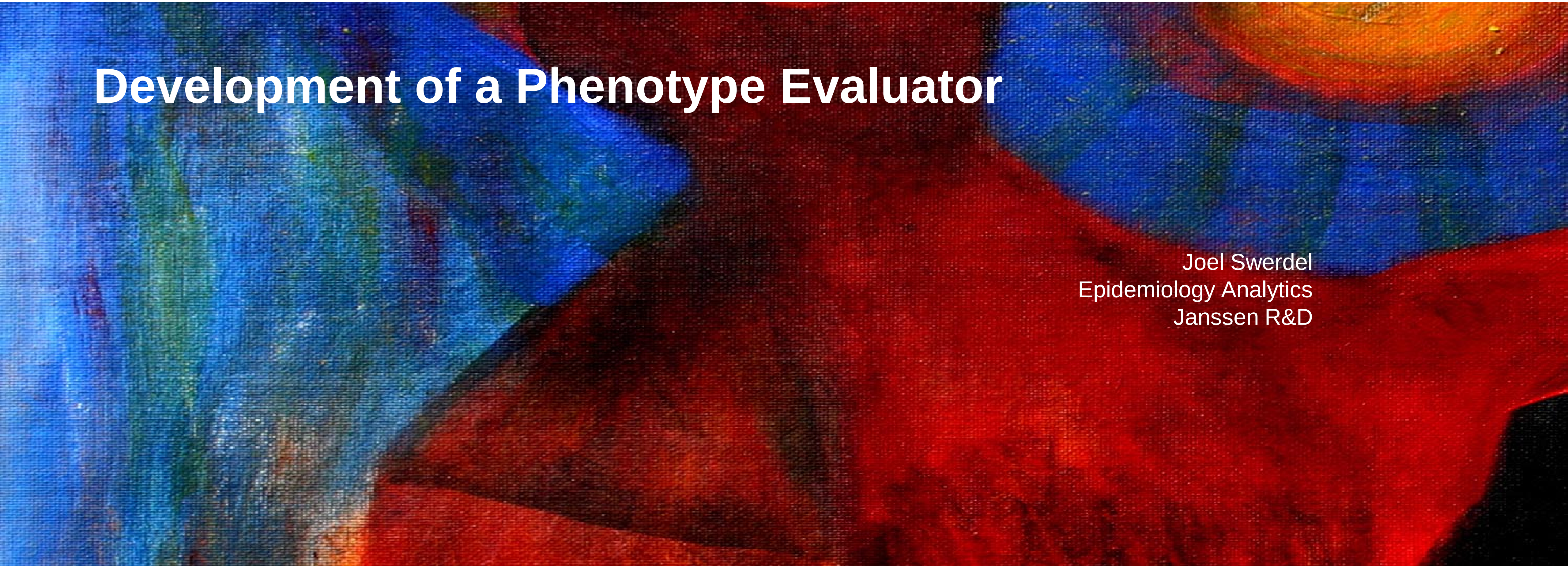


Development of a Phenotype Evaluator

Joel Swerdel
Epidemiology Analytics
Janssen R&D

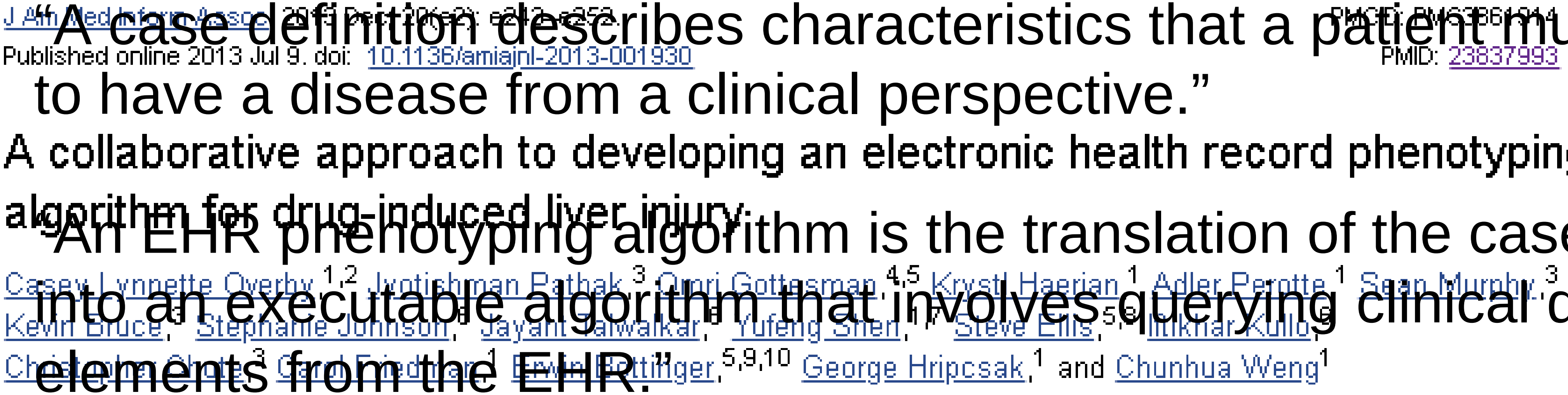
janssen | PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Agenda

- What is a phenotype and why do we need them?
- Why do we need a phenotype evaluator?
- Development of the evaluator
- Results from the evaluation

Case Definitions and Phenotyping Algorithms

- “A case definition describes characteristics that a patient must possess to have a disease from a clinical perspective.”
- “An EHR phenotyping algorithm is the translation of the case definition into an executable algorithm that involves querying clinical data elements from the EHR.”

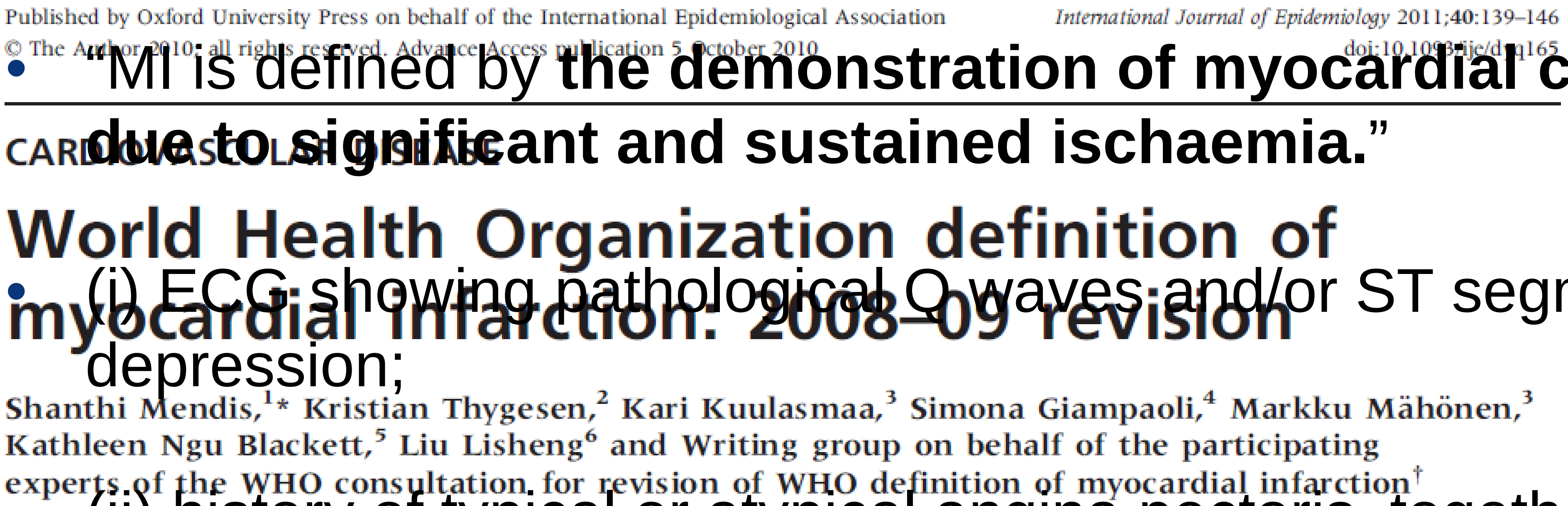


J Am Med Inform Assoc. 2013 Dec; 20(e2): e243–e252. PMCID: PMC3861914
Published online 2013 Jul 9. doi: 10.1136/amiajnl-2013-001930 PMID: 23837993

