

Data differences between VHA sites reduce machine learning model transportability.

 **Kimberley ES Cox**, Carlos G Grijalva, Amanda S Mixon, Candace D McNaughton, Michael E Matheny, Thomas A Lasko

Challenges in model transportability between VHA sites

INTRO

Medical machine learning models have performance drops when used outside the healthcare system they were trained in. Does this result hold between sites within the same healthcare system? We tested this in the Veterans Health Administration.

METHODS

Population: Patients [18–101 y.o.] with study site visits in 2015-2019 • **Model:** XGBoost

Predictors: Demographics, meds, labs, vitals, conditions, ED visits

Target: Non-ED outpatient visits/year • **Unit of analysis:** Patient/year

We build models for four single sites using OMOP CDM² data then tested them on data from the other sites.

RESULTS

		Test Site			
		508	528	626	660
Training Site	508	[3.6, 3.7]	[5.1, 5.3]	[3.7, 3.9]	[2.8, 2.9]
	528	[4.7, 4.9]	[5.0, 5.1]	[4.4, 4.8]	[3.3, 3.6]
	626	[3.8, 4.0]	[5.1, 5.3]	[3.3, 3.4]	[2.7, 2.8]
	660	[4.4, 4.6]	[5.5, 5.8]	[4.1, 4.4]	[2.7, 2.7]
Yearly site visits/patient mean (SD) / median (IQR)		7.8 (16.1) / 2 (9)	10.6 (23.8) / 3 (11)	7.6 (16.3) / 2 (9)	5.9 (16.5) / 0 (4)

Values are Mean Absolute Error. 95% Student's t confidence intervals (DOF: 9) derived from 10-fold cross-validation.

DISCUSSION

Even when using data from a single national healthcare system using a common data model, model users should consider cross-site effects on model transportability.

REFERENCES

- Lasko TA, Strobl EV, Stead WW. Why do probabilistic clinical models fail to transport between sites. *npj Digit Med*. 2024 Mar 1;7(1):1–8.
- FitzHenry F, Resnic FS, Robbins SL, Denton J, Nookala L, Meeker D, et al. Creating a Common Data Model for Comparative Effectiveness with the Observational Medical Outcomes Partnership. *Appl Clin Inform*. 2015 Aug 26;6(3):536–47.

SITE KEY

- 508 – Atlanta, Georgia
- 528 – Western New York
- 626 – Tennessee Valley
- 660 – Salt Lake City

CHALLENGES¹

- Different patient populations
- Different data representations
- Different data distributions

SITE DIFFERENCES

VISITS

	508	528	626	660
Valid visits per patient	44	60	42	33
Patients with valid visits	92.1%	93.6%	86.2%	87.8%
Primary care	76.3%	81.8%	72.3%	44.3%
Long-term care	1,294 (0.83%)	7,679 (4.41%)	1,536 (1.11%)	435 (0.38%)
In-patient stay / >90 days	20,772 (13.3%) / 32 (0.02%)	31,637 (18.2%) / 83 (0.02%)	21,294 (15.4%) / 24 (0.02%)	12,400 (10.9%) / 22 (0.02%)
Patient-years >100 visits	2.21%	6.37%	2.34%	2.78%

