Welcome to the Sentinel Innovation and Methods Seminar Series

The webinar will begin momentarily

- Please visit www.sentinelinitiative.org for recordings of past sessions and details on upcoming webinars.
- Note: closed-captioning for today's webinar will be available on the recording posted at the link above.

Sentinel

Data leakage due to care provided outside of the study electronic health record system: potential biases and solutions

Joshua K. Lin, MD, ScD, MPH

Executive Director, Mass General Brigham Center for Integrated Healthcare Data Research
Assistant Professor of Medicine, Harvard Medical School, Boston, MA
Pharmacoepidemiologist, Division of Pharmacoepidemiology and Pharmacoeconomics
Department of Medicine, Brigham and Women's Hospital, Boston, MA
Attending Physician, Hospital Medicine Group, Massachusetts General Hospital, Boston, MA

March 2022

Disclosures

- I have no conflict of interest to disclose.
- This study is funded by NIH-National Library of Medicine 5R01LM012594-03
- I also received funding from NIH-National Institute on Aging (1R01AG062713-01, R01AG067507, 3RF1AG063381-01S1, and 1RF1AG063381-01) and NIH-National Library of Medicine (1R01LM013204-01A1).

Information bias due to electronic health record (EHR)-discontinuity

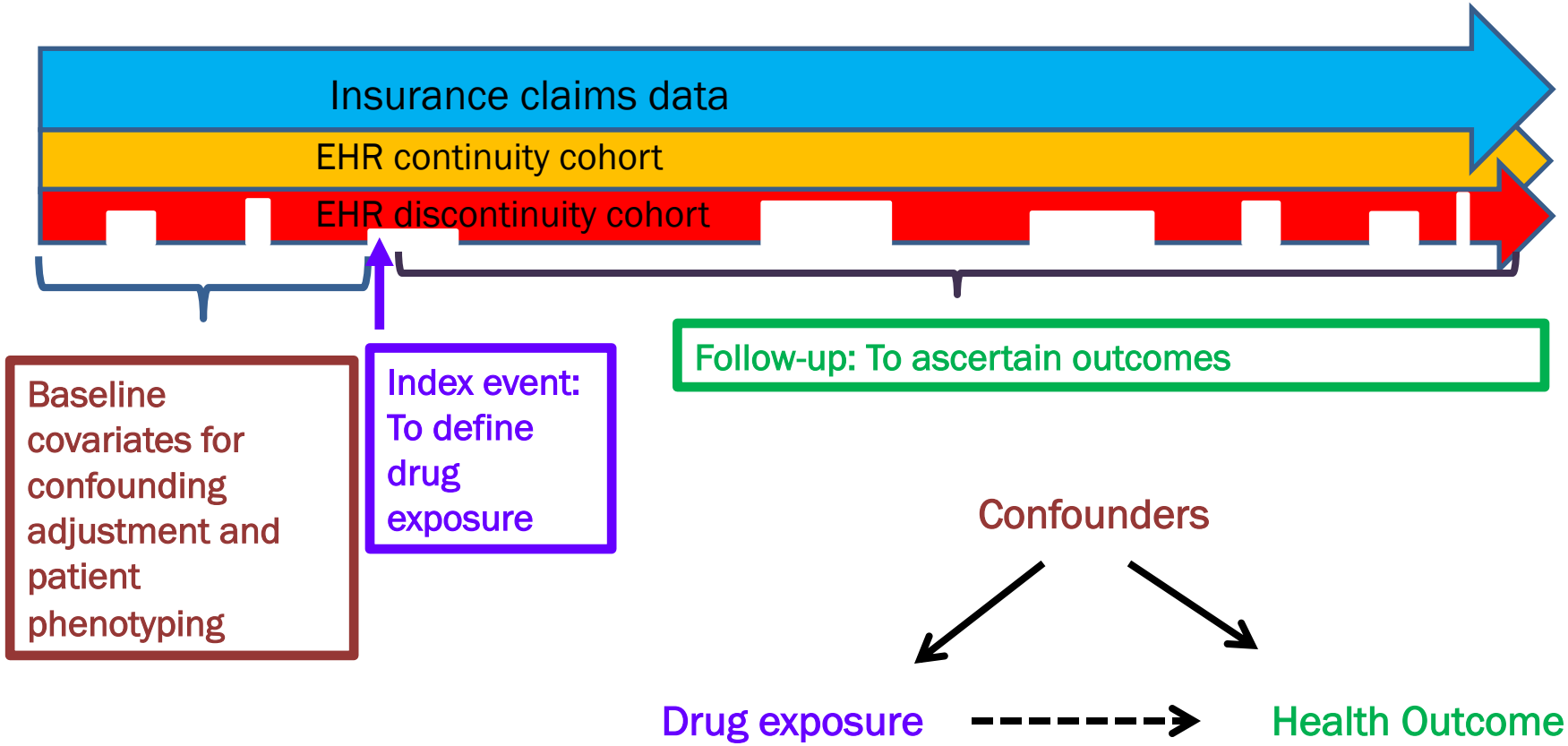
- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Information bias due to electronic health record (EHR)-discontinuity

- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Information bias due to EHR-discontinuity

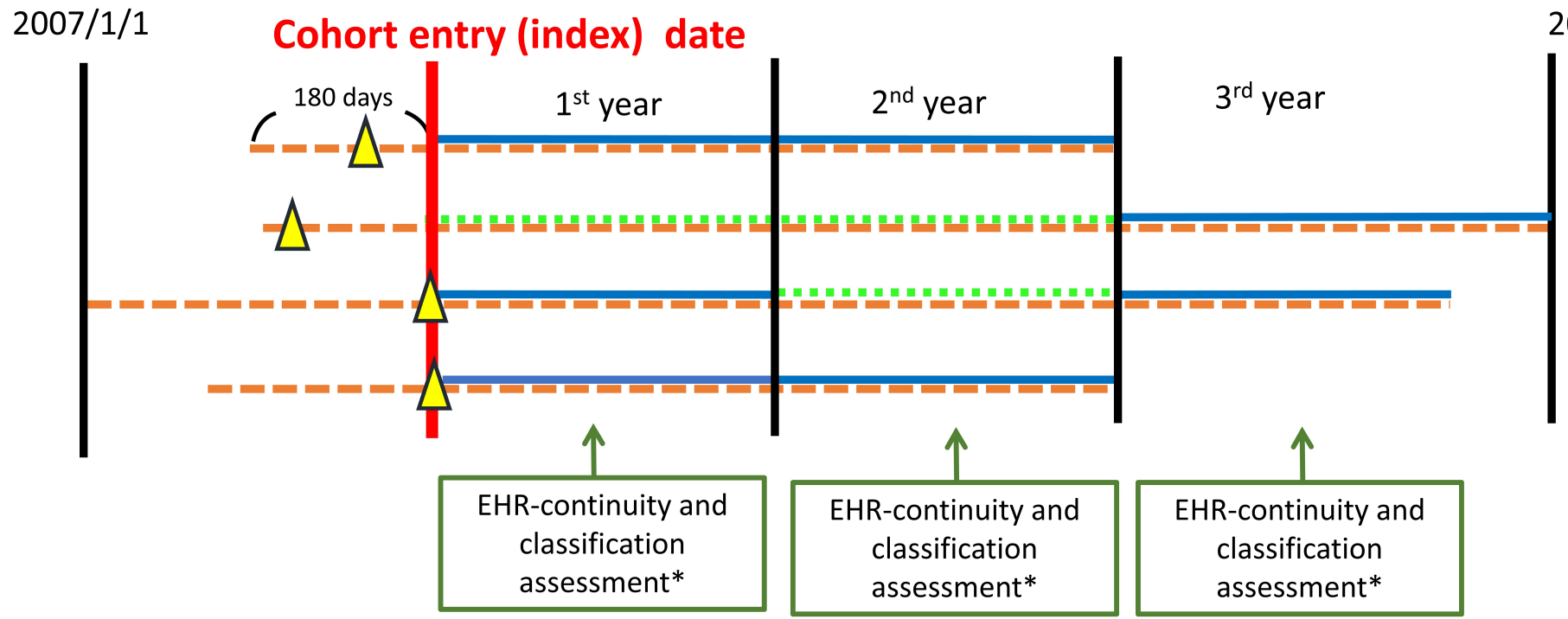
- Most US EHR systems are subject to data incompleteness due to EHR-discontinuity, defined as “receiving care outside of study EHR”



