Listening Session 3: Addressing Challenges Regarding Data, Benchmarking, and Risk Adjustment

Presenters:

Subject Matter Experts

- **Robert Saunders, PhD** – Senior Research Director, Health Care Transformation, Adjunct Associate Professor and Core Faculty Member, Duke-Margolis Institute for Health Policy, Duke University
- **Randall P. Ellis, PhD** – Professor, Department of Economics, Boston University
- **Aneesh Chopra, MPP** – President, CareJourney
- **John Supra, MS** – Chief Digital Health & Analytics Officer, Value-Based Care Institute, Cone Health

Listening Session 3: Addressing Challenges Regarding Data, Benchmarking, and Risk Adjustment

Robert Saunders, PhD

Senior Research Director, Health Care Transformation, Adjunct Associate Professor and Core Faculty Member, Duke-Margolis Institute for Health Policy, Duke University

Accelerating Adoption of Accountable Care: Setting Benchmarks & Determining Financial Risk

September 17, 2024

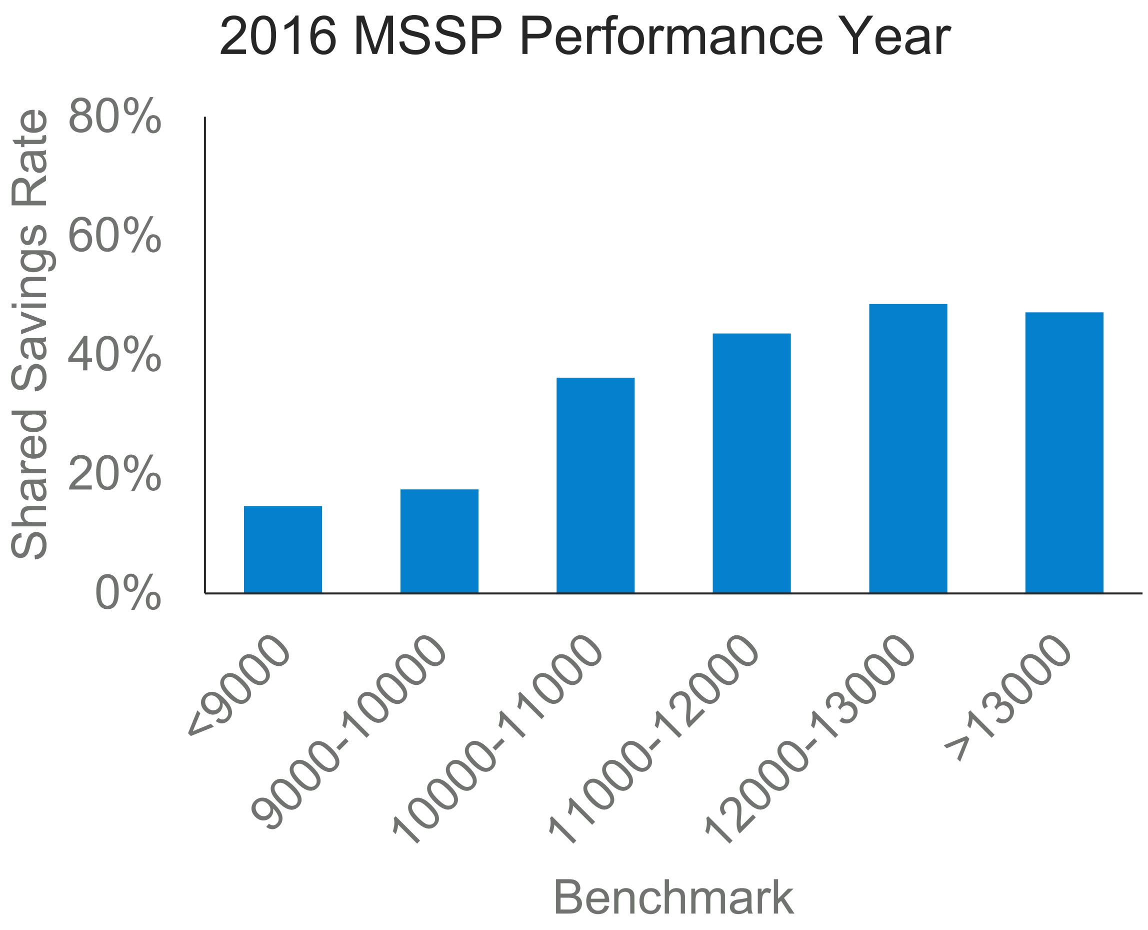
Dr. Rob Saunders, Senior Research Director, Health Care Transformation

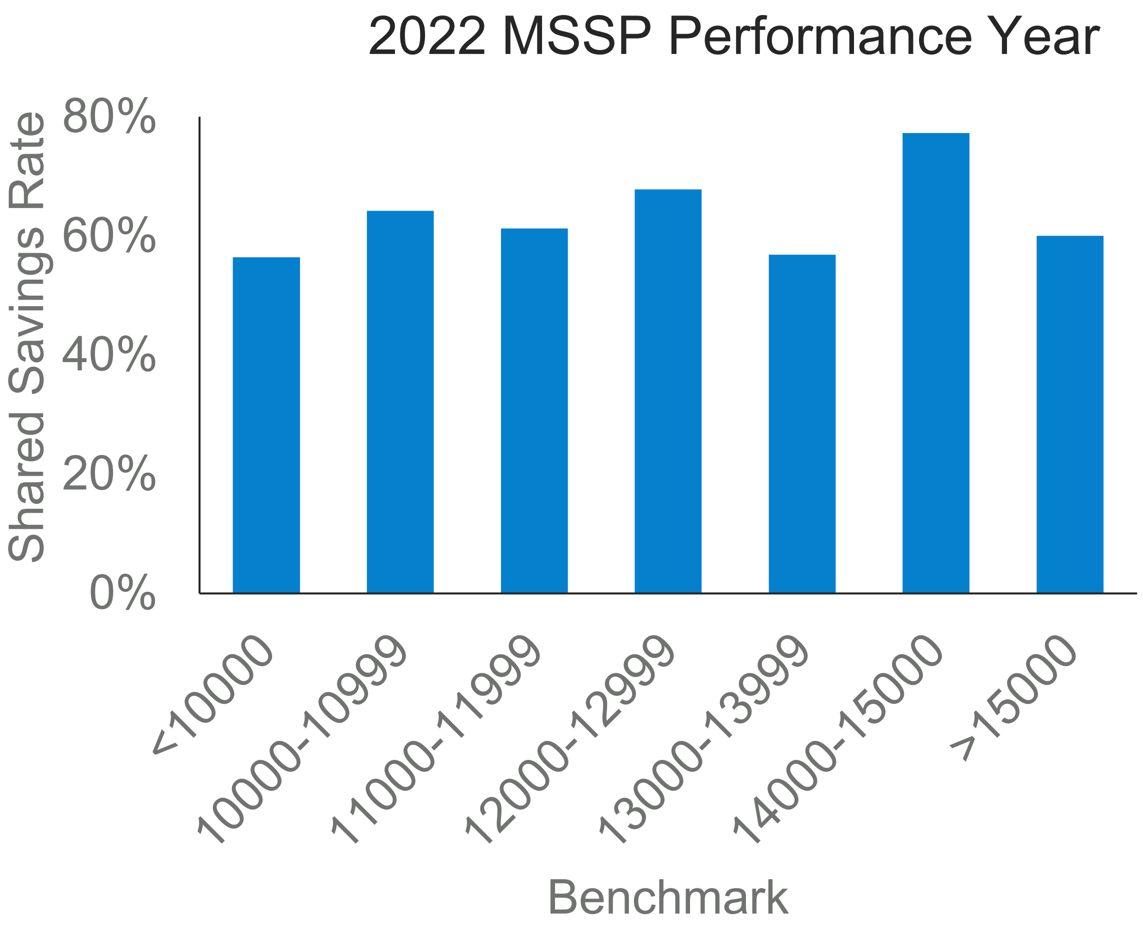
Duke | MARGOLIS INSTITUTE *for* Health Policy

Lessons about benchmarking

- Benchmark is traditionally strong predictor of VBP performance (although less true in recent ACO results).
- Benchmark normally strongly related to long-term participation in ACO and other VBP programs (likely because of shared savings performance).
- Benchmark (and effects of benchmark) varies for different programs and types of organizations (e.g., hospital vs physician-led ACOs, safety net)
- New data and technical approaches can improve benchmarking's accuracy and overall incentives, but there will always be policy tradeoffs.

Benchmark continues to be important, but less of a predictor of shared savings for ACOs

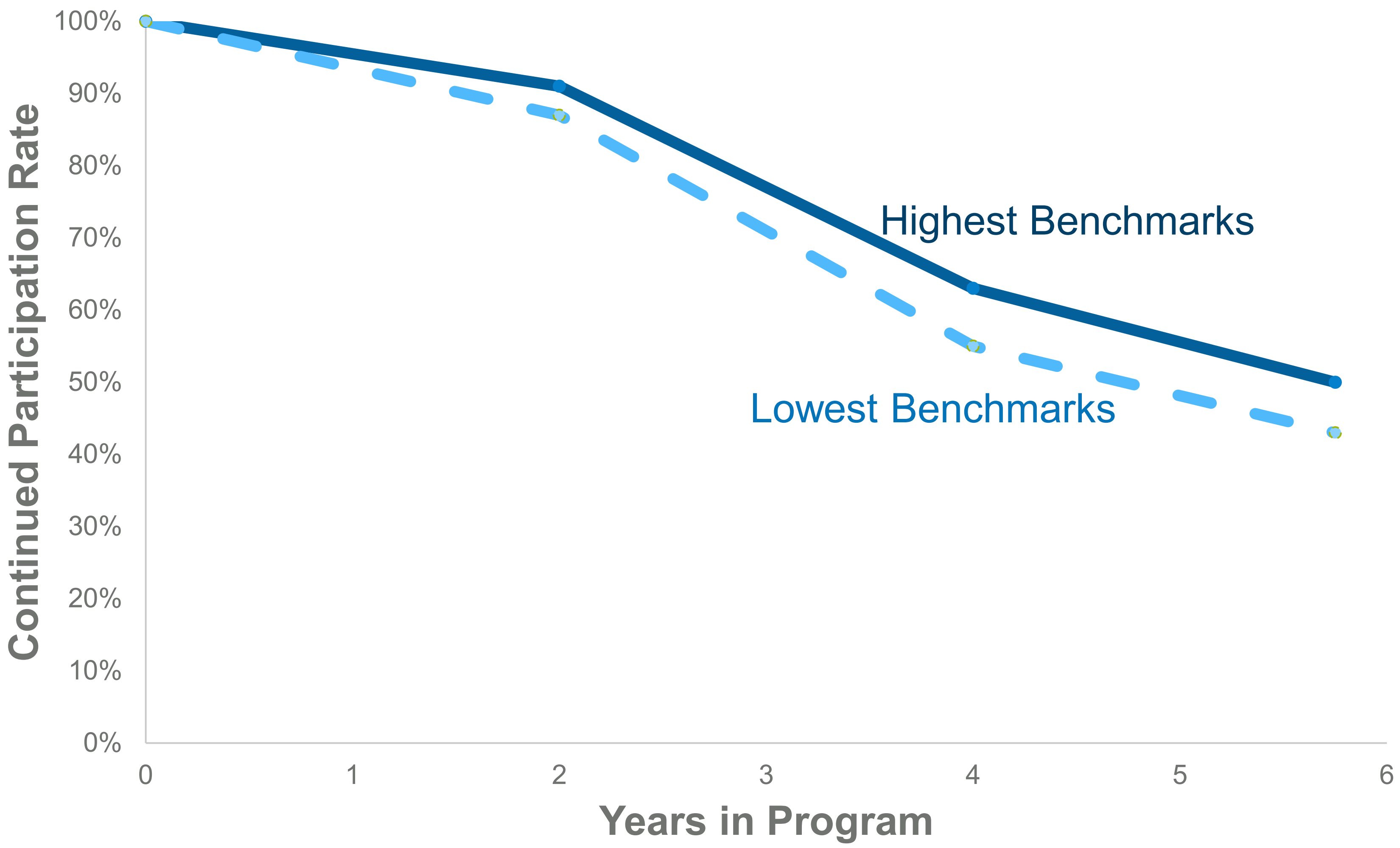




***2022 MSSSP benchmark calculated differently than 2016.

Duke | MARGOLIS INSTITUTE for Health Policy

Research has shown organizations with higher benchmarks more likely to stay in VBP programs



Bleser et al. Why Do Accountable Care Organizations Leave the Medicare Shared Savings Program? Health Affairs 38(5): 794-803. 2019.

Duke | MARGOLIS INSTITUTE for Health Policy

Research highlights different benchmarking challenges for different programs and organizations

- From research, organizations may not join VBP model if benchmark is low/challenging to meet, and organizations may leave VBP models if benchmark rebases/ratchets down over time (and therefore becomes increasingly challenging to meet).
- Benchmarking can affect participation differently for physician-led vs hospital-led ACOs based on benchmarks for their local providers.
- Safety net organizations and similar may not have culture of coding, which can lower their effective benchmark.
- Only some VBP programs account for factors related to medically and socially underserved populations in their benchmark (e.g., ACO REACH's Health Equity Benchmark Adjustment).
- Differences in overall incentives (combination of benchmark, risk adjustment, stop-loss provisions) can make some programs (like MA) more financially sustainable than many VBP programs.

Technical factors and new data can improve benchmarking and overall financial incentives

Domain	Challenges and Opportunities
Data Collection and Quality	• Social factors: Given the nascent collection of SDoH data, there are multiple legal, regulatory, and practical obstacles to better data quality and use. Many heath related social needs screening/referral tools that are not standardized, so payers are challenged with incorporating data and advising on its collection. • Risk adjustment: New approaches could move from self-reported condition coding to leveraging eCQMs, EHR data, and other data that show management of conditions (not just coding).
Capturing Population's True Health Care Needs	• Health equity benchmark adjustments are starting to be used, and early versions (leveraging geographic level ADI) may make VBP financially sustainable for several types of safety net organizations (especially in urban areas). • Seriously ill populations and older, complex patients often not well captured under current risk adjustment (e.g., w/o frailty adjustments) or may be excluded (like through risk truncation).
Organizational Competency	• Benchmarking is one tool for ongoing financial incentives, but upfront capital needed for building organizational capabilities to manage populations and improve care.

Lessons about benchmarking

- Benchmark is traditionally strong predictor of VBP performance (although less true in recent ACO results).
- Benchmark normally strongly related to long-term participation in ACO and other VBP programs (likely because of shared savings performance).
- Benchmark (and effects of benchmark) varies for different programs and types of organizations (e.g., hospital vs physician-led ACOs, safety net)
- New data and technical approaches can improve benchmarking's accuracy and overall incentives, but there will always be policy tradeoffs.

Thank You!

Contact Us

healthpolicy.duke.edu

Subscribe to our monthly newsletter at dukemargolis@duke.edu

1201 Pennsylvania Avenue, NW, Suite 500 • Washington, DC 20004

DC office: 202-621-2800
Durham office: 919-419-2504

Follow Us

DukeMargolis

@dukemargolis

@DukeMargolis

Duke Margolis

Duke-Margolis Institute For Health Policy

Listening Session 3: Addressing Challenges Regarding Data, Benchmarking, and Risk Adjustment

Randall P. Ellis, PhD

Professor, Department of Economics, Boston University

Risk Adjustment (RA) for Population-Based Total Cost of Care (PB-TCOC) Models

Randall P Ellis, Ph.D.

Professor of Economics

Boston University

September 17, 2024

BOSTON UNIVERSITY

My RA Background

- **Co-inventor:** HCC risk adjustment framework used for MA Part C, Part D, and ACA Marketplace, 1987 to 2000
- **PI:** AHRQ-funded a new disease classification system (DXI) 2018-2022
- **Co-creator:** New machine learning algorithm that automates RA formulas, JAMA HF 2024
- **Co-I:** Primary Care Payment Model (PCAL) approved for MassHealth ACOs 2024

BOSTON UNIVERSITY

Specific questions asked to address

1. What are the most appropriate risk-adjustment methods to use for PB-TCOC models?
2. What are the most important concerns to address in order to encourage increased provider participation in PB-TCOC models?
3. How should the optimal risk-adjustment approaches differ for different types of organizations and/or performance measures?

BOSTON UNIVERSITY

#1 Most appropriate risk-adjustment (RA) methods for PB-TCOC models?

- Concurrent models, not prospective
- Use ACA not MA risk equalization
- Use multiple, not one, RA formula, to refine incentives across different dimensions
- Estimate on very large samples
- Use a very detailed diagnosis classification systems that capture distinctions in illness
- Include adjustments for Social Drivers of Health (SDOH)
- Update regularly, including *ex post* adjustments.

BOSTON UNIVERSITY

#2 How to best encourage increased provider participation?

- *Don't make it optional?*
- Build in higher rewards for participating than for not participating (as in Traditional Medicare)
- Minimize administrative burdens on providers
- Tilt payments to favor continuity of care for complex, high-cost patients
  - Avoid overpaying very healthy
  - Appropriately RA for complex, chronic patients (See #1)
  - Reinsurance
  - FFS for prevention and other necessary work
  - Adjust payments when patients change PCPs
- Make performance pay >10% of total, not 1%

BOSTON UNIVERSITY

#3 How to RA different organizations and/or performance measures?

- Calculate RA models for each performance measure
- Use different contracts for organizations
  - E.g., ACA has different formulas for Platinum/Gold/Silver/Bronze
  - E.g., MassHealth uses RA for three ACO contracts
    - Medical care only
    - Medical + OP Behavioral health
    - Medical + IP and OP Behavioral health
  - Medicare Advantage uses 11 RA formulas
    - Community, new enrollee, LTC facility, ESRD, ….
- Use unified DXI+SDOH RA predictors
- Use relevant measures of model performance and fairness
  - Focus on O:E Ratios of observed to expected for outcomes and population subgroups of interest
- Standardize approach, keep it simple

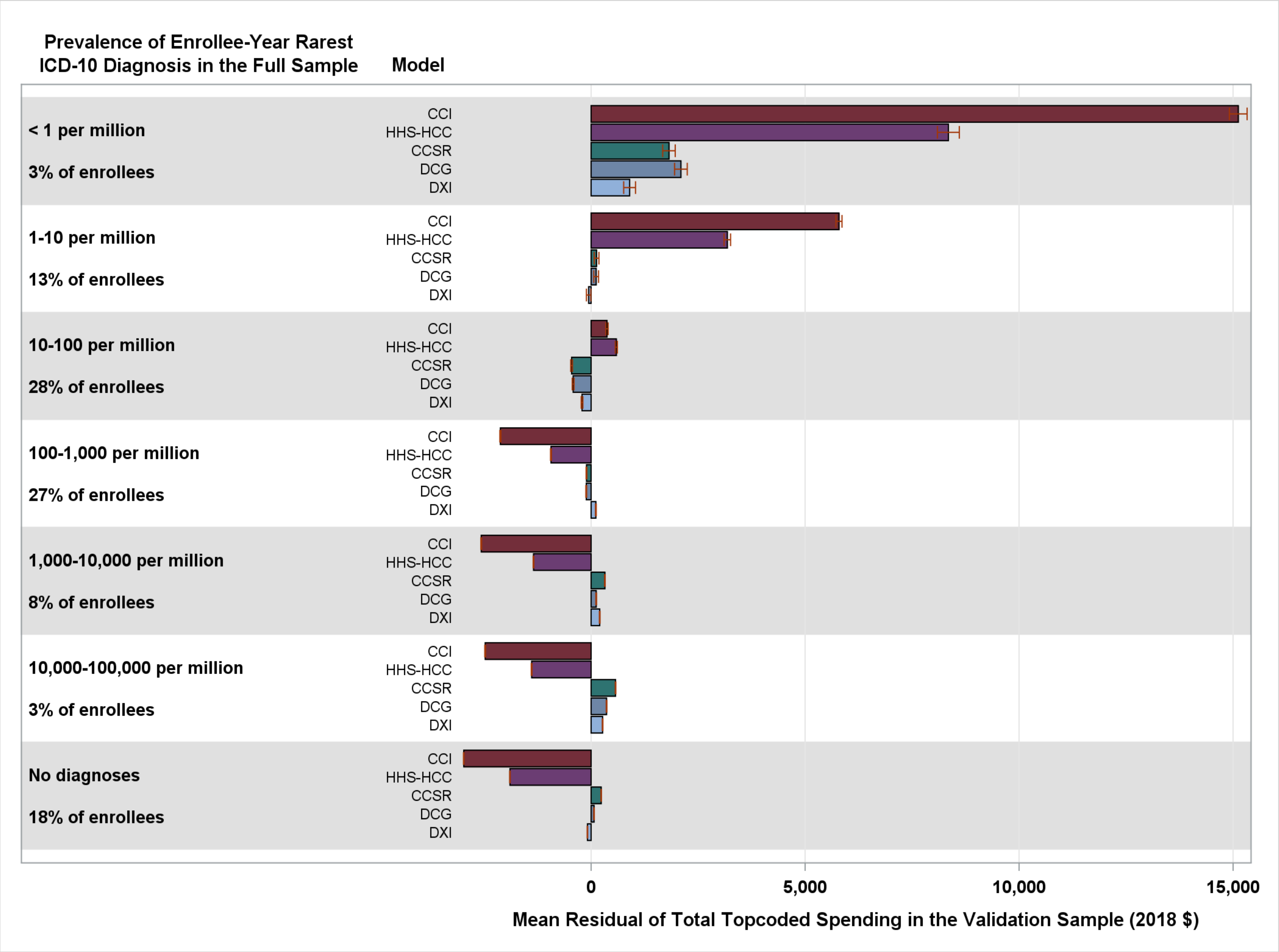
BOSTON UNIVERSITY

Why Use a New RA Framework?