A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury

Casey Lynnette Overby,[1,2] Jyotishman Pathak,[3] Omri Gottesman,[4,5] Krystl Haerian,[1] Adler Perotte,[1] Sean Murphy,[3] Kevin Bruce,[5] Stephanie Johnson,[6] Jayant Talwalkar,[6] Yufeng Shen,[1,7] Steve Ellis,[5,8] Iftikhar Kullo,[6] Christopher Chute,[3] Carol Friedman,[1] Erwin Bottinger,[5,9,10] George Hripcsak,[1] and Chunhua Weng[1]

Case Definition – Myocardial Infarction



Published by Oxford University Press on behalf of the International Epidemiological Association *International Journal of Epidemiology* 2011;**40**:139–146
© The Author 2010; all rights reserved. Advance Access publication 5 October 2010 doi:10.1093/ije/dyq165

CARDIOVASCULAR DISEASE

World Health Organization definition of myocardial infarction: 2008–09 revision

Shanthi Mendis,[1]* Kristian Thygesen,[2] Kari Kuulasmaa,[3] Simona Giampaoli,[4] Markku Mähönen,[3] Kathleen Ngu Blackett,[5] Liu Lisheng[6] and Writing group on behalf of the participating experts of the WHO consultation for revision of WHO definition of myocardial infarction[†]

- "MI is defined by **the demonstration of myocardial cell necrosis due to significant and sustained ischaemia.**"
- (i) ECG showing pathological Q waves and/or ST segment elevation or depression;
- (ii) history of typical or atypical angina pectoris, together with changes on the ECG and elevated enzymes;
- (iii) history of typical angina pectoris and elevated enzymes with no changes on the ECG or not available

PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Phenotyping Algorithm

Abstract

Purpose—To validate an algorithm based upon International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) codes for acute myocardial infarction (AMI) documented within the Mini-Sentinel Distributed Database (MSDD).

Methods—**Using an ICD-9-CM-based algorithm (hospitalized patients with 410.x0 or 410.x1 in primary position),** we identified a random sample of potential cases of AMI in 2009 from 4 Data Partners participating in the Mini-Sentinel Program. Cardiologist reviewers used information abstracted from hospital records to assess the likelihood of an AMI diagnosis based on criteria from the joint European Society of Cardiology and American College of Cardiology Global Task Force. Positive predictive values (PPVs) of the ICD-9-based algorithm were calculated.

Results—Of the 153 potential cases of AMI identified, hospital records for 143 (93%) were retrieved and abstracted. Overall, the PPV was 86.0% (95% confidence interval; 79.2%, 91.2%). PPVs ranged from 76.3% to 94.3% across the 4 Data Partners.

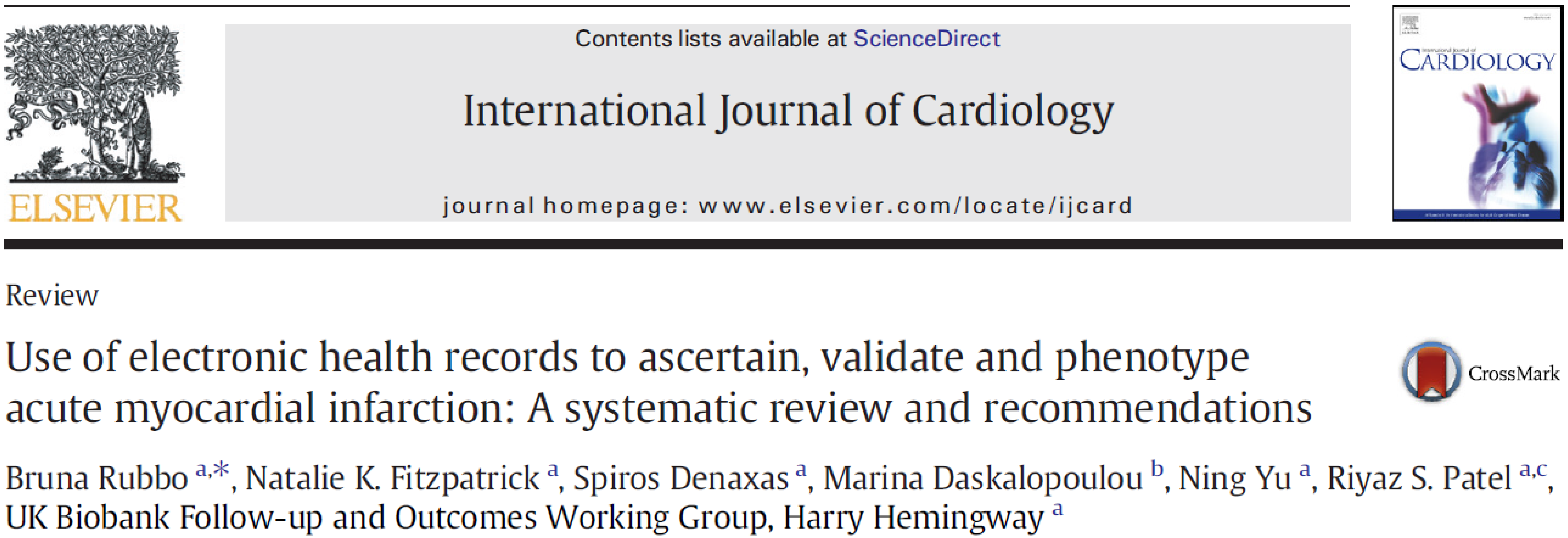
Conclusions—The overall PPV of potential AMI cases, as identified using an ICD-9-CM-based algorithm, may be acceptable for safety surveillance; however, PPVs do vary across Data Partners. This validation effort provides a contemporary estimate of the reliability of this algorithm for use in future surveillance efforts conducted using the FDA's MSDD.

What is a phenotype and why do we need them

- Tendency to equate the case definition with the phenotype algorithm (or the cohort definition) – the algorithm is the coded *approximation* of the case definition.
- Case definitions must be translated into algorithms for working with observational datasets
- But many properties of case definitions are lost in an algorithm causing imprecision when using an algorithm
- How much imprecision? → Need for validation

Validating Algorithms

Many research studies have attempted to validate algorithms



Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

Review

Use of electronic health records to ascertain, validate and phenotype acute myocardial infarction: A systematic review and recommendations

Bruna Rubbo [a,*], Natalie K. Fitzpatrick [a], Spiros Denaxas [a], Marina Daskalopoulou [b], Ning Yu [a], Riyaz S. Patel [a,c], UK Biobank Follow-up and Outcomes Working Group, Harry Hemingway [a]

- Examined 33 studies
- Found significant heterogeneity in algorithms used, validation methods, and results

Validating an Algorithm

		Truth	
		Positive	**Negative**
Test	**Positive**	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Test – Comes from the algorithm/cohort definition

Truth – Some form of “gold standard” reference

Ex.: True Positive (TP) – Test and Truth agree Positive

For a complete validation of the algorithm we need:

1) Sensitivity: TP / (TP + FN)
2) Specificity: TN / (TN + FP)
3) Positive Predictive Value: TP / (TP + FP)

Evaluating Performance of Algorithm - Examples

Abstract

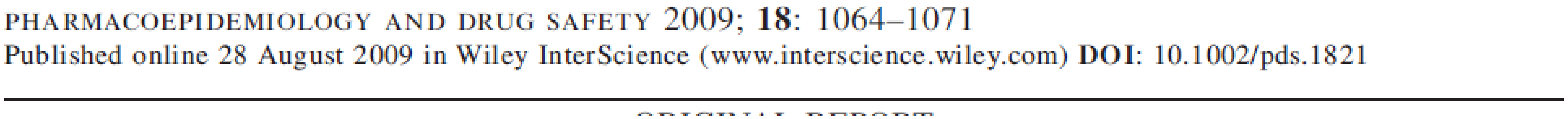
Purpose—To validate an algorithm based upon International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) codes for acute myocardial infarction (AMI) documented within the Mini-Sentinel Distributed Database (MSDD).