Information bias due to EHR-discontinuity

- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Study design



EHR=electronic health record; discontinuity=receiving care outside of the study EHR

Encounters in EHR

Medicare claims data for Part A+B+D

With high EHR-continuity

With low EHR-continuity

*Assess the completeness of the records captured by the EHR and classification of key variables during the same period

EHR-continuity metric

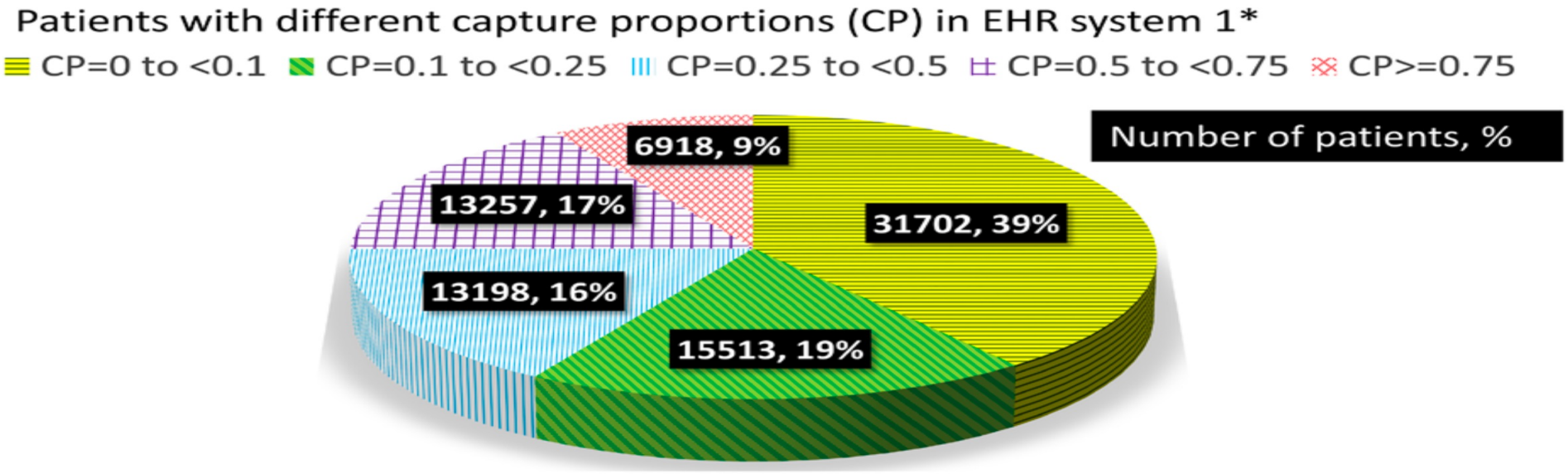
Clinical encounters	ID=1		ID=2		ID=3		…
	Claims	EHR	Claims	EHR	Claims	EHR	…
1	1	1	1	1	1	1	…
2	1	1	1	1	1	1	…
3	1	1	1	1	1	1	…
4	1	1	1	0	1	1	…
5	1	1	1	0	1	1	…
6	1	1	1	0	1	1	…
7	1	1	1	0	1	0	…
8	1	1	1	0	1	0	…
9	1	1	1	0	1	0	…
10	1	0	1	0	1	0	…
Capture %	9/10 = 90% **High** EHR-continuity		3/10 = 30% **Low** EHR-continuity		6/10 = 60% "in-between"		…

= services recorded in claims AND EHR

Mean Proportions of Encounters Captured (**MPEC**) = $\left(\frac{\text{Outpatient encounters recorded in EHR}}{\text{Outpatient encounters recorded in claims data}} + \frac{\text{Inpatient encounters recorded in EHR}}{\text{Inpatient encounters recorded in claims data}}\right)/2$

Proportion of encounters captured by electronic health record systems

A Patients with different capture proportions (CP) in EHR system 1*



B Mean capture proportions by EHR system and year after cohort entry

Year after cohort entry	1	2	3	4	5	6	7
Mean capture proportion in EHR system 1	0.27	0.22	0.22	0.22	0.23	0.24	0.26
Mean capture proportion in EHR system 2	0.22	0.16	0.16	0.16	0.16	0.18	0.20

EHR= electronic health record, CP=capture proportion, * Proportions were based on data in the EHR system 1 in the first year, but the pattern was similar in the subsequent years and in EHR system 2.

Epidemiology. 2018;29(3):356-363

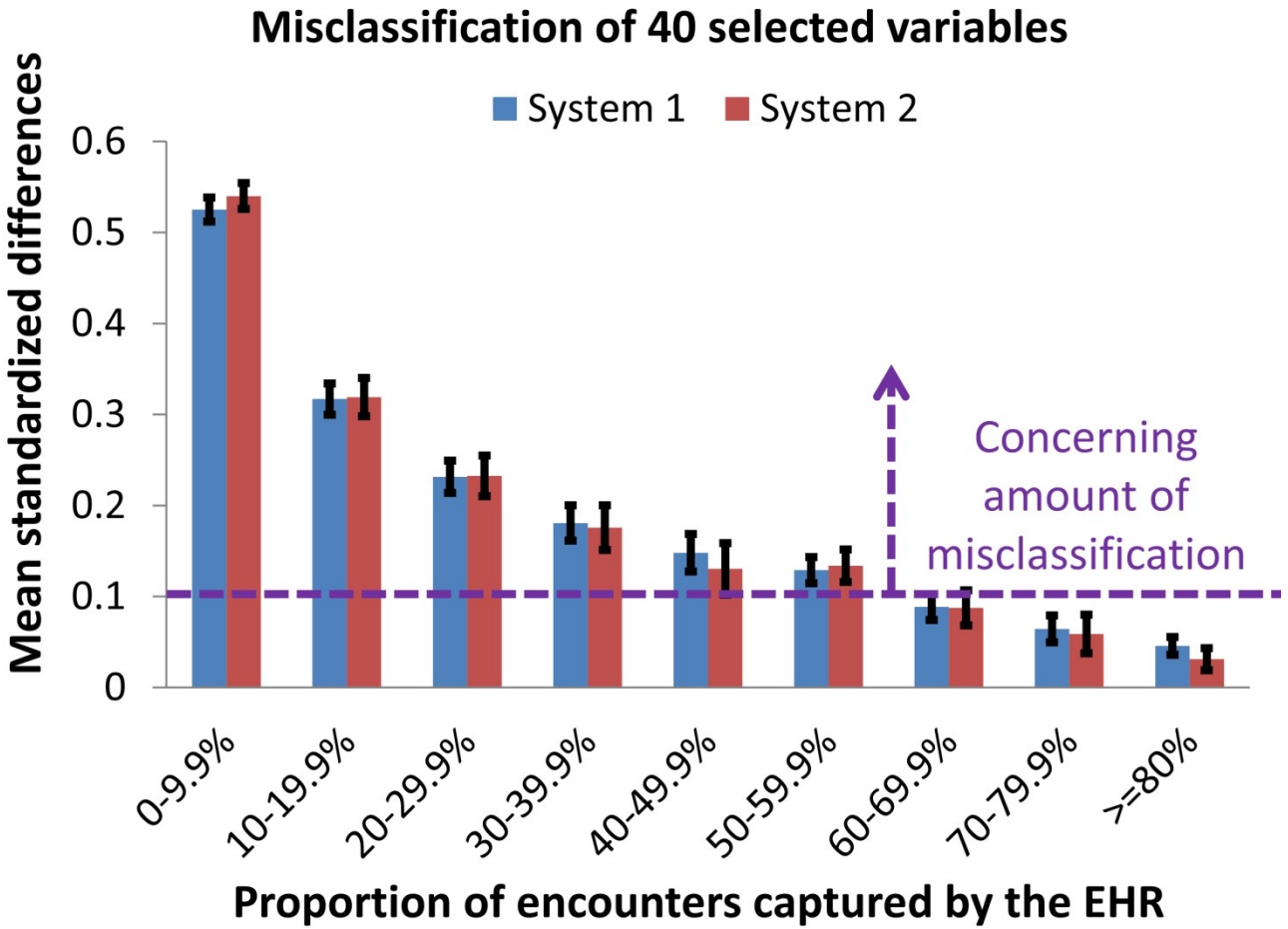
Misclassification metric

	EHR	Claims-EHR	Standardized difference
X1	P1(X1=1)	P1' (X1=1)	D1
X2	P2 (X2=1)	P2' (X2=1)	D2
X3	P3 (X3=1)	P3' (X3=1)	D3
…	…	…	…
X40	P40 (X40=1)	P40' (X40=1)	D40