- US HCC system largely unchanged for 20 years
- Growing problems of fraud and gaming
- Current methodology not well-documented
- New ICD-10-CM coding from 2015 not fully used
- Need for flexibility and routine, speedy updates
- Better data, faster computers, better algorithms
- New RA models can do much better on people with multiple conditions, rare diseases, or in special population subsamples.

BOSTON UNIVERSITY

Figure 4: Model Residuals When Enrollees Are Grouped by Their Rarest Diagnosis



CCI=Charlson Comorbidity Index; CCSR = AHRQ Clinical Classification Software Refined; DXI = Diagnostic Items; DCG=Diagnostic Cost Groups

BOSTON UNIVERSITY

Footnote to Figure 4

Notes: CCI is the Quan et al. (2005) version of the Charlson Comorbidity Index, HHS-HCC is the Department of Health and Human Services Hierarchical Condition Category model, CCSR is the Clinical Classifications Software Refined model, DCG is the Diagnostic Costs Groups algorithm, and DXI is the Diagnostic Items model. All models include age-sex dummy variables. We calculated enrollee-weighted mean residuals in the validation sample using the binned frequencies of diagnoses in the full sample, with frequency intervals determined by powers of ten per million. Plot whiskers correspond to 95% confidence intervals, corrected for clustering at the patient level.

BOSTON UNIVERSITY

Relevant References

1. Andriola, Corinne, Randall P Ellis, et al (2024). "A Novel Machine Learning Algorithm for Creating Risk-adjusted Payment Formulas" *JAMA Health Forum*, Apr 5; 5(4):e240625. doi.org/10.1001/jamahealthforum.2024.0625.

2. Ellis, Randall P., Heather E. Hsu, Jeffrey J. Siracuse, et al (2022) "Development and Assessment of a New Framework for Disease Surveillance, Prediction, and Risk Adjustment: The Diagnostic Items Classification System" JAMA Health Forum. vol. 3, no. 3, pp. e220276-e220276. doi.org/10.1001/jamahealthforum.2022.0276

3. Ellis, Randall P, Heather E Hsu, Chenlu Song, et al. (2020) "Diagnostic Category Prevalence in 3 Classification Systems Across the Transition to the International Classification of Diseases, Tenth Revision, Clinical Modification" *JAMA Network Open*. April 8. 3(4), e202280-e202280.

4. Ash, Arlene S., Eric O. Mick,Randall P. Ellis, et al., (2017) Social Determinants of Health in Managed Care Payment Formulas. *Journal of the American Medical Association – Internal Medicine*. Published online August 07, 2017. doi:10.1001/jamainternmed.2017.3317. Supplementary material.

5. Pope, Gregory C,Kautter, John, Ellis, Randall P., Ash, Arlene S., Ayanian, John Z.,. Iezzoni, Lisa I. Ingber, Melvin J., Levy, Jesse M,. and Robst, John (2004) "Risk adjustment of Medicare capitation payments using the CMS-HCC model." Health Care Financing Review. Summer 25(4): 119-141.

6. Chen, Danrong, Corinne Andriola, and Randall P. Ellis, (2024) "Insurance Adjustments to Risk Adjustment Payment Models," Danrong Chen PhD Dissertation Chapter 3, July 2024. (Unpublished)

BOSTON UNIVERSITY

Risk Adjustment (RA) for Population-Based Total Cost of Care (PB-TCOC) Models

Randall P Ellis, Ph.D.

Thank you!

September 17, 2024

BOSTON UNIVERSITY

Risk Adjustment (RA) for Population-Based Total Cost of Care (PB-TCOC) Models

Randall P Ellis, Ph.D.
Department of Economics
Boston University

SUPPLEMENTARY SLIDES

BOSTON UNIVERSITY

Outline

- Background
- Data
- DXI disease classification system
- Primary Care PCAL model
- Machine Learning DCG algorithm
- Conclusions

BOSTON UNIVERSITY

Variables Used for Risk Adjustment

- Age and sex
- Diagnoses
- Pharmaceuticals
- Survey Information
- Eligibility information
  - Health plan
  - Employment

- SOCIAL DRIVERS OF HEALTH
  - Individual-specific or neighborhood information?
  - Education, crowding
  - Chemicals in air, food, water
  - Homelessness
  - Behavioral health/Substance abuse
- Summarized into one Neighborhood Stress Score (Ash et al, 2017, 2024)

BOSTON UNIVERSITY

Risk Adjustment in the News

The New York Times

By Reed Abelson and Margot Sanger-Katz

Published March 22, 2023

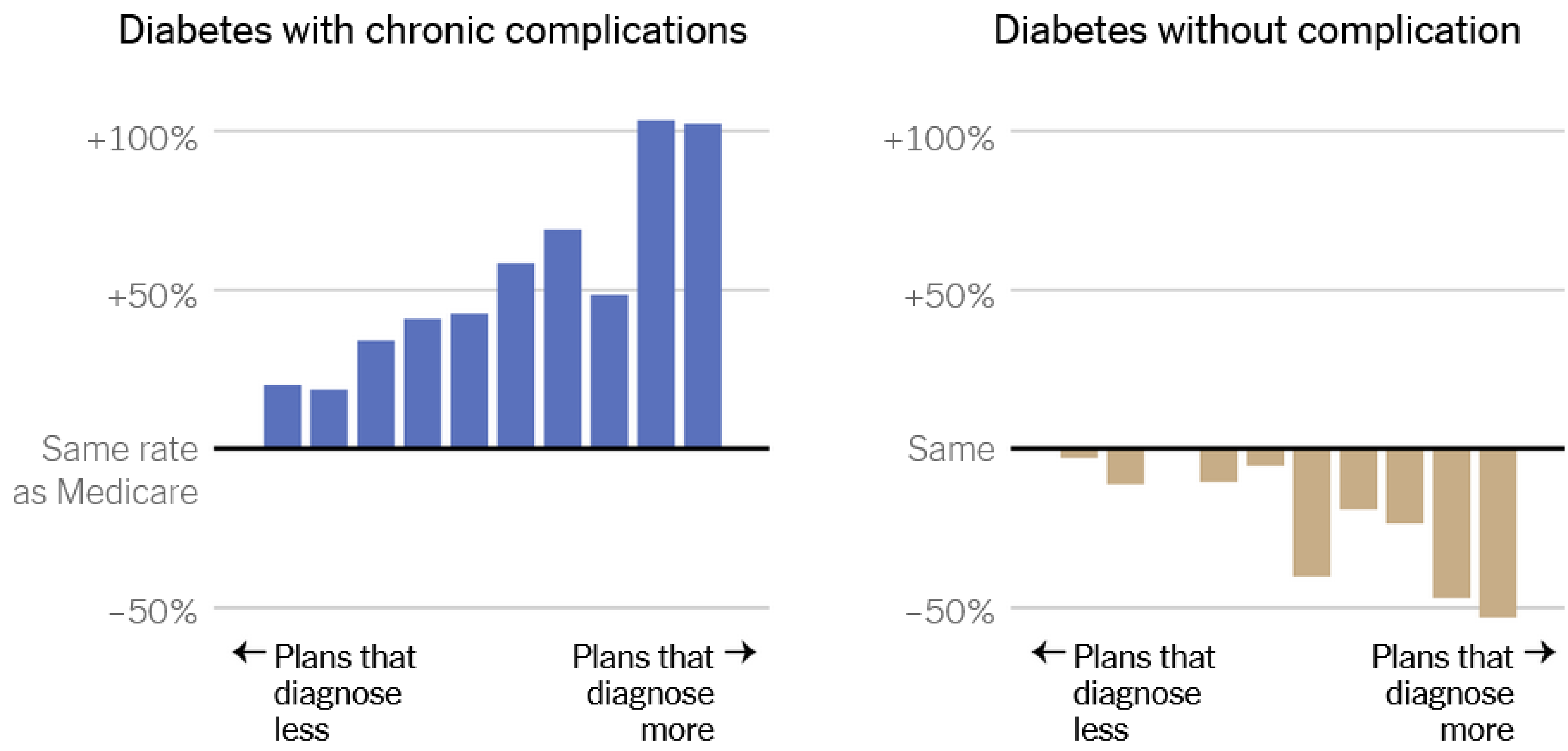
Billions in Medicare Fraud Ignites Lobbying Frenzy

BOSTON UNIVERSITY

The New York Times

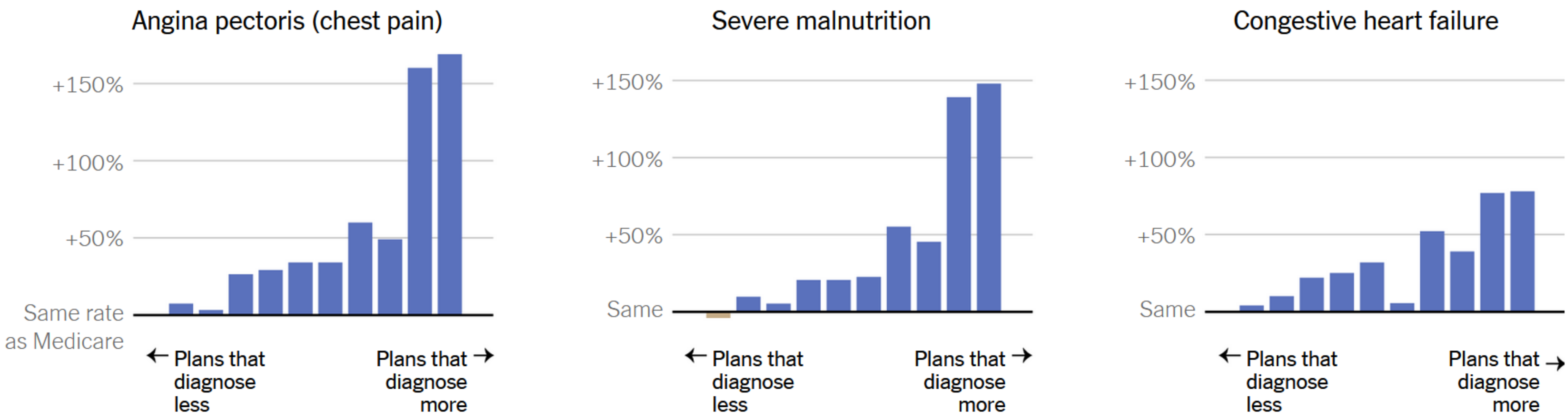
By Reed Abelson and Margot Sanger-Katz

Published March 22, 2023



These Diagnoses Are Much More Common in Medicare Advantage Than Traditional Medicare

Medicare is proposing to remove bonus payments for patients diagnosed with these conditions.



Each bar represents 10 percent of Medicare Advantage contracts, adjusted for enrollment size, sorted by those that diagnose the fewest illnesses to those that code the most. • Source: Medicare Payment Advisory Commission • By Alicia Parlapiano

These three categories of disease are dropped from the payment formula altogether in FY2024.

All types of diabetes are being put in one category.

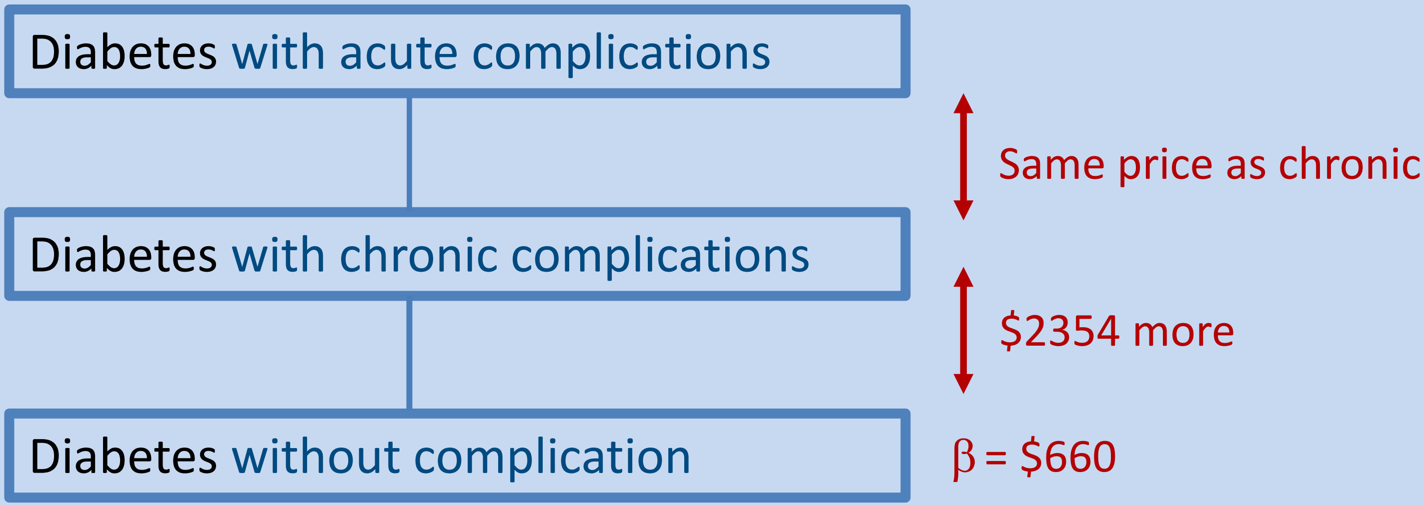
Peripheral artery disease home tests are a big problem

Two questions

- Why is this upcoding happening?
- How can we design better risk adjustment payment formulas?

BOSTON UNIVERSITY

Existing CMS-HCC Diabetes Hierarchy



For 2024: CMS is constraining all diabetes diagnoses to be the same $2148

HCC Payment formula rewards even unspecified type of diabetes with unspecified complications.

Using acute vs chronic vs with complication is vague and highly gameable

Over 400 diagnoses for diabetes with complications in 2020 formula

ICD10	ICD10 Label
E13621	Other specified diabetes mellitus with foot ulcer
E13622	Other specified diabetes mellitus with other skin ulcer
E13628	Other specified diabetes mellitus with other skin complications
E13630	Other specified diabetes mellitus with periodontal disease
E13638	Other specified diabetes mellitus with other oral complications
E13649	Other specified diabetes mellitus with hypoglycemia without coma
E1365	Other specified diabetes mellitus with hyperglycemia
E1369	Other specified diabetes mellitus with other specified complication
E138	Other specified diabetes mellitus with unspecified complications

Modifiers

Only six diagnoses for diabetes without complications

ICD10	ICD10 Label
E089	Diabetes mellitus due to underlying condition without complications
E099	Drug or chemical induced diabetes mellitus without complications
E109	Type 1 diabetes mellitus without complications
E119	Type 2 diabetes mellitus without complications
E139	Other specified diabetes mellitus without complications
Z794	Long term (current) use of insulin

BOSTON UNIVERSITY

For 2024, all types of diabetes are in one plan payment category.

Similar problems with chest pain, severe malnutrition, and congestive heart failure, which were dropped from the payment formula altogether in FY2024.

Existing HCC risk adjustment formula makes it too easy for health plans to change coding patterns to increase revenue.

No easy way to update or change the payment formula to respond to incentives and inequities.