Methods—**Using an ICD-9-CM-based algorithm (hospitalized patients with 410.x0 or 410.x1 in primary position),** we identified a random sample of potential cases of AMI in 2009 from 4 Data Partners participating in the Mini-Sentinel Program. **Cardiologist reviewers used information abstracted from hospital records to assess the likelihood of an AMI diagnosis based on criteria from the joint European Society of Cardiology and American College of Cardiology Global Task Force.** Positive predictive values (PPVs) of the ICD-9-based algorithm were calculated.

Results—Of the 153 potential cases of AMI identified, hospital records for 143 (93%) were retrieved and abstracted. **Overall, the PPV was 86.0% (95% confidence interval; 79.2%, 91.2%).** PPVs ranged from 76.3% to 94.3% across the 4 Data Partners.

Conclusions—The overall PPV of potential AMI cases, as identified using an ICD-9-CM-based algorithm, may be acceptable for safety surveillance; however, PPVs do vary across Data Partners. This validation effort provides a contemporary estimate of the reliability of this algorithm for use in future surveillance efforts conducted using the FDA's MSDD.

Evaluating Performance of Algorithm - Examples



PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2009; **18**: 1064–1071
Published online 28 August 2009 in Wiley InterScience (www.interscience.wiley.com) **DOI**: 10.1002/pds.1821

ORIGINAL REPORT

SUMMARY
Purpose Studies of non-steroidal anti-inflammatory drugs (NSAIDs) and cardiovascular events using administrative data require identification of incident acute myocardial infarctions (AMIs) and information on whether confounders differ by NSAID status.
Methods **We identified patients with a first AMI hospitalization from Tennessee Medicaid files as those with primary ICD-9 discharge diagnosis 410.x and hospitalization stay of>2 calendar days.** Eligible persons were non-institutionalized, aged 50–84 years between 1999–2004, had continuous enrollment and no AMI, stroke, or non-cardiovascular serious medical illness in the prior year. Of 5524 patients with a potential first AMI, a systematic sample (n¼350) was selected for review. **Using defined criteria, we classified events using chest pain history, EKG, and cardiac enzymes, and calculated the positive predictive value (PPV) for definite or probable AMI.**
Results 337 of 350 (96.3%) charts were abstracted and 307 (91.1%), 6 (1.8%), and 24 (7.1%) events were categorized as definite, probable, and no AMI, respectively. **PPV for any definite or probable AMI was 92.8% (95% CI 89.6–95.2); for an AMI without an event in the past year 91.7% (95% CI 88.3–94.2), and for an incident AMI was 72.7% (95% CI 67.7–77.2).** Age-adjusted prevalence of current smoking (46.4% vs. 39.1%, p¼0.35) and aspirin use (36.9% vs. 35.9%, p¼0.90) was similar among NSAID users and non-users
Conclusions ICD-9 code 410.x had high predictive value for identifying AMI. Among those with AMI, smoking and aspirin use was similar in NSAID exposure groups, suggesting these factors will not confound the relationship between NSAIDs and cardiovascular outcomes.

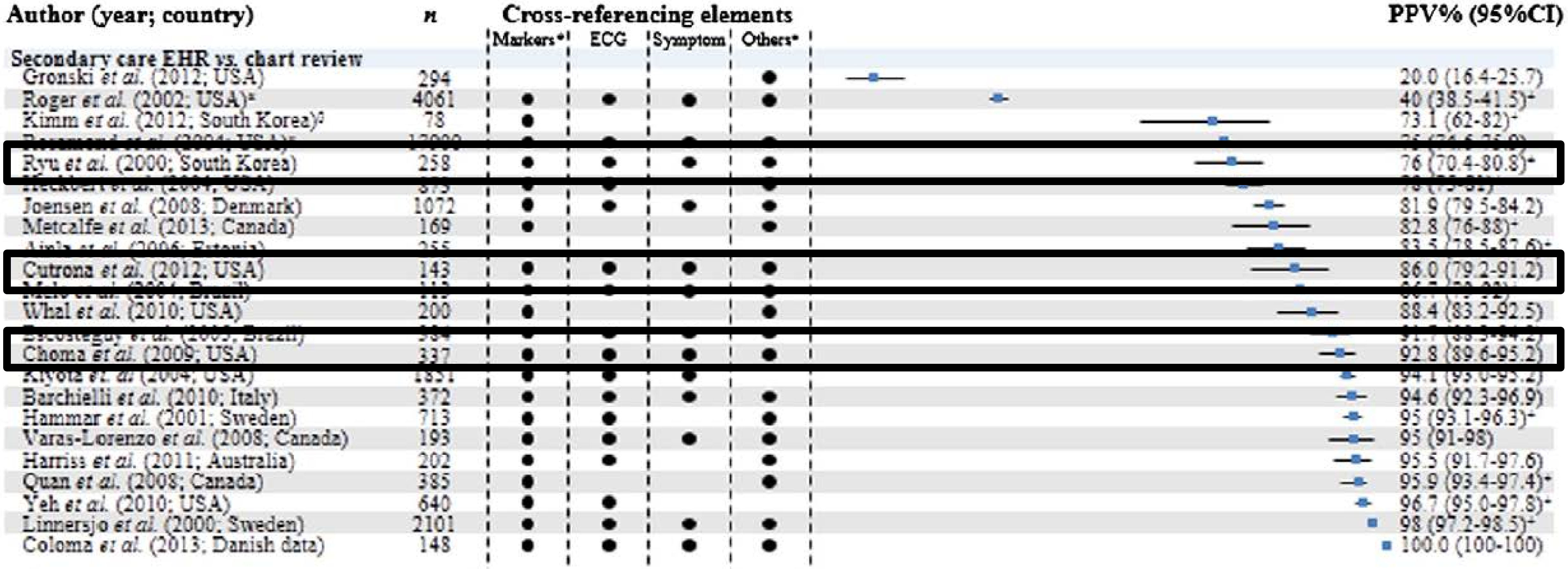
Evaluating Performance of Algorithm - Examples

Yonsei Medical Journal
Vol. 41, No. 5, pp. 570—576, 2000

Abstract

We attempted to assess the accuracy of the International Classification of Diseases (ICD) codes for myocardial infarction (MI) in medical insurance claims, and to investigate the reasons for any inaccuracy. This study was designed as a preliminary study to establish a surveillance system for cardiovascular diseases in Korea. A sample of 258 male patients who were diagnosed with MI from 1993 to 1997 was selected from the Korea Medical Insurance Corporation cohort (KMIC cohort: 183,461 people). The registered medical record administrators were trained in the survey technique, and gathered data by investigating the medical records of the study subjects from March 1999 to May 1999. **The definition of MI for this study included symptoms pursuant to the diagnostic criteria of chest pain, electrocardiogram (ECG) findings, cardiac enzyme and results of coronary angiography or nuclear scan.** We asked the record administrators for the reasons of incorrectness for cases where the final diagnosis was 'not MI'. **The accuracy rate of the ICD codes for MI in medical insurance claims was 76.0% (196 cases) of the study sample,** and 3.9% (ten cases) of the medical records were not available due to hospital closures, non-computerization or missing information. Nineteen cases (7.4%) were classified as insufficient due to insufficient records of chest pain, ECG findings, or cardiac enzymes. The major reason of inaccuracy in the disease code for MI in medical insurance claims was 'to meet the review criteria of medical insurance benefits (45.5%)'. The department responsible for the inaccuracy was the department of inspection for medical insurance benefit of the hospitals.

