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

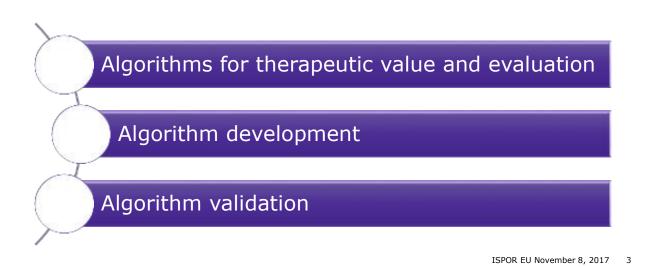
Schiffon L Wong, MPH Monica G Kobayashi, PhD, MBMA / Hoa V Le, MD, PhD Aaron WC Kamauu, MD, MS, MPH

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

Topic Leaders



Workshop Overview



Optimized Patient Access Requires Life Cycle Evidence Generation



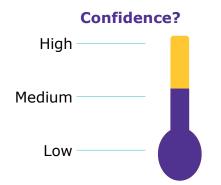
Accurate Case Ascertainment and Health Outcomes Identification is Critical



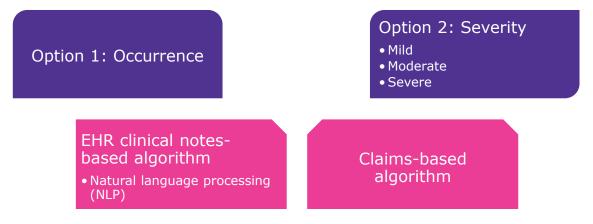
ISPOR EU November 8, 2017 6

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



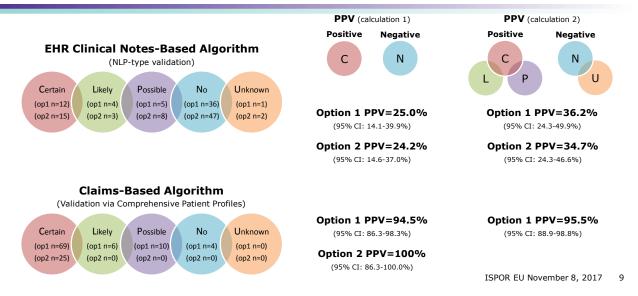
Health Outcome Example: Multiple Sclerosis (MS) Relapse Episodes



Positive predictive value (PPV) calculations for validation

ISPOR EU November 8, 2017 8

Health Outcome Example: MS Relapse Episodes



Health Outcome Example: MS Relapse Episodes Key Findings

EHR Clinical Notes-Based Algorithm

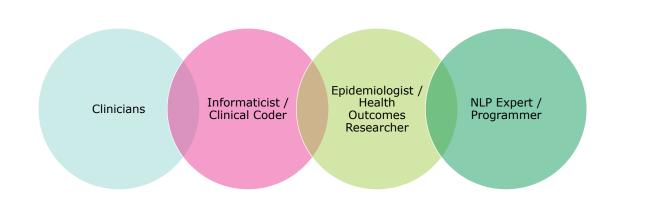
- Relapses were not explicitly recorded within the clinical notes
- Search terms were too general, not limited to MS and/or relapse, and therefore returned false positives

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)