BOSTON UNIVERSITY

DXI/DCG project's contribution

- Rich new classification created 3000 Diagnostic Items (DXIs) with strong clinical foundation
- Two new clinically-derived disease metrics created
- New machine learning (ML) algorithm for variable selection in risk adjustment formulas
  - Computationally feasible on very large samples
  - Transparent and replicable
  - Gaming incentives are mitigated
  - Rare but potentially high-cost conditions incorporated
- Small reduction in model fit from worrying about incentives
- Enormous improvement on payment accuracy for rare diseases
- ML reduces the number of variables by 73%
- Publicly posted software

BOSTON UNIVERSITY

Outline

- Background
- Data
- DXI disease classification system
- Primary Care PCAL model
- Machine Learning DCG algorithm
- Conclusions

BOSTON UNIVERSITY

Table 1 Summary Statistics for diverse outcomes and key demographic variables in Development and Validation Samples, 2016-2018

	Development Sample (N = 59,297,201)		Validation Sample (N = 6,604,259)	
	Mean	Std Dev	Mean	Std Dev
Total Spending	$6,167	$26,405	$6,146	$26,336
Total Spending Topcoded at $250,000	$5,862	$18,159	$5,847	$18,156
Plan Paid	$5,318	$25,888	$5,297	$25,823
Plan Paid Topcoded at $250,000	$5,020	$17,582	$5,006	$17,579
Out-of-pocket (OOP) Spending	$850	$1,485	$849	$1,491
Emergency department (ED) visits	0.24	0.83	0.24	0.84
Inpatient days	0.23	2.78	0.23	2.75
Age	33.93	17.09	33.92	17.09
Female	0.515		0.515	
Months eligible in prediction year	11.34	1.78	11.34	1.78

DXI creation

- Physicians in 20 specialties clustered 94,000 ICD10-CM diagnoses into 3000 Diagnostic Items (DXIs)
- Included all root codes to make compatible with WHO ICD10.
- Base model predicting annualized spending, top-coded at $250,000
- Used data from 2016-2017

BOSTON UNIVERSITY

DXI classification structure

Disease chapters

BLD	Blood
CIR	Circulatory
DIG	Digestive
EAR	Ear
END	Endocrine
EXT	External_causes
EYE	Eye
FAC	Factors_influencing
GEN	Genito-urinary
INF	Infections
INJ	**Injuries**
MAL	Malformations
MBD	Mental_behav_devel
MSK	Muscular_skeletal
NEO	Neoplasm
NVS	Nervous
PNL	Perinatal
PRG	Pregnancy
RSP	Respiratory
SKN	Skin_Connective
SPL	Special
SYM	Symptoms

BOSTON UNIVERSITY

DXI classification structure

Disease chapters

BLD Blood
CIR Circulatory
DIG Digestive
EAR Ear
END Endocrine
EXT External_causes
EYE Eye
FAC Factors_influencing
GEN Genito-urinary
INF Infections
INJ Injuries
MAL Malformations
MBD Mental_behav_devel
MSK Muscular_skeletal
NEO Neoplasm
NVS Nervous
PNL Perinatal
PRG Pregnancy
RSP Respiratory
SKN Skin_Connective
SPL Special
SYM Symptoms

Hierarchies

INJ_Head_neck_eye
INJ_Thoracic
INJ_Abdominal
INJ_Spine_back
INJ_Fracture
INJ_Minor
INJ_Foreign_body
INJ_Burn
INJ_Frostbite_hypotherm
INJ_Poisoning
INJ_Abuse
INJ_Allergies
INJ_Complic
INJ_Nerves
INJ_Traumatic_injuries
INJ_Vascular
INJ_Self_harm
INJ_Vague

BOSTON UNIVERSITY

DXI classification structure

Disease chapters

- BLD Blood
- CIR Circulatory
- DIG Digestive
- EAR Ear
- END Endocrine
- EXT External_causes
- EYE Eye
- FAC Factors_influencing
- GEN Genito-urinary
- INF Infections
- **INJ Injuries**
- MAL Malformations
- MBD Mental_behav_devel
- MSK Muscular_skeletal
- NEO Neoplasm
- NVS Nervous
- PNL Perinatal
- PRG Pregnancy
- RSP Respiratory
- SKN Skin_Connective
- SPL Special
- SYM Symptoms

Hierarchies

- INJ_Head_neck_eye
- INJ_Thoracic
- INJ_Abdominal
- **INJ_Spine_back**
- INJ_Fracture
- INJ_Minor
- INJ_Foreign_body
- INJ_Burn
- INJ_Frostbite_hypotherm
- INJ_Poisoning
- INJ_Abuse
- INJ_Allergies
- INJ_Complic
- INJ_Nerves
- INJ_Traumatic_injuries
- INJ_Vascular
- INJ_Self_harm
- *INJ_Vague*

DXI main effects

- Concussion
- Dislocation
- Fracture_oth
- Fracture_spondylolysis_and_spondylolisthesis
- Fracture_stable_burst
- Injury_nerves
- Lesion_spinal_cord

DXI classification structure

Disease chapters

BLD Blood
CIR Circulatory
DIG Digestive
EAR Ear
END Endocrine
EXT External_causes
EYE Eye
FAC Factors_influencing
GEN Genito-urinary
INF Infections
INJ Injuries
MAL Malformations
MBD Mental_behav_devel
MSK Muscular_skeletal
NEO Neoplasm
NVS Nervous
PNL Perinatal
PRG Pregnancy
RSP Respiratory
SKN Skin_Connective
SPL Special
SYM Symptoms

Hierarchies

INJ_Head_neck_eye
INJ_Thoracic
INJ_Abdominal
INJ_Spine_back
INJ_Fracture
INJ_Minor
INJ_Foreign_body
INJ_Burn
INJ_Frostbite_hypotherm
INJ_Poisoning
INJ_Abuse
INJ_Allergies
INJ_Complic
INJ_Nerves
INJ_Traumatic_injuries
INJ_Vascular
INJ_Self_harm
INJ_Vague

DXI main effects

Concussion
Dislocation
Fracture_oth
Fracture_spondylolysis_and_spondylolisthesis
Fracture_stable_burst
Injury_nerves
Lesion_spinal_cord

CCSR

Dislocations, initial encounter
Dislocations, subsequent encounter
Fracture of the spine and back, initial encount
Fracture of the spine and back, subseq encount
Spinal cord injury (SCI), initial encount
Spinal cord injury (SCI), subseq encount

BOSTON UNIVERSITY

DXI Structure for Diabetes

Disease chapters

ENDocrine

Hierarchies

Diabetes

- Abn_glucose_in_preg_chldbrth_and_puerperium
- Diabetes_gestational_in_preg_chldbrth_and_puerperium
- Diabetes_mellitus_drug_or_chemical_induced
- Diabetes_mellitus_oth
- Diabetes_mellitus_secondary
- Diabetes_mellitus_Type_1
- Diabetes_mellitus_Type_2
- Diabetes_pre-existing_in_preg_chldbrth_and_puerperium
- Diabetes_unsp_in_preg_chldbrth_and_puerperium
- Postprocedural_hypoinsulinemia
- Stable_prolif_diabetic_retinopathy

BOSTON UNIVERSITY

DXI Structure for Diabetes

Disease chapters

ENDocrine

Hierarchies

END_DM_Type_1
END_DM_Type_2
END_DM_Drug_Chem
END_DM_Other

DXI1 Main effects

Abn_glucose_preg_chldbrth_puerperium
Diabetes_gestational
DM_secondary
DM_pre-existing_preg_chldbrth_puerp
DM_unsp_in_preg_chldbrth_and_puerp
Postprocedural_hypoinsulinemia
Stable_prolif_diabetic_retinopathy

DXI2 Modifiers

right
left
Bilateral

diet_controlled
insulin_controlled
controlled_by_oral_hypoglyc_drugs_preg
unsp_control

w_coma
w_hyperglycemia
w_ketoacidosis
w_kidney_complications
w_neurological_manifestations
w_ophthalmic complications
w_periperhal_circulatory_manifestations

intraop_and_postproc
moderate

BOSTON UNIVERSITY

DXI/CCSR classification structure

Disease chapters

BLD Blood
CIR Circulatory
DIG Digestive
EAR Ear
END Endocrine
EXT External_causes
EYE Eye
FAC Factors_influencing
GEN Genito-urinary
INF Infections
INJ Injuries
MAL Malformations
MBD Mental_behav_devel
MSK Muscular_skeletal
NEO Neoplasm
NVS Nervous
PNL Perinatal
PRG Pregnancy
RSP Respiratory
SKN Skin_Connective
SPL Special
SYM Symptoms

Hierarchies

MBD_Anxiety
MBD_Dementia
MBD_Eating_Disorder
MBD_Gender_Sexuality
MBD_Mood_Disorder
MBD_Neuro_Physio_Develop
MBD_Personality_Behavioral_Other
MBD_Psychosis
MBD_Schizophrenia
MBD_Sleep
MBD_Stress_Trauma
MBD_Substance_Abuse
MBD_Suicide
MBD_Symptoms
MBD_Anxiety
MBD_Dementia
MBD_Eating_Disorder
MBD_Gender_Sexuality

DXIs

Alcohol-related_disorders
Opioid-related_disorders
Cannabis-related_disorders
Sedative-related disorders
Stimulant-related_disorders
Hallucinogen-related_disorders
Inhalant-related_disorders
Tobacco-related_disorders
Other_specified substance-related_disorders

BOSTON UNIVERSITY

Diagnostic Items

- Project physicians mapped all 94k diagnoses (with illegal roots) into Diagnostic Items (DXI)
- Current version 1.5 has 3407 DXIs
  - 2446 DXI1 main effects
  - 961 DXI2 modifiers
  - 17 DXI3 continuous scale measures

BOSTON UNIVERSITY

Table 2 R-Square in the Validation Sample

	Age-sex OLS	Age-sex + HCC OLS	Age-Sex + CCSR OLS	Age-Sex + CCSR + DXI OLS	Age-Sex + CCSR + DXI Stepwise
Annualized Total Spending	1.37%	34.75%	40.16%	47.35%	47.35%
Annualized Total Spending Topcoded at $250,000	2.54%	42.68%	52.14%	56.98%	56.98%
Annualized Plan Paid	1.18%	33.91%	38.94%	46.30%	46.30%
Annualized Plan Paid Topcoded at $250,000	2.21%	41.99%	50.84%	55.80%	55.79%
Annualized Out-of-pocket (OOP) Spending	3.95%	18.86%	30.66%	32.49%	32.48%
N =6,604,259					
Number of explanatory variables	30	166	569	3,015	(2,079 to 2,061)

Note: All dependent variables were annualized and then weighted by the fraction of the year each enrollee is eligible to reflect values per annual period.[4] Models were estimated using the development sample with N=59,297,201, These validation sample measures use N =6,604,259.

Overfitting has a Minor Impact on R-Square in Development and Validation Samples

	Age-sex OLS	Age-sex + HCC OLS	Age-Sex + CCSR OLS	Age-Sex + CCSR + DXI OLS	Validation Age-Sex + CCSR + DXI Stepwise	Develop-ment sample, Same model
Annualized Total Spending	1.37%	34.75%	40.16%	47.35%	47.35%	47.79
Annualized Total Spending Topcoded at $250,000	**2.54%**	**42.68%**	**52.14%**	**56.98%**	**56.98%**	**57.19**
Annualized Plan Paid	1.18%	33.91%	38.94%	46.30%	46.30%	46.74
Annualized Plan Paid Topcoded at $250,000	2.21%	41.99%	50.84%	55.80%	55.79%	56.01
Annualized Out-of-pocket (OOP) Spending	3.95%	18.86%	30.66%	32.49%	32.48%	32.79
N =6,604,259						
Number of explanatory variables	30	166	569	3,015	(2,079 to 2,061)	(2,079 to 2,061)

Note: All dependent variables were annualized and then weighted by the fraction of the year each enrollee is eligible to reflect values per annual period.[4] Models were estimated using the development sample with N=59,297,201. Validation sample measures use N =6,604,259.

Table 3 Comparison of HCC, CCSR, and DXI Classification Systems

WHO Chapter	ICD Code Range	Labels	Billable ICD-10	HHS-HCCs	CCSR Categories	Total DXI_1 Items	Billable Diagnoses per DXI	Statistically Sign. DXI_1 in Model 1.2
1	A00-B99	Infectious and parasitic diseases	1,058	5	12	110	10	73
2	C00-D49	Neoplasms	1,661	6	74	145	11	100
3	D50-D89	Blood, blood-forming organs and immunity diseases	247	9	10	47	5	39
4	E00-E89	Endocrine, nutritional and metabolic diseases	908	9	17	85	11	71
5	F01-F99	Mental, behavioral and neurodevelopmental disorders	747	9	32	153	5	127
6	G00-G99	Nervous system diseases	622	13	22	129	5	96
7	H00-H59	Eye and adnexa diseases	2,606	0	12	270	10	141
8	H60-H95	Ear and mastoid process diseases	656	0	6	38	17	30
9	I00-I99	Circulatory system diseases	1,350	11	39	95	14	84
10	J00-J99	Respiratory system diseases	341	4	17	64	5	56
11	K00-K95	Digestive system diseases	799	9	25	107	7	86
12	L00-L99	Skin and subcutaneous tissue diseases	845	1	7	93	9	55
13	M00-M99	Musculoskeletal system and connective tissue diseases	6,487	6	38	222	29	182
14	N00-N99	Genitourinary system diseases	669	3	26	109	6	90
15	O00-O9A	Pregnancy, childbirth and the puerperium	2,267	6	30	136	17	86
16	P00-P96	Perinatal period conditions	443	0	15	54	8	44
17	Q00-Q99	Congenital malformations, deformations, chromosomal abnormalities	817	3	10	26	31	23
18	R00-R99	Symptoms, signs and abnormal clinical and lab findings	720	2	17	178	4	134
19	S00-T88	Injury, poisoning and other consequences of external causes	40,570	7	76	176	231	118
	U00-U99	Emergency code additions	3	0	0	3	1	
20	V00-Y99	Factors influencing health status and health service contacts	6,865	0	30	31	221	11
21	Z00-Z99	Infectious and parasitic diseases	1,253	11	25	175	7	135
		Totals	**71,934**	**114**	**540**	**2,446**	**29**	**1,781**

Outline

- Background
- Data
- DXI disease classification system
- Primary Care PCAL model
- Machine Learning DCG algorithm
- Conclusions

BOSTON UNIVERSITY

Primary care is underpaid and underprovided in the US

- FFS payment rewards treatment, not prevention
- Higher fees for surgery and imaging than primary care
- Too few primary care providers (PCP)
- Specialists collude to keep PCP fees low
- Capitation may make this worse because it does not pay for extra burdens of complex patients
- Patient needs depend not only on diseases but Social Drivers of Health (SDoH)

Ash and Ellis work on PCAL

1. Arlene S. Ash, Matthew J. Alcusky, Ellis et al. Fixing Primary Care Through Payment Reform, May, 2024. Working paper: *Do not cite/quote*.
2. Alcusky MJ, Mick EO, Allison JJ, et al. Paying for Medical and Social Complexity in Massachusetts Medicaid. JAMA Netw Open 2023;6:e2332173. doi: 10.1001/jamanetworkopen.2023.32173. PMID: 37669052.
3. Vats, Sonal, Arlene S. Ash and Randall P. Ellis. (2013) "Bending the Cost Curve? Results from a Comprehensive Primary Care Payment Pilot." *Medical Care*. 51(11):964-969. DOI: 10.1097/MLR.0b013e3182a97bdc
4. Ash, Arlene, and Randall P. Ellis (2012) “Risk-Adjusted Payment and Performance Assessment for Primary Care.” *Medical Care*. Aug. 2012 50(8):643–53 doi: 10.1097/MLR.0b013e3182549c74
5. Ellis, Randall P., and Arlene S. Ash (2012) “Payments in Support of Effective Primary Care for Chronic Conditions.” *Nordic Economic Policy Review*. 2:191-210.https://blogs.bu.edu/ellisrp/files/2012/08/2012-EllisAsh_PracticalPCMHPayment_NEPR_20120815.pdf

BOSTON UNIVERSITY

Problem #1: Many primary care services are poorly paid using FFS payments

- Prevention services
- Follow-up care to hospital, ED, specialists, prescription drugs
- Counseling, behavioral health
- End-of-life planning, hospice care
- Group therapy
- Email, telemedicine, remote care
- Physician referrals
- Much more

BOSTON UNIVERSITY

Problem #2: Additional burdens of complex patients badly captured by their diagnoses

- Homelessness
- Low education
- Language barriers
- Race and discrimination
- Environmental factors: air, water, food, insects, crime, illegal drugs
- Behavioral problems, family stress, scarcity
- Low income

BOSTON UNIVERSITY

Problem #3: Approach assumes each enrollee can be assigned to an identified PCP

- People move around between different facilities and doctors
- Need to assign based on prior year data
- What to do with new enrollees?
- What to do when patients go outside of their assigned network of providers?
- What to do if a specialist acts as the PCP?
- What to do when patients do not see any PCP?

BOSTON UNIVERSITY

Ash et al PCAL Solution

- Use a separate payment formula for Primary Care called the Primary Care ~~Burden~~ Activity Level (PCAL)
- Recognize and pay for all diagnoses relevant
- Recognize Social Drivers of Health (SDoH)
- Come up estimated resources needed to treat complex patients using Proxies of spending on hospitals, drugs, specialists, and emergency departments
- Use principal components to collapse multiple SDoH into a single index to avoid overfitting
- Validate model predictions by PCP review of credibility
- Choose a PCP assignment that accommodates PCP switchers but rewards keeping the same one.

BOSTON UNIVERSITY

What is in PCAL? Dependent variable

Infer the extra primary care resources needed to *prevent or manage* the other kinds of care episodes likely to experience.

The model sums:

- All primary care service costs
- Fractions of the dollars spent on other services, such as
  - Specialty care
  - Hospital care
  - ED care
  - Prescription drugs

Principle: Members expected to incur these other health care costs may be likely to need more attention from their primary care teams

What is in PCAL? Dependent variable

Type of Activity	% of All Such Costs Contributing to Constructed PCAL	% of PCAL
Primary care activities	100%	64%
Specialty care related	6%	20%
Hospital care	6%	1%
ED spending	30%	5%
Rx spending	9%	10%

Note: Spending on services also subject to a maximum and minimum.

BOSTON UNIVERSITY

What is in PCAL? Independent variables of MassHealth SDH 3.2

PCAL 2022 Model		
taPCAL_22	**Coef.**	**t**
RxCG RRS	288	262
RxCG-spline-5	-233	-147
DxCG RRS	523	419
DxCG-spline-5	-124	-65
DxCG-spline-20	-288	-61
Serious Mental Illness	180	56
Opioid Use Disorder	400	82
Alcohol Use Disorder	117	22
Other Substance Use Disorder	224	41
Serious Emotional Disturbance	54	10
Other Disabled	50	14
DDS (not DMH)	557	79
DMH Client	1115	94
NSS7+ X DxCG	5	
Rural Area	22	5
Housing Problems x DxCG x BH	41	90

PCAL independent variables / model design

- Largely consistent w/ SDH3.2, with three important differences:
  - Non-linear relationship with DxCG and RxCG ("spline"): moving from "healthy" to "sick" leads to larger revenue increase than moving from "sick" to "very sick"*
  - Fewer BH/SUD variables: better correlated with PCAL
  - Several variables excluded (e.g., newborn complexity) that lacked clear correlation with PCAL
- Overall model concurrent R-squared is 68.8%

* The RxCG and DxCG scores have a declining slope; higher scores will increase the coefficient but at a lower rate. For example, DxCG scores up to 5 will get $523 per DxCG point. If a member has a DxCG score of 10, their coefficient will be (523*4) + ((523-124)*6) . These "splines" are for better fit with the data

BOSTON UNIVERSITY

Data used to capture SDoH

Table 1. MassHealth Members, Age 0-64 years, 2019 (1,014,625 person-years)		
Member Characteristics	**Person-years**	**%**
Age 18 years and younger	512,955	46.9
Female	579,891	53.1
Housing Problems		
Homeless	21,010	1.9
Unstably Housed (not homeless)	99,495	9.1
Neither of the above	972,242	89.0
Disability Status		
DMH Client	6,919	0.6
DDS Client (not DMH)	19,441	1.8
Other Disability	108,791	10.0
None of the above	957,596	87.6

Data used to capture SDoH - 2

Table 1. MassHealth Members, Age 0-64 years, 2019 (1,014,625 person-years)		
Member Characteristics	**Person-years**	**%**
Behavioral Health Comorbidity		
Serious Mental Illness	145,440	13.3
Opioid Use Disorder	55,516	5.1
Other Substance Use Disorder	48,011	4.4
Severe Emotional Disorder (in children)	38,381	3.5
Rural*	45,423	4.5
	Mean	**SD**
Age	25.9	18.4
Neighborhood Stress Score (NSS)	0.00	1.0
Medical morbidity (Rx-based) score	0.99	2.1
Medical morbidity (Dx-based) score	1.00	2.1

Data used to capture SDoH - 3

Notes

- Requires 183+ days of managed care eligibility (MassHealth as primary insurer).
- Homeless requires a Z59.0 ICD-10 code; unstably housed is having 3 or more addresses during 2019.
- Disability status indicates MassHealth eligibility as a client of the Department of Mental Health (DMH), or, if not DMH, then as a client of the Department of Developmental Services (DDS), or, if neither DMH or DDS, entitled to Medicaid due to disability ("Other disability").
- DxCG is the v4.2 concurrent Model 88 risk score; RxCG, the v4.2 concurrent Model 86 risk score, each normalized to have mean = 1 in the full MassHealth population.
- The NSS is standardized (mean = 0; SD = 1); higher scores indicate greater socioeconomic stress.