Evaluating Performance of Algorithm



Author (year; country)	n	Markers*	ECG	Symptom	Others*	PPV% (95%CI)
Secondary care EHR vs. chart review						
Gronski *et al.* (2012; USA)	294				●	20.0 (16.4-25.7)
Roger *et al.* (2002; USA)[a]	4061	●	●	●	●	40 (38.5-41.5)+
Kimm *et al.* (2012; South Korea)[3]	78	●				73.1 (62-82)+
Rosamond *et al.* (2004; USA)[5]	17000	●	●	●	●	75 (74.6-75.0)
Ryu *et al.* (2000; South Korea)	258	●	●	●	●	76 (70.4-80.8)+
Heckbert *et al.* (2004; USA)	375	●	●		●	78 (73-81)
Joensen *et al.* (2008; Denmark)	1072	●	●	●	●	81.9 (79.5-84.2)
Metcalfe *et al.* (2013; Canada)	169	●			●	82.8 (76-88)+
Ainla *et al.* (2006; Estonia)	255					83.5 (78.5-87.6)+
Cutrona *et al.* (2012; USA)	143	●	●	●	●	86.0 (79.2-91.2)
Melo *et al.* (2004; Brazil)	113	●	●	●	●	86.7 (79-92)
Whal *et al.* (2010; USA)	200	●			●	88.4 (83.2-92.5)
Escosteguy *et al.* (2003; Brazil)	384	●	●	●	●	91.7 (88.5-94.2)
Choma *et al.* (2009; USA)	337	●	●	●	●	92.8 (89.6-95.2)
Kiyota *et al.* (2004; USA)	1851	●	●	●		94.1 (93.0-95.2)
Barchielli *et al.* (2010; Italy)	372	●	●	●	●	94.6 (92.3-96.9)
Hammar *et al.* (2001; Sweden)	713	●	●		●	95 (93.1-96.3)+
Varas-Lorenzo *et al.* (2008; Canada)	193	●	●	●	●	95 (91-98)
Harriss *et al.* (2011; Australia)	202	●	●		●	95.5 (91.7-97.6)
Quan *et al.* (2008; Canada)	385	●			●	95.9 (93.4-97.4)+
Yeh *et al.* (2010; USA)	640	●	●			96.7 (95.0-97.8)+
Linnersjo *et al.* (2000; Sweden)	2101	●	●	●	●	98 (97.2-98.5)+
Coloma *et al.* (2013; Danish data)	148	●	●	●	●	100.0 (100-100)

Evaluating Performance of Algorithm

- Conclusion – for MI → no "gold standard" algorithm available
- Process is very costly and time consuming
- Small sample sizes → wide variation in estimates with wide confidence intervals

- In 33 studies "validating" algorithms, all reported PPV but:
  - Only 11 reported sensitivity
  - Only 5 reported specificity
  - **Is this really validation?**

The Value of Positive Predictive Value

- PPV is almost always reported in validation studies – easiest to assess
- Sensitivity and Specificity much less frequently reported
  - High cost and time to evaluate
- BUT – sensitivity and specificity are the actual characteristics of the test
  - PPV is a function of sensitivity, specificity and **prevalence** of Heath Outcome of Interest (HOI)

PPV Example – 1 Test, 2 Populations

Test Characteristics:

Sensitivity = 75%

Specificity = 99.9%

Population = 10,000

Prevalence = 1%		Truth	
		Positive	Negative
Test	Positive	75	10
	Negative	25	9890
	Total	100	9900

PPV =
75 / (75 + 10) =
88%

Prevalence = 5%		Truth	
		Positive	Negative
Test	Positive	375	10
	Negative	125	9490
		500	9500

PPV =
375 / (375 + 10) =
97%

PPV Example – 1 Population, 2 Tests

PPV = 90%

Population = 10,000

Prevalence = 5%		Truth	
		Positive	Negative
Test	Positive	90	10
	Negative	410	9490
	Total	500	9500

PPV = 90/(90+10) = 90%

Sens = 90/500 = 18%

Spec = 9490/9500 = 99.9%

Prevalence = 5%		Truth	
		Positive	Negative
Test	Positive	360	40
	Negative	140	9460
		500	9500

PPV = 360/(360+40) = 90%

Sens = 360/500 = 72%

Spec = 9460/9500 = 99.6%

Living with Algorithms

- Algorithms are used in most research with observational data
- Many ways to define an algorithm for any health outcome
- Each definition will have its own performance characteristics
  - Need to validate the algorithm to understand these characteristics
- Validation of an algorithm to be used in an observational dataset through chart review is likely not possible
  - Costly
  - Time consuming
  - Data is usually not available

Validating Algorithms in Observational Data

		Truth	
		Positive	Negative
Test	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Test – Comes from the algorithm/cohort definition
Truth – Some form of “gold standard” reference

Possible alternative for finding the “Truth”

Diagnostic Predictive Models

Prediction models used to estimate the probability of having a particular disease or outcome.

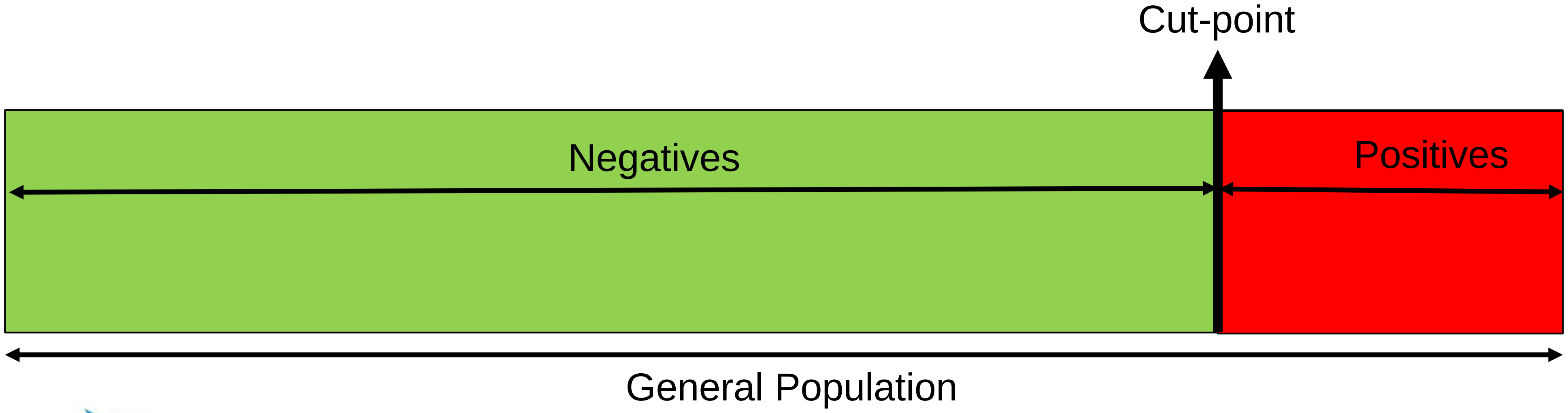
Finding the Truth – using Diagnostic Predictive Models

Step 1: Find a Gold standard of subjects for the HOI

Step 2: Develop the predictive model

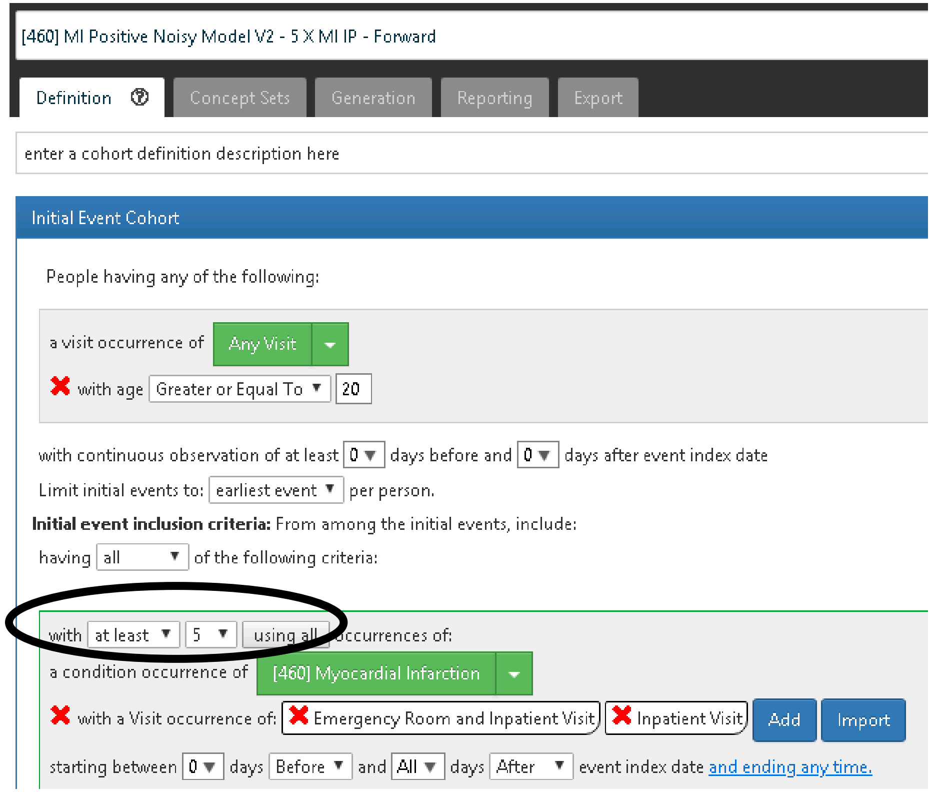
Step 3: Apply the model to a general population

