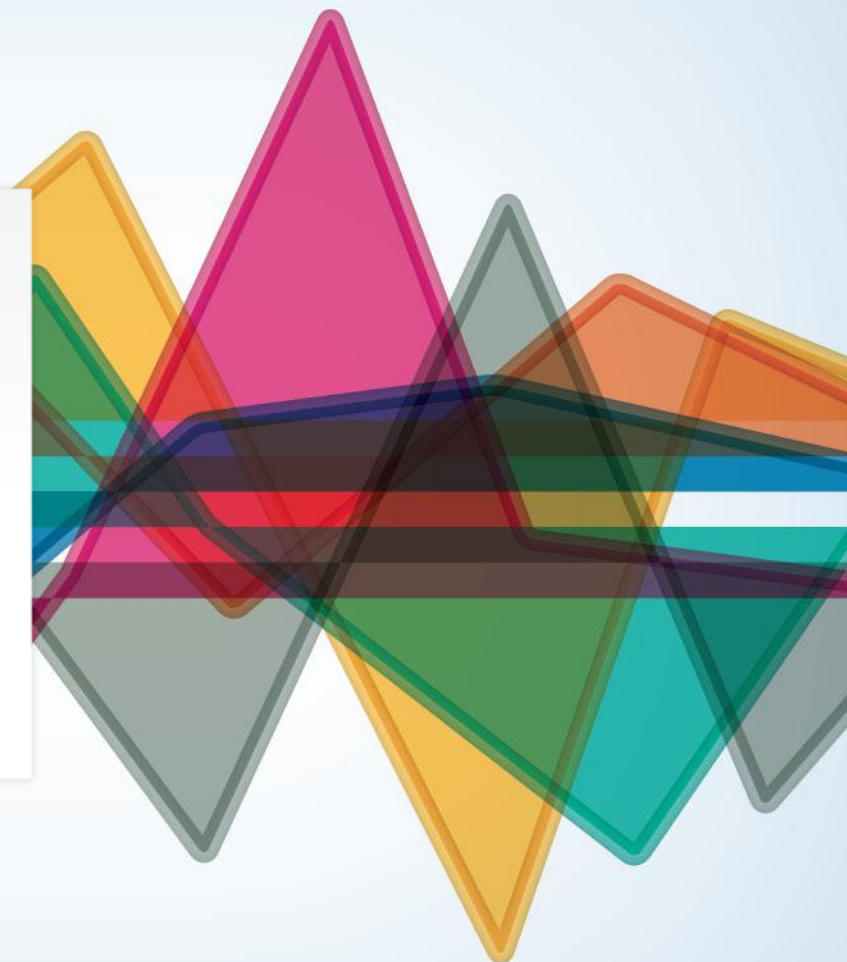


Risk Adjustment in Primary Care

Arlene Ash, Ph.D.

April 13th, 2011



Acknowledgements

My talk today draws upon a large body of work conducted over decades with substantial contributions from

- Randall P. Ellis (Boston University Dept. of Economics)
- Other colleagues at BU
- Colleagues at the University of Massachusetts Medical School and Medical Center (Worcester, MA)
- Scientists at Verisk Health, Inc. (formerly, DxCG, Inc.)
- Many, many other national and international colleagues



- *History*
- *DxCG Classification and Modeling*
- *Lessons from Medicare Implementation*
- *Risk-Based Comprehensive Payment for the Patient-Centered Medical Home – Paying for Primary Care*
- *Risk-based Bonus Payments*

Medicare Risk Adjustment Timeline

1982	TEFRA enables Medicare beneficiaries to enroll in HMOs
1984 - 1994	HCFA funds several teams and projects (DCGs, ACGs, CDPS, etc)
1993	Clinton health initiative- Additional funds to expand risk assessment to non-Medicare populations
1996	First SOA Study compares Models DxCG founded by DCG researchers (later, Verisk Health)
1997	BBA mandates risk adjustment for Medicare HMOs by 2000
2002	Principal Inpatient DCG model implemented (10%) Second SOA study
2004	All encounter DCG model (30%) implemented
2007	Fully phased-in (100%)

Additional History

- 2001- Pharmacy models developed in cooperation with Kaiser Permanente – RxGroups models use NDC codes to predict costs
- 2005 – Models adapted to include ICD-10 diagnoses and ATC drug coding
- 2006 –German Physician Payment Commission selects DxCG for health risk assessment
- 2008 –German Insurance Office selects DxCG models for risk adjusting its new national health care fund
- 2009 – DxCG Models used for risk adjustment in all sickness funds in Germany
- 2010 - Development of primary care payment and performance assessment models

