Improving Performance of Algorithms to Power Unmet Need and Effectiveness in Health Economics and Outcomes Research Using Electronic Health Records and Healthcare Claims Data Sources

Schiffon L Wong, MPH
Monica G Kobayashi, PhD, MBMA / Hoa V Le, MD, PhD
Aaron WC Kamauu, MD, MS, MPH

ISPOR 20th Annual European Congress, Glasgow
November 8, 2017

Topic Leaders

Schiffon L. Wong
MPH
Franchise Head Neurology
Global Research & Development
EMD Serono, Inc., Billerica, MA USA

Monica G. Kobayashi
PhD MBMA
Consultant
PAREXEL International, Durham, NC USA

Aaron W. C. Kamauu
MD MS MPH
VP, RWDS
PAREXEL International, Waltham, MA USA

Hoa V. Le
MD PhD
Senior Consultant
PAREXEL International, Durham, NC USA
Workshop Overview

Algorithms for therapeutic value and evaluation

Algorithm development

Algorithm validation

Optimized Patient Access Requires Life Cycle Evidence Generation

Real-World Evidence

1. Inform and enhance clinical development and regulatory decision-making; characterize unmet need, support product differentiation; demonstrate effectiveness and safety

2. Inform options for innovative pricing; access agreements with payers; pay-for-performance

3. Fulfill post-marketing commitments

4. Healthcare quality metrics
Accurate Case Ascertainment and Health Outcomes Identification is Critical

- Electronic Health/Medical Records
- Healthcare Claims Data
- Prescription Drug Database
- Disease Registries
- Hospitalization Database
- Other (e.g., Primary Data Collection)

Misclassification is a Risk to Sound Inferences & Healthcare Decision Making

- Commonly use algorithms
  - Ad-hoc
  - Inconsistent
  - May not be fit-for-purpose
  - May not be apt for the data source
- Validity non-commonly assessed

Confidence?

- High
- Medium
- Low
Health Outcome Example: Multiple Sclerosis (MS) Relapse Episodes

Option 1: Occurrence

EHR clinical notes-based algorithm
• Natural language processing (NLP)

Claims-based algorithm

Positive predictive value (PPV) calculations for validation

Health Outcome Example: MS Relapse Episodes

EHR Clinical Notes-Based Algorithm (NLP-type validation)

<table>
<thead>
<tr>
<th>Certain (op1 n=12)</th>
<th>Likely (op1 n=4)</th>
<th>Possible (op1 n=5)</th>
<th>No (op1 n=36)</th>
<th>Unknown (op1 n=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(op2 n=15)</td>
<td>(op2 n=3)</td>
<td>(op2 n=8)</td>
<td>(op2 n=47)</td>
<td>(op2 n=2)</td>
</tr>
</tbody>
</table>

Claim-Based Algorithm (Validation via Comprehensive Patient Profiles)

<table>
<thead>
<tr>
<th>Certain (op1 n=69)</th>
<th>Likely (op1 n=6)</th>
<th>Possible (op1 n=10)</th>
<th>No (op1 n=4)</th>
<th>Unknown (op1 n=0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(op2 n=25)</td>
<td>(op2 n=0)</td>
<td>(op2 n=0)</td>
<td>(op2 n=0)</td>
<td>(op2 n=0)</td>
</tr>
</tbody>
</table>

PPV (calculation 1)

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certain</td>
<td>N</td>
</tr>
<tr>
<td>L</td>
<td>P</td>
</tr>
<tr>
<td>No</td>
<td>U</td>
</tr>
</tbody>
</table>

PPV (calculation 2)

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certain</td>
<td>N</td>
</tr>
<tr>
<td>L</td>
<td>P</td>
</tr>
<tr>
<td>No</td>
<td>U</td>
</tr>
</tbody>
</table>

Option 1 PPV=25.0% (95% CI: 14.1-39.9%)

Option 2 PPV=24.2% (95% CI: 14.6-37.0%)

Option 1 PPV=36.2% (95% CI: 24.3-49.9%)

Option 2 PPV=34.7% (95% CI: 24.3-46.6%)

Option 1 PPV=94.5% (95% CI: 86.3-98.3%)

Option 2 PPV=100% (95% CI: 86.3-100.0%)

Option 1 PPV=25.0% (95% CI: 14.1-39.9%)

Option 2 PPV=24.2% (95% CI: 14.6-37.0%)

Option 1 PPV=36.2% (95% CI: 24.3-49.9%)

Option 2 PPV=34.7% (95% CI: 24.3-46.6%)

Option 1 PPV=94.5% (95% CI: 86.3-98.3%)