Table 3. Ratios of Observed to Expected (O:E) PCAL for Select Patient Subgroups

		Observed Mean PCAL $	O:E ratio with E predicted by		
Person-Years			Average	Age-Sex Model	PCAL model
Age Groups in years					
0-6	142,749	$787	0.80	1.00	1.00
>6-12	160,978	$528	0.54	1.00	1.01
>12-18	150,483	$579	0.59	1.00	0.99
>18-26	106,538	$783	0.80	0.98	1.00
>26-44	245,682	$1,197	1.22	1.00	1.00
>44-64	199,203	$1,619	1.64	1.00	1.00
Rurality					
Level 2 Rural	45,423	$913	0.93	0.90	1.01
Non-Rural	969,176	$989	1.00	1.00	1.00
Housing Problems					
Homeless	19,501	$3,378	3.43	2.75	0.95
Unstably Housed Only	92348	$1,188	1.21	1.30	1.02
None	902,407	$913	0.93	0.92	1.00

Table 3 (cont.). Ratios of Observed to Expected (O:E) PCAL Select Patient Subgroups (continued)

				O:E ratio with E predicted by		
	Person-Years		Observed Mean PCAL $	Average	Age-Sex Model	PCAL model
	Total	1,014,625	985	1.00	1.00	1.00
	Race					
	White, Non Hispanic	349,900	$1,114	1.13	1.06	1.01
	Black, Non-Hispanic	103,418	$883	0.90	0.93	0.96
	Hispanic	78,776	$1,045	1.06	1.07	1.00
	Other	57,017	$701	0.71	0.72	0.97
	Unknown	425,514	$932	0.95	0.99	1.00
	NSS Quintile					
	Least Stressed	202,519	$965	0.98	0.95	1.01
	2nd Quintile	202,567	$984	1.00	0.98	1.01
	3rd Quintile	202,150	$979	0.99	0.99	1.01
	4th Quintile	203,316	$988	1.00	1.03	0.99
	Most Stressed	204,072	$1,010	1.03	1.05	0.97
	None	902,407	$913	0.93	0.92	1.00

How MassHealth will apply PCAL

<u>Illustrative ACO monthly payment flow</u>

April 2023 illustration

ACO

Receives capitation from MassHealth

PCE 1

PCE 2

PCE 3

		PIDSL 1	PIDSL 2	PIDSL 3
A	PCE rate (PMPM)	$45	$50	$55
B	Tier Add-On	$7	$11	$5
C=A+B	PIDSL rate	$52	$61	$60
D	PCAL score	1.1	0.9	0.95
E=C*D	Health status-adjusted PIDSL rate	$57.2	$54.9	$57

PCE funding is **allocated to PID/SLs** and must be based on tier and member health status

For RY23, the health status adjustment will not affect the rate paid by MassHealth to ACOs or the rate ACOs pay to PCEs; PCEs should use the PCAL scores by RC to distribute dollars between PID/SLs accordingly; in future years, MassHealth anticipates using PCAL to apply a health status adjustment to the market rate by PCE

SDH 4.0 Model

CONTEXT

- Over the last several years, MassHealth has refined its ACO/MCO risk adjustment model to achieve better predictive results and its policy objectives, particularly to supplement the model with additional SDoH variables and add coefficients for disease states to improve prediction
- In 2022, MassHealth started a project to revisit its risk adjustment model more holistically for RY24 rates

OBJECTIVES

SDH 4.0 will be the risk adjustment model that MassHealth implements in RY24, with the following objectives:

1. **Decide on base medical model.** MassHealth has been using Cotiviti's DxCG, but wanted to evaluate other more commonly used Medicaid risk adjustment models
2. **Assess and align on the SDoH** variables that we risk-adjust for in SDH 4.0, including Z codes and indices/composite metrics like NSS7
3. **Evaluate other aspects of model**. Pharmacy v. medical model, core medical model "splines," etc.
4. **Evaluate data sources we use for risk adjustment**, (e.g. alternative data sources). *Will be assessed for RY25*

Outline

- Background
- Data
- DXI disease classification system
- Primary Care PCAL model
- Machine Learning DCG algorithm
- Conclusions

BOSTON UNIVERSITY

Warning

- What three things will you enjoy less if you see when they are being made?
- Sausages
- Econometric estimates
  - Risk adjustment models

eTable 2A: Modeling Principles used in this project expanded from Ash et al. (2000)[6] as in CMS-CIOO (2021)[i]

1. Diagnostic categories should be clinically meaningful.
2. Diagnostic categories should predict medical (including drug) expenditures.
3. Diagnostic categories that will affect payments should have adequate sample sizes to permit accurate and stable estimates of expenditures.
4. In creating an individual's clinical profile, hierarchies should be used to characterize the person's illness level within each disease process, while the effects of unrelated disease processes accumulate.
5. The diagnostic classification should encourage specific coding.
6. The diagnostic classification should not reward coding proliferation.
7. Providers should not be penalized for recording additional diagnoses (monotonicity).
8. The classification system should be internally consistent (transitive).
9. The diagnostic classification should assign all ICD-10-CM codes (exhaustive classification).
10. Discretionary diagnostic categories should be excluded from payment models.

Two new principles were added in this project:

11. Models should do well even on sets of rare diagnoses and demographics.
12. Parsimonious models with fewer parameters are preferred.

BOSTON UNIVERSITY

Difference between predictive and payment models

- Worry about incentives
  - To control costs
  - To upcode by coding more serious codes
  - To reward vague coding
- Simplicity and explainability to policy makers
- Desirable to be able to recalibrate on samples of 1 million

BOSTON UNIVERSITY

Machine Learning RA Techniques

Major concerns

- **Computationally challenging for N >**1 million. Most researchers use, N < 1 million, K < 200
- **Diagnoses are complex.** Many people have >10 in a year
- **Black box:** Difficult to interpret results
- *Enforcing nonnegative predictions?*
- *Ease of updating?*
- *Stability over time?*

BOSTON UNIVERSITY

Machine Learning (Neural Networks)?

ICD10

codes

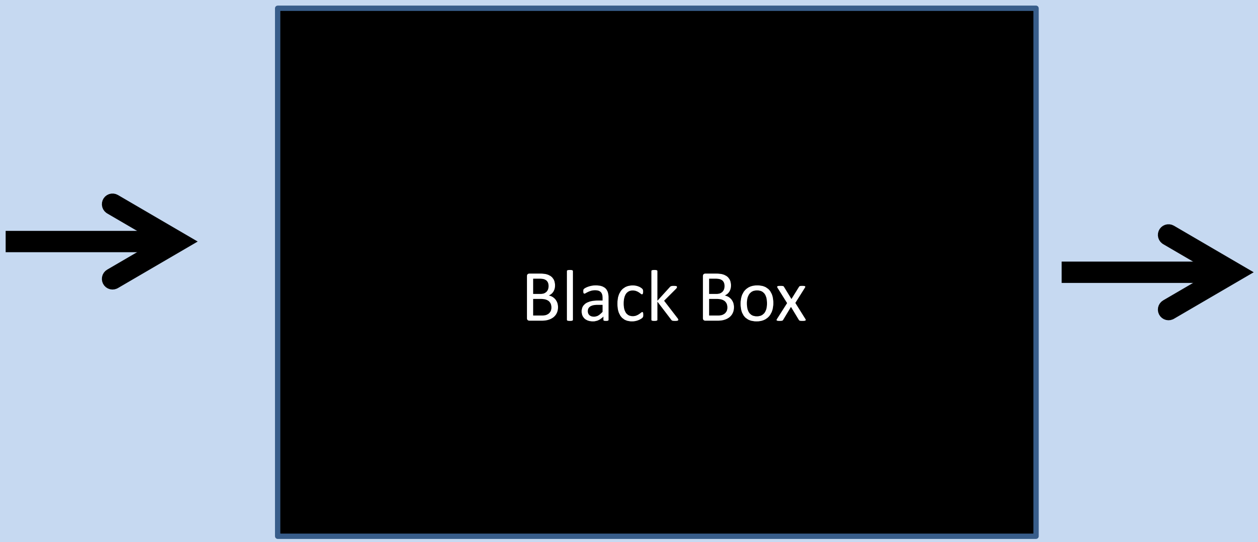
A01

A02

A03

…

Z998



Predictions/

interpretation

$\widehat{Y}_i$

Boston University

We built upon the DCG and HCC approaches

- **Clinical input**
- **Very big data**
- **Hierarchies**
- **Explainable**
- **Automated estimation**

BOSTON UNIVERSITY

DXI model specification

$$Y_i = A_{ia} * \alpha_a + DXI_{ij} * \beta_j + \varepsilon_i \qquad (1)$$

DCG model specification

$$Y_i = A_{ia} * \alpha_a + \sum_h \sum_g DCG_{ihg} * \beta_{hg} + \varepsilon_i \qquad (2)$$

Index notation:
t = time (year) (omitted)
i = person-year observation
a =age-sex group
j = DXI items

Y_i = dependent variable
A_{ia} = age-sex groups
DXI_{ij} = Diagnostic Items
DCG_{hg} = Diagnostic Cost Groups

BOSTON UNIVERSITY

Gains in incentives from four things

- Aggregating into DCGs to avoid rewarding slight coding variation
- Hierarchies ignore less serious conditions
- Ignoring low cost, common, vague or gameable information
- Avoiding the underpayment of rare diagnoses

BOSTON UNIVERSITY

Having a rare disease is not so rare!

Frequency Table for Rarest Diagnosis Category per Enrollee Year in the Validation Sample, 2006-2008

Bin	Count	% of enrollees
No diagnoses present	1,479,306	22%
< 1 per million	197,181	3%
1-10 per million	769,591	12%
10-100 per million	1,730,611	26%
100-1,000 per million	1,726,323	26%
1,000-10,000 per million	513,109	8%
10,000-100,000 per million	188,138	3%
Total	6,604,259	100%

BOSTON UNIVERSITY

Table 3 Sensitivity Analysis: Validation sample measures of alternative specifications

	R-Square	Mean absolute error	Number of parameters	Rare disease mean error: enrollee-year mean residual of people with any diagnosis rarer than 100 per million
Prediction models				
HCC	0.428	$5,227	166	$1,927
DXI additive model	0.589	$3,785	2929	-$82
Payment Models				
Appropriateness to Include (ATI)\ Scores				
DCG ATI=0	0.4689	$4,520	445	$609.99
DCG ATI<2	0.5032	$4,313	526	$296.03
DCG ATI<3	0.5264	$4,151	619	-$4.18
DCG ATI<4 (Base)	**0.5345**	**$4,113**	**661**	**-$70.50**
DCG Base, omitting CCSR	0.5253	*	691	*

BOSTON UNIVERSITY

Outline

- Background and methodology
- Data
- New DXI disease classification system
- New machine learning algorithm
- Results from DXI/DCG estimation
- Conclusions and ideas for future work

BOSTON UNIVERSITY

DXI model specification

$$Y_i = A_{ia} * a_a + DXI_{ij} * \beta_j + \varepsilon_i \qquad (1)$$

DCG model specification

$$Y_i = A_{ia} * a_a + \sum_h \sum_g DCG_{ihg} * \beta_{hg} + \varepsilon_i \qquad (2)$$

Index notation:

t = time (year) (omitted)

i = person-year observation

a =age-sex group

r = CCSR categories

j = DXI items

Y_i = dependent variable

A_{ia} = age-sex groups

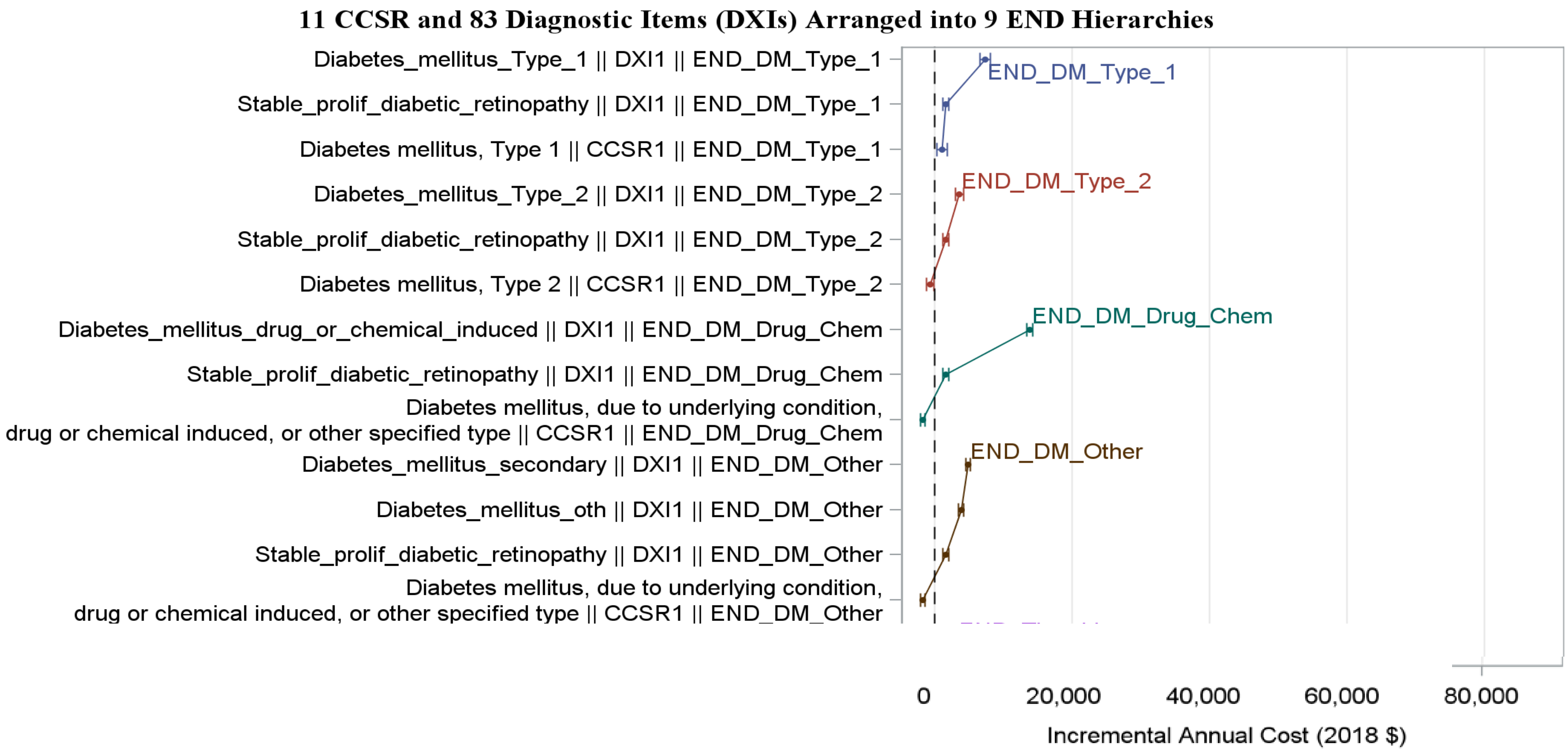
$CCSR_{ir}$ = Clinical Classification System, Refined categories

DXI_{ij} = Diagnostic Items

DCG_{hg} = Diagnostic Cost Groups

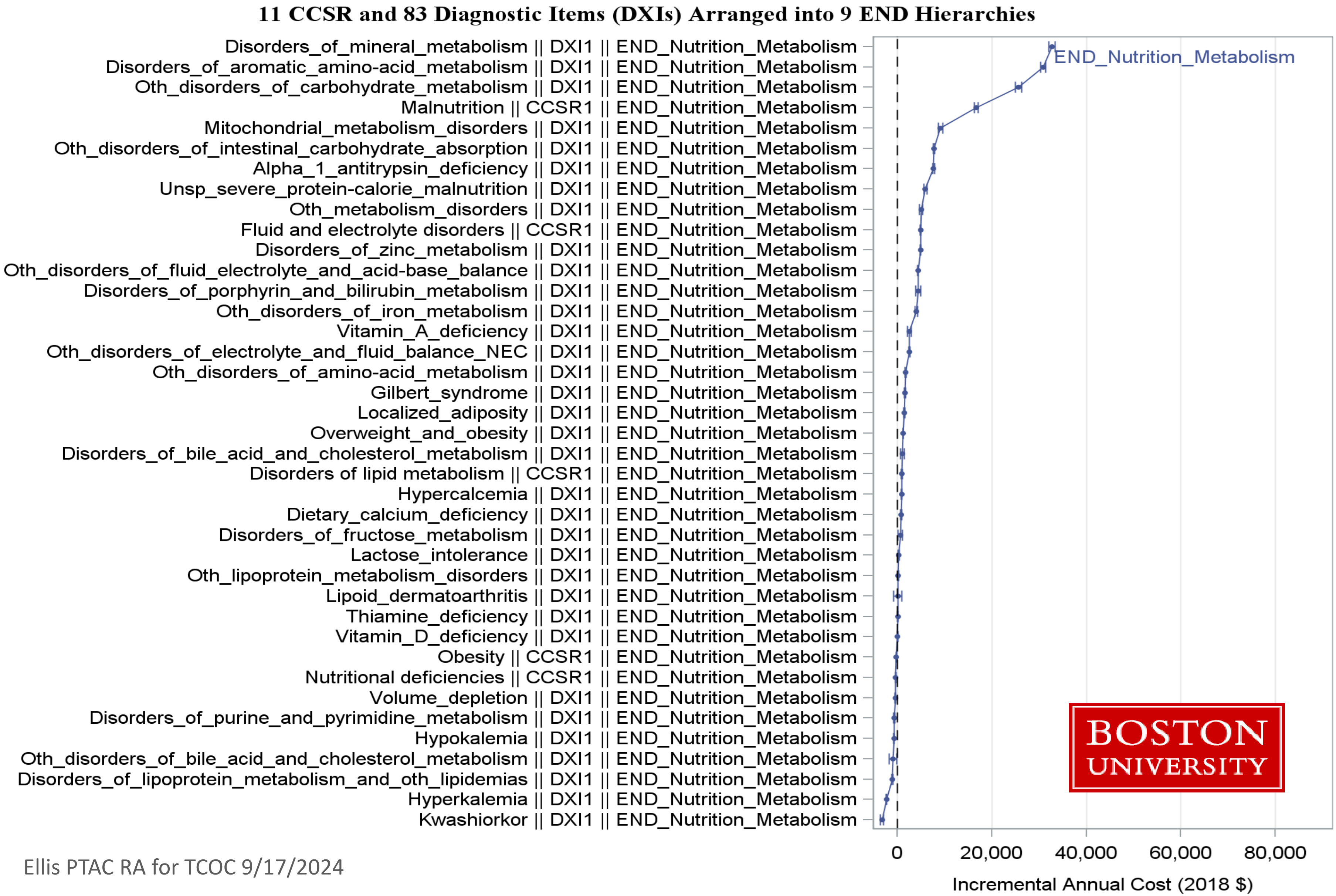
BOSTON UNIVERSITY

Endocrine system **DXI** regression coef - 1



BOSTON UNIVERSITY

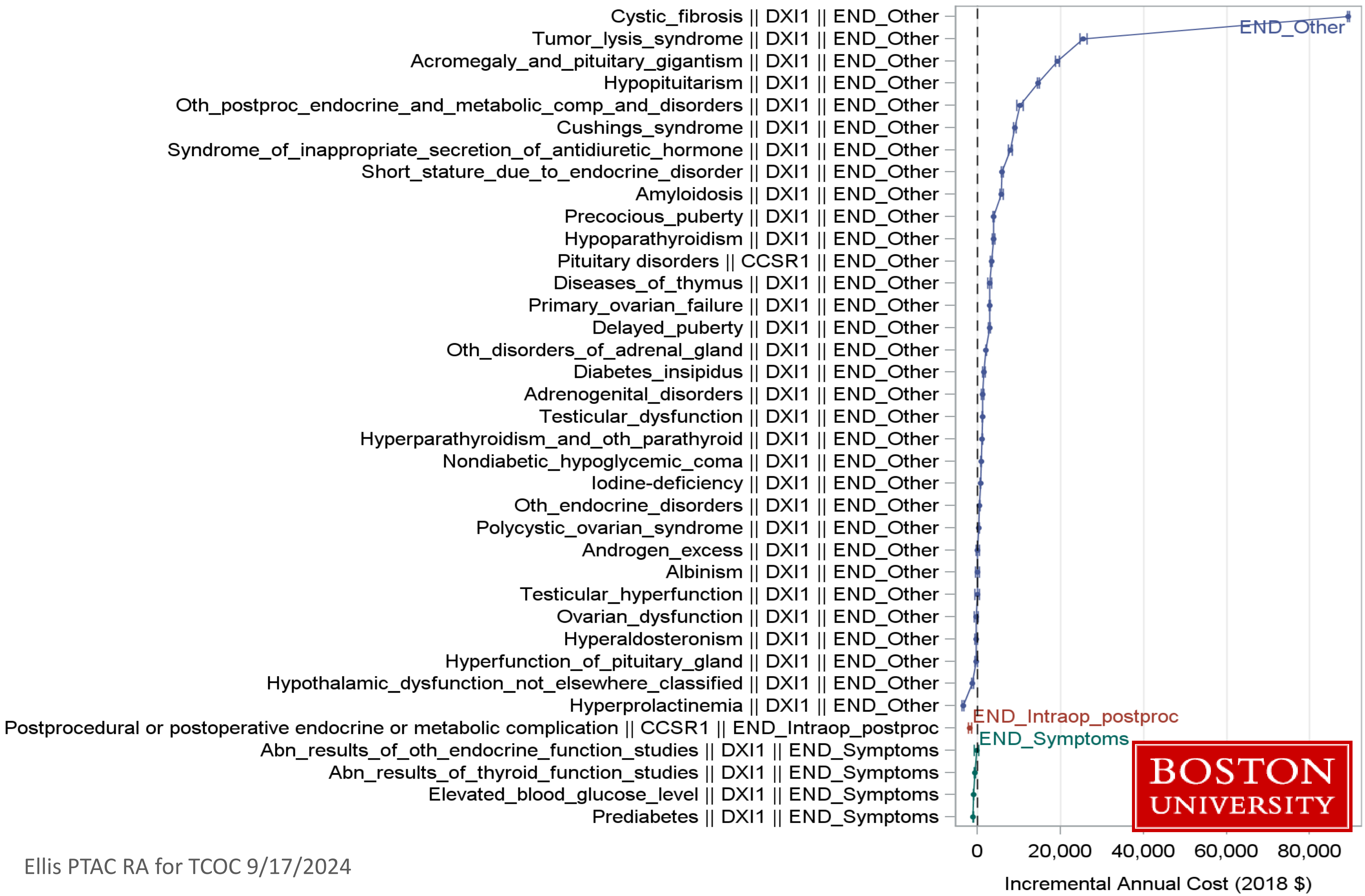
Endocrine system **DXI** regression coef - 2