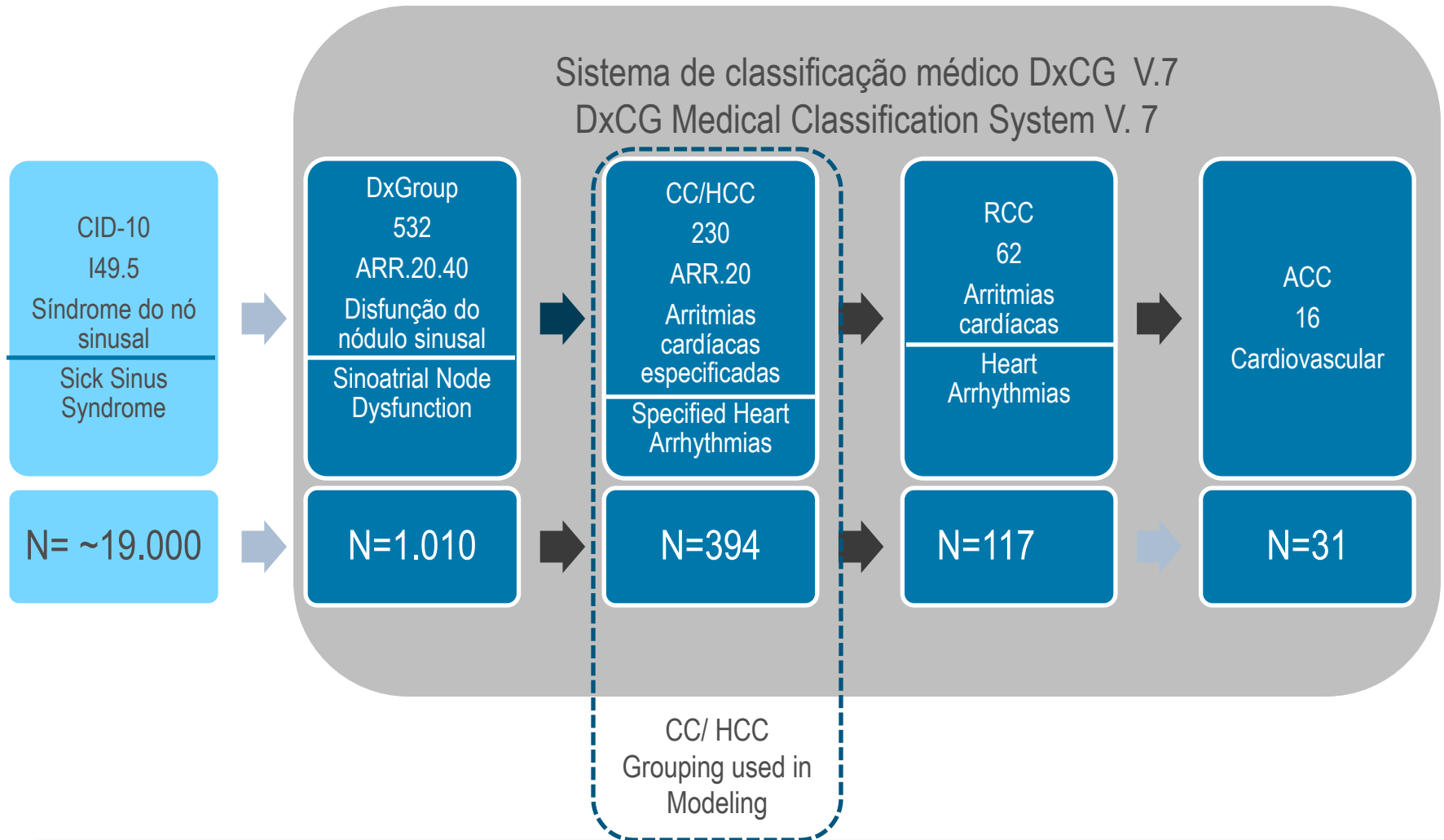
DxCG Modeling Principles

- Principles for diagnostic categories and payment models:
 - Be clinically meaningful
 - Predict health care costs
 - Have adequate sample size
 - Impose dominance relationships only among related conditions
 - Encourage specific coding
 - Not reward coding proliferation
 - Not penalize for recording an additional diagnosis
 - Classify all diagnostic codes
 - Omit discretionary or ambiguous categories from payment models
- Hierarchical structure is transitive
 - If A dominates B and B dominates C, then A dominates C)

DxGroups & Condition Categories

- A DxGroup is a collection of ICD-9/ICD-10 codes that describe closely related clinical conditions
- A single person could have codes that map to 0, 1 or many DxGroups
- DxGroups are the building blocks of DxCG models
- CCs are groups of clinically-related DxGroups with similar implications for resource use