Step 4: Determine a cut-point from the model



Finding a Gold Standard

- It turns out that having a very good set of positives is good enough – a “noisy” model
- We use an “extremely specific” (xSpec) cohort



[460] MI Positive Noisy Model V2 - 5 X MI IP - Forward

enter a cohort definition description here

Initial Event Cohort

People having any of the following:

a visit occurrence of Any Visit

with age Greater or Equal To 20

with continuous observation of at least 0 days before and 0 days after event index date

Limit initial events to: earliest event per person.

Initial event inclusion criteria: From among the initial events, include:

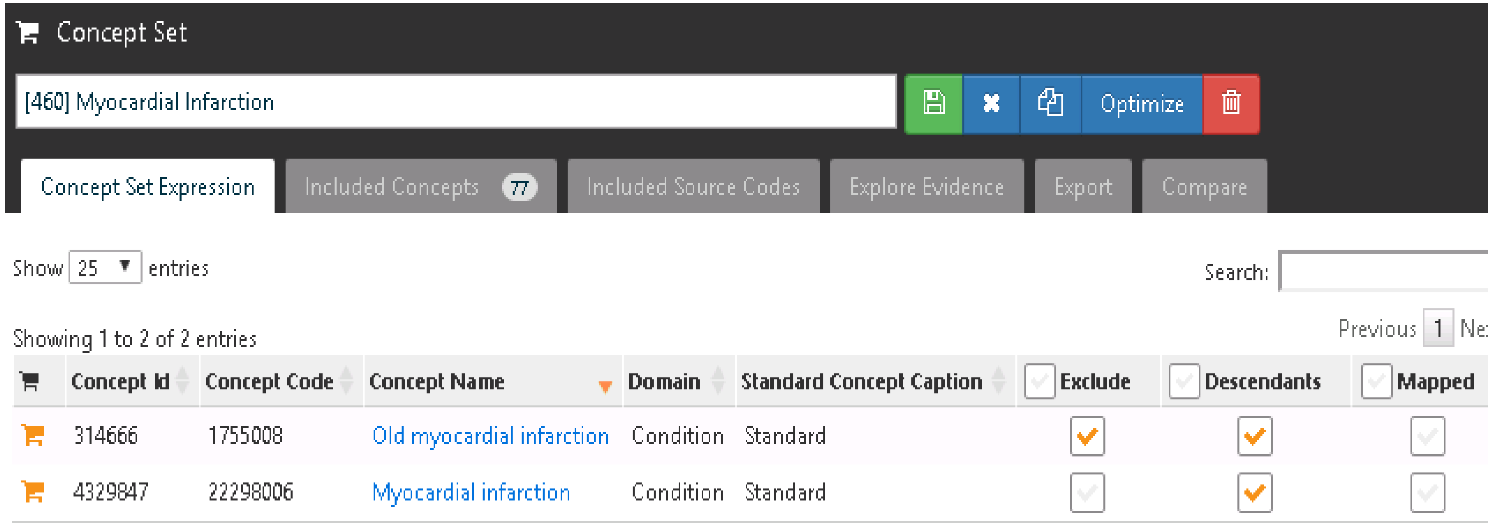
having all of the following criteria:

with at least 5 using all occurrences of:

a condition occurrence of [460] Myocardial Infarction

with a Visit occurrence of: Emergency Room and Inpatient Visit Inpatient Visit

starting between 0 days Before and All days After event index date and ending any time.



Concept Set

[460] Myocardial Infarction

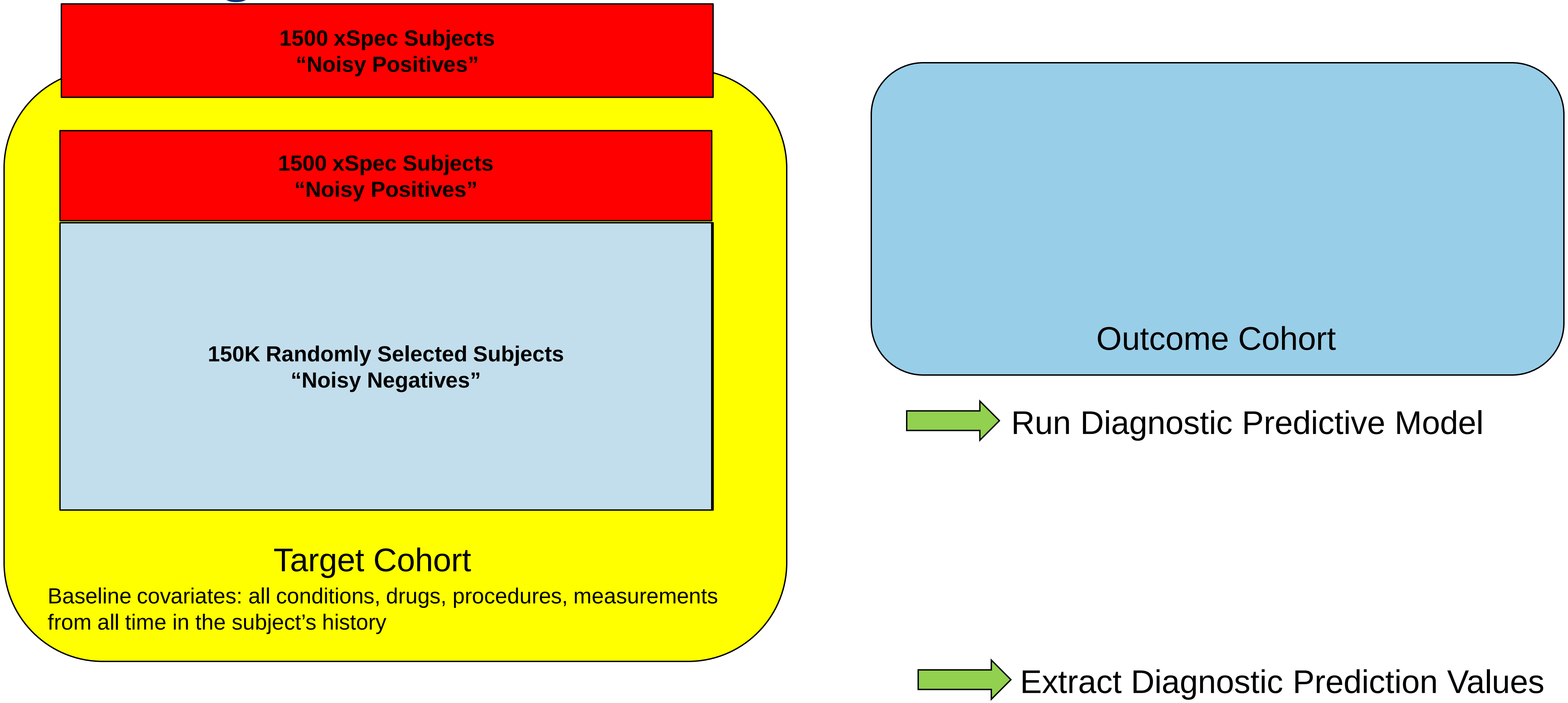
Show 25 entries

Search:

Showing 1 to 2 of 2 entries

	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
	314666	1755008	Old myocardial infarction	Condition	Standard	☑	☑	□
	4329847	22298006	Myocardial infarction	Condition	Standard	□	☑	□

Running the Model



1500 xSpec Subjects
“Noisy Positives”

1500 xSpec Subjects
“Noisy Positives”

150K Randomly Selected Subjects
“Noisy Negatives”

Target Cohort

Baseline covariates: all conditions, drugs, procedures, measurements
from all time in the subject’s history

Outcome Cohort

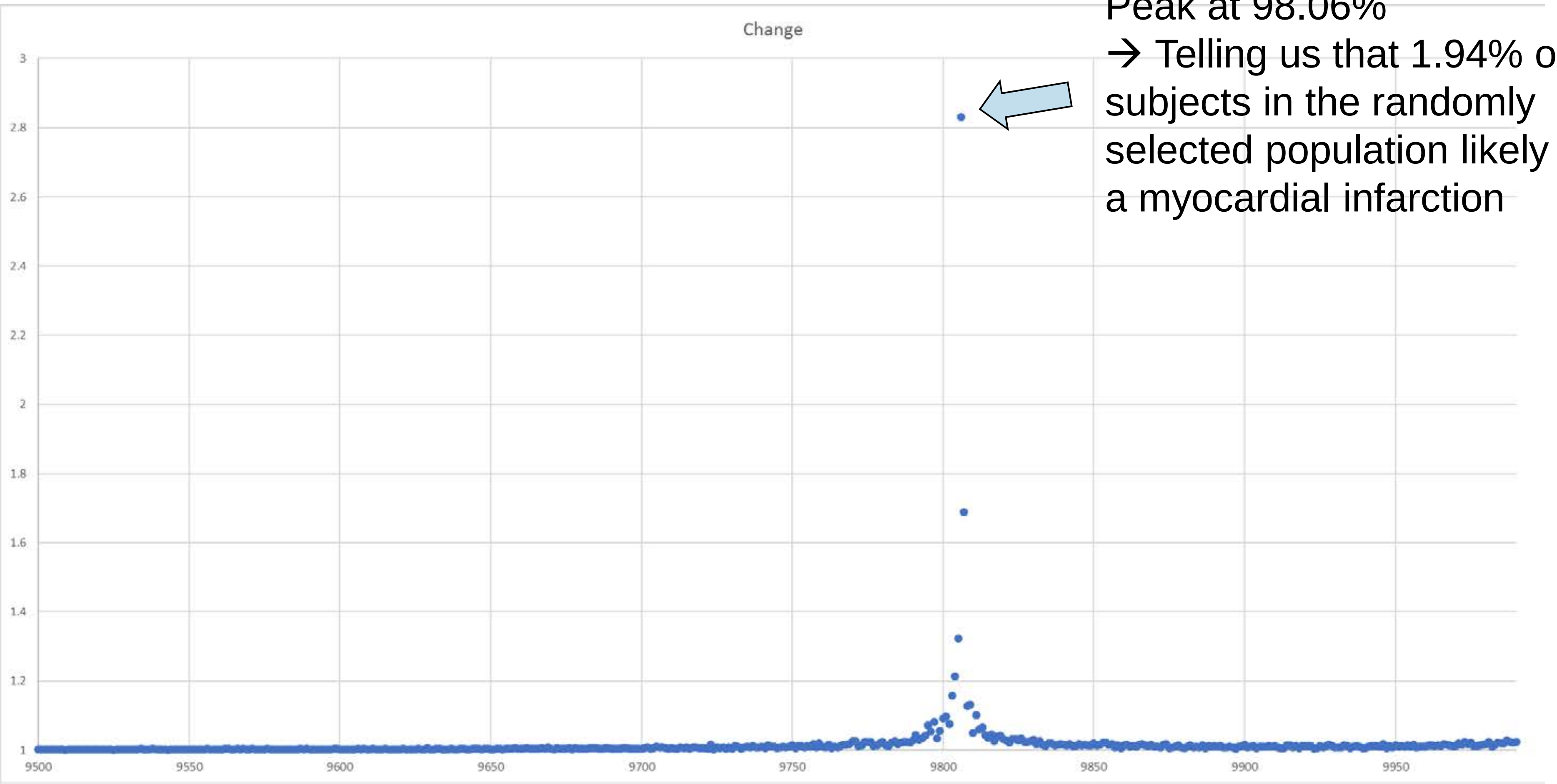
Run Diagnostic Predictive Model

Extract Diagnostic Prediction Values

Determining the Cut-Point

- We hypothesized that there should be a obvious change in the predictive values if you have the outcome or you don't
  - i.e., a subject doesn't "sorta" have a myocardial infarction
- We take the randomly chosen subjects and order them by predictive value
- Extract 10,000 subjects evenly spaced (by count) – each 0.01%

Prediction Curves – Myocardial Infarction



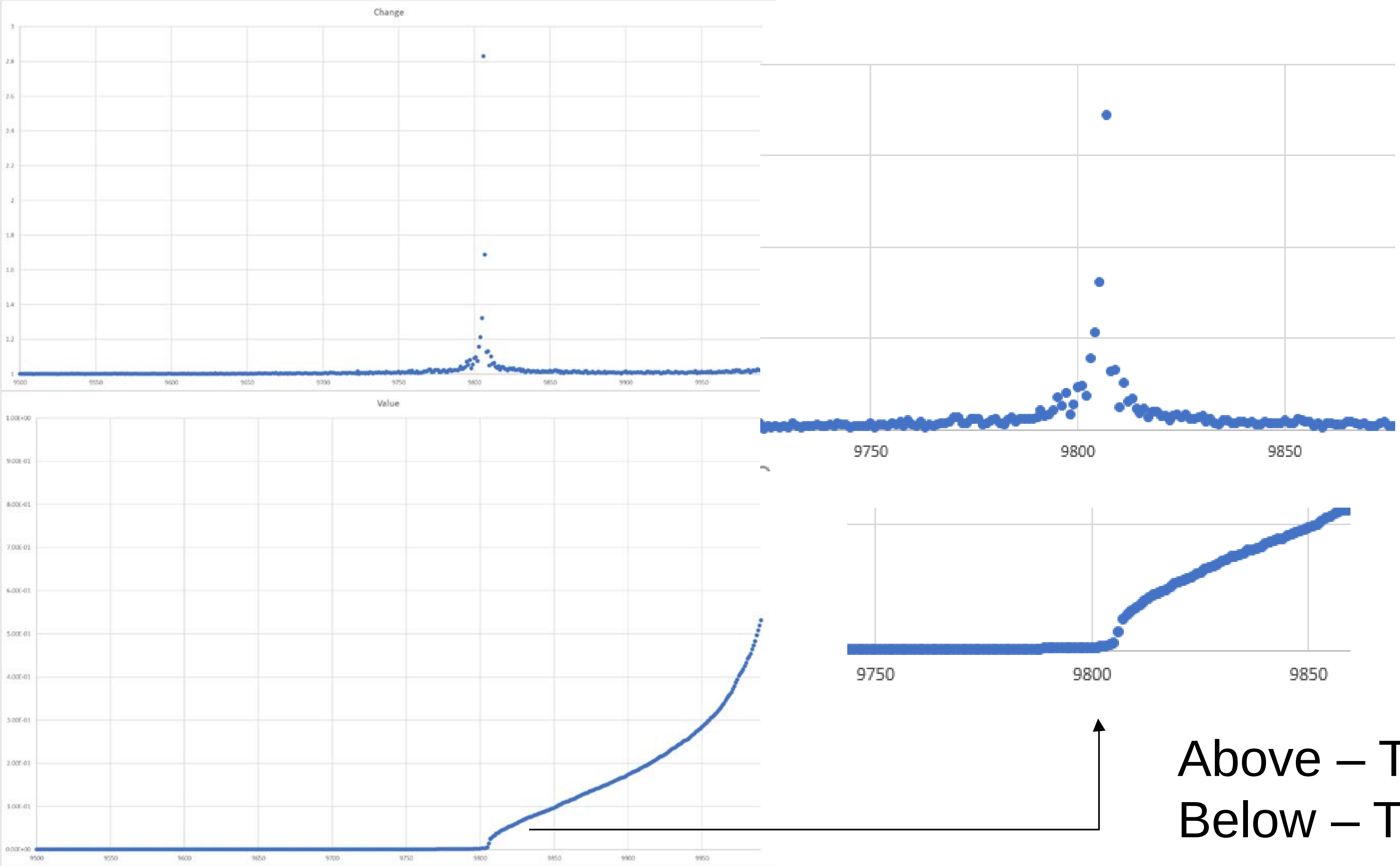
Peak at 98.06%
→ Telling us that 1.94% of the subjects in the randomly selected population likely had a myocardial infarction