BOSTON UNIVERSITY

Endocrine system **DXI** regression coef - 3

11 CCSR and 83 Diagnostic Items (DXIs) Arranged into 9 END Hierarchies



DXI/DCG ML algorithm parameters

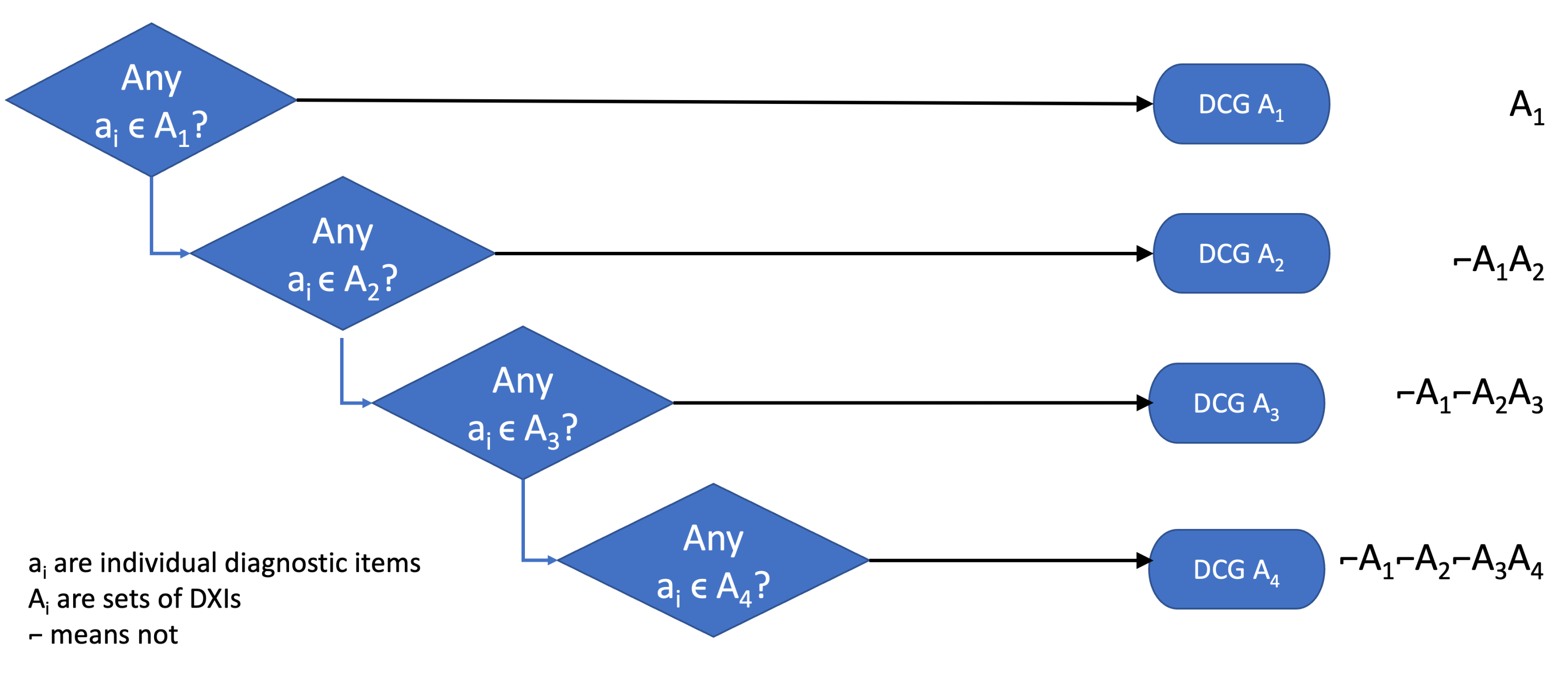
	Base values
Minimum N for a DCG	2000
Maximum percent difference between coefficients put together in one DCG	30%
Minimum size needed for residual DXI	0
Statistical significance required for including a DCG in final model	$p<0.0001$
Whether to allow negative DCGs	no

DXI/DCG algorithm details

- MDs cluster DXIs into 218 HIER groups
- Run OLS using all DXIs to predict residual spending to get incremental cost coefficients.
- Sort DXIs from highest to lowest cost coefficients not yet assigned to any HIER
- Group DXIs into DCG_i using:
  1. Highest coefficients
  2. Reasonably similar coef (base case: <30% difference)
  3. Require minimum N for DCG_i (base case: >2000 people)
  4. Disregard statistical significance of individual DXIs except for stopping rule
- Reset any assigned DXIs in any DCG_i to zero.
- Rerun regression while including the DCG_i variables and new conditional DXIs.
- Iterate until no more nonnegative DXIs are available to group
- Drop all DXI with negative or insignificant coefficients
- Rerun regression using only DCGs
- Drop DCGs with negative coefficients
- Use stepwise regression to keep only DCGs with statistically significant positive coefficients

BOSTON UNIVERSITY

Figure 1: Flow chart of assignment of DXIs to DCGs in Hierarchy A



Notes: DXI is a Diagnostic Item, and DCG is a Diagnostic Cost Group.

BOSTON UNIVERSITY

Single hierarchy DCG algorithm (Ash et al. 1989) can also be modeled as a regression tree.

DCG A_1	A_1
DCG A_2	$\neg A_1 A_2$
DCG A_3	$\neg A_1 \neg A_2 A_3$
DCG A_4	$\neg A_1 \neg A_2 \neg A_3 A_4$

Patient with conditions in A_2, and A_3 get paid only for DCG A_2

BOSTON UNIVERSITY

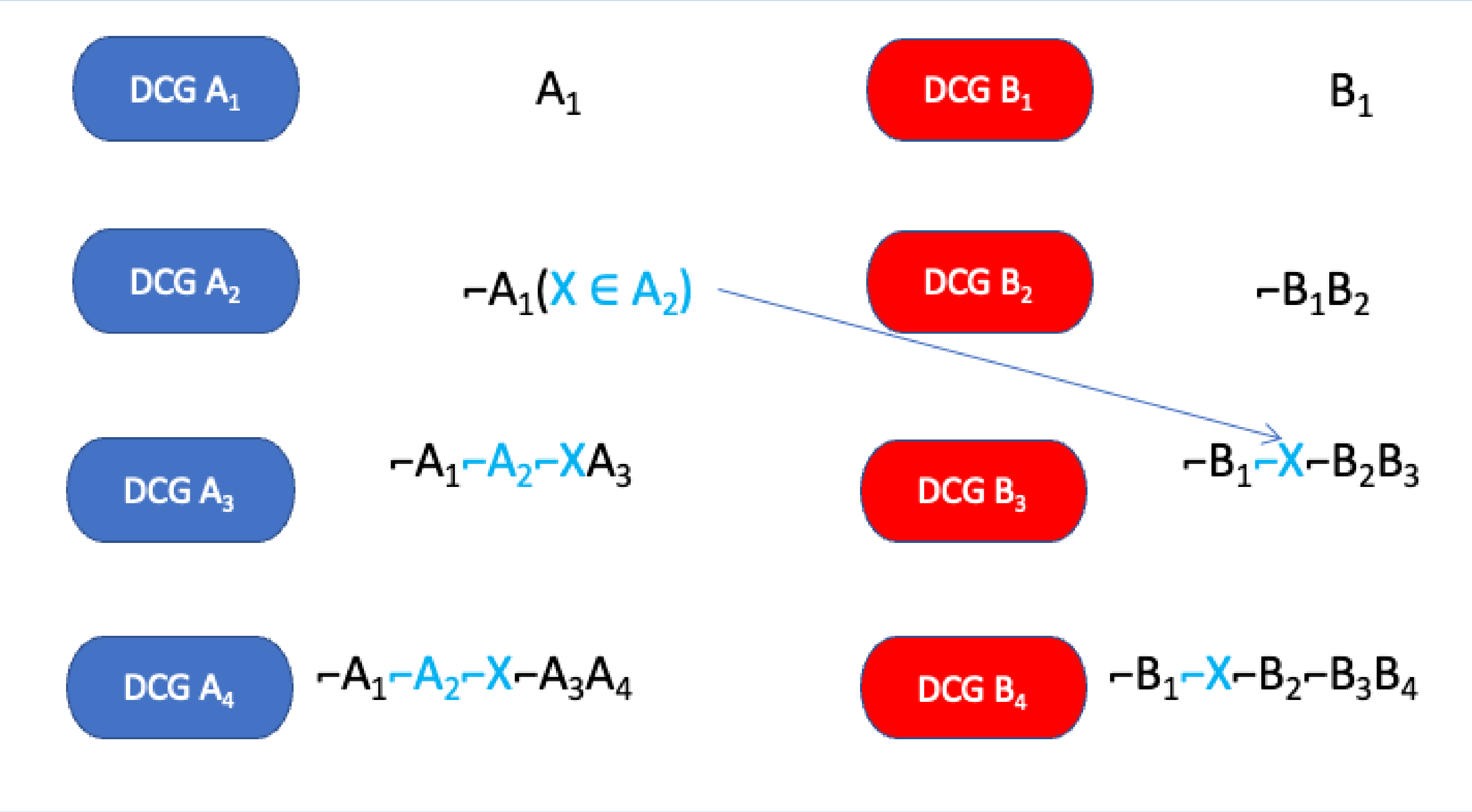
The DXI/DCG algorithm uses coefficients, not sample averages, to calculate payments contribution for each leaf, and uses multiple hierarchies. Here there are two HIER groups, A and B, with each DXI assigned uniquely to one HIER.

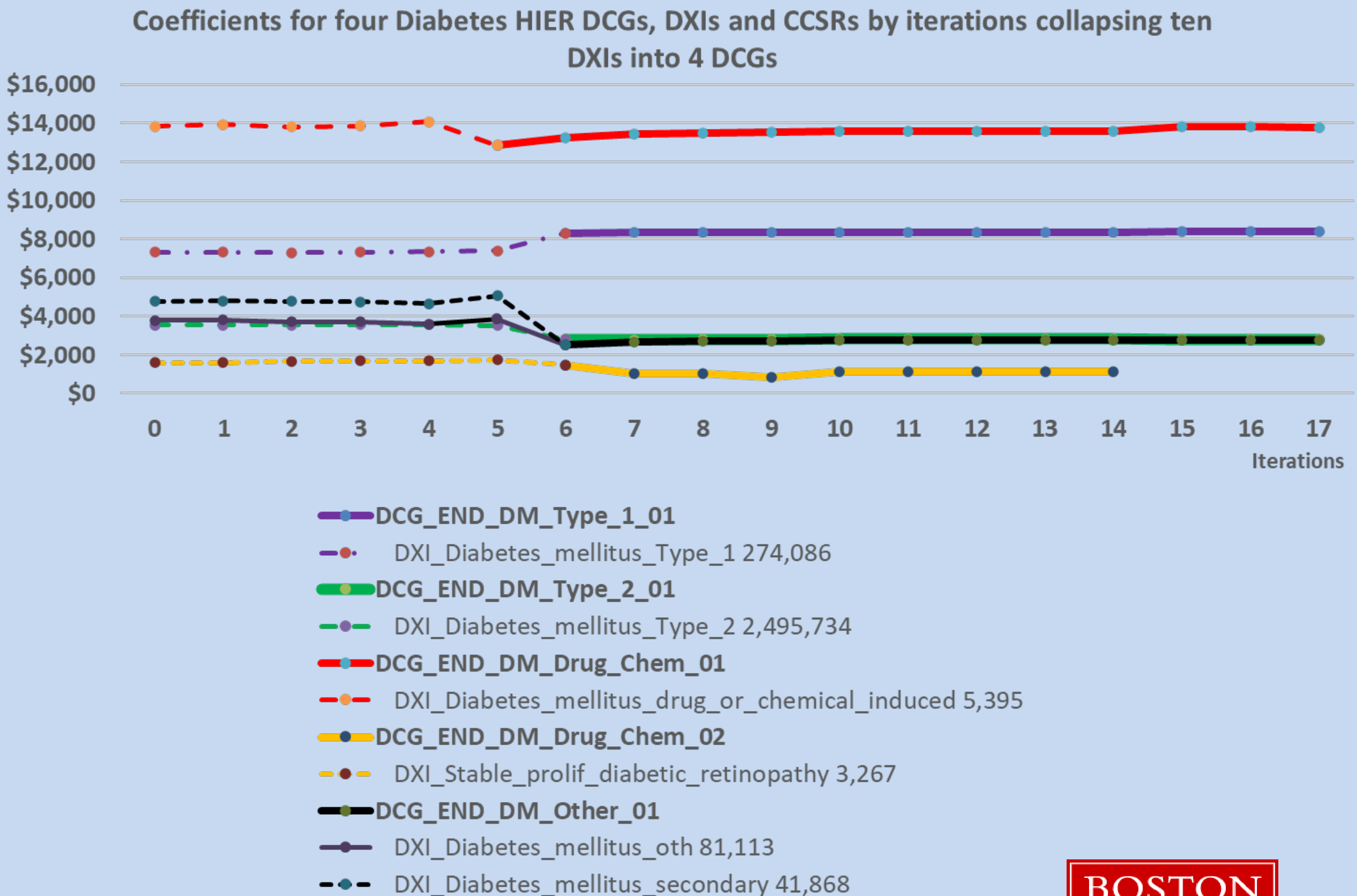
DCG A_1	A_1	DCG B_1	B_1
DCG A_2	$\neg A_1 A_2$	DCG B_1	$\neg B_1 B_2$
DCG A_3	$\neg A_1 \neg A_2 A_3$	DCG B_3	$\neg B_1 \neg B_2 B_3$
DCG A_4	$\neg A_1 \neg A_2 \neg A_3 A_4$	DCG B_4	$\neg B_1 \neg B_2 \neg B_3 B_4$

Patient with conditions in A_2, A_3, and B_3 get paid for sum of coefficients on DCG A_2 and DCG B_3 while DCG A3 gets ignored.

BOSTON UNIVERSITY

eFigure 2: Flow Chart of Hypothetical DCG Algorithm assignment when DXI_X maps to both Hierarchies A and B

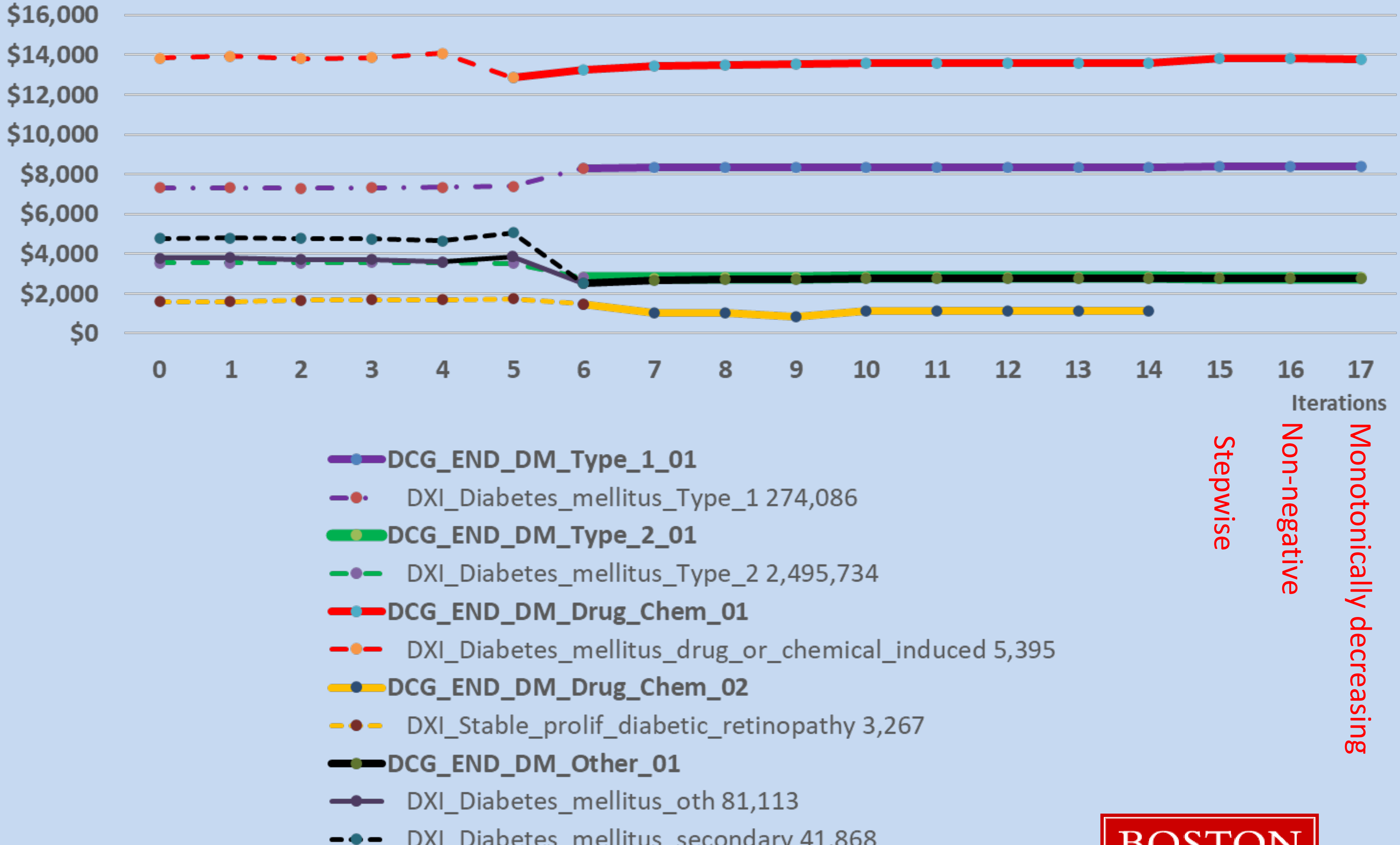




Four additional DXIs/CCSRs included but dropped when not significant

BOSTON UNIVERSITY

Coefficients for four Diabetes HIER DCGs, DXIs and CCSRs by iterations collapsing ten DXIs into 4 DCGs