Option 2 PPV=100% (95% CI: 86.3-100.0%)

ISPOR EU November 8, 2017
Health Outcome Example: MS Relapse Episodes Key Findings

**EHR Clinical Notes-Based Algorithm**
- Relapses were not explicitly recorded within the clinical notes
- Search terms were too general, not limited to MS and/or relapse, and therefore returned false positives

**Claims-Based Algorithm**
- Option 1 identified more than three times as many relapse episodes and about 50% more patients (n=11,362 relapses), than Option 2, designed to categorize severity among relapses (n=3,444 relapses)

Algorithm Development Starts With a Team

Clinicians  
Informaticist / Clinical Coder  
Epidemiologist / Health Outcomes Researcher  
NLP Expert / Programmer

NLP = Natural language processing
Consider Your Data Source

- Structured or Unstructured data
- Population captured
- Single or multiple source
- What data will not be captured?
- Where is data captured?
- Follow-up & look-back periods
- Validation

Validation Measures

- Sensitivity
- Specificity
- Positive predictive value (PPV)
- Negative predictive value (NPV)
- Area under curve
- Likelihood ratio
- Youden's index
- Diagnostic odds ratio (DOR)
Case Ascertainment Algorithm
Example: Subtype Identification

Inclusion And Exclusion Criteria

**INCLUSIONS**
Multiple sclerosis (MS) patient identification by combinations of:
- MS diagnosis
- Specific MS symptoms during a neurology visit
- Use of disease-modifying therapy (DMT), or
- Brain/spinal magnetic resonance imaging (MRI)

**EXCLUSIONS**
Patients with progressive disease were excluded by one of the following options:

**Option A:** Change of Expanded Disability Status Scale (EDSS) scores based on conversion of Kurtzke Functional Systems Scores (KFSS) into ICD-9-CM

**Option B:** Use of medication often used for progressive disease

**Option C:** Supportive therapy use (nursing home, home health, selected rehabilitation/durable medical equipment) over 12 months*

**RRMS COHORTS**
- Cohort A
- Cohort B
- Cohort C

* Adapted from Gilden et al. 2011
RRMS = Relapsing-remitting multiple sclerosis

ISPOR EU November 8, 2017  16
### Criteria Contributions

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Count</th>
<th>Percentage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of patients who met inclusion criteria (n=2,960)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1 Diagnosis + 1 of 5 following</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1 MS indicated DMT</td>
<td>1,545</td>
<td>52%</td>
<td></td>
</tr>
<tr>
<td>≥1 MS specific symptom therapy during a neurology visit</td>
<td>1,483</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>≥1 MS specific symptom during a neurology visit ≥ 30 days apart</td>
<td>1,112</td>
<td>38%</td>
<td></td>
</tr>
<tr>
<td>≥1 Brain or Spinal MRI before index</td>
<td>872</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>≥1 ICD-9 code 378.86 at least 30 days apart</td>
<td>5</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>≥1 DMT + diagnosis history + 1 of 3 following</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1 Brain or Spinal MRI</td>
<td>714</td>
<td>24%</td>
<td></td>
</tr>
<tr>
<td>≥1 MS specific symptom therapy during a neurology visit</td>
<td>479</td>
<td>16%</td>
<td></td>
</tr>
<tr>
<td>≥1 MS specific symptom during a neurology visit ≥ 30 days apart</td>
<td>275</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td>≥1 Brain or Spinal MRI before index</td>
<td>676</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>≥1 ICD-9 code 378.86 at least 30 days apart</td>
<td>5</td>
<td>&lt;1%</td>
<td></td>
</tr>
</tbody>
</table>

ICD-9 378.86: Internuclear ophthalmoplegia

### Progressive MS Identification

689 total patients excluded based on one of the 3 options

<table>
<thead>
<tr>
<th>Option</th>
<th>Count</th>
<th>Percentage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option A exclusion (n=607, 88%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease progression based on a specific change of EDSS scores in the last 12 months of the patient’s most recent year of care coverage after the index date and during the study period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Option B exclusion (n=60, 9%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medications often used for progressive disease (mitoxantrone, cyclophosphamide, or methotrexate)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Option C exclusion (n=44, 6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least 12 months of recorded MS history and one of the following:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• At least 10 of the last 12 months at the exacerbation level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The last 12 months at the plateau/stable level with a final therapy type of nursing home, home health, selected rehabilitation/DME</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total number of patients who met inclusion criteria and excluding patients with progressive disease (n=2,271)
EHR Clinical Notes-based Case Ascertainment Algorithm Example: Subtype Identification