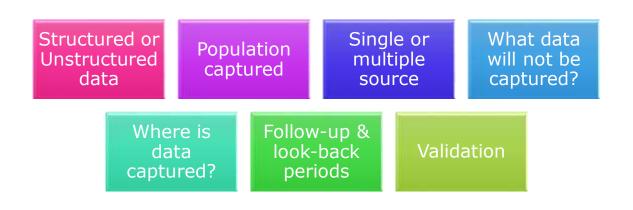
ISPOR EU November 8, 2017 10

Algorithm Development Starts With a Team



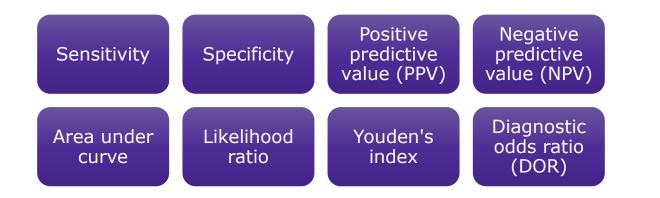
NLP = Natural language processing

Consider Your Data Source



ISPOR EU November 8, 2017 13

Validation Measures



Case Ascertainment Algorithm Example: Subtype Identification

Inclusion And Exclusion Criteria



Multiple sclerosis (MS) patient identification by combinations of:

- MS diagnosis

 Specific MS symptoms during a neurology visit

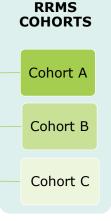
 Use of diseasemodifying therapy (DMT), or

 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

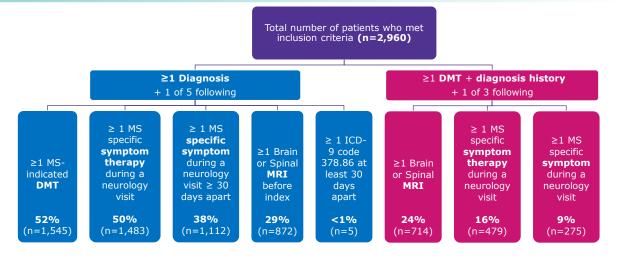
Option B: Use of medication often used for progressive disease

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



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

Criteria Contributions



ICD-9 378.86: Internuclear ophthalmoplegia

ISPOR EU November 8, 2017 17

Progressive MS Identification

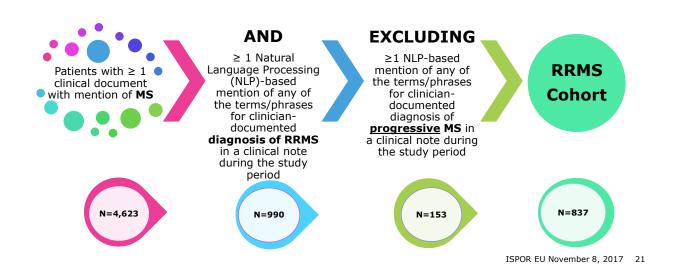
	689 total patients excluded based on one of the 3 options				
	Option A exclusion (n=607, [88%])				
Total number of	Disease progression based on a specific change of EDSS scores in the last 12 months	Option B exclusion (n=60, [9%] Medications often used for	Option C exclusion (n=44, [6%])		number of patients who met inclusion
patients who met inclusion criteria	of the patient's most recent year of care coverage after the index date and during the study period	progressive disease (mitoxantrone, cyclophosphamide, or methotrexate)	At least 12 months of recorded MS history and one of the following:		criteria and excluding patients with
(n=2,960)			• At least 10 of the last 12 months at the exacerbation level		progressive disease
			 The last 12 months at the plateau/stable level with a final therapy type of nursing home, home health, selected rehabilitation/DME. 		(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

Inclusion and Exclusion Criteria



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

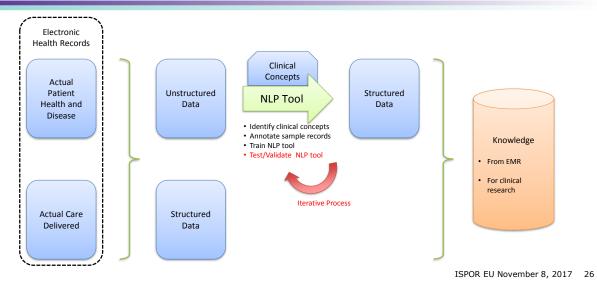
- What has been your experience with developing algorithms?
 - Intended purpose of the algorithm
 - Challenges encountered
 - Lessons learned
 - Impact

ISPOR EU November 8, 2017 23

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)