Standardized difference
Di = (Pi - Pi')/pooled standard deviation; i=1-40. *Epidemiology. 2018;29(3):356-363*

Patient characteristics to be assessed for accuracy of classification	
25 co-morbidity variables	a) **15 variables commonly used as covariates:** dementia, atrial fibrillation, chronic lung disease, chronic liver disease, chronic kidney disease, cancer, diabetes, hypertension, anemia, psychosis, depression, pneumonia, HIV, fracture, and rheumatoid arthritis b) **10 variables with validated algorithm commonly used as outcome variables:** ischemic stroke, intracranial hemorrhage, congestive heart failure, acute kidney injury, myocardial infarction, pulmonary embolism, deep vein thrombosis, hepatotoxicity, GI Bleeding, major bleeding
15 medication use variables	antiplatelet agents, antidiabetics, antihypertensives, nonsteroidal anti-inflammatory drugs, opioids, antidepressants, antipsychotics, anticonvulsants, proton pump inhibitors, antiarrhythmics, statins, dementia, hormone therapy, antibiotics, and oral anticoagulants

Decreasing misclassification associated with increasing EHR continuity



Mean standardized difference in patients with MPEC <10 % (=0.53) was 11.4 fold (95% CI: 9.4-14.6) greater than that for MPEC ≥ 80% (=0.05)

Epidemiology. 2018;29(3):356-363

Information bias due to EHR-discontinuity

- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Correlation between predicted vs. observed EHR continuity

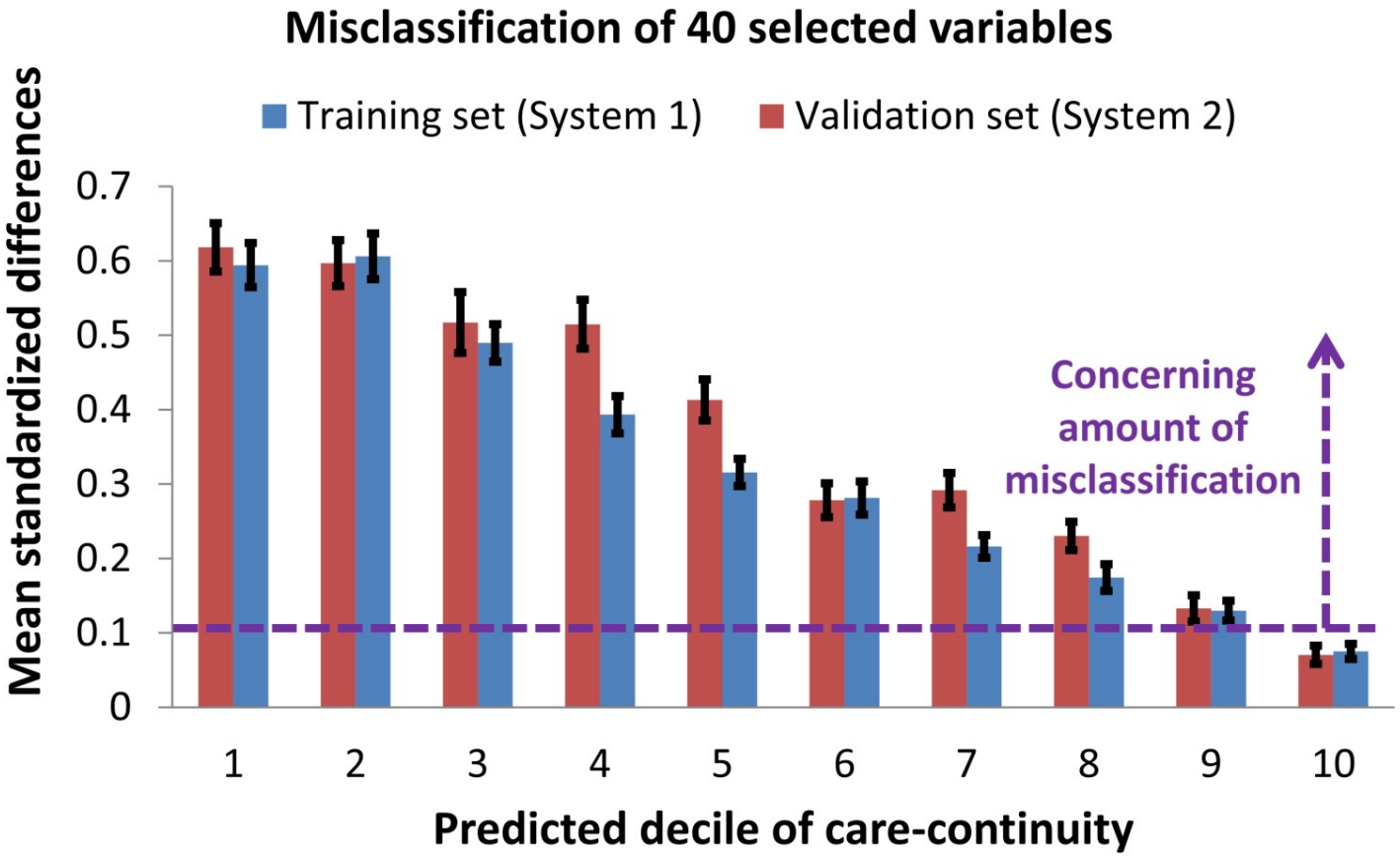
Variable	Coefficient
Intercept	-0.010
Having seen the same provider twice	0.049
Having seen the same provider >=3 times	0.087
Having general medical exam*	0.078
Mammography*	0.075
Pap smear*	0.009
PSA Test*	0.103
Colonoscopy*	0.064
Fecal occult blood test*	0.034
Influenza vaccine*	0.102
Pneumococcal vaccine*	0.031
Having BMI recorded*	0.017
Having 2 of the above routine care facts**	0.049
With any one medication use record	0.002
With at least 2 medication use records	0.074
Having A1C ordered or value recorded*	0.018
Having at least one inpatient or outpatient encounter	0.091
Having at least two outpatient encounters	0.050
With 1 diagnosis recorded in the EHR	-0.026
With at least 2 diagnoses recorded in the EHR	0.037
Having any ED visit in the EHR	0.078
** having 2 of the facts followed by* PSA= prostate specific antigen	

	Training	Validation
AUC for predicting MPEC ≥ 60%	0.86	0.86
Spearman coefficient with measured MPEC	0.78	0.73

AUC=Area under the ROC curve;
MPEC=Mean Proportions of Encounters Captured

Clin Pharmacol Ther. 2018;103(5):899-905.
Clinical Epidemiology. Volume 2020:12 Pages 133—141

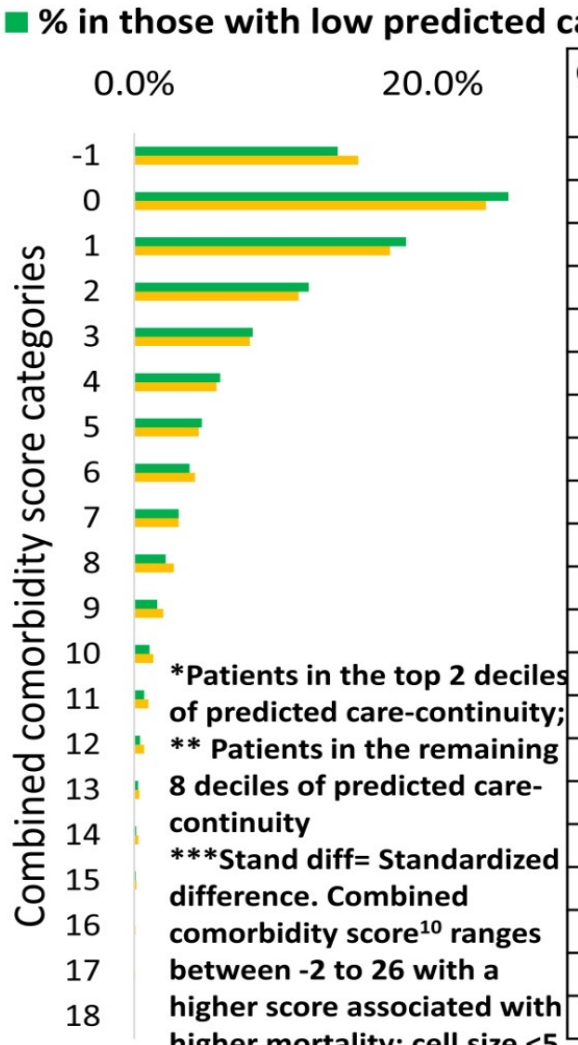
Decreasing misclassification associated with increasing predicted EHR-continuity



In the validation set, the mean standardized difference between the proportions of the 40 selected variables based on EHR alone vs. the linked claims-EHR data in the lowest decile of predicted EHR-continuity (=0.62) was 8.8 fold greater than that in the highest predicted EHR-continuity decile (=0.08).