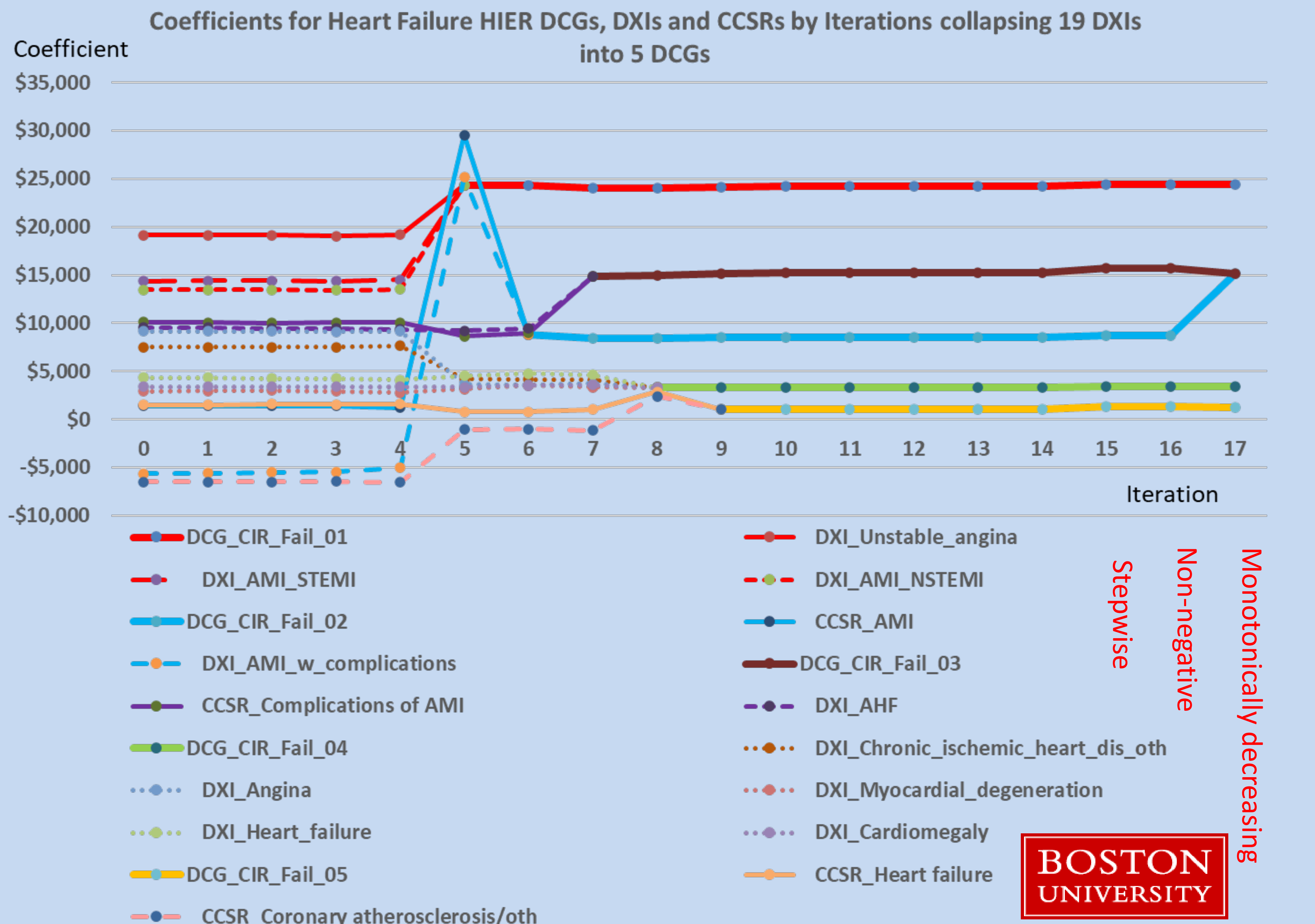
Stepwise

Non-negative

Monotonically decreasing

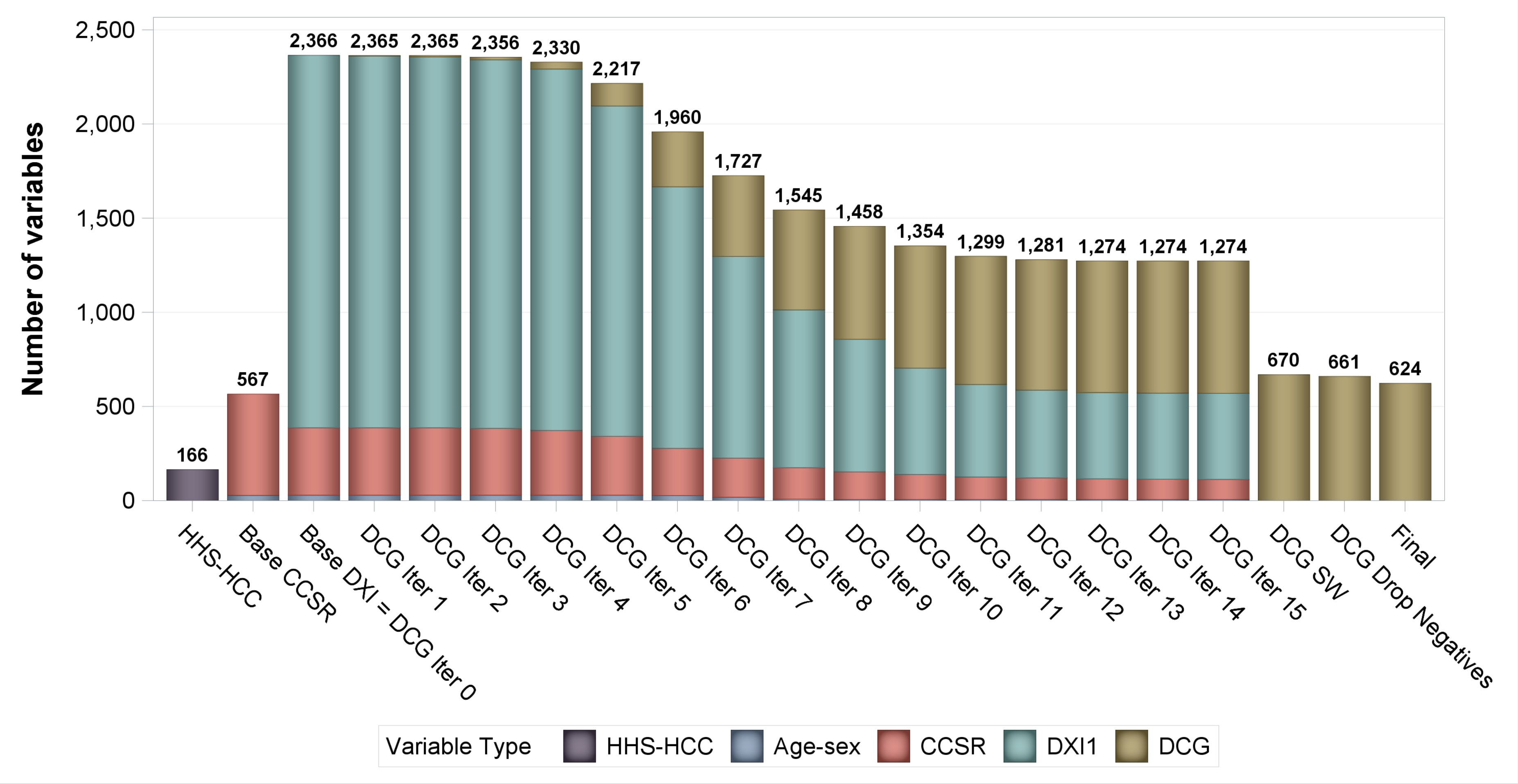
Four additional DXIs/CCSRs included but dropped when not significant

BOSTON UNIVERSITY



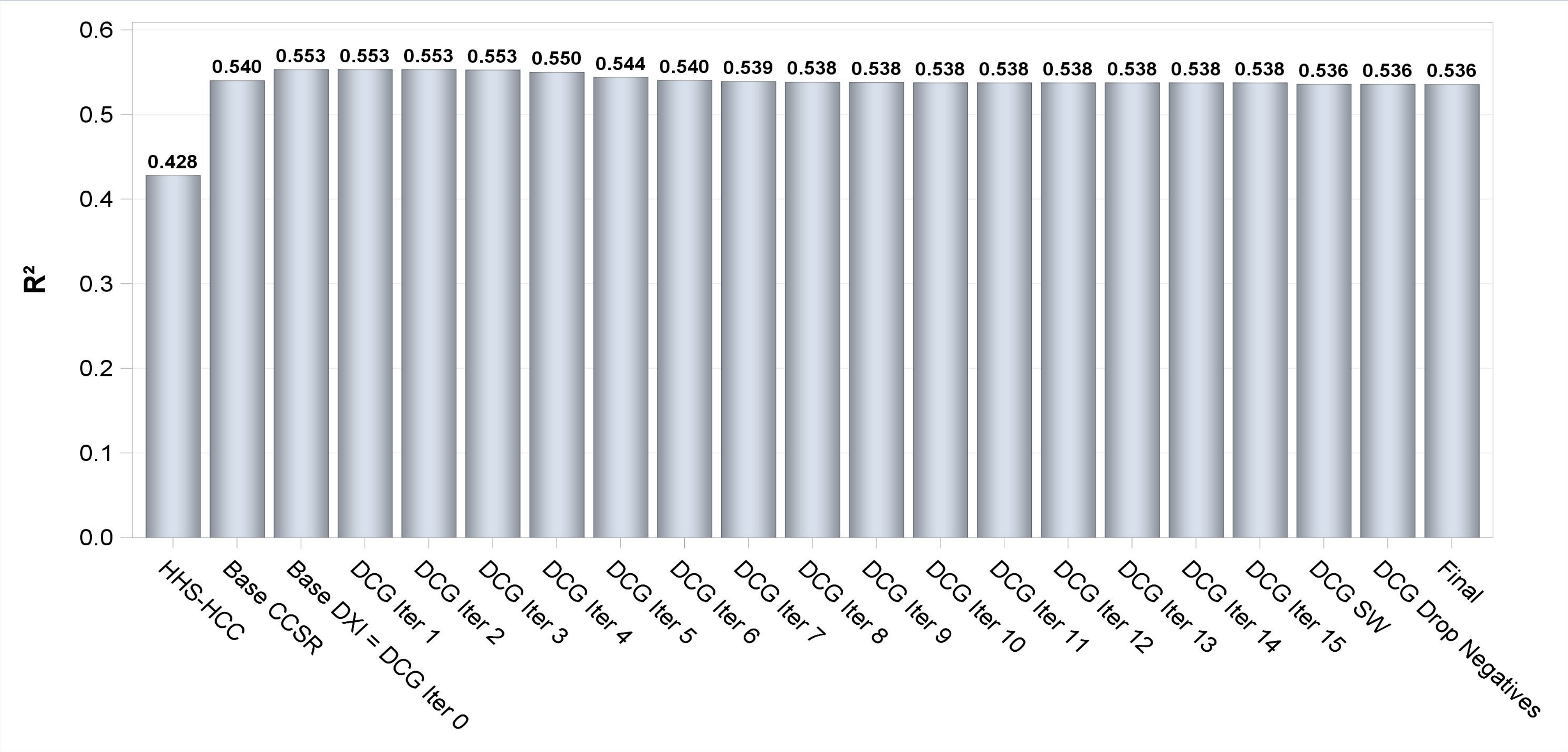
BOSTON UNIVERSITY

Figure 3: Model Parameter Counts across DCG iterations for the Base Model



BOSTON UNIVERSITY

Figure 3: Model R^2 across DCG iterations for the Base Model



BOSTON UNIVERSITY

Footnote to Figure 2

Notes: HHS-HCC is the Department of Health and Human Services Hierarchical Condition Category model, CCSR is the Clinical Classifications Software Refined model, DXI is the Diagnostic Items model, and DCG is the Diagnostic Costs Groups algorithm. The HHS-HCC model uses the combined set of HHS-HCCs included in the adult, child or infant models in a single regression. Base Case CCSR uses OLS on 538 observed CCSR categories, while Base case DXI uses CCSR plus DXIs. As DCGs are created, DXI and CCSR which fall into them are dropped from the model. After all DCGs have been found, the DCG SW iteration estimates a stepwise regression that omits any all remaining DXI and CCSR variables not assigned to DCGs and includes only statistically significant DCGS. The run labeled Final excludes any variables with negative coefficients. All models include age-sex dummy variables.

BOSTON UNIVERSITY

DXI classification structure

Disease chapters

BLD Blood
CIR Circulatory
DIG Digestive
EAR Ear
END Endocrine
EXT External_causes
EYE Eye
FAC Factors_influencing
GEN Genito-urinary
INF Infections
INJ Injuries
MAL Malformations
MBD Mental_behav_devel
MSK Muscular_skeletal
NEO Neoplasm
NVS Nervous
PNL Perinatal
PRG Pregnancy
RSP Respiratory
SKN Skin_Connective
SPL Special
SYM Symptoms

Hierarchies

INJ_Head_neck_eye
INJ_Thoracic
INJ_Abdominal
INJ_Spine_back
INJ_Fracture
INJ_Minor
INJ_Foreign_body
INJ_Burn
INJ_Frostbite_hypotherm
INJ_Poisoning
INJ_Abuse
INJ_Allergies
INJ_Complic
INJ_Nerves
INJ_Traumatic_injuries
INJ_Vascular
INJ_Self_harm
INJ_Vague

DXI main effects

{Next slide}

CCSR

Dislocations, initial encounter
Dislocations, subsequent encounter
Fracture of the spine and back, initial encounter
Fracture of the spine and back, subseq encounter
Spinal cord injury (SCI), initial encounter
Spinal cord injury (SCI), subseq encount

DXIs main effects for INJ_poisoning

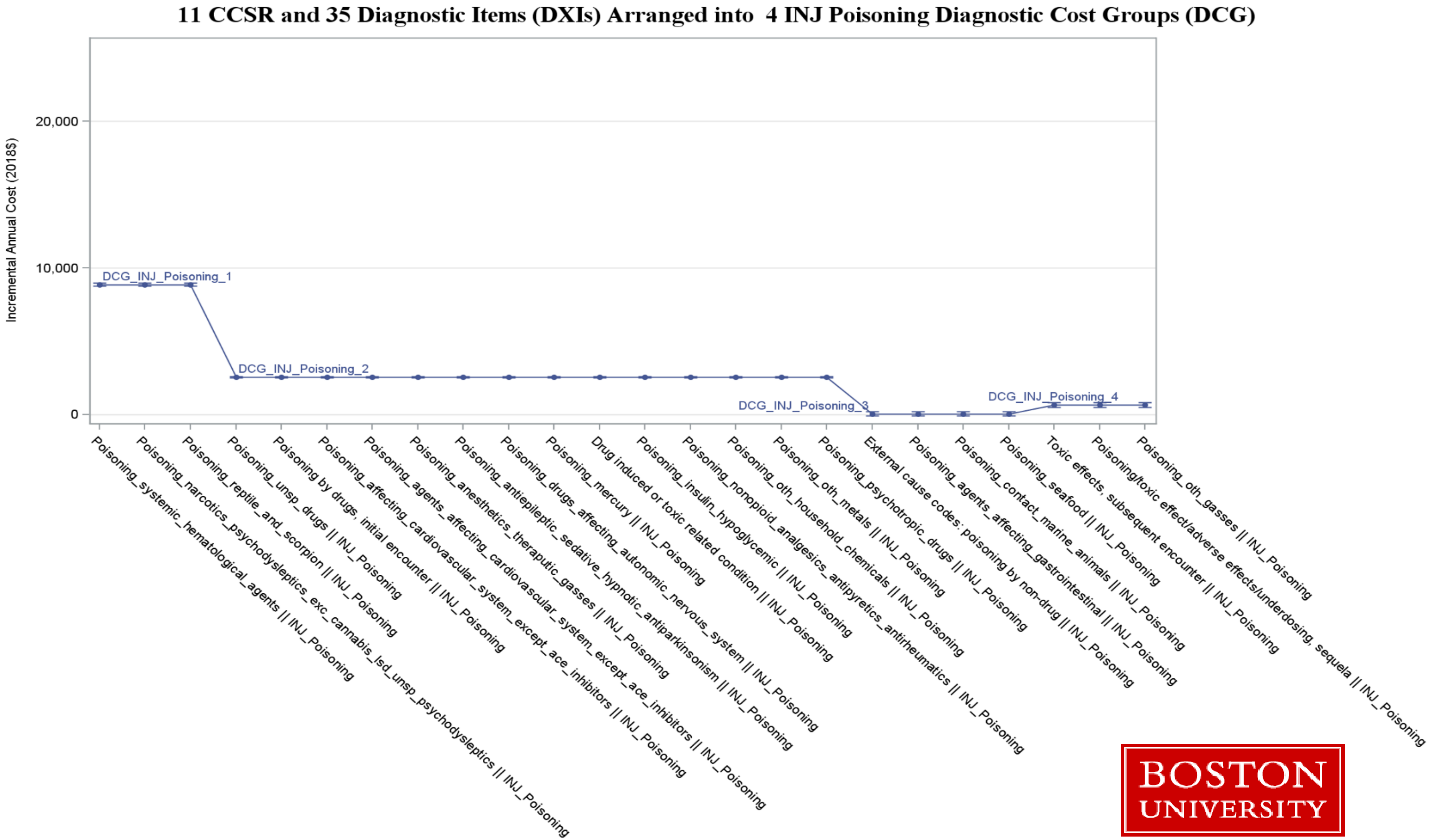
Poison_antibiotics_anti-infect_antiparas_ hormones_exc_insul_hypoglyc
Poisoning_ace_inhibitors
Poisoning_affecting_cardiovascular_ system_except_ace_inhibitors
Poisoning_agents_affecting_cardiovasc_ syst_exc_ace_inhibitors
Poisoning_agents_affect_gastrointestinal
Poisoning_agents_affecting_smooth_ and_skeletal_musc_and_resp_sys
Poisoning_anesthetics_theraputic_gasses
Poisoning_antiepileptic_sedative_ hypnotic_antiparkinsonism
Poisoning_cannabis
Poisoning_carbon_monoxide
Poisoning_contact_marine_animals
Poisoning_contact_plant_or_oth_animals
Poisoning_diuretics
Poisoning_drugs_affect_autonom_ nervous_ system
Poisoning_food_except_seafood
Poisoning_foreign_body
Poisoning_insulin_hypoglycemic
Poisoning_lead

Poisoning_lsd_unsp_psychodysleptics
Poisoning_mercury
Poisoning_narcs_psychodysl_exc_cannabis_lsd_unsp_ psychodys
Poisoning_nonopioid_analgesics_antipyretics_ antirheumatics
Poisoning_oth_arthropods
Poisoning_oth_gasses
Poisoning_oth_household_chemicals
Poisoning_oth_metals
Poisoning_pesticides
Poisoning_psychotropic_drugs
Poisoning_reptile_and_scorpion
Poisoning_seafood
Poisoning_spider
Poisoning_systemic_hematological_agents
Poisoning_topical_skin_eye_ent_dental_drugs
Poisoning_toxic_effects_oth_unsp
Poisoning_unsp_drugs

BOSTON UNIVERSITY

DCG Model Coefficients

(K=4 using 24 out or 47 DXIs)



BOSTON UNIVERSITY

Table 1: Sensitivity Analysis: Validation Sample Measures of Alternative Specifications

	R-Square	Mean absolute error	Number of parameters	Rare disease mean error: enrollee-year mean residual of people with any diagnosis rarer than 100 per million
DCG: Base model	**0.535**	**$4,114**	**624**	**-$73**
Panel A: Alternative Model Structures				
Charlson Comorbidity Index (CCI)	0.227	$6,116	18+30 = 48	$3,055
HHS-HCC Marketplace using hierarchies	0.428	$5,227	136+30 = 166	$1,927
CCSR additive model	0.539	$4,140	567	-$114
DXI+CCSR additive model	0.589	$3,786	2,929	-$83
Disease chapters additive model	0.201	$6,226	52	$556

All models include 30 age*sex dummy variables
The CCI index has been used in 12,800 articles indexed in Google scholar since 2023, despite being a very weak predictor.