Change – Difference in predictive value between each point and the previous point

Comparing Curves

Change Curve

Predicted Value Curve



Above – True Positives
Below – True Negatives

Testing the Phenotypes

Typical Phenotypes for MI:

- 1 X MI (Myocardial Infarction - SNOMED concept ID 22298006)
- 2 X MI, second MI diagnosis within 5 days of first MI diagnosis
- 1 X MI, In-patient
- 1 X MI, In-patient in first position
- Mini-Sentinel – ICD-9 410.x0 or 410.x1, In-patient in first position

Diagnostic testing:

- DRG codes (Optum only) – discharge codes not in concept set
- 5 X MI (xSpec) - acts as a positive control
- Pneumonia – acts as a negative control

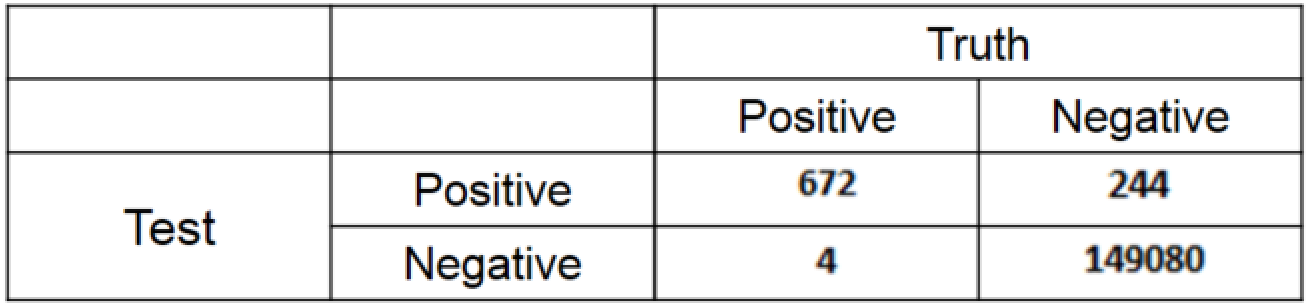
Comparing Results from Multiple Datasets

CDM	Pheno_Cohort_Name	Sens	PPV	Spec
dod	1 x MI	0.993	0.785	0.995
ccae	1 x MI	0.994	0.734	0.998
mdcr	1 x MI	0.984	0.84	0.99
mdcd	1 x MI	0.983	0.732	0.994
dod	2 x MI	0.597	0.913	0.999
ccae	2 x MI	0.713	0.896	> 0.999
mdcr	2 x MI	0.555	0.922	0.998
mdcd	2 x MI	0.558	0.847	0.998
dod	1 x MI - In-Patient	0.839	0.908	0.999
ccae	1 x MI - In-Patient	0.896	0.899	> 0.999
mdcr	1 x MI - In-Patient	0.78	0.918	0.996
mdcd	1 x MI - In-Patient	0.752	0.824	0.997
dod	1 x MI, IP - 1st Position	0.709	0.952	0.999
ccae	1 x MI, IP - 1st Position	0.834	0.934	> 0.999
mdcr	1 x MI, IP - 1st Position	0.693	0.952	0.998
mdcd	1 x MI, IP - 1st Position	0.59	0.89	0.999

CDM	Pheno_Cohort_Name	Sens	PPV	Spec
dod	1 x MI DRG	0.123	0.941	> 0.999
dod	Mini-Sentinel	0.704	0.953	0.999
ccae	Mini-Sentinel	0.833	0.934	> 0.999
mdcr	Mini-Sentinel	0.689	0.952	0.998
mdcd	Mini-Sentinel	0.586	0.89	0.999
dod	Pos. control (5 X MI IP)	0.108	> 0.999	> 0.999
ccae	Pos. control (5 X MI IP)	0.173	> 0.999	> 0.999
mdcr	Pos. control (5 X MI IP)	0.091	> 0.999	> 0.999
mdcd	Pos. control (5 X MI IP)	0.1	> 0.999	> 0.999
dod	Neg. control (Pneumonia)	0.452	0.108	0.938
ccae	Neg. control (Pneumonia)	0.206	0.029	0.969
mdcr	Neg. control (Pneumonia)	0.483	0.154	0.859
mdcd	Neg. control (Pneumonia)	0.495	0.091	0.914

Is the Cut-point the “Truth”

- The cut-point is critical for the analysis
- Is there a way to test it’s validity?

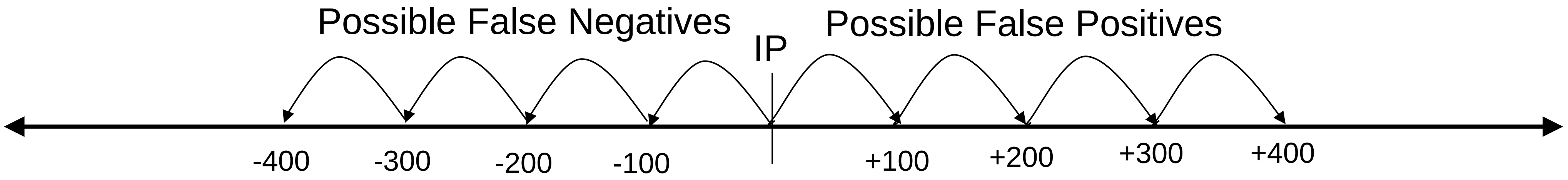


		Truth	
		Positive	Negative
Test	Positive	672	244
	Negative	4	149080

The “truth” says there are 676 (672 + 4) Positives and 149,324 (244 + 149,080) Negatives

That’s a lot of testing!

Prioritized Testing

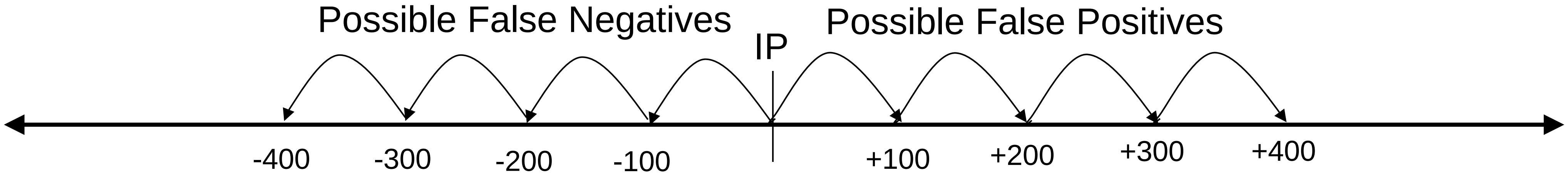


Using 100 subject increments above and below the inflection point (IP)

- Find Possible False Positives (from Model) – test subjects above the IP for **lack of MI concepts** from the concept set

- Find Possible False Negatives (from Model) – test subjects below the IP for **presence of MI concepts** from the concept set