DxCG Diagnosis Grouping v7



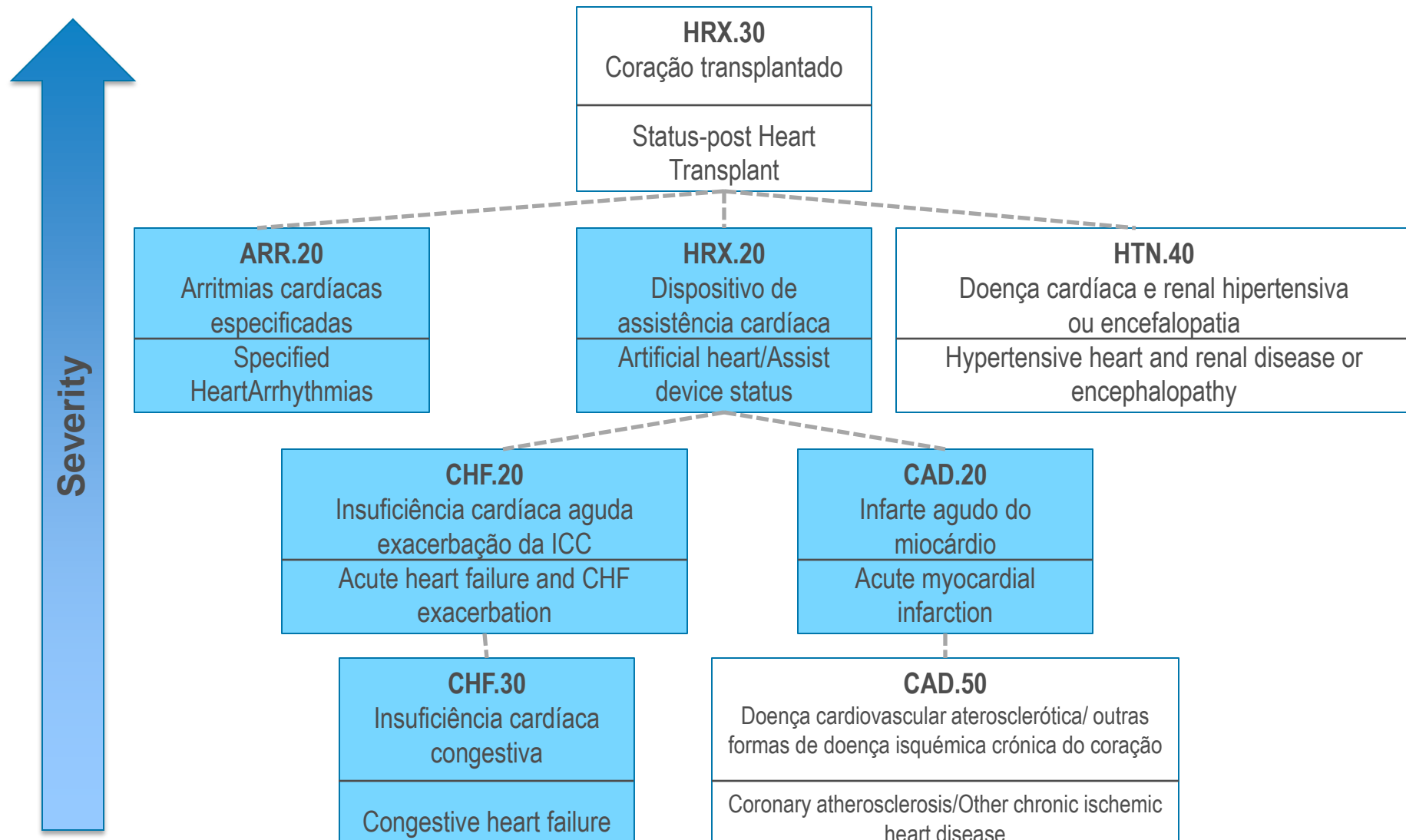
HCCs versus CCs

- There is one and only one HCC for each CC

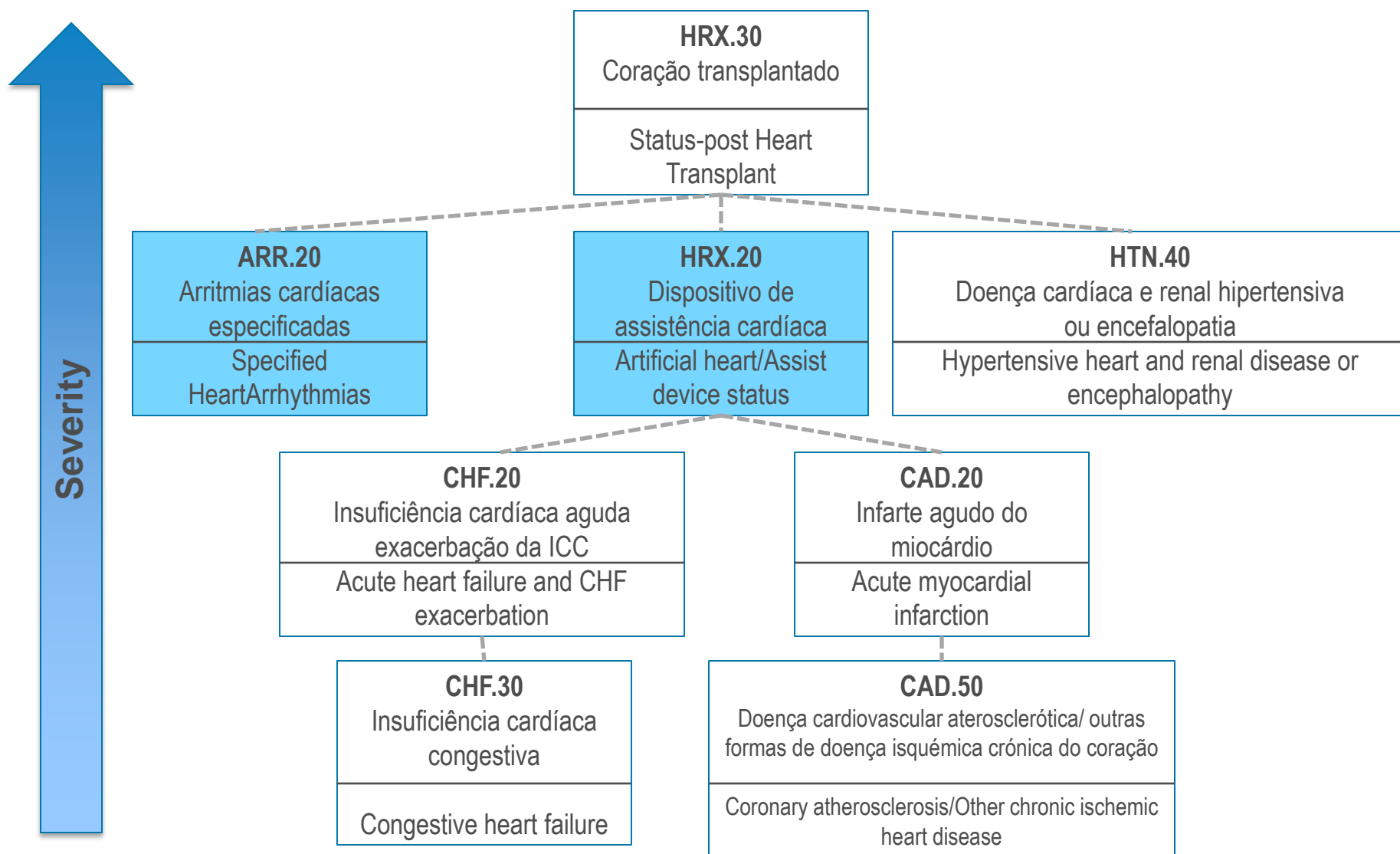


- Differences:
 - HCCs result when *hierarchies are applied to a person's **vector of CCs***
 - HCCs drive the analytic model
 - HCCs retain the most severe version of a disease process for modeling, for example, when a woman has both chronic obstructive lung disease (COLD) and cough :
 - Both her CCs are coded as 1
 - However, COLD > cough, so her HCC for COLD is 1 but her HCC for cough is 0

Heart CCs – Hierarchy Has Not Been Applied



Heart HCCs – Hierarchy Has Been Applied



Condition Categories Also Classified by Chronicity

Chronic

All chronic diagnosis codes were identified based on AHRQ designated “chronic” diagnosis codes.

Non-chronic

- Acute condition, no lasting effect
- Acute exacerbations of chronic conditions
- Initial event leading to chronic condition
- Other

Examples:

Migraine and tension headache

Other headache

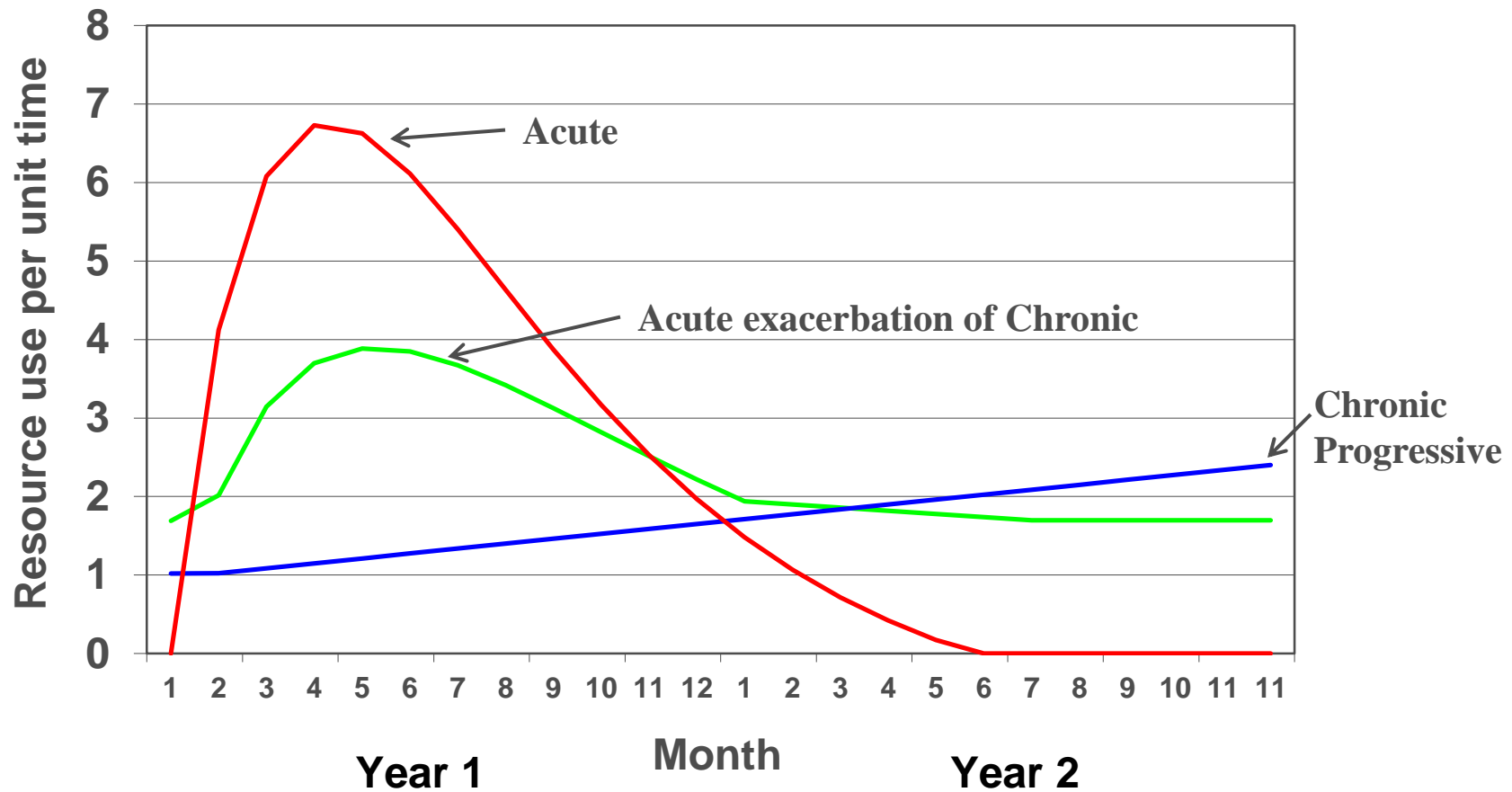
Chronic valvular & rheumatic heart disease

Acute rheumatic heart disease and chorea

Congenital heart and circulatory anomalies

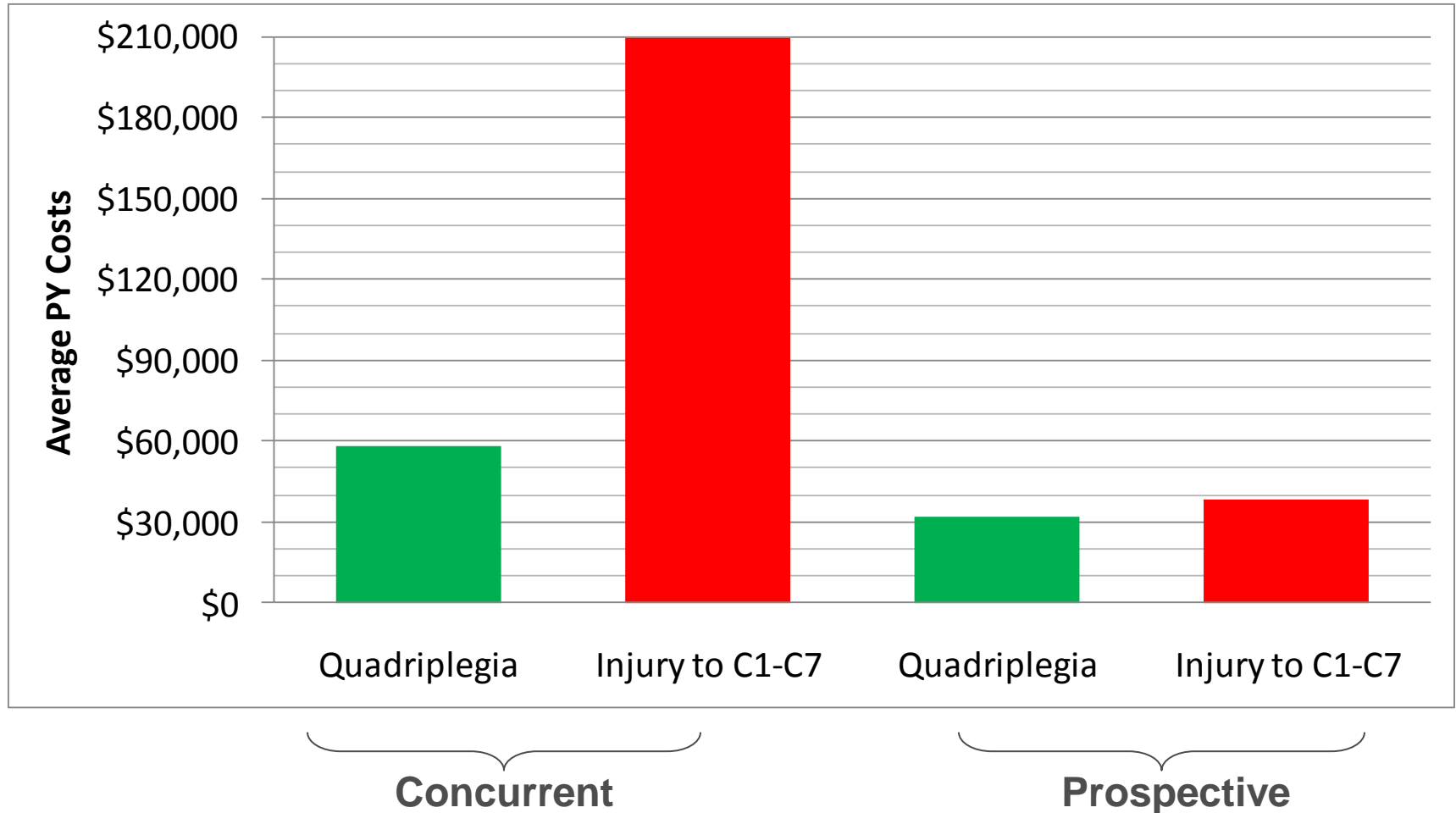
Transient neonatal circulatory conditions