Natural Language Processing Search

Search terms and % hit of 62,909 documents

- multiple' 97%
- sclerosis' 96%
- relapsing' 4.7%
- remitting' 4.2%
- progressive' 5.6%
- subtype' 0.055%
- RRMS' 0.098%
- multiple sclerosis' 94.8%
- relapsing remitting' 2.79%
- NEAR( (multiple, sclerosis) , 3, TRUE ) ' 95.2%
- NEAR( (relapsing, remitting) , 4, FALSE ) ' 4%
- NEAR( (relapsing, remitting, multiple, sclerosis) , 12 , FALSE ) ' 2.8%
- NEAR( (multiple, sclerosis, relapsing, remitting, subtype ) , 12, FALSE ) ' 0.023%
- NEAR( (remittent, progressive , multiple, sclerosis ) , 12, FALSE ) ' 0.003%
- NEAR( (multiple, sclerosis, relapsing, remitting, type) , 12, FALSE ) ' 0.063%
- NEAR( {{not}, multiple, sclerosis} , 20, FALSE ) ' 0.95%
- NEAR( (unlikely, multiple, sclerosis) , 15, FALSE ) ' 0.12%
Inclusion and Exclusion Criteria

AND

≥ 1 Natural Language Processing (NLP)-based mention of any of the terms/phrases for clinician-documented diagnosis of RRMS in a clinical note during the study period

EXCLUDING

≥ 1 NLP-based mention of any of the terms/phrases for clinician-documented diagnosis of progressive MS in a clinical note during the study period

RRMS Cohort

N=4,623
N=990
N=153
N=837

Challenges & Considerations

• Case definitions and data capture

• Availability of data recorded in clinician’s documentation
  – Explicit documentation
  – Detail on image report vs clinical notes

• Measure(s) for validation and data availability
Audience Participation

• What has been your experience with developing algorithms?
  – Intended purpose of the algorithm
  – Challenges encountered
  – Lessons learned
  – Impact

Healthcare Data

• Diagnoses, medications/prescriptions, procedures
• Observations (including vital signs)
• Problem lists with symptoms
• Laboratory and microbiology/pathology results
• Imaging studies (PACS images & radiologist notes)
• Clinical documents
  – Clinician notes, radiology reports, microbiology/pathology reports
  – Medical test results, symptoms, disease characteristics/qualities
• Advanced state-of-the-art Natural Language Processing (NLP) methods extract meaningful information from text notes
Natural Language Processing (NLP)

Electronic Health Records

Actual Patient Health and Disease

Actual Care Delivered

Unstructured Data

NLP Tool

Structured Data

Clinical Concepts

Structured Data

Knowledge

• From EMR
  • For clinical research

Iterative Process

NLP Tool

• Identify clinical concepts
• Annotate sample records
• Train NLP tool
• Test/Validate NLP tool

NLP Example: Relapsing Multiple Sclerosis

Evidence of RRMS

<table>
<thead>
<tr>
<th>RRMS Terms</th>
<th># of Unique Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>relapsing remitting</td>
<td>839</td>
</tr>
<tr>
<td>relapsing</td>
<td>970</td>
</tr>
<tr>
<td>remitting</td>
<td>862</td>
</tr>
</tbody>
</table>

Evidence of Progressive MS

<table>
<thead>
<tr>
<th>Progressive MS Terms</th>
<th># of Unique Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>contains(document_text,' NEAR( (progressive, multiple, sclerosis) , 6, FALSE ) ',' ,18 )&gt; 0</td>
<td>153</td>
</tr>
<tr>
<td>contains(document_text,'progressive', 5 )&gt;0</td>
<td>522</td>
</tr>
<tr>
<td>Not used: proved too broad, resulting in false positives</td>
<td></td>
</tr>
</tbody>
</table>
Why Validate Algorithms?

- Determine measurement characteristics (accuracy) of the algorithm
- Provides baseline understanding to support interpretation of results
- Support improvement of the algorithm for case ascertainment or study measures
- Higher accuracy → step toward standardization of case ascertainment or study measures

Some Types of Validation

"Traditional" Manual Paper Chart Review

NLP-type Validation

- Run Algorithm
- Healthcare Data
- Patients "positive" for algorithm
- Random Sample
- Patients "positive" for algorithm
- Random Sample
- Create Patient Profiles
- Longitudinal
- Comprehensive
- Encounters
- Dx, Rx, Proc, Tests, Notes
- Manually review the documents to determine if the "positive hit" represents the intended target concept
- Manually review the documents to determine if the patient truly is positive for the intended target concept
- Certain
- Likely
- Possible
- No
- Unknown
- Calculate PPVs
Case Ascertainment Algorithm Example: Multiple Sclerosis Subtype RRMS