NLP Example: Relapsing Multiple Sclerosis

Evidence of RRMS

RRMS Terms	# of Unique Patients
relapsing remitting	839
relapsing	970
remitting	862

Evidence of Progressive MS

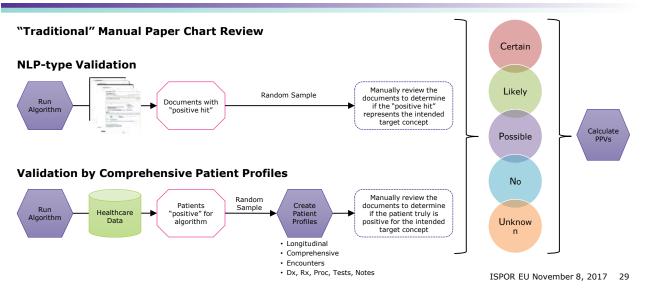
Progressive MS Terms	# of Unique Patients
contains(document_text,' NEAR((progressive, multiple, sclerosis) , 6, FALSE) ' ,18)> 0	153
contains(document_text,'progressive', 5)>0 Not used: proved too broad, resulting in false positives	522

Why Validate Algorithms?

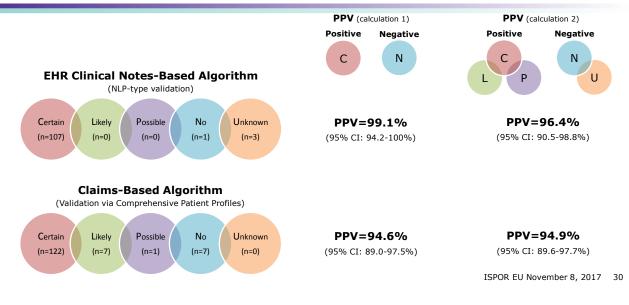
- Determine measurement characteristics (accuracy) of the algorithm
- Provides baseline understanding to support interpretation of results
- Support improvement of the algorithm for case ascertainment or study measures
- Higher accuracy → step toward standardization of case ascertainment or study measures

ISPOR EU November 8, 2017 28

Some Types of Validation



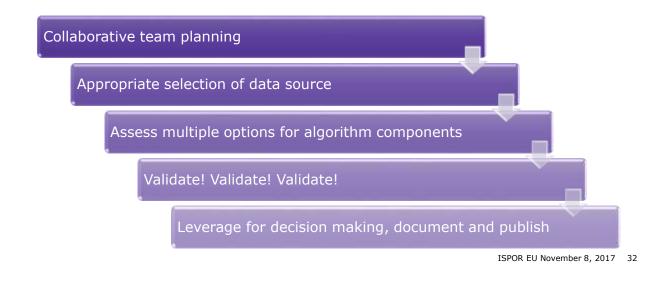
Case Ascertainment Algorithm Example: Multiple Sclerosis Subtype RRMS



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



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

ISPOR EU November 8, 2017 34

Back-up slides

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

Title	Conference	Date
Lessons Learned in Identifying Relapsing-Remitting Multiple Sclerosis (RRMS) in United States Integrated Delivery Network Healthcare Claims and Electronic Health Record (EHR) Data	ISPOR 2017 (Podium Presentation)	May 2017
Preliminary performance of EHR-based algorithm to identify relapsing-remitting multiple sclerosis (RRMS) in United States integrated delivery network electronic health record data	ICPE 2017 (Podium Presentation)	August 2017
Identifying Relapsing-Remitting Multiple Sclerosis (RRMS) in United States Integrated Delivery Network Healthcare Claims Data	ICPE 2017 (Poster)	August 2017
Creating a Claims-Based Adaptation of Kurtzke Functional Systems Scores for MS Severity/Progression	ECTRIMS/ACTRIMS (ePoster)	October 2017
Using algorithms to identify High Disease Activity Relapse-Remitting Multiple Sclerosis patients using electronic health record data with natural language processing	ECTRIMS/ACTRIMS (Poster)	October 2017
Lessons Learned Using United States Integrated Delivery Network (IDN) Claims-Based Algorithms to identify relapses in Relapse-Remitting Multiple Sclerosis (RRMS) Patients	ECTRIMS/ACTRIMS (Poster)	October 2017
Identifying Relapses in Relapsing-Remitting Multiple Sclerosis Patients in United States Integrated Delivery Network Healthcare Electronic Health Record Data	ECTRIMS/ACTRIMS (Poster)	October 2017
Considerations in the use of EHR- and Claims-based Algorithms to Identify RRMS and Relapse in an US IDN database	AMIA (Podium Presentation)	November 2017
	ISPO	R EU November 8, 201