Hypothetical Time Course of Resource Consumption

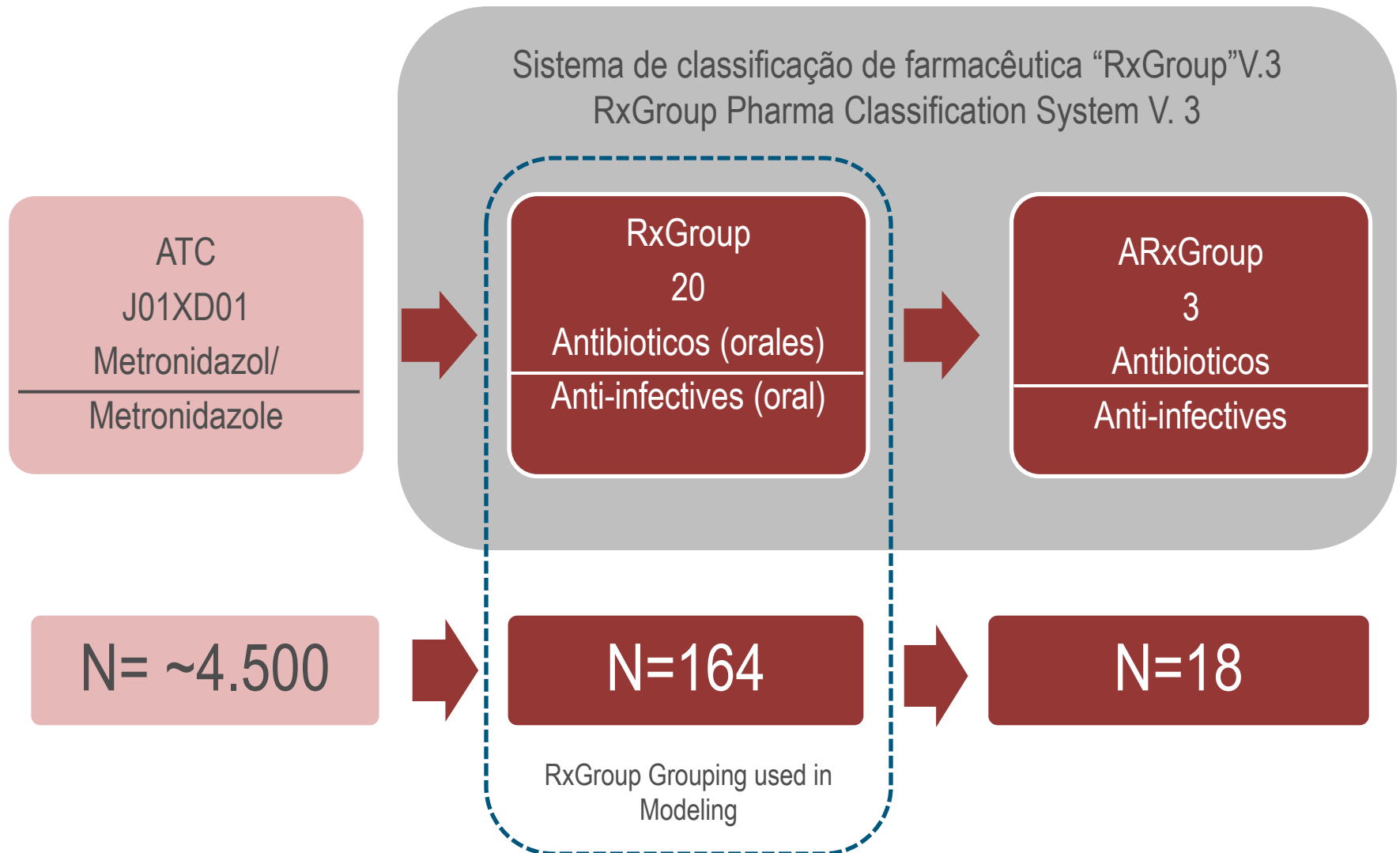


Non-Chronic Code Examples

Incident Case Leads to Chronic Condition?



DxCG Pharmacy Grouping



The Value of the CC or HCC Profile

- Compactly represents millions of distinct comorbidity patterns
- Retains full information for individuals with:
 - Multiple, clinically un-related conditions
 - E.g., asthma and breast cancer
 - Clinically related co-morbid conditions
 - E.g., diabetes and heart disease
 - Unrelated complications and uncommon condition pairings
- HCCs eliminate redundancy and “noise”
 - Focus on the serious conditions that drive health and utilization

Multiple Uses for Diagnostic Code Groupings

- Populating disease registries
- Selecting people for care management programs
- Predicting utilization and cost, e.g.,
 - Hospital admissions
 - Emergency room visits
 - Use of high-tech imaging
 - Use of high-cost “antibiotics of concern”
 - Total, outpatient, pharmacy or other “bundles of” cost
 - Use of primary care services “primary care activity level (PCAL)”
- Predicting processes or outcomes, e.g.,
 - Mammography use
 - Receipt of flu vaccination
 - Level of glucose control for people with diabetes

HCCs linear regression model form

$$\begin{aligned} \text{PCAL} = & \beta_{0_{agecat}} + \beta_1 HCC_1 + \beta_2 HCC_2 + \beta_3 HCC_3 + \dots + \beta_n HCC_n \\ & + (Age < 2) * (\beta_{i0} + \beta_{i1} HCC_1 + \dots + \beta_{in} HCC_n) \\ & + (Age < 18) * (\beta_{k0} + \beta_{k1} HCC_1 + \dots + \beta_{kn} HCC_n) \\ & + (Age > 65) * (\beta_{e0} + \beta_{e1} HCC_1 + \dots + \beta_{en} HCC_n) \\ & + (\text{Selected interactions of HCCs}) \end{aligned}$$