Clin Pharmacol Ther. 2018;103(5):899-905.

High representativeness: Comorbidity in patients with high vs. low EHR-continuity



*Patients in the top 2 deciles of predicted care-continuity; ** Patients in the remaining 8 deciles of predicted care-continuity

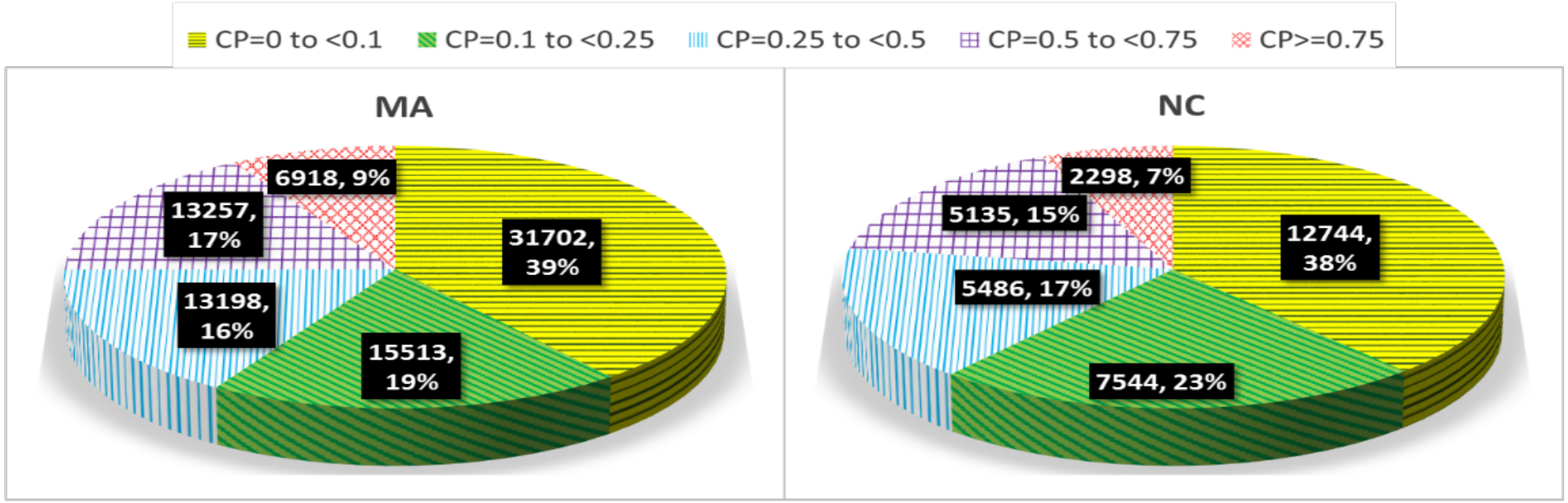
***Stand diff= Standardized difference. Combined comorbidity score[10] ranges between -2 to 26 with a higher score associated with higher mortality; cell size <5 were not presented here.

Comorbidity score	Low care-continuity N (%)*	High care-continuity N (%)**	Stand diff***
-1	8813 (13.6)	2407 (15.0)	0.04
0	16214 (25.1)	3780 (23.6)	0.04
1	11779 (18.2)	2750 (17.2)	0.03
2	7562 (11.7)	1767 (11.0)	0.02
3	5137 (8.0)	1242 (7.8)	0.01
4	3729 (5.8)	884 (5.5)	0.01
5	2926 (4.5)	693 (4.3)	0.01
6	2394 (3.7)	655 (4.1)	0.02
7	1931 (3.0)	479 (3.0)	0.00
8	1362 (2.1)	426 (2.7)	0.04
9	992 (1.5)	311 (1.9)	0.03
10	661 (1.0)	205 (1.3)	0.02
11	436 (0.7)	151 (0.9)	0.03
12	253 (0.4)	108 (0.7)	0.04
13	170 (0.3)	54 (0.3)	0.01
14	83 (0.1)	45 (0.3)	0.03
15	66 (0.1)	25 (0.2)	0.01
16	27 (0.04)	13 (0.1)	0.02
17	16 (0.02)	10 (0.1)	0.02
18	12 (0.02)	7 (0.04)	0.01
	Total=64569	Total=16019	Mean=0.02

Clin Pharmacol Ther. 2018;103(5):899-905.