NLP Example: Mild Cognitive Impairment

A.O.O.	Study Mat	eest - MCI
exception MC1 (Comp) Patients	1.1.8 Salas 231.9	to Annual Au
ante tendi : Onder gant Vicini Nicini (Binne) (18]tament (Vine)	(topet.)	AATENT WARE AND
State Mass Community 0 64:07 13 0 64:07 13 0 14:04 15 0 14:04 15 0 14:04 16 0 14:04 16 0 14:04 16 0 14:04 16 0 14:04 16 0 14:04 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16 14:04 16 16		and Bulketis in occurs particular of phases. The online sease air hypersection, Severandowski, Karlin Karlin, Karlin Karlin Karlin, Karlin Karlin Karlin, Karlin Karl
View View		AMPEGISCRA AND PLANE Services.metro robushi mundoge. Neurology will be runnyddor or he Mar Carlon Fall Mar Car

NLP Example: Prostate Cancer

C b Lofinson	Keernak Date 2013	
IP- phases gendered had array symptoms will de in soluble to the right. Or modelson the right road right base toolsaroot de right road right base conserved about the same	anth pass mediaal houses (aphtopent) for the samed ar ways flow. The patient and infer to the optimum and the same same and right At the Boopt, muscles training for (PA meniood a), the reletioned the inferent #1 meniood A. The reletioned and inferent #1 meniood as the PIA.	Trag
 and right Revenues must remark with the pattern his heatbacks at release through the seven state of the seven seven at restrict seven as a seven seven at Rept MCDCCL PRIMARY As given which tablicate	mm with a Chicago scane of Y (): o consist prevalue can concel to the M-Y Links. We want and taken this door body. The particle policies large body particular methods are access in produce of the program much in the models. The particle of the produce of the f models involves and the particle of the f models involves and the particle of the f models.	ing the different of the latest energy with a theory scale of a 1-11 sevening the optimizations and main right latest, with scale scales of the scale of the scal
DOLDS, METTINY, Married, 3 (1988) YAMEN INSTORY, Associationships HEVER OF SISTING. As above Devices, Praidment Institute.	ner, 1. A horizonta el basalneg disentens. Reners, Mai Gener, Rabibers ante - ne sectemento Dener, Nationactór, 5	Select Cough in Innersphyrit Devisit Devisit Breatt Breatt - Devisit Reside, remaining,
NAM WO-mailed art-tree	Skilk. Ma	ga ophataan. Inal exam Mal lapis, mising, isan inee removal sin an analasi in ite maisa waats. Inal kanan Mal Banaranyi kana ikara if A.A. Cartar, kasale radim, kanan (U. S. Komurner or sahar) kanati kasadada
Contra patiphonally. Addresser: 3 Contral tervels	of, inclusionspheromegals Brown Witte	artner raissa or dachenja. Terevellar: No jada' alema er defigerenare, Asarahge: O'tered —
Also to depend and call and plate	et sole and while differential, the po- m blood county that would inequire immedi	rengine IT1, parent IT1, PLCUIL - KTR compare than could will differential an united to conset a bite, frengine a display of the second
Definition (Labora, Lett., Lett.	In more mitter and the provided states of the point the character of the states of the character of the states of the character and the states of the states	— We valid the putter transport