Models predict annualized spending and weight by fraction of the year eligible

LESSONS FROM IMPLEMENTING DCGS FOR MEDICARE

Lessons from Medicare Implementation

- The Health Care Financing Administration (HCFA then, CMS now)
 - Was far ahead of its time
- In 1984 other payers' data were much inferior
 - Family contracts did not have positive enrollment of individual family members
 - You needed to pull paper records to find, say, people with diabetes
 - Diagnostic coding, even during hospitalizations, was incomplete and unreliable (DRG payment soon changed that)
- Broad lessons
 - Coding improves when it is used
 - Systems must be monitored to identify and address changes in both coding and medical practice

Lessons: We Know How Deal With “Code Creep”

- Facilitate *appropriate* upcoding
- **Track/adjust for (differential) trends in coding**
- Monitor (investigate/punish?) suspicious/non-credible coding patterns
- Create disincentives to over-coding
 - Hold entities accountable for treating coded illness
 - Provide rewards for favorable trends in population health status

Lessons: The Need for Uniform Protocols for Reporting New Kinds of Data

- Patient-supplied information
 - Key health concerns, HRQOL
 - Who is their provider?
 - Satisfaction with care
- Information from office visits (or patients)
 - Blood pressure, height/weight
 - Smoking status, functional status
 - Preferred language/native language proficiency
- Laboratory findings

Lessons: More New Kinds of Data

- Demographics and barriers/facilitators
 - Education, race, ethnicity
 - Social support, income, food insecurity
 - Insurance type, access to needed services
 - DSS involvement, welfare
 - Immigration status
- Conditions of life and work
 - Housing status (stability, home health hazards, homelessness)
 - Safety (domestic/neighborhood violence)
 - Place of home, work (zip code or census block)

Lessons: Data on (Dis)Enrollment

- Patients and providers should know that “they belong to each other” and not just for
 - Making proper (non-FFS) payments and assigning responsibility for quality, but
 - Hypothesis: enrollment itself may improve health outcomes
- Even when enrollment is clear cut, is there a formal process for getting in or out, so that
 - Clinicians assume responsibility for a patient (can be proactive)
 - Patient knows “where to start looking” for care
 - We will know why (death, dissatisfaction, etc.)

Lessons: Universal Requirements for Data

- Paying for data reporting works
- Everyone must report
 - It is “their business” if data are incomplete
 - Risk adjusted payments and quality reporting will rapidly improve the data
 - Example to NOT follow: Medicare Advantage (plans allowed to report very collapsed summaries)
- All-payor data should be available at the patient-level
 - Encrypted, to researchers and regulatory authorities
 - Unencrypted, to providers for their patients

Lessons: Data and Phase-In

- **NOW**

- Institute payment reforms with existing (limited) data
- Develop expanded uniform reporting protocols
- Announce a timetable for near- and long-term reliance on specified data for payment and quality oversight

- **NEAR FUTURE**

- Implement reforms that use the expanded data, using phase-in periods with non-symmetric risk corridors (**especially important to limit short-term downside potential**)
- Simulate what will happen/Study what does happen
- GRADUALLY remove the downside protections

Lessons: Uniform Data Reporting Was/Is Key

- **Thank goodness for “Uniform Billing” forms and standardized electronic transaction formats!**
 - UB82/UB92/UB04; HCFA-1500/CMS-1500 – paper formats (facility/ professional)
 - ANSI 837 Professional and Institutional – electronic transaction formats
 - Very good models can be built on such data
- Payment reforms began when many organizations did not yet have good electronic data capture but
 - **There should be a structure for richer, uniform reporting**
- Most insurers/providers today have electronic data capture of billing information
 - Transaction Standards section of 1996 HIPAA act has ensured that since Oct of 2003, the vast majority of claims have been submitted electronically (ANSI 837 P/ I)
- **Today’s models and data are good enough to risk adjust payments to insurers**
- **Crucial to build a better data infrastructure** (in parallel) to support CQI in risk-adjusted payment and quality assessment

Risk-Based Comprehensive Payment for the Patient-Centered Medical Home



Paying for Primary Care

- Existing FFS payments for primary care
 - Do not adequately reward innovations
 - Email, phone calls, group visits, non-physician care, etc.
 - Create perverse incentives to increase costs
 - Are not attracting new MDs into primary care
 - Do not directly reveal the cost of high-quality primary care
- Our idea: create a proxy variable, Y , in available claims data; fit a model to predict Y ; pay on model-predicted Y
- PCMH is paid the *expected primary care cost* for each of its patients = **Primary Care Activity Level = PCAL**
 - This is partial capitation, or bundled payment (not-FFS)

Weights Used to Create PCAL Proxy in Commercial (and Medicare) Samples

Type of Activity	% of All Such Costs Contributing to PCAL Proxy	% of PCAL Commercial	% of PCAL Medicare
Primary care core	100%	59%	69%
Specialty-care related	6%	7%	6%
Hospital-care related	6%	8%	18%
ER spending	17%	2%	2%
Rx spending	12%	13%	0%*
Fixed cost component	\$65	11%	5%

Regressions to Predict PCAL

	Commercial Concurrent HCC model	Medicare Concurrent HCC model
Dependent variable mean	\$659	\$1674
R-Square (individual level)	67%	51%
Standard Error	\$536	\$1906

Notes:

- 1) Each WLS regression predicts annualized PCAL spending for individual patients using 34 age-sex + 394 HCC dummy variables + interactions
- 2) Commercial model had $R^2 = 43\%$ when applied to 2006 5% SAF Medicare data