**EHR Clinical Notes-Based Algorithm**
(NLP-type validation)

- Certain (n=107)
- Likely (n=0)
- Possible (n=0)
- No (n=1)
- Unknown (n=3)

**PPV (calculation 1)**
Positive: C
Negative: N

- PPV = 99.1% (95% CI: 94.2-100%)

**Claims-Based Algorithm**
(Validation via Comprehensive Patient Profiles)

- Certain (n=122)
- Likely (n=7)
- Possible (n=1)
- No (n=7)
- Unknown (n=0)

**PPV (calculation 2)**
Positive: C
Negative: N

- PPV = 96.4% (95% CI: 90.5-98.8%)

- PPV = 94.6% (95% CI: 89.0-97.5%)

- PPV = 94.9% (95% CI: 89.6-97.7%)

**ISPOR EU November 8, 2017**

Audience Participation

- What is your perspective on algorithm validation as represented in the literature?
  - Quality of the evidence?
  - Quality of the described methods?

- Experience with algorithm validation
  - Methods used
  - Challenges encountered
  - Lessons learned
  - Impact

- Any experience using NLP or another advanced methodology?

- Other thoughts?
Concluding Power Points for Algorithms

Collaborative team planning

Appropriate selection of data source

Assess multiple options for algorithm components

Validate! Validate! Validate!

Leverage for decision making, document and publish

Acknowledgements

The study was supported by EMD Serono, Inc., Rockland, MA, USA (a business of Merck KGaA, Darmstadt, Germany)

• Canter Martin, Camelia Graham, Hannah Crooke, Nicole Bailey & Andrew Wilson (PAREXEL International)

• Chi T. L. Truong (MedCodeWorld)

• Meritxell Sabidó-Espin (Merck KGaA)

• John R. Holmen, Christopher L. Fillmore, Justin Mundt & Jason Gagner (Intermountain Healthcare)
Questions?

Thank you!

Contacts:
- Schiffon Wong: schiffon.wong@emdserono.com
- Monica Kobayashi: monica.kobayashi@parexel.com
- Hoa Le: hoa.le@parexel.com
- Aaron Kamauu: aaron.kamauu@parexel.com

Back-up slides
Claims and EHR clinical note-based algorithms to support Multiple Sclerosis research

<table>
<thead>
<tr>
<th>Title</th>
<th>Conference</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lessons Learned in Identifying Relapsing-Remitting Multiple Sclerosis (RRMS) in United States Integrated Delivery Network Healthcare Claims and Electronic Health Record (EHR) Data</td>
<td>ISPOR 2017 (Podium Presentation)</td>
<td>May 2017</td>
</tr>
<tr>
<td>Preliminary performance of EHR-based algorithm to identify relapsing-remitting multiple sclerosis (RRMS) in United States integrated delivery network electronic health record data</td>
<td>ICPE 2017 (Podium Presentation)</td>
<td>August 2017</td>
</tr>
<tr>
<td>Identifying Relapsing-Remitting Multiple Sclerosis (RRMS) in United States Integrated Delivery Network Healthcare Claims Data</td>
<td>ICPE 2017 (Poster)</td>
<td>August 2017</td>
</tr>
<tr>
<td>Creating a Claims-Based Adaptation of Kurtzke Functional Systems Scores for MS Severity/Progression</td>
<td>ECTRIMS/ACTRIMS (ePoster)</td>
<td>October 2017</td>
</tr>
<tr>
<td>Using algorithms to identify High Disease Activity Relapse-Remitting Multiple Sclerosis patients using electronic health record data with natural language processing</td>
<td>ECTRIMS/ACTRIMS (Poster)</td>
<td>October 2017</td>
</tr>
<tr>
<td>Lessons Learned Using United States Integrated Delivery Network (IDN) Claims-Based Algorithms to identify relapses in Relapse-Remitting Multiple Sclerosis (RRMS) Patients</td>
<td>ECTRIMS/ACTRIMS (Poster)</td>
<td>October 2017</td>
</tr>
<tr>
<td>Identifying Relapses in Relapsing-Remitting Multiple Sclerosis Patients in United States Integrated Delivery Network Healthcare Electronic Health Record Data</td>
<td>ECTRIMS/ACTRIMS (Poster)</td>
<td>October 2017</td>
</tr>
<tr>
<td>Considerations in the use of EHR- and Claims-based Algorithms to Identify RRMS and Relapse in an US IDN database</td>
<td>AMIA (Podium Presentation)</td>
<td>November 2017</td>
</tr>
</tbody>
</table>

NLP Example: Mild Cognitive Impairment
NLP Example: Prostate Cancer