External validation: very similar EHR continuity pattern in NC vs MA

A Patients with different capture proportions (CP) in MA and NC EHR system

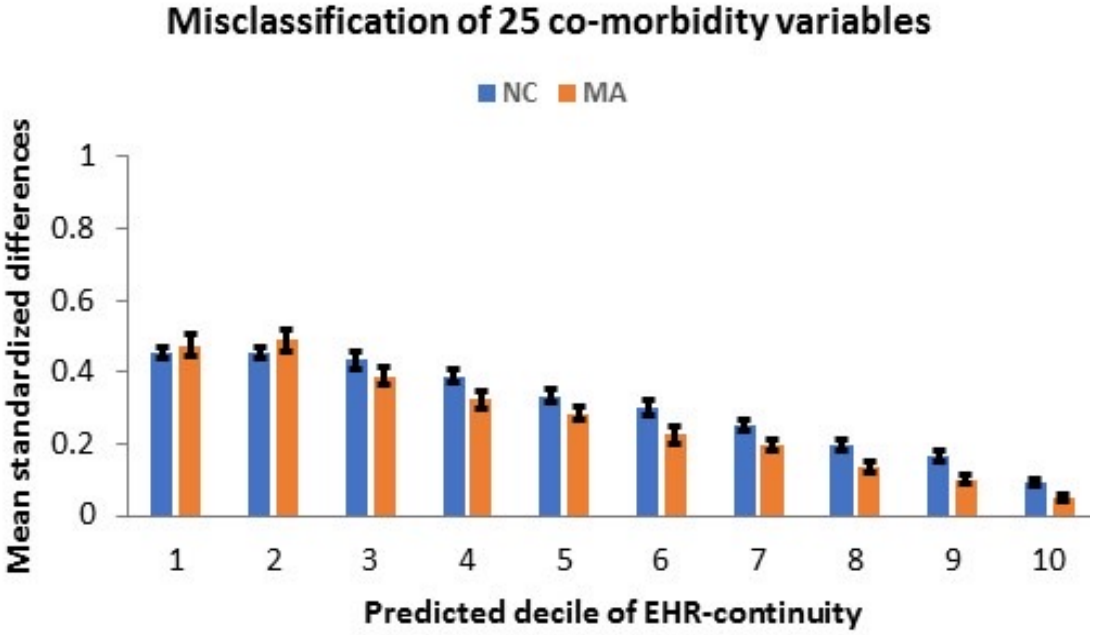


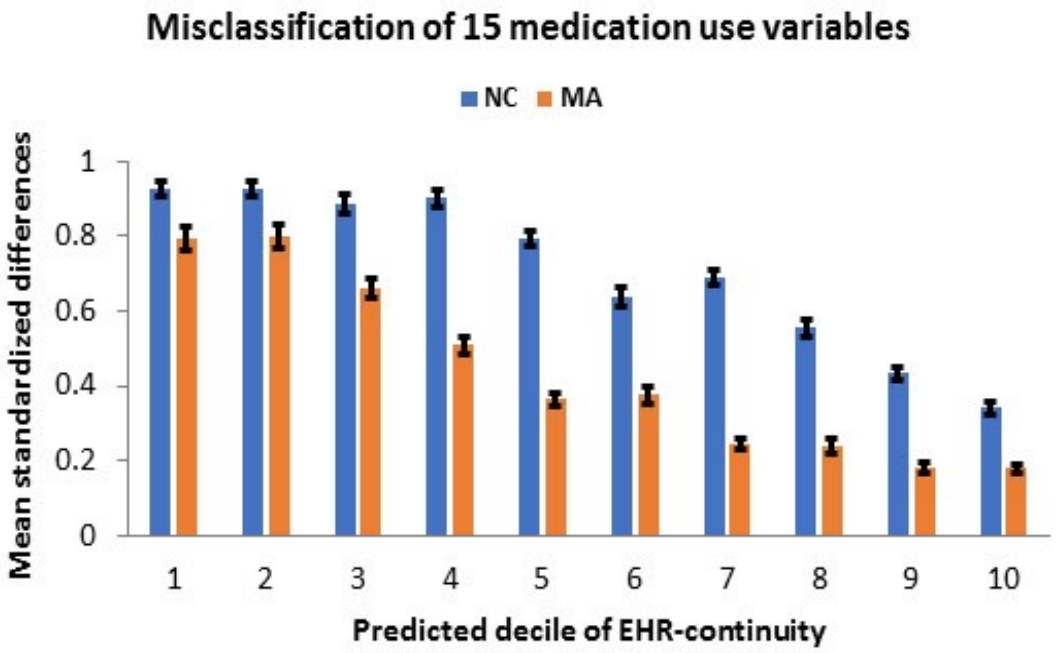
B Mean capture proportions (CP) by EHR system and year after cohort entry

Year after cohort entry	1	2	3	4	5	6	7
Mean CP in MA EHR	0.27	0.22	0.22	0.22	0.23	0.24	0.26
Mean CP in NC EHR	0.26	0.21	0.21	0.21	0.22	0.22	0.25

Clinical Epidemiology. Volume 2020:12 Pages 133—141

External validation: Decreasing misclassification associated with increasing predicted EHR-continuity

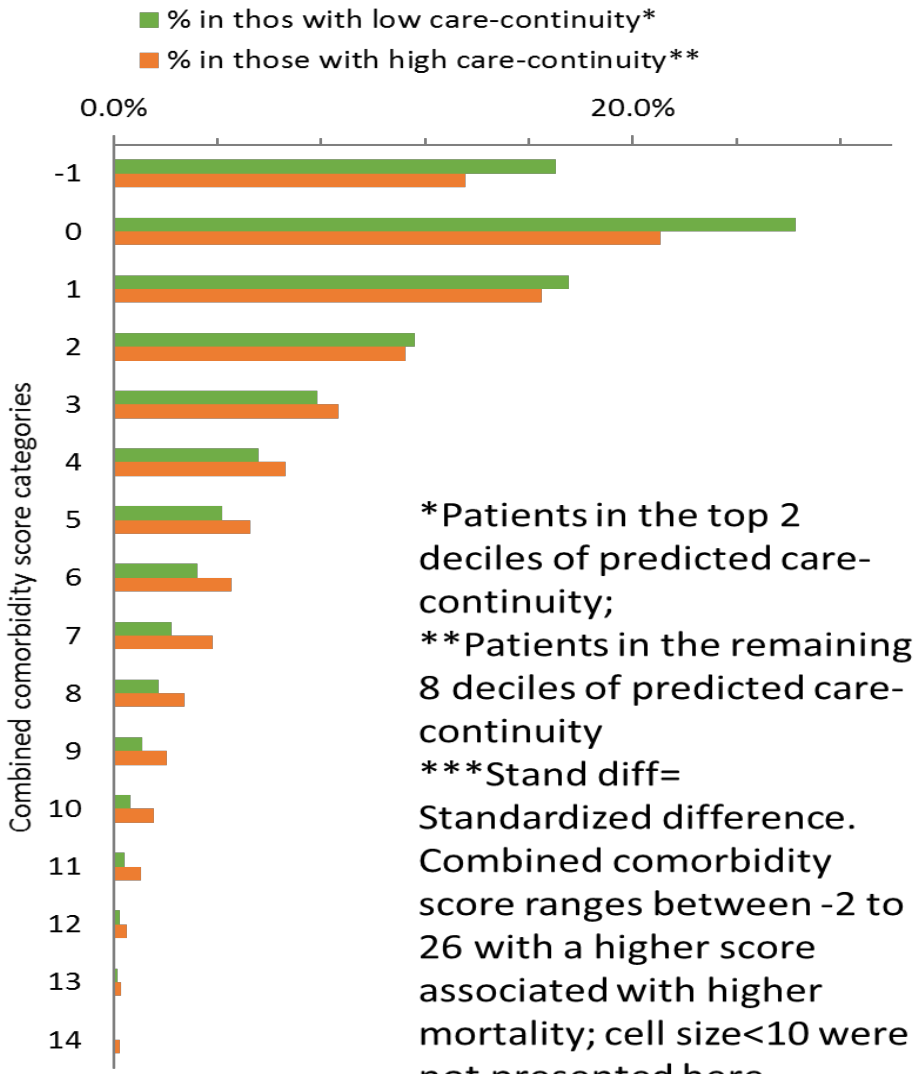




- The sources of medication information in EHR:
  - Prescribing (order entry) data
  - Medication reconciliation
  - Electronic medication administration data
  - Dispensing (mostly only inpatient dispensing).
- Electronic medication administration data were not available in the MA EHR research database and medication reconciliation information was not available in the NC EHR data.

Clinical Epidemiology. Volume 2020:12 Pages 133—141

Representativeness of the EHR-continuity cohort



*Patients in the top 2 deciles of predicted care-continuity;
**Patients in the remaining 8 deciles of predicted care-continuity
***Stand diff= Standardized difference. Combined comorbidity score ranges between -2 to 26 with a higher score associated with higher mortality; cell size<10 were not presented here.

Comorbidity score categories	Low care-continuity N (%)*	High care-continuity N (%)**	Stand. Diff.***
-1	4538 (26.3)	896 (13.6)	0.07
0	6994 (17.6)	1393 (27.7)	0.06
1	4670 (11.6)	1089 (16.5)	0.03
2	3082 (7.8)	742 (10.3)	0.05
3	2085 (5.6)	572 (7.3)	0.03
4	1490 (4.2)	439 (4.9)	0.04
5	1117 (3.2)	347 (4)	0.03
6	862 (2.2)	301 (3.2)	0.02
7	596 (1.7)	251 (2.7)	0.01
8	462 (1.1)	181 (1.7)	0.02
9	289 (0.6)	135 (1.1)	0.02
10	170 (0.4)	102 (0.6)	0.03
11	116 (0.2)	70 (0.5)	0.01
12	61 (0.1)	35 (0.3)	0.01
13	33 (0.1)	19 (0.1)	0.01
14	17 (0)	17 (0.1)	0.01
	Total N = 26,599 (100)	Total = 6,608 (100)	Mean stand diff = 0.05

Clinical Epidemiology. Volume 2020:12 Pages 133—141

Information bias due to EHR-discontinuity

- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Impact of EHR continuity on risk score classification

- We calculated four commonly used risk scores:
  - $CHAD_2DS_2$-VASc in patients with atrial fibrillation
  - HAS-BLED in patients with atrial fibrillation
  - Combined co-morbidity score (CCS) in the general population
  - Claims-based frailty index (CFI) in the general population
  - Reference standard: scores assessed using the linked EHR-claims data.

Misclassification by 2 categories					
	Risk scores	$CHAD_2DS_2$-VASc	HAS-BLED	CCS	Frailty Index
EHR System 1	High EHR-continuity*	16.48%	11.91%	16.14%	1.64%
	Low EHR-continuity**	55.15%	54.90%	36.94%	10.40%
	RR low vs. high EHR-continuity (95% CI)	3.35 (3.14 - 3.58)	4.61 (4.27 - 5.00)	2.29 (2.14 - 2.45)	6.34 (5.17 - 8.16)
EHR System 2	High EHR-continuity*	16.22%	13.56%	18.81%	1.76%
	Low EHR-continuity**	54.50%	55.14%	41.37%	9.51%
	RR low vs. high EHR-continuity (95% CI)	3.36 (3.09 - 3.68)	4.07 (3.71 - 4.50)	2.20 (2.03 - 2.39)	5.40 (4.18 - 7.57)

*High vs. low EHR-continuity: predicted EHR-continuity ≤0.3 vs. >0.3, which corresponds to the cut-off for the top 20% predicted EHR continuity in the original training set.