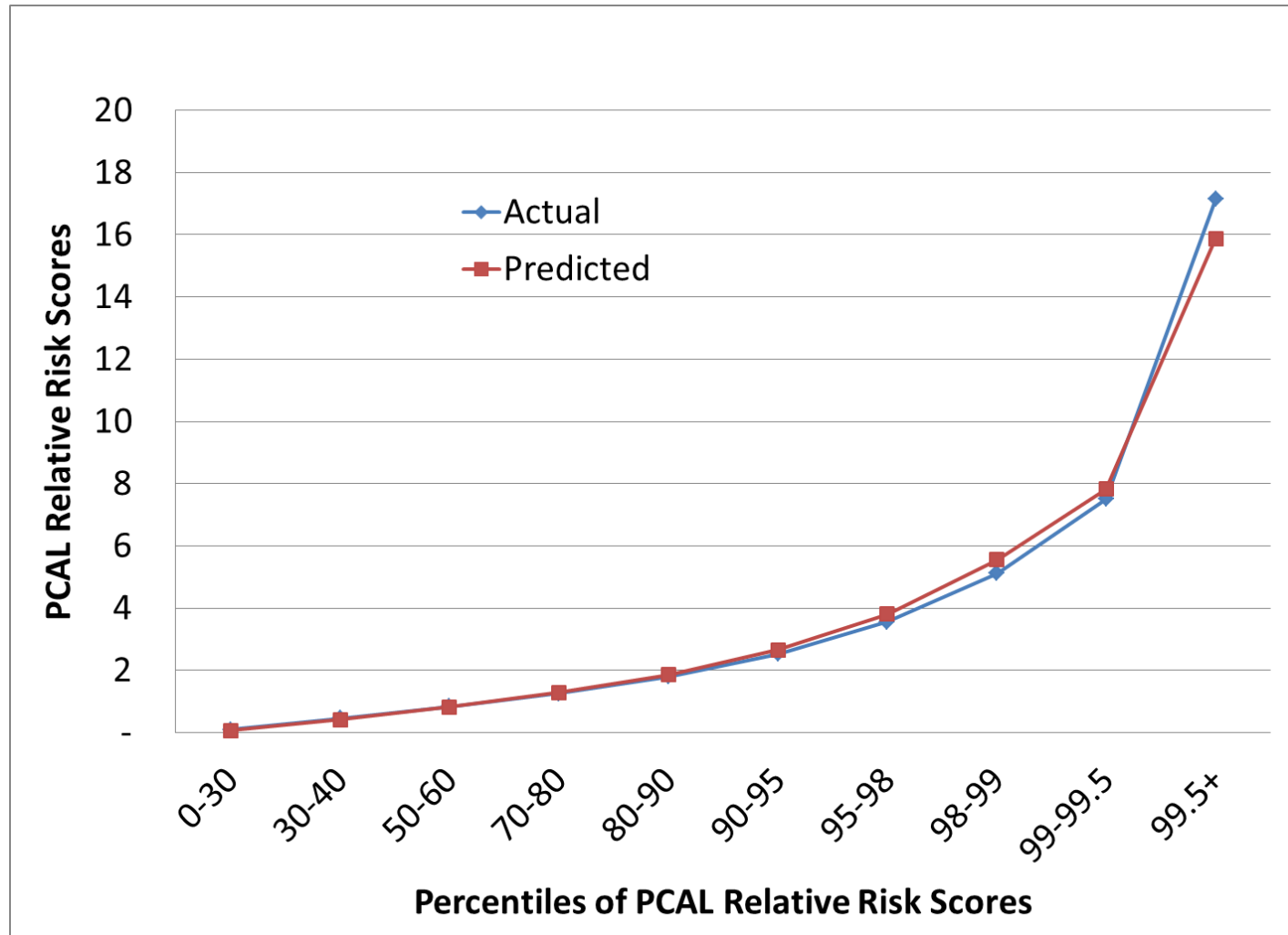
Converting An Expected Outcome to a Relative Risk Score (RRS)

- A model is fit to a benchmark population and produces a prediction (PRED) for each person
 - PRED could be in any units, e.g.,
 - Euros, dollars, # of visits, # of antibiotic prescriptions
 - Expected cost of appropriate primary care
 - Units of self-reported health status
 - Let M_0 = mean of PRED in the benchmark data
 - The RRS for each person = $PRED/M_0$
- Thus, in the benchmark data, the mean RRS is 1
- To convert RRSs to predictions in a new population, just multiply them by M_{new} , the mean in that population

PCAL Payment Examples (Commercially Insured Patients)

Patient	PCAL RRS	Annual PCAL Payment
Male, Age 16	.134	\$65
Female, Age 11 Other Non-Chronic Ear, Nose, Throat, & Mouth Disorders Other Dermatological Disorders Screening/Observation/Special Exams	.411	\$197
Male, Age 56 Benign Digestive or Urinary Neoplasm Diabetes with no complication Fluid/Electrolyte/Acid-Base Imbalance Ulcer with Perforation/Obstruction Screening/Observation, History of Disease	3.360	\$1,613
Female, Age 50 Diabetes with Ophthalmologic Manifestation , Hyperlipidemia Endocrine/Metabolic Disorder , Lower Back Pain Pelvic/Uterine Inflammation ,Rehab Screening , Surgical Misadventure or Complication	4.791	\$2,300

Average Actual Vs. Average Predicted PCAL Relative Risk Score (RRS)



Conclusions for Base Payment Model

- Claims can be used to calculate resources for primary care
 - Our models have used all encounter claims
 - When more complete data are available, it is easy to explore the performance of the models using outpatient data only, or data that are otherwise incomplete
- A practice's PCAL (that is, resource needs for enhanced primary care) is fairly predictable
- Practice-level payments are relatively stable from one year to the next compared to initial FFS or full capitation

Risk-Based Bonus Payments



Bonus for Performance in 3 Domains

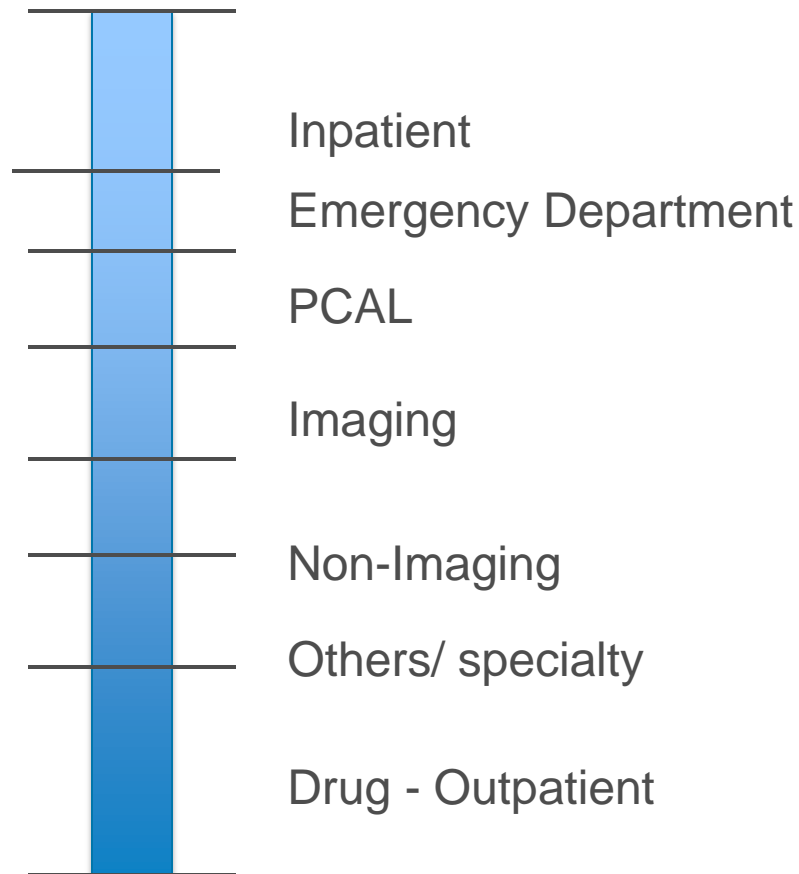
- Goal is to reward:
 - Efficiency (utilization and cost)
 - Quality (processes and patient health)
 - The patient experience (access and satisfaction)
- Bundled base payments (non-FFS)
 - Allows practices to use their judgment to allocate resources
 - Does not explicitly reward desirable behaviors wrt cost, quality and the patient experience
- Key idea:
 - Because performance measures with absolute targets discourage enrolling challenging or complex patients
 - Reward for performance *relative to expected*