LABS

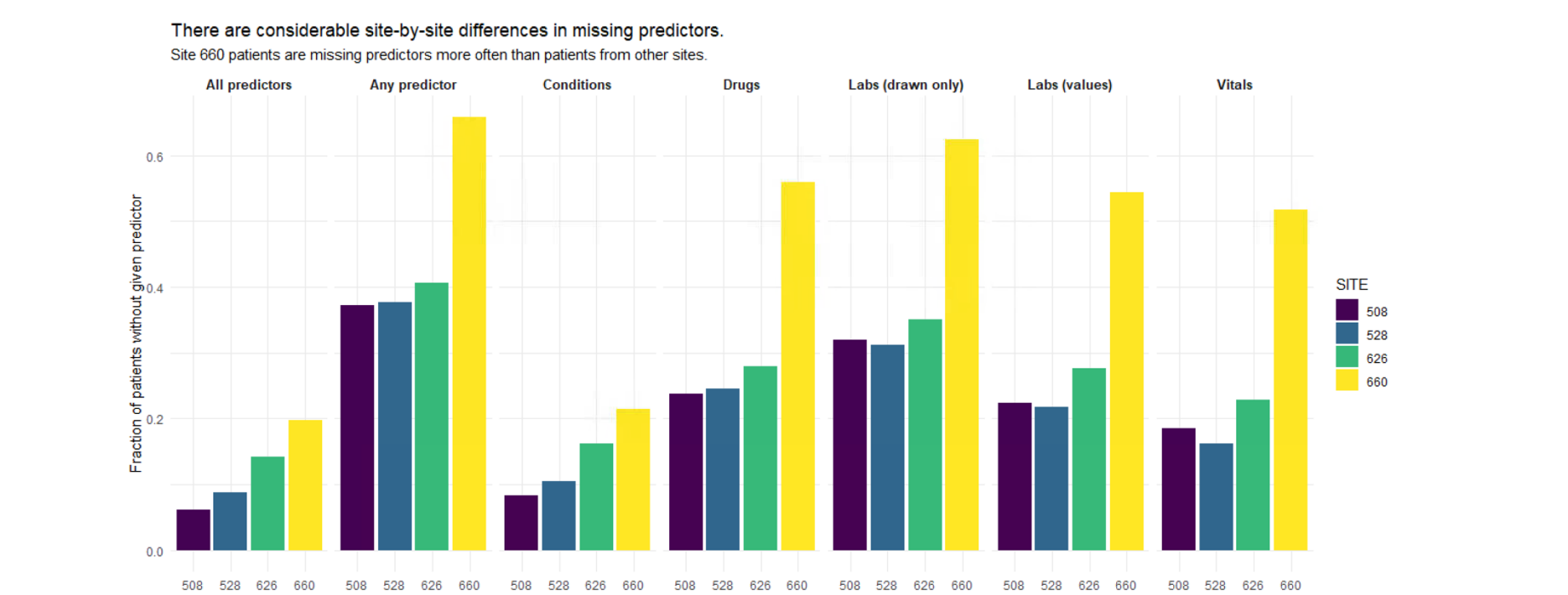
Rank	508	528	626	660
1	Glucose_BSP_mg (2345-7)	Monoclonal_BSP (5905-5)	MCV_BSP (785-6)	Creat_BSP_mg (2160-0)
2	Creat_BSP_mg (2160-0)	Iron_mgdl (5159-1)	K_BSP_mmol (2160-0)	BUN_BSP_mg (3094-0)
3	K_BSP_mmol (2160-0)	Creat_BSP_mg (2160-0)	Creat_BSP_mg (2160-0)	Ca_BSP_mg (17861-6)
4	Glucose_BSP_mg (2345-7)	GFR_Calc_MDRD (77147-7)	Na_BSP_mmol (2951-2)	K_BSP_mmol (2823-3)
5	Na_BSP_mmol (2951-2)	K_BSP_mmol (2823-3)	Glucose_BSP_mg (2345-7)	GFR_Calc_MDRD (77147-7)
6	GFR_Calc_MDRD (48643-3)	BUN_BSP_mg (3094-0)	Ca_BSP_mg (17861-6)	Na_BSP_mmol (2951-2)
7	GFR_Calc_MDRD (48643-3)	Glucose_BSP_mg (2345-7)	BUN_BSP_mg (3094-0)	Glucose_BSP_mg (2345-7)
8	Ca_BSP_mg (17861-6)	Na_BSP_mmol (2951-2)	Cl_BSP_mmol (2075-0)	CO2_BSP_mmol (2028-9)
9	BUN_BSP_mg (3094-0)	Ca_BSP_mg (17861-6)	CO2_BSP_mmol (2028-9)	Cl_BSP_mmol (2075-0)
10	CO2_BSP_mmol (2028-9)	CO2_BSP_mmol (2028-9)	GFR_Calc_MDRD (77147-7)	PK_BSP_ont (777-3)