BOSTON UNIVERSITY

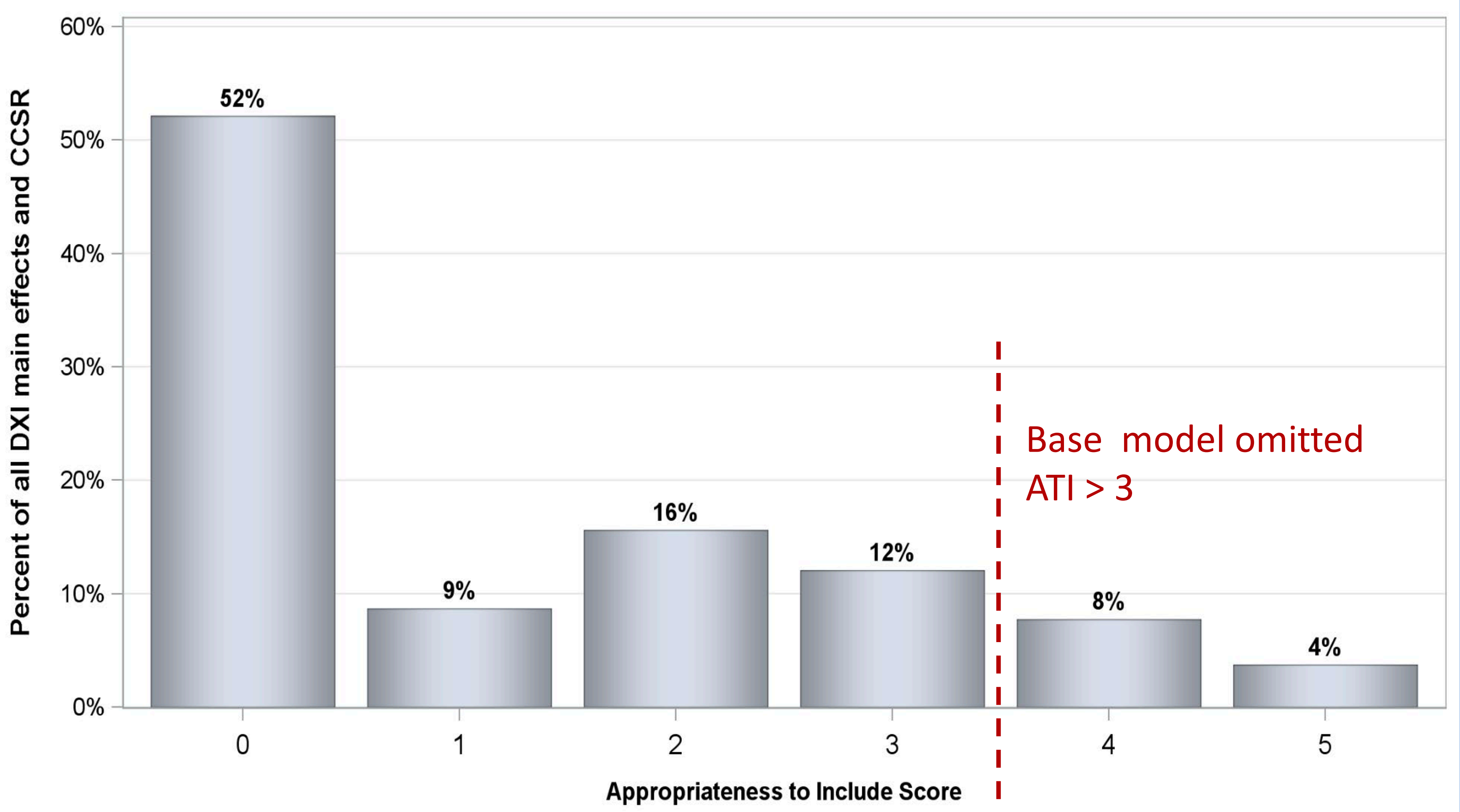
Physicians also rated DXIs by their Appropriateness to Include scores

- 0 => no concerns about using for payment
- 1 Trivial concerns …
- 2 Minor concerns …
- 3 Meaningful concerns …
- 4 Serious concerns …
- 5 Major concerns: avoid using for payment

Later added

- *6 DXI/CCSR is too collinear with other DXIs*

Figure 3: Percent Distribution of Appropriateness to Include (ATI) scores in DXI Main Effects and CCSRs



BOSTON UNIVERSITY

Table 1: Sensitivity Analysis: Validation Sample Measures of Alternative Specifications Panel B

	R-Square	Mean absolute error	Number of parameters	Rare disease mean error: enrollee-year mean residual of people with any diagnosis rarer than 100 per million
DCG: Base model	**0.535**	**$4,114**	**624**	**-$73**
Panel B: Appropriateness to Include (ATI) Score (0 = least gameability/vagueness concerns, 5 = most concerns)				
DCG: ATI = 0	0.469	$4,520	445	$610
DCG: ATI < 2	0.503	$4,313	526	$296
DCG: ATI < 3	0.526	$4,151	619	-$4
DCG: ATI < 4 Base without forcing monotonicity	0.535	$4,113	661	-$71
DCG: ATI < 5	0.536	$4,134	667	-$115
DCG: ATI All Values	0.539	$4,112	683	-$109

All models include 30 age*sex dummy variables

BOSTON UNIVERSITY

Table 1: Sensitivity Analysis: Validation Sample Measures of Alternative Specifications Panel C

	R-Square	Mean absolute error	Number of parameters	Rare disease mean error: enrollee-year mean residual of people with any diagnosis rarer than 100 per million
DCG: Base model	**0.535**	**$4,114**	**624**	**-$73**
Panel C: Alternative Information Sets+				
DCG: Including EXT, FAC chapters	0.568	$3,910	710	-$139
DCG: Allow negative/insignificant coefficients	0.534	$4,114	672	-$74
DCG: No exclusions imposed within hierarchies (DCC model)	0.541	$4,071	687	-$87
DCG: Single hierarchy for each chapter	0.495	$4,339	202	-$26
DCG: Single hierarchy	0.315	$5,031	28	$898
DCG: Base model using only CCSR variables	0.461	$4,514	248	$212
DCG: Base model using only DXI variables	0.524	$4,170	676	$22

All models include 30 age*sex dummy variables

BOSTON UNIVERSITY

Commitment to public posting of models and software

- DXI classification system and formulas posted on the web, linked at Ellis et al JAMA 2022 DXI article. (Open access)
- SAS software implementing DXI models (and eventually BU-DCG models) posted at

  http://tinyurl.com/DXI-Software
- DXI models have been successfully tested on Belgian and South Korean data!

BOSTON UNIVERSITY

Conclusions and Limitations

Limitations

- Do not currently used diagnostic modifier information
- More complex than existing HCC models
- Have not applied to additional populations (e.g. Medicaid and Medicare)

Contribution

We have developed a new Machine Learning algorithm that is

- Automated
- Readily interpreted
- Highly predictive
- Avoids negative predictions
- Resistant to upcoding
- Downweights vague and inappropriate diagnoses

BOSTON UNIVERSITY

Many possible extensions

- Better machine learning algorithms
- Use diagnostic modifiers and interactions
- Focus on insurance type (Medicaid, Medicare, Marketplace)
- Population subgroups children, women, racial groups
- Data from other countries
- Add in social drivers of health info
- Further examine measures of equity
- Develop prospective models, add dynamics, lags
- Empirical measures of gameability and vagueness
- Use for detecting fraud, errors, overpricing, technology choices
- Redo the analysis of 1000s of papers that used inferior risk adjustment models (HCC, Charleson, Elixhauser)
- Calibrate for diverse outcomes and performance measures
- Use for epidemiological work

Diagnostic Item (DXI) and Diagnostic Cost Group (DCG) Formulas for Healthcare Payment and Decision-making

Randall P Ellis, Corinne Andriola, Jeffrey J Siracuse, Alexander Hoagland, Tzu-Chun Kuo, Heather E Hsu, Allan Walkey, Karen E. Lasser, Arlene S Ash

Thank you!

BOSTON UNIVERSITY

Listening Session 3: Addressing Challenges Regarding Data, Benchmarking, and Risk Adjustment

Aneesh Chopra, MPP

President, CareJourney

PTAC Listening Session:

Addressing Challenges Regarding Data, Benchmarking, Risk Adjustment

Aneesh Chopra
@aneeshchopra

CareJourney
A Hunch Analytics Company

Skating to the Puck, Converging on FHIR

PROPOSED STRATEGY FOR EXECUTION OF THE HEALTH INFORMATION TECHNOLOGY INVESTMENT PROGRAM

Draft, February 24, 2009

EXECUTIVE SUMMARY

The $19 billion health information technology (HIT) investment authorized in the American Recovery and Reinvestment Act (ARRA) represents a landmark opportunity to improve health care. In considering how best to execute on this opportunity, it is critical to understand that to treat the HIT investment program as a pure technology implementation program is to effectively guarantee its failure. HIT is not magic. In the absence of provider payment reform and care delivery innovation, it is all too easy to imagine spending $19 billion on HIT adoption and producing little tangible social benefit. However, there is a clear path to victory:

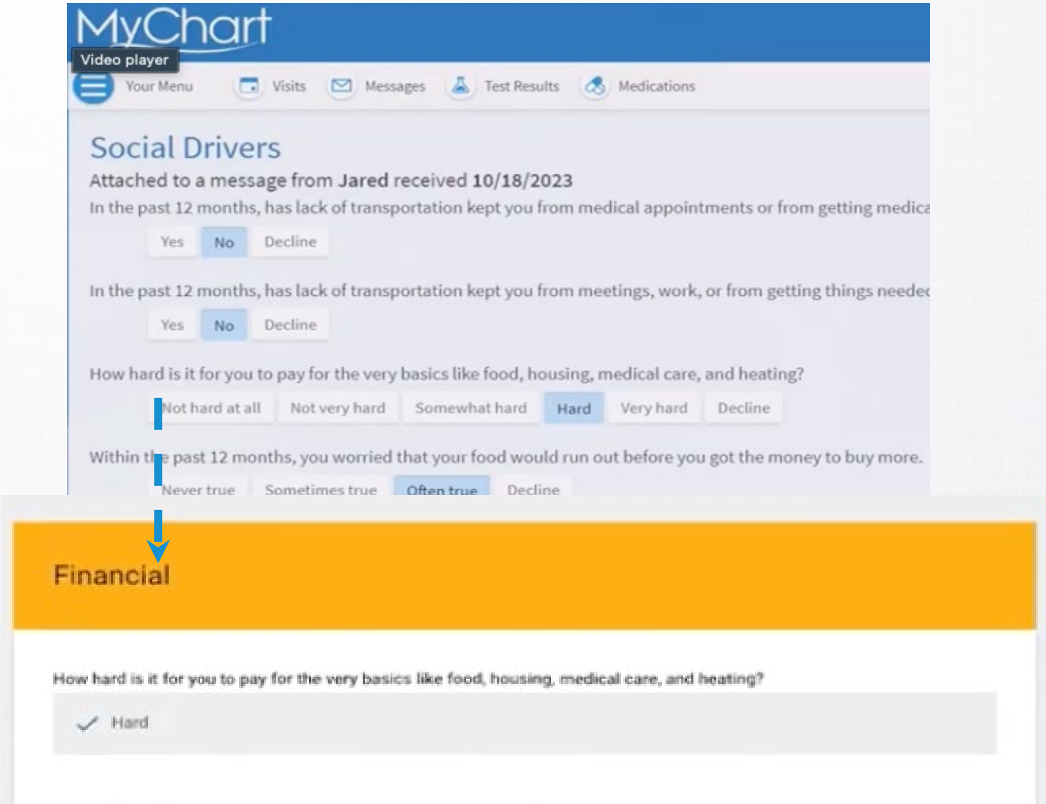
- If we avoid focusing the HIT investment program narrowly on HIT adoption and instead focus it explicitly on the actual improvement of population health, and
- If we use the HIT investment to catalyze a "virtuous cycle" of (1) provider payment reform, (2) care delivery innovation, and (3) HIT adoption
- Then: the HIT investment can literally transform health care as we know it.

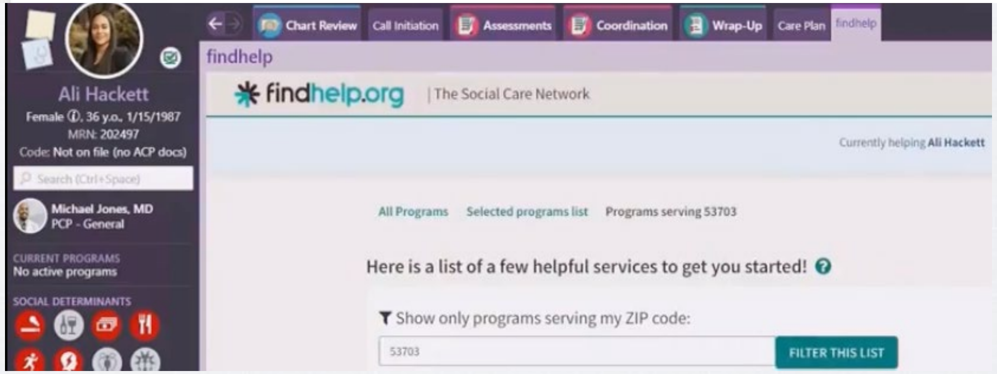
"Low HIT adoption cripples the ability to pursue provider payment reform..."

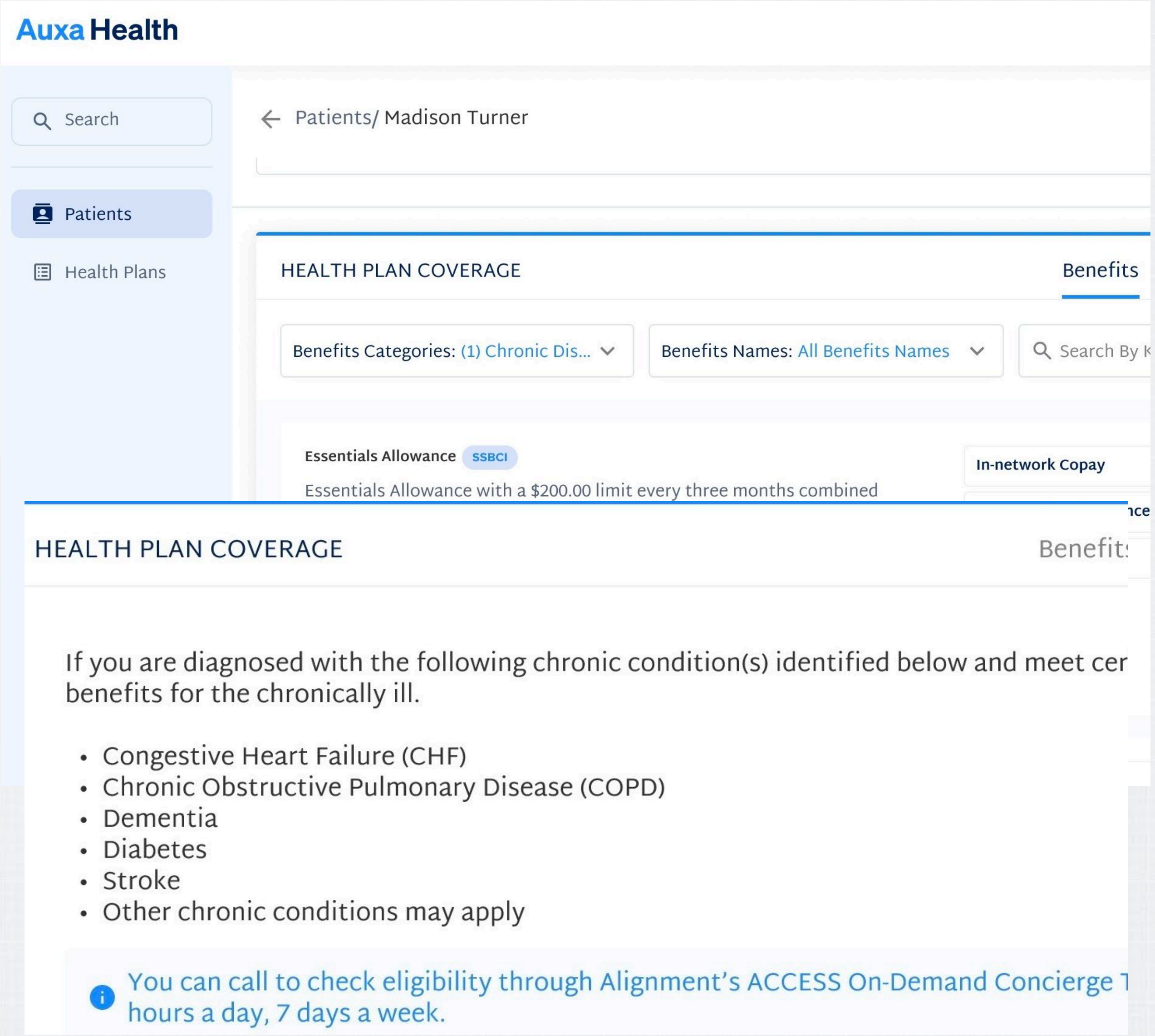
1. Coverage (Benefits)
2. Clinical Data (Payment)
3. Bulk FHIR Networks
4. Price Transparency
5. Consumer Navigation

Notes: "FHIR" = Fast Health Interoperability Resources

#1: SDOH Data Standards





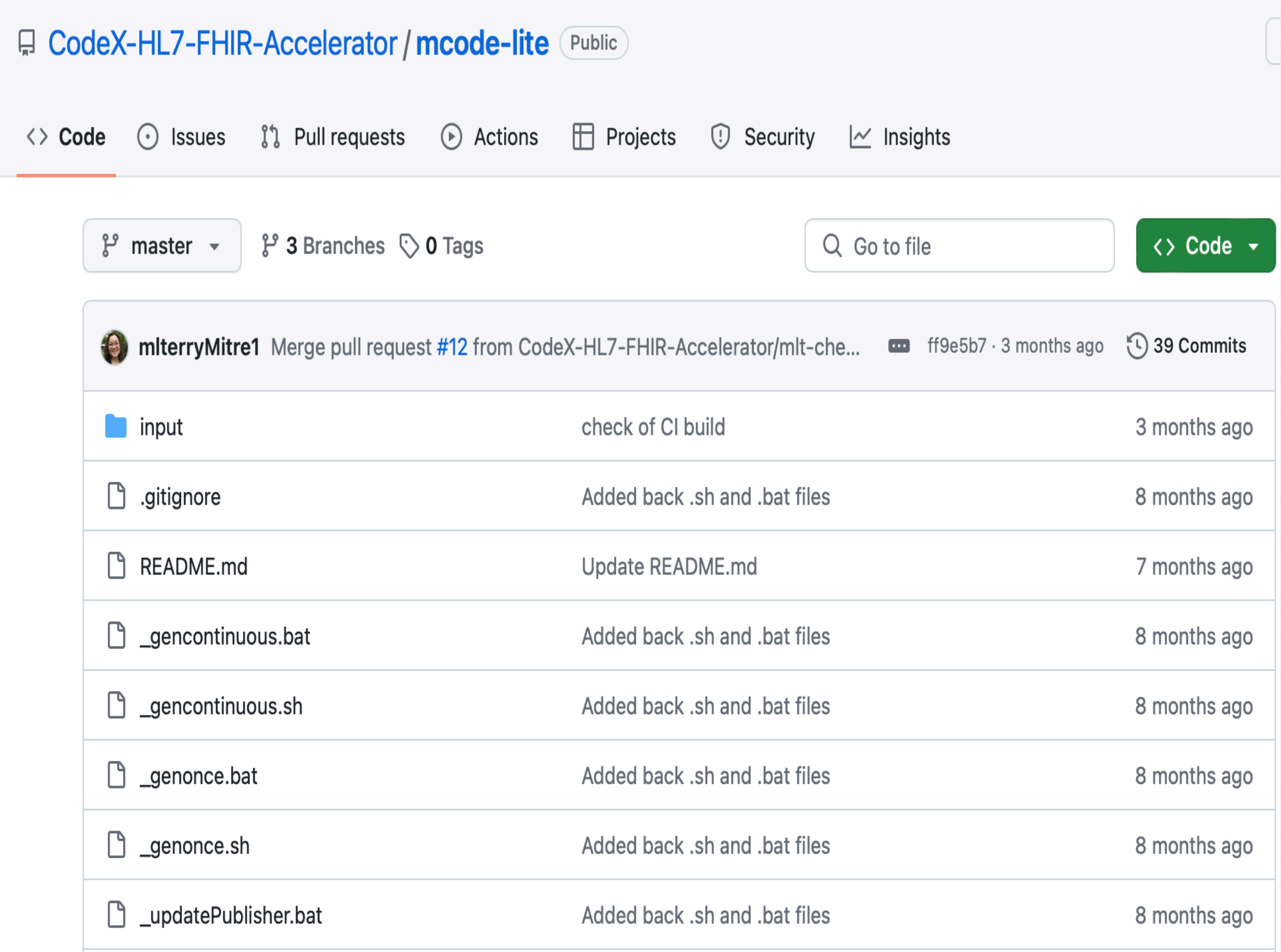


Source: https://health.gov/sites/default/files/2024-02/White%20House%20Challenge%20Commitments.pdf; Epic-FindHelp Demo, 2023; Auxa Health

#2: Clinical Data for VBC Payment

MARCH 05, 2024

Improving Cancer Care Through Better Electronic Health Records: Voluntary Commitments and Call to Action



"Commitments to adopt the core EOM data elements…were made by Epic; Oracle; Ontada, a McKesson business; Meditech; Flatiron; and ThymeCare. CVS Health and Athenahealth are working to promote these steps in their work as well."

