia technique. This returns type of anesthetic used within the possibilities listed in the value types.

Limitation

This phenotype has several limitations. Tracheostomy is not considered. ETT and LMA single notes will return as general - unknown if there is only a single ETT or single LMA note in the case., meaning if there are no accompanying general, ETT, or LMA note. This phenotype also does not differentiate type of intubation (ex. fiberoptic or glidescope). Also, there is no differentiation of success, only if the type was documented as attempted. A negative return ("no") means there was either no general anesthetic attempted in the case (ex. a MAC / sedation case) or not enough documentation to determine a general anesthetic. Neuromuscular Blocker Only and Inhaled Anesthetic Only were used as distinct categories as they may convey meaning if isolated such as in pediatric mask induction cases or ECT cases in which succinylcholine was used.

Value Type

Categorical
## Enumeration

<table>
<thead>
<tr>
<th>Value</th>
<th>Value Code</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0</td>
<td>No general, ETT, or LMA note and no sedative medications or inhaled anesthetics or paralytics associated with the case.</td>
</tr>
<tr>
<td>General - both ETT and LMA</td>
<td>1</td>
<td>There were ETT and LMA notes associated with this case.</td>
</tr>
<tr>
<td>General - ETT</td>
<td>2</td>
<td>There was at least one ETT note, with another general or ETT note associated with this case. There were no LMA notes.</td>
</tr>
<tr>
<td>General - LMA</td>
<td>3</td>
<td>There was at least one LMA note, There were no ETT notes.</td>
</tr>
<tr>
<td>General - Inhaled Anesthetic Only</td>
<td>4</td>
<td>There were inhaled anesthetics associated with this case. There were no ETT or LMA notes.</td>
</tr>
<tr>
<td>General - Neuromuscular Blocker Only</td>
<td>5</td>
<td>There were neuromuscular blockers associated with this case. There were no ETT or LMA notes.</td>
</tr>
<tr>
<td>General - Unknown</td>
<td>6</td>
<td>There were general notes associated with this case. There were both neuromuscular blockers and inhaled anesthetics associated with this case without ETT or LMA notes.</td>
</tr>
</tbody>
</table>
CASE RESULT

IMPORTANT: ‘ETT’ result takes precedence over all other results except ‘ETT and LMA’

Here is the hierarchical order of return for this phenotype:

General - Both ETT and LMA: ETT and LMA notes both exist
General - ETT: ETT note with any other inhaled anesthetic, neuromuscular blocker or general unknown note
General - LMA: LMA note exists, without any ETT notes
General - Inhaled anesthetic only: Inhaled anesthetic documented without any ETT or LMA notes
General - neuromuscular blocker only: Paralytic administered without inhaled anesthetic, ETT or LMA note.
General - unknown: Note from General - Unknown Concept Subset without ETT, LMA, paralytic or inhaled anesthetic notes.

No: No general notes, sedative medications, inhaled anesthetics or paralytics associated with the case.

*General - Unknown Concept Subset:

- 50099 Intubation - Nasal approach note
- 50100 Intubation - Videolaryngoscopy View
- 50115 Intubation Laryngoscopy Blade Type and Size
- 50116 Intubation Laryngoscopy Cricoid Pressure or BURP Applied
- 50117 Intubation - ETT Placed
- 50118 Intubation Number of Attempts
- 50119 Intubation Direct Laryngoscopy View
- 50120 Intubation Bougie Introducer Used
- 50127 Intubation Extubated Awake or Deep
- 50129 Intubation Fiberoptic Asleep or Awake
- 50130 Intubation Fiberoptic Nasal or Oral
- 50131 Intubation Fiberoptic Topicalization Type
- 50132 Intubation Fiberoptic Transtracheal block
Phenotypes – Future Directions

Phenotype validity must be maintained over time – tremendous effort
- New sites & documentation patterns
- New projects
- Changing needs (sensitive versus specific)
- Improved specs

…but you can help!
Phenotypes – Key Takeaways

The EHR is messy

Phenotypes combine messy data sources to create a robust clinical inference

Phenotypes = building blocks for Research & QI

Website: http://phenotypes.mpog.org/

Questions? mpog-research@med.umich.edu