JAMIA 2022 (in press)

Impact of EHR continuity on risk score classification

Area under the ROC curve of the risk scores when predicting the target outcome				
EHR-continuity*	CHADS_EHR	CHADS_EHR+C	HASBLED_EHR	HASBLED_EHR+C
Q1	0.605	0.809	0.549	0.753
Q2	0.757	0.802	0.615	0.758
Q3	0.814	0.815	0.72	0.765
Q4	0.757	0.792	0.75	0.781
Using CHAD2DS2-VASc to predict 1-year risk of stroke and HASBLED to predict 1-year risk of major bleeding				

Area under the ROC curve of the risk scores when predicting the target outcome				
EHR-continuity*	CCS_EHR	CCS_EHR+C	Frailty_EHR	Frailty_EHR+C
Q1	0.562	0.806	0.527	0.748
Q2	0.678	0.824	0.568	0.713
Q3	0.769	0.848	0.651	0.750
Q4	0.838	0.866	0.699	0.748
Using CCS (Combined co-morbidity score) and frailty index to predict 1-year mortality				

_EHR: based on electronic health records (EHR); _EHR+C: based on EHR and claims data
*Predicted EHR continuity cut-off that corresponds to the 1st to 4th quartiles

JAMIA 2022 (in press) and a study in progress

Information bias due to EHR-discontinuity

- What is the issue?
- Quantify the bias
- Potential solution
- Impact on patient phenotyping by risk scores
- Impact on treatment effect estimates

Impact of EHR continuity on treatment effect estimates

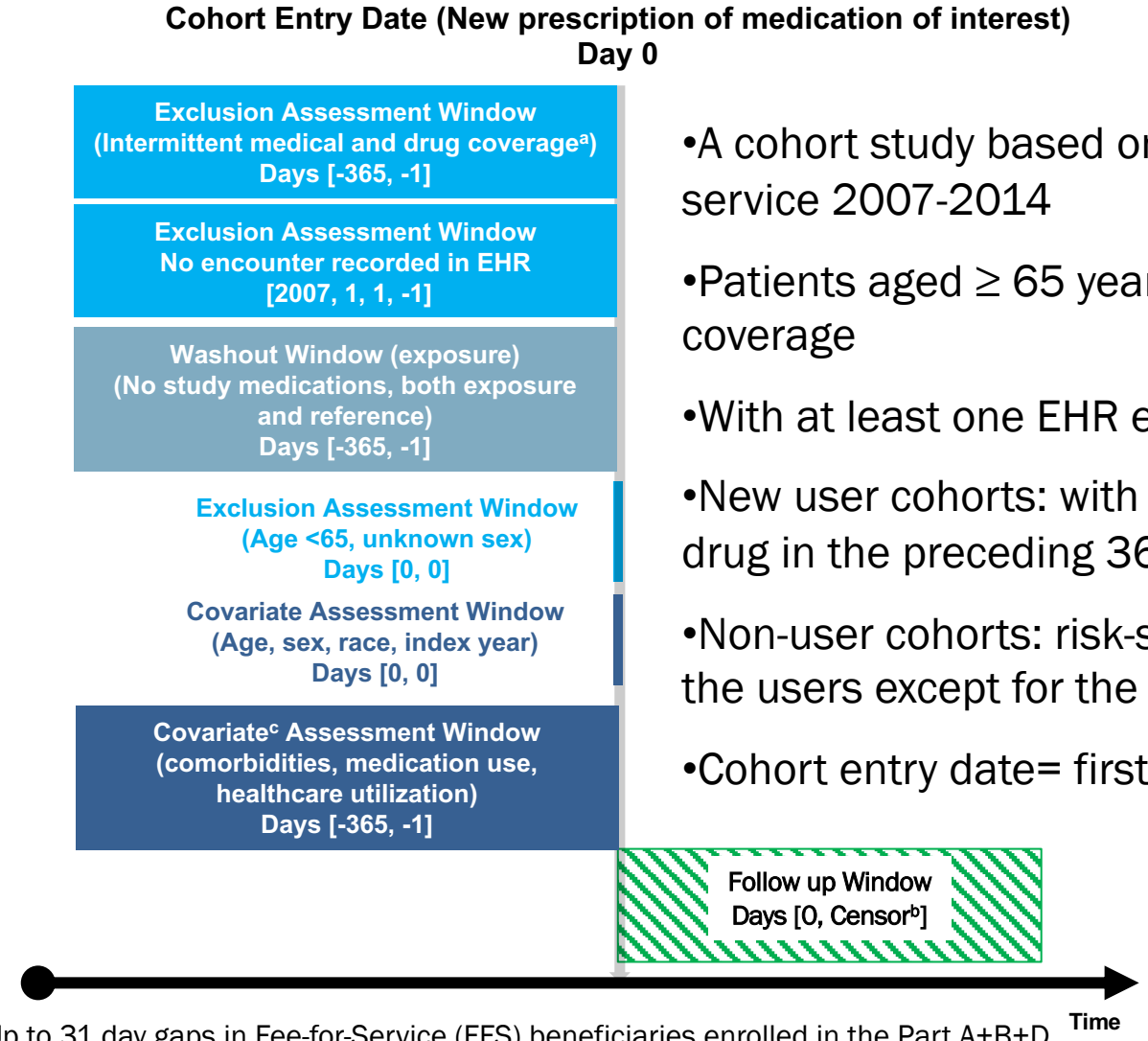
Comparison type	Outcome event	Exposure group	Referent group
Acute medication→ short-term outcomes	Hyperkalemia in 30 days	Bactrim	Cephalexin
Acute medication→ long-term outcome	Clostridium difficile infection (CDI) in 1 year	Bactrim	Cephalexin
Chronic medication→ chronic outcomes, non-use comparison	Pneumonia in 1 year	PPI	Non-PPI
Chronic medication A vs. B→ long-term outcome	Pneumonia in 1 year	PPI	H2RA

PPI= proton pump inhibitor; H2RA: H2 receptor antagonist

The point is not estimating causal effect of these examples but to quantify differences in estimates based on EHR alone vs. that based on EHR-claims data.

Clin Pharmacol Ther. 2022 Jan;111(1):243-251.
BMJ **343**, d5228 (2011)
Archives of internal medicine **170**, 1045-1049 (2010)
Curr Opin Gastroenterol **28**, 1-9 (2012)
JAMA **292**, 1955-1960 (2004)
Intensive Care Med **46**, 1987-2000 (2020)

Study Design

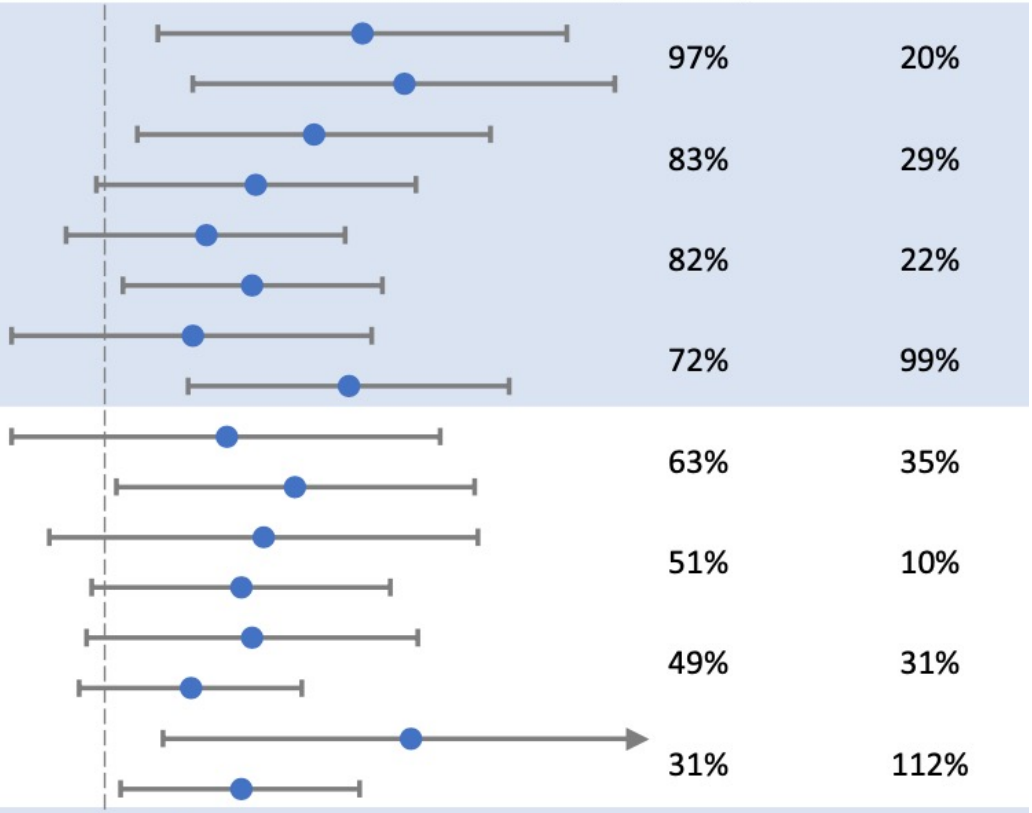


- A cohort study based on administrative claims data from Medicare fee-for-service 2007-2014
- Patients aged ≥ 65 year with at least 365 days continuous Medicare coverage
- With at least one EHR encounter in the baseline after 2007
- New user cohorts: with a dispensing of drug of interest without using the drug in the preceding 365 days
- Non-user cohorts: risk-set sampling of those with the same eligible criteria as the users except for the drug use
- Cohort entry date= first dispensing date or sampling date for the non-users