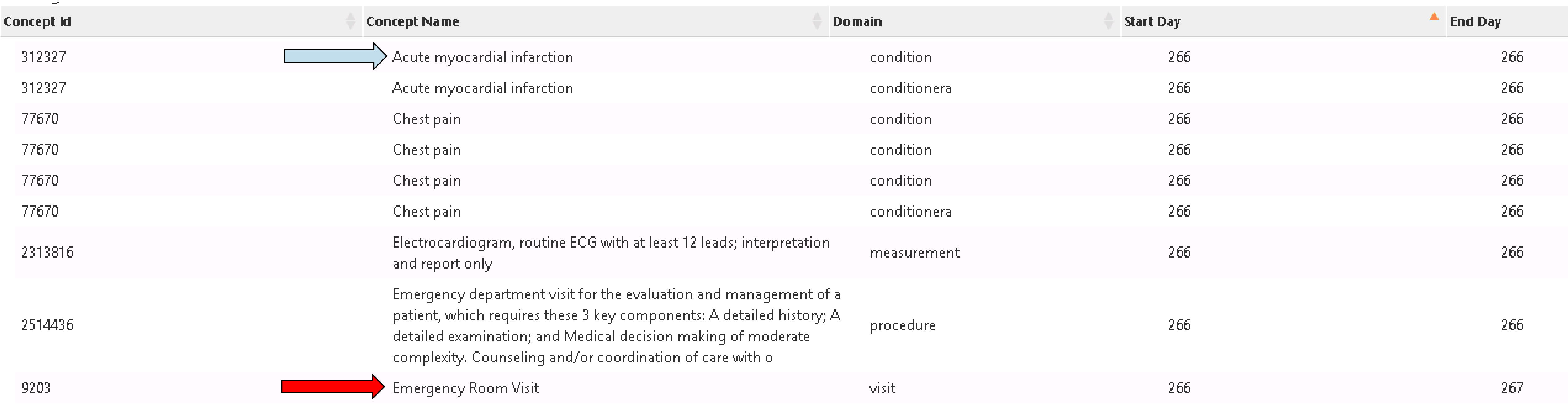
Prioritized Testing



Error Type	Low Subj	High Subj	startPoint	Possible Err Count	Subject_1	Value_1	Subject_2	Value_2	Subject_3	Value_3
Possible False Negatives	-1	-100	0.025571175	87	899054601	0.02129999	2062583901	0.020999708	683751301	0.013315725
Possible False Positives	1	100	0.025774348	4	26475227701	0.033513567	26455412501	0.039489521	27875525001	0.030372463
Possible False Negatives	-101	-200	0.011291329	65	2311562001	0.00681076	27565896701	0.007426981	2057505501	0.007620929
Possible False Positives	101	200	0.049078328	0						
Possible False Negatives	-201	-300	0.006568489	45	715751801	0.006281902	2211537001	0.005296667	27225203701	0.006412863
Possible False Positives	201	300	0.093735495	0						
Possible False Negatives	-301	-400	0.003750125	19	2225555001	0.002821464	2309592402	0.003343347	1863528802	0.003361777
Possible False Positives	301	400	0.172181149	0						
Possible False Negatives	-401	-500	0.002625932	14	2299532401	0.002554089	27905650103	0.002432534	25335922202	0.001981134
Possible False Positives	401	500	0.285868877	0						
Possible False Negatives	-501	-600	0.001967144	5	26205233201	0.001773443	27005952201	0.00169774	27335010001	0.001783614
Possible False Positives	501	600	0.515169065	0						
Possible False Negatives	-601	-700	0.001581028	3	27435344401	0.00150035	614652402	0.001472033	26355015701	0.001368673
Possible False Negatives	-701	-800	0.001292544	2	27565608601	0.001197725	27335711601	0.001231248		
Possible False Negatives	-801	-900	0.001087654	0						

Testing for False Negatives

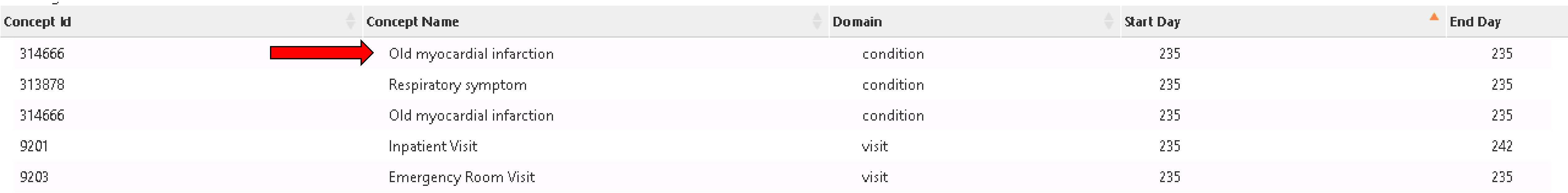
Subject ID: 899054601



Concept Id	Concept Name	Domain	Start Day	End Day
312327	Acute myocardial infarction	condition	266	266
312327	Acute myocardial infarction	conditionera	266	266
77670	Chest pain	condition	266	266
77670	Chest pain	condition	266	266
77670	Chest pain	condition	266	266
77670	Chest pain	conditionera	266	266
2313816	Electrocardiogram, routine ECG with at least 12 leads; interpretation and report only	measurement	266	266
2514436	Emergency department visit for the evaluation and management of a patient, which requires these 3 key components: A detailed history; A detailed examination; and Medical decision making of moderate complexity. Counseling and/or coordination of care with o	procedure	266	266
9203	Emergency Room Visit	visit	266	267

Testing for False Positives

Subject ID: 26475227701



Concept Id	Concept Name	Domain	Start Day	End Day
314666	Old myocardial infarction	condition	235	235
313878	Respiratory symptom	condition	235	235
314666	Old myocardial infarction	condition	235	235
9201	Inpatient Visit	visit	235	242
9203	Emergency Room Visit	visit	235	235

Other Disease Phenotypes Tested

Acute Diseases:

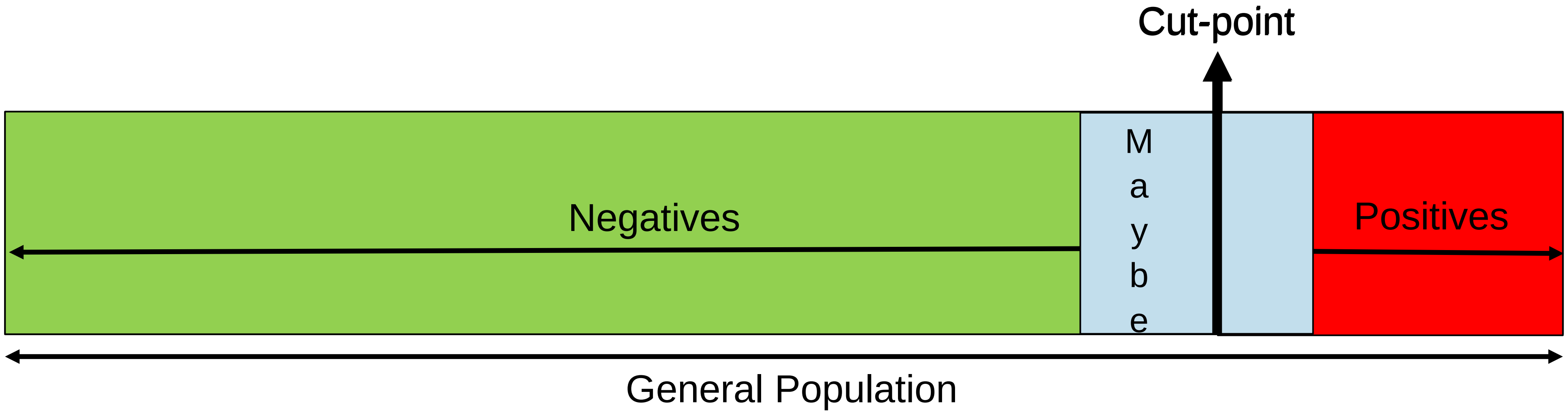
- Hemorrhagic Stroke
- GI Hemorrhage
- Ischemic Stroke
- Acute Respiratory Failure

Chronic Disease:

- Type 2 Diabetes
- Rheumatoid Arthritis
- Heart Failure
- Psoriasis
- Multiple Myeloma

Limitations

- Sparse data for subjects
- Databases vary with overall level of detail
- Complex coding for conditions, e.g., MI v. T2DM



- Cutrona – 10% of patients with insufficient evidence
- Ryo – 7.5% of patients with insufficient evidence

Conclusion

- Using diagnostic predictive models to assess algorithm performance appears promising
- Having metrics for phenotype performance increases confidence in the use of observational data in research.
- Potential to use results of phenotype evaluation to correct/adjust our estimates
- Next steps: methods to reduce the indeterminants
    - Testing adjusting the xSpec cohort

Questions