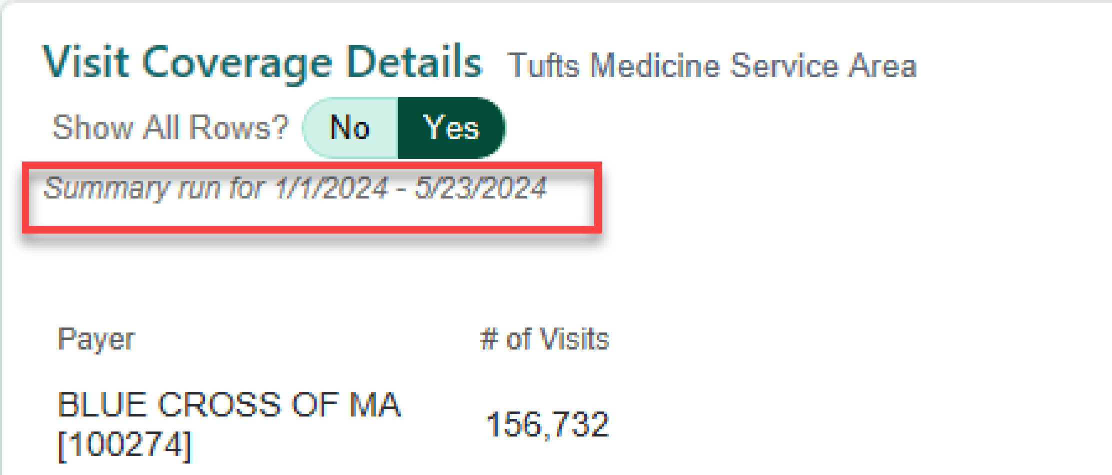
Source: https://www.whitehouse.gov/ostp/news-updates/2024/03/05/improving-cancer-care-through-better-electronic-health-records-voluntary-commitments-and-call-to-action/
https://github.com/CodeX-HL7-FHIR-Accelerator/mcode-lite

CareJourney

#3: TEFCA for Population Health

Health Care Operations (HCO) SubXP-1: means transactions for any of the following activities, under TEFCA Exchange, to the extent permitted by Applicable Law and the Common Agreement:

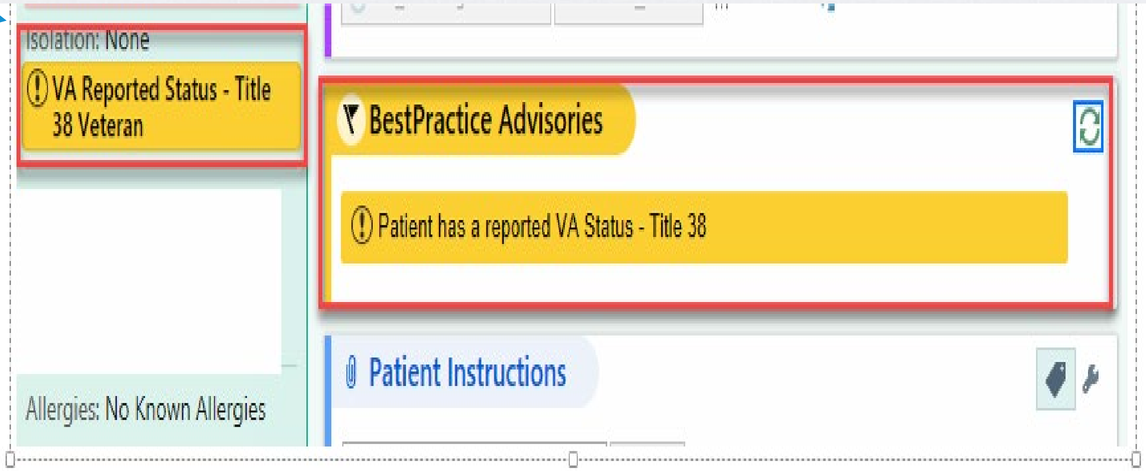
> Conducting quality assessment and improvement activities, including outcomes evaluation and development of clinical guidelines, provided that the obtaining of generalizable knowledge is not the primary purpose of any studies resulting from such activities; patient safety activities (as defined in 42 CFR 3.20); population-based activities relating to improving health or reducing health care costs, protocol development, case management and care coordination, contacting of health care providers and patients with information about treatment alternatives; and related functions that do not include treatment.[1]



Veteran Interoperability Pledge

The Veteran Interoperability Pledge works toward developing a framework to allow VA and community providers to securely exchange information to assist in the care of Veterans receiving treatment inside and outside VA.

> *"With commitments to transfer vital information and records electronically between VA and signatory health systems, we also hope that this this pledge will make it seamless for our partner health systems to identify Veterans at the point of care," **said VA Under Secretary for Health Dr. Shereef Elnahal**. "That is inherently valuable for the Veteran receiving care, but it will also allow us to send helpful information to our partner health systems that they can then offer to Veterans in their care — to include information about new benefits we are offering under the PACT Act and other resources that assist with suicide prevention and identifying social risk factors"*
>
> – VA Under Secretary for Health, Dr. Shereef Elnahal



Source: https://rce.sequoiaproject.org/wp-content/uploads/2024/01/Draft-SOP-XP-Implementation-Health-Care-Operations-SubXP1-508-Compliant.pdf; https://www.ncqa.org/bulk-fhir-api-quality-coalition/; https://www.va.gov/health/veteran-interoperability-pledge.asp; Tufts Medicine; "TEFCA" = Trusted Exchange Framework & Common Agreement

CareJourney

#4: Request Price (Bundle) Estimates

OMB Control Number [XXXX-XXXX]
ExpirationDate [MM/DD/YYYY]

[NAME OF PROVIDER OR FACILITY]

Good Faith Estimate for Health Care Items and Services

Patient		
Patient First Name	Middle Name	Last Name
Pa		
Pa		
Pa		
Str		
Cit		
Ph		
En		
Pa		
Pa		
Primary Service or Item Requested/Scheduled		
Patient Primary Diagnosis	Primary Diagnosis Code	
Patient Secondary Diagnosis	Secondary Diagnosis Code	

MS033

Total Hip Replacement with Optional Grafting Surgeries

PROJECT CLARITY SUPPORTED ■ MSK

SSP Beta Short Consumer-Friendly Description

Total Hip Replacement with Optional Grafting Surgeries

Facility Fees

27 Fees

Total Hip Replacement with Optional Grafting Surgeries 100% Association Index $17818 Estimated Charge	27130 CPT
Dynamic One-on-one Therapeutic Activity to Improve Functioning, 15 Minutes Each 100% Association Index $490 Estimated Charge	97530 CPT
Implantable Joint Device For Motion Restoration 89% Association Index $19119 Estimated Charge	C1776 HCPCS
Operating Room Services - General 84% Association Index $20766 Estimated Charge	0360 Revenue Code
Medical/surgical Supplies and Devices (also See 062x, an Extension of 027x) - Other Implants 82% Association Index $13596 Estimated Charge	0278 Revenue Code

Professional Fees

2 Fees

Total Hip Replacement with Optional Grafting Surgeries 100% Association Index $5057 Estimated Charge	27130 CPT

"FHIR is already being used to support electronic data exchanges among providers, payers, and patients, and **may allow a consumer friendly AEOB** to be produced that could encourage important discussions between patients and their care teams regarding cost and value." – Administrator Brooks-LaSure

Source: Project Clarity, https://www.cms.gov/blog/interoperability-and-connected-health-care-system; "AEOB" = Advanced Explanation of Benefits

6

#5: "Opt-In" for Navigation, Alignment

MARCH 08, 2024

FACT SHEET: Biden Cancer Moonshot Announces Commitments from Leading Health Insurers and Oncology Providers to Make Navigation Services Accessible to More than 150 Million Americans

Navigators are **"opt-in"** services that **connect** patients with community resources, care transition services, behavioral health support, **identify** appropriate providers for clinical care, and helps secure appointments, **track health outcomes** such as **ER & urgent care visits**, **patient-reported outcomes**, and other **care quality and experience** measures.

"…consumers have access to their own health data – and to the applications and services that can safely and accurately analyze it…" – President Obama (January 2015)

Source: https://www.whitehouse.gov/ostp/news-updates/2024/03/08/fact-sheet-biden-cancer-moonshot-announces-commitments-from-leading-health-insurers-and-oncology-providers-to-make-navigation-services-accessible-to-more-than-150-million-americans/

Listening Session 3: Addressing Challenges Regarding Data, Benchmarking, and Risk Adjustment

John Supra, MS

Chief Digital Health & Analytics Officer,
Value-Based Care Institute, Cone Health

Pathways on the Value-Based Care Journey toward Open Data Exchange and Shared Analytics

John Supra

Chief Digital Health & Analytics Officer

Value-Based Care Institute, Cone Health

A Data Path to Value-Based Participation

Where do I Start with Data & Analytics?

- Receive data from multiple sources, often new unfamiliar ones
- Report data back to various sources
- Engage with vendors – selection, integration, and learning
- Understanding of terms and concepts related to value-based care and contracts

Quality Reporting	Financial Reporting	Operational Reporting

Clinical Data	Payor Data	Program Data	3rd Party Data

A Data Path to Value-Based Participation

Where do I Start with Data & Analytics?

- Receive data from multiple sources, often new unfamiliar ones
- Report data back to various sources
- Engage with vendors – selection, integration, and learning
- Understanding of terms and concepts related to value-based care and contracts

Requires Artisan Craftsmanship

- Requires building expert skills and knowledge in a variety of data and analytics areas
  - Data types and methodologies not common in the practice of medicine and traditional FFS billing
- Demands investments that require significant upfront costs
- Develops a reliance on a single or patchwork of vendors (experts)

Unintended Consequences

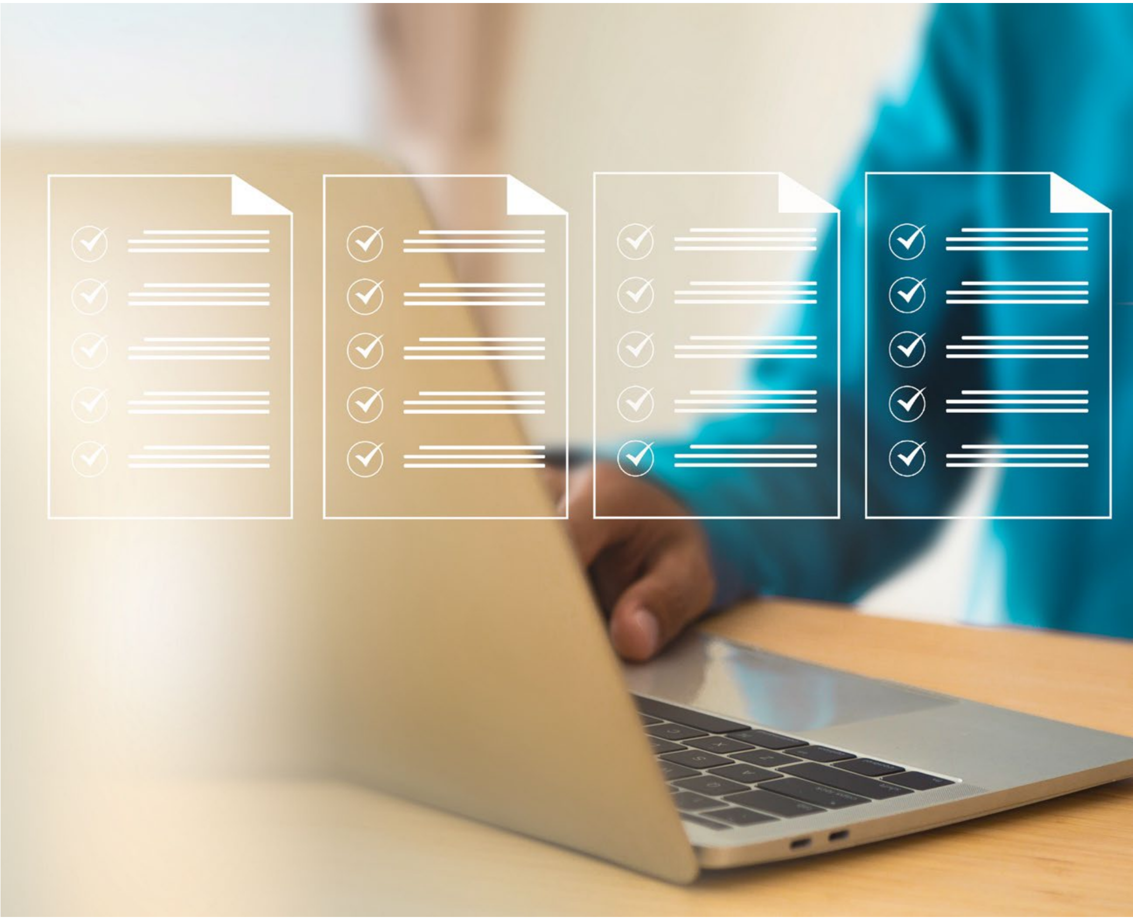
> *"Value-based contracting is intended to incentivize care improvement, but it is unlikely a clinician or practice can reasonably optimize against 50 or more measures at a time."*

> *"We found saturation of the quality measure environment as a possible explanation: average physicians were incentivized to meet 57.08 different quality measures annually."*

Value-Based Contracting in Clinical Care
Claire Boone, PhD; Anna Zink, PhD; Bill J. Wright, PhD; Ari Robicsek, MD
JAMA Health Forum. 2024;5(8):e242020. https://jamanetwork.com/journals/jama-health-forum/fullarticle/2822685

Standardization – A Strong First Step

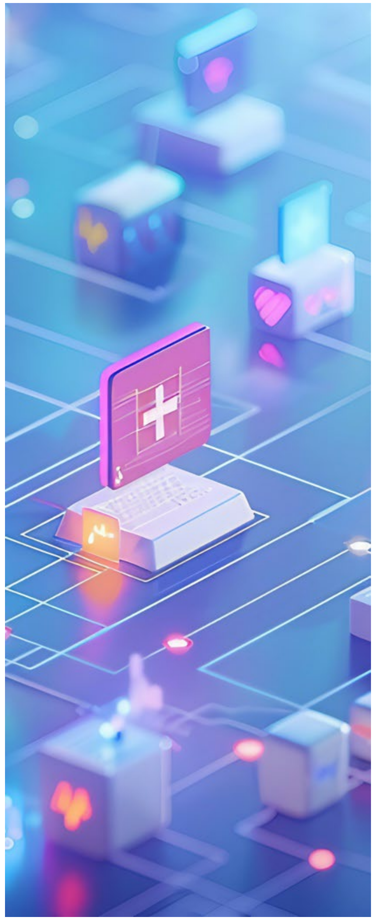
- Standard Data Models
  - United States Code Data for Interoperability (USCDI)
- Standard Data Exchange Specifications
  - Fast Healthcare Interoperability Resources (FHIR)
- Common Framework for Data Exchange
  - Trusted Exchange Framework and Common Agreement (TEFCA)



Data Services – A Good Second Step

- Meaningful Progress
  - Beneficiary Claims Data API (BCDA) https://bcda.cms.gov
  - 4 Innovation API (4i) https://developer.4innovation.cms.gov
  - Claims Data to Part D Sponsors (AB2D) https://ab2d.cms.gov
- Accelerate Access to Data
- Enable System-level Integration
- Still Require "Craftsmanship" to Leverage Benefits

Need for Health Data & Analytics Ecosystems



- **Data Sharing Approaches** – Healthcare data exchange remains dominated by point-to-point (sharing of specific files). Shift toward enabling open, standards-based, secure frameworks to replace the point-to-point exchange
- **Modernize** –Use of modern technologies and cloud data platforms to reduce and eliminate the reliance on an ETL/ELT mindset
- **Easy-Onboarding** – Reduce the burden and ”ramp-up” on providers and ACOs to get engaged in value-based care programs and non-value-add duplicative efforts

CMS Innovation Center Key Takeaways

- Timing and Frequency Valued More Than Perfect Data
- Participant Heterogeneity
- Data-Sharing Heterogeneity
- Context is Key
- Learning Data System Needed
- Data As a Burden

Improving Participation in Value-Based Care–The CMS Innovation Center's Data-Sharing Strategy Initiative
William J. Gordon, Zoe Hruban, Velda McGhee, Todd Couts, Purva Rawal, Elizabeth Fowler
Health Affairs Forefront. August 21, 2024. *https://www.healthaffairs.org/content/forefront/improving-participation-value-based-care-cms-innovation-center-s-data-sharing-strategy*

Opportunities for Alignment

CMS Data Availability	Open Model Metrics

• Accelerate the speed at which data is made available to VBC model participants
• Shift toward data-system-ready reporting
  • Shift CMS standard reporting to include data ingest formats
  • Availability of CMS standard reporting and data feeds as secure data shares
	• Require value-based care model metrics to have an open-source code to run over specific data sets, these may include operational proxies
	• Facilitate and/or incentivize open-source data applications (tools) that leverage standard data models and sources