a. Up to 31 day gaps in Fee-for-Service (FFS) beneficiaries enrolled in the Part A+B+D
b. Earliest of: outcome of interest, death, 365 days after the index date, end of the study period (31 Dec 2014). In in the Claims_based cohorts, additional censoring event is loss of Medicare FFS A or B or D enrollment
c. LASSO-selected 72 baseline covariates adjusted in a COX proportional hazard regression

Comparison between patients with high vs. low EHR-continuity (system 2)

Empirical example	EHR-continuity	Data	Adjusted HR (95% CI)	% event captured by EHR	% bias
A-STO	Top 25%	EHR only	3.11 (1.26, 7.68)	97%	20%
		EHR+ claims	3.74 (1.47, 9.49)		
	Top 25-50%	EHR only	2.51 (1.15, 5.48)	83%	29%
		EHR+ claims	1.94 (0.96, 3.94)		
	Top 50-75%	EHR only	1.56 (0.84, 2.88)	82%	22%
		EHR+ claims	1.91 (1.08, 3.40)		
	Lower 25%	EHR only	1.47 (0.66, 3.24)	72%	99%
		EHR+ claims	2.93 (1.44, 5.95)		
A-LTO	Top 25%	EHR only	1.71 (0.66, 4.39)	63%	35%
		EHR+ claims	2.31 (1.05, 5.11)		
	Top 25-50%	EHR only	2.01 (0.78, 5.18)	51%	10%
		EHR+ claims	1.82 (0.94, 3.52)		
	Top 50-75%	EHR only	1.91 (0.92, 3.97)	49%	31%
		EHR+ claims	1.46 (0.89, 2.38)		
	Lower 25%	EHR only	3.85 (1.29, 11.49)	31%	112%
		EHR+ claims	1.82 (1.08, 3.07)		



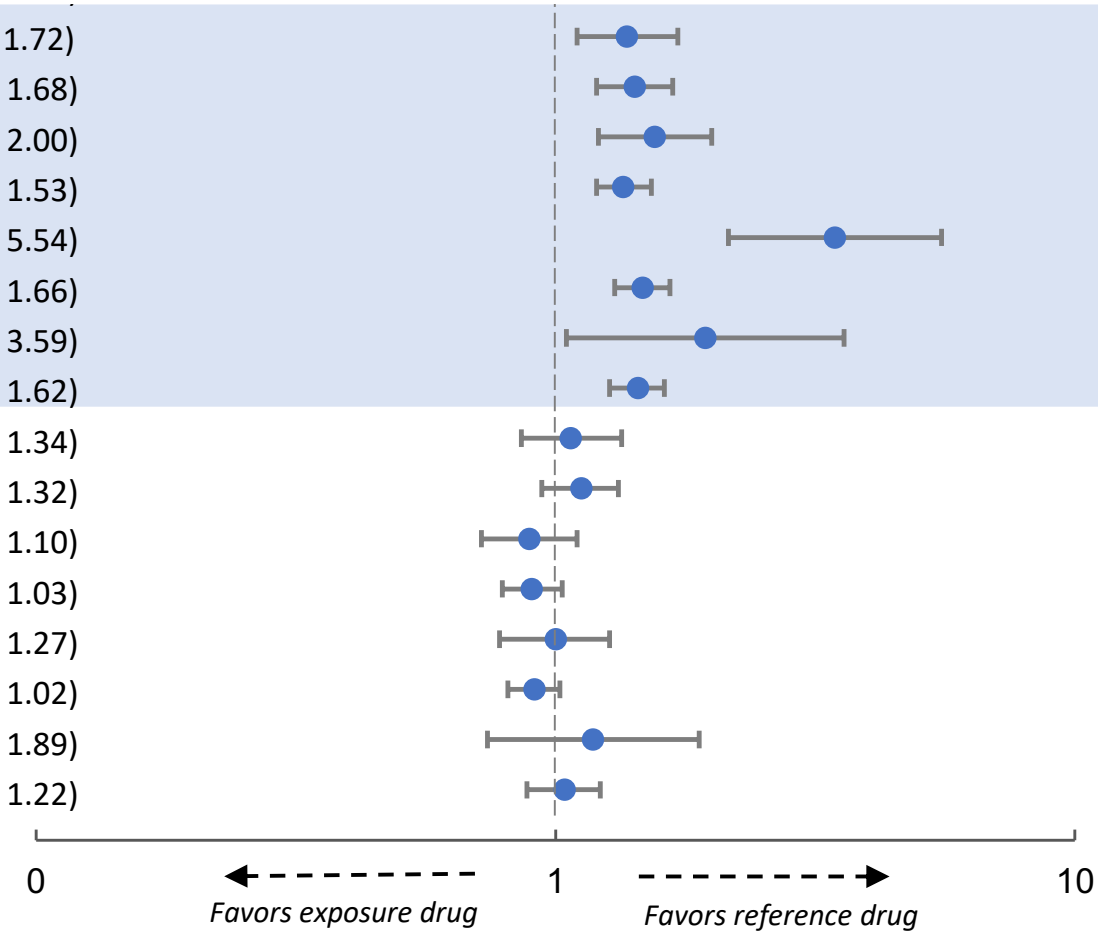
A-STO: acute medication effect on a short-term outcome: Bactrim vs. cephalexin→ 30-day hyperkalemia

A-LTO: acute medication effect on a long-term outcome: Bactrim vs. cephalexin→ 1-year clostridium difficile infection

Clin Pharmacol Ther. 2022 Jan;111(1):243-251.

Comparison between patients with high vs. low EHR-continuity (system 2)

Empirical example	EHR-continuity	Data	Adjusted HR (95% CI)	% event captured by EHR	% bias
GN-LTO	Top 25%	EHR only	1.37(1.10, 1.72)	57%	4%
		EHR+ claims	1.42 (1.20, 1.68)		
	Top 25-50%	EHR only	1.55 (1.21, 2.00)	30%	15%
		EHR+ claims	1.35 (1.20, 1.53)		
	Top 50-75%	EHR only	3.45 (2.15, 5.54)	17%	135%
		EHR+ claims	1.47 (1.30, 1.66)		
	Lower 25%	EHR only	1.94 (1.05, 3.59)	10%	35%
		EHR+ claims	1.44 (1.27, 1.62)		
GG-LTO	Top 25%	EHR only	1.07 (0.86, 1.34)	59%	5%
		EHR+ claims	1.12 (0.94, 1.32)		
	Top 25-50%	EHR only	0.89 (0.72, 1.10)	38%	1%
		EHR+ claims	0.90 (0.79, 1.03)		
	Top 50-75%	EHR only	1.00 (0.78, 1.27)	22%	10%
		EHR+ claims	0.91 (0.81, 1.02)		
	Lower 25%	EHR only	1.18 (0.74, 1.89)	12%	13%
		EHR+ claims	1.04 (0.88, 1.22)		



GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome: proton pump inhibitors (PPI) vs. no PPI→ 1-year pneumonia

GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome: PPI vs. histamine type-2 receptor antagonists → 1-year pneumonia

Clin Pharmacol Ther. 2022 Jan;111(1):243-251.