Glucose: 2345-7; Glucose [Mass/volume] in Blood by Automated test strip • 2345-7; Glucose [Mass/volume] in Blood eGFR: 48643-3; Glomerular filtration rate ... Among non-blacks • 48643-3; GFR ... Among blacks • 77147-7; GFR

DEMOGRAPHICS

	508	528	626	660
Total patients	155,656	174,289	138,230	113,222
Age in 2015 (mean / median)	56 / 57	60 / 64	54 / 59	58 / 61
Sex				
Male	132,348 (85%)	156,193 (89.6%)	122,738 (88.8%)	98,099 (86.6%)
Female	23,308 (15%)	18,096 (10.4%)	15,492 (11.2%)	15,423 (13.4%)
Race				
African	696 (0.4%)	669 (0.37%)	511 (0.37%)	700 (0.62%)
Black	75,341 (48.4%)	15,799 (9.0%)	20,869 (15.1%)	8,037 (7.0%)
Native Am.	527 (0.3%)	460 (0.26%)	428 (0.31%)	1,208 (1.0%)
Pacific Islander	600 (0.3%)	428 (0.24%)	428 (0.31%)	473 (0.4%)
White	67,422 (43.1%)	143,894 (82.6%)	97,161 (70.1%)	94,775 (83.5%)
Unknown	10,874 (6.9%)	14,367 (8.2%)	18,569 (13.4%)	7,309 (6.4%)
Ethnicity				
Hispanic	3,460 (2.2%)	3,492 (2.0%)	3,073 (2.2%)	6,132 (5.4%)
Non-Hispanic	149,221 (91.1%)	169,196 (96.9%)	135,157 (97.8%)	107,093 (94.6%)
Unknown	6,975 (4.4%)	15,555 (8.9%)	9,891 (7.1%)	4,157 (3.7%)

MISSING PREDICTORS



VARIABLE IMPORTANCES

508	528	626	660
MentalDisorder (1.75)	MentalDisorder (2.5)	INR_BldPPP (1.25)	MentalDisorder (0.9)
AcuteDiseaseInjury (0.8)	Non_opioid_analgesics (1.5)	MentalDisorder (1.0)	Weight (0.77)
Creat_BSP_mg (0.8)	VisualDisorder (1.2)	VisualDisorder (0.7)	AcuteDiseaseInjury (0.75)
VisualDisorder (0.76)	AcuteDiseaseInjury (1.0)	Respiration (0.7)	INR_BldPPP (0.73)
Respiration (0.75)	OtherCondition (0.8)	OtherCondition (0.6)	Other_GL_medications (0.62)
Pain (0.55)	Respiration (0.8)	AcuteDiseaseInjury (0.73)	Opioid_analgesics (0.57)
Antidepressants (0.51)	SkinDisorder (0.8)	MuscularDisorder (0.55)	VisualDisorder (0.48)
Opioid_analgesics (0.49)	MuscularDisorder (0.77)	Non_opioid_analgesics (0.55)	OtherCondition (0.4)
OtherCondition (0.48)	NeuroDisorder (0.74)	NeuroDisorder (0.48)	Pain (0.39)
DrugAndSubstance (0.47)	Pain (0.6)	Antidepressants (0.45)	ER_Visits (0.37)

Key: Variable name (SHAP value) – units: outpatient visits/year. Cell color = unique variable.

VANDERBILT UNIVERSITY
MEDICAL CENTER

Department of
Biomedical Informatics

 kimberley.kondratieff@vumc.org

VA U.S. Department of Veterans Affairs

Funded under the research priority to
Put VA Data to Work for Veterans (VA ORD 22-D4V).



Scan for full abstract and
bonus material.