Treatment effect heterogeneity by EHR-continuity (System 2)

Example	EHR-continuity	HR_{E+C} (95% CI)	Ratio of HR_{E+C} (95% CI) *	p for interaction**
A-STO	Top 25%	3.74 (1.47,9.49)	Ref	ref
	Top 25-50%	1.94 (0.96,3.94)	0.60 (0.20,1.81)	0.3685
	Top 50-75%	1.91 (1.08,3.40)	0.66 (0.24,1.81)	0.4231
	Lower 25%	2.93 (1.44,5.95)	0.91 (0.31,2.70)	0.8687
A-LTO	Top 25%	2.31 (1.05,5.11)	Ref	ref
	Top 25-50%	1.82 (0.94,3.52)	0.80 (0.28,2.24)	0.6663
	Top 50-75%	1.46 (0.89,2.38)	0.61 (0.24,1.55)	0.302
	Lower 25%	1.82 (1.07,3.07)	0.73 (0.28,1.87)	0.5102
GN-LTO	Top 25%	1.42 (1.20,1.68)	Ref	ref
	Top 25-50%	1.35 (1.20,1.53)	0.93 (0.76,1.13)	0.4671
	Top 50-75%	1.47 (1.30,1.66)	0.94 (0.77,1.14)	0.5187
	Lower 25%	1.44 (1.27,1.62)	0.93 (0.76,1.14)	0.4888
GC-LTO	Top 25%	1.12 (0.94,1.32)	Ref	ref
	Top 25-50%	0.90 (0.79,1.03)	0.81 (0.66,0.99)	0.0393
	Top 50-75%	0.91 (0.81,1.02)	0.83 (0.68,1.01)	0.0631
	Lower 25%	1.04 (0.88,1.22)	0.94 (0.75,1.18)	0.5792

E+C= EHR + Claims data
A-STO: comparing the effect of two Antibiotics on a short-term outcome; A-LTO: comparing the effect of two Antibiotics effect on a long-term outcome; GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome; GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome

We did not find evidence of treatment effect heterogeneity by EHR-continuity when results are based on EHR plus claims data.

Clin Pharmacol Ther. 2022 Jan;111(1):243-251.

Discussion

- We observed a trend that the information bias due to EHR-discontinuity appears more pronounced for long-term (e.g., assessed over a year) than short-term outcomes (e.g., evaluated in the first 30 days).
- The information bias due to EHR-discontinuity appears more pronounced for the non-use comparison than an active comparator design: requiring a medication use at cohort entry → more likely that follow-up visits will be observable in the same system
- Patients in the lower 25-50% of predicted EHR continuity have more misclassification in subgroup classification and their treatment estimates tend to have more bias and less precision when compared to estimates based on EHR plus claims data.

Limitations

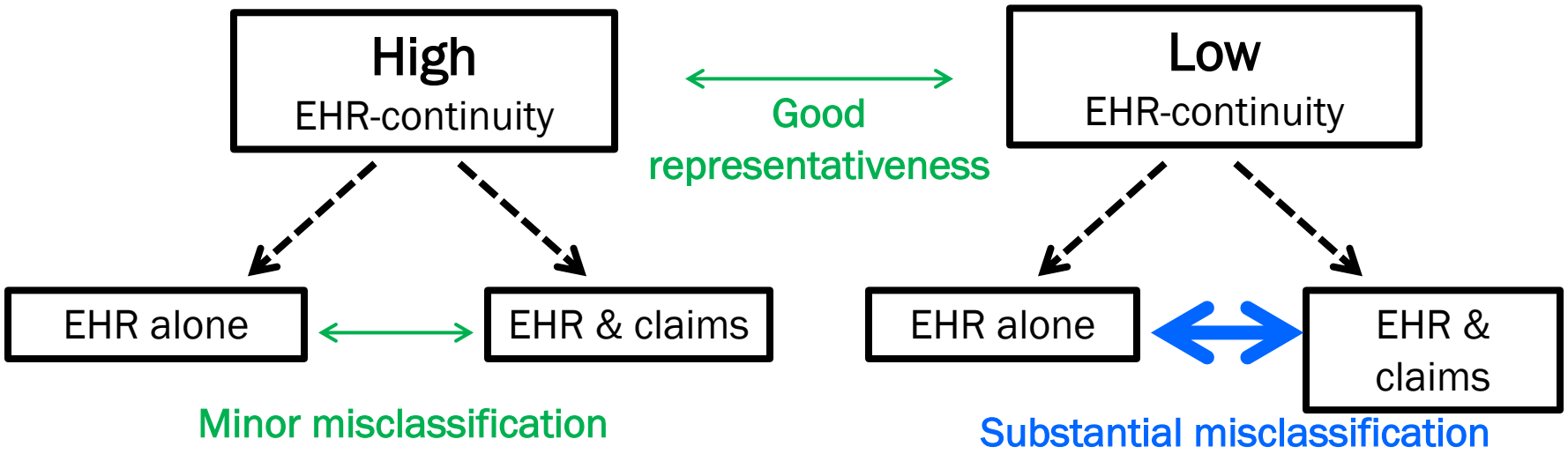
- The impact of EHR-discontinuity on CER is context specific:
  - ➢ Depends on research questions (outcome, exposure, confounders, etc.
  - ➢ Only 4 risk scores and 4 CER examples → Further investigations in a wider range of research questions are needed
- Performance may depend on health system and its EHR penetration
  - ➢ Based on 3 academic EHR systems in MA and NC→ Validation in other types of care delivery systems is needed.

Limitations

- Findings based on older adults >= 65 years→ NOT intended to generalize to younger populations
  - ➢ Validation in the Medicaid is ongoing.
- Generalizability to special population may be limited;
  - ➢ Cancer patients (validation in a oncology population is ongoing)
  - ➢ Pregnant women
  - ➢ Pediatric populations

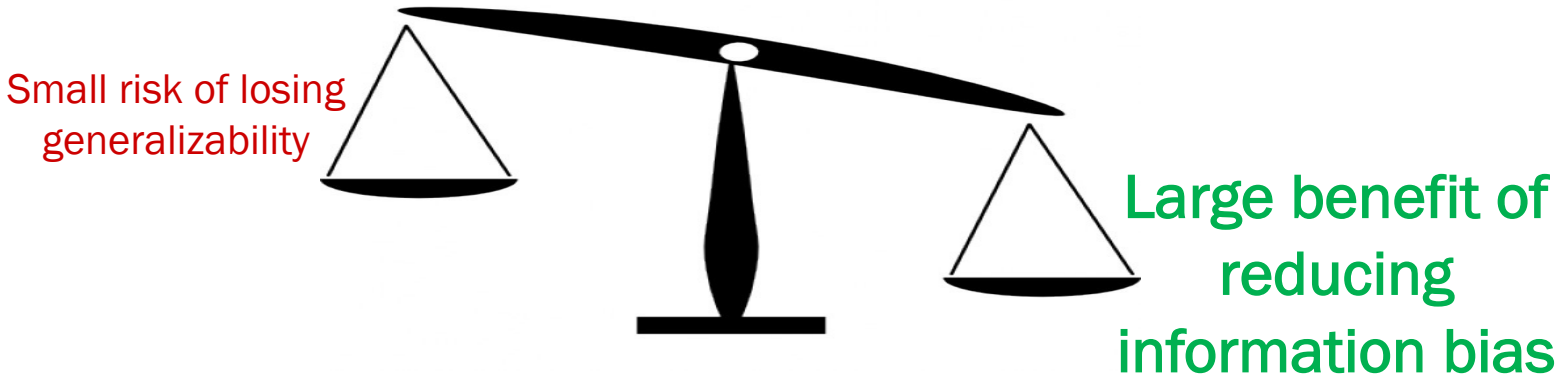
Conclusions

- EHR-continuity is low in majority of patients seen in three US academic EHR systems.
- EHR-discontinuity can lead to substantial amount of information bias.
- Patients with high EHR-continuity were found to have much reduced variable misclassification based on EHR alone with acceptable representativeness.



Conclusions

- Restrict a CER study to patients with high EHR-continuity, which may confer a favorable benefit (reducing information bias) to risk (losing generalizability) ratio.



Acknowledgement

- Sebastian Schneeweiss
- Daniel Singer
- Joshua Gagne
- Shawn Murphy
- Supporting staff/students: Yinzhu Jin/Dave Merola/Jun Liu/Joyce Lii
- Christopher Herrick and the entire RPDR team

Thank you!

Questions?

jklin@bwh.harvard.edu