Principles for Measuring Performance & Calculating Bonuses

1. Judge performance on observed outcomes in comparison to what is expected given patient characteristics (O vs. E, not just O)
2. Provide opportunities for larger bonuses for larger or more complex panels
3. Weight bonuses to reflect practice case-mix
4. Calculate measures using largest feasible denominators
5. Allow more stable measures to contribute more
6. In groups of correlated measures, down-weight each individual measure
7. Give higher value activities more weight
8. Limit both payer and provider risk
9. Account for extra-medical factors
10. Enable actionable, transparent feedback based on bonus calculations

Performance Assessment Models Framework

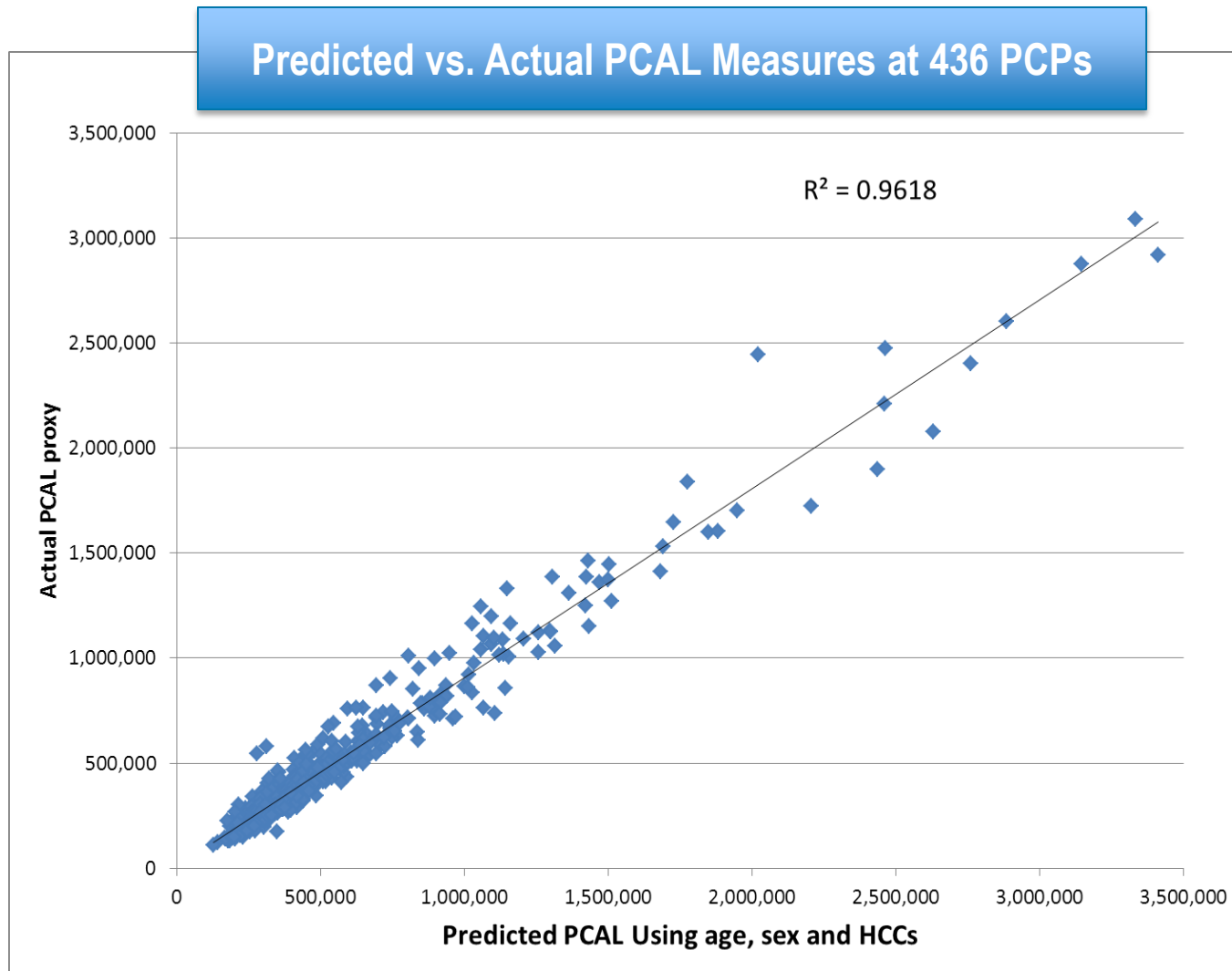
Models are all *concurrent*– determine expected values for current year
Weightings change depending on predicted dependent variable



How Base and Bonus Payments Can Be Scaled to a Practice's Effective Size

Quantity	Notation	Hypothetical Practices	
		B	C
Number of patients	N	1000	1500
Mean relative risk	RRS	1.5	1.0
Effective practice size	$N \times RRS$	1500	1500
Base payment	$N \times RRS \times B$	1500B	1500B
Maximum bonus available for Domain	D	10%	10%
Share of maximum earned	S	80%	60%
Bonus paid for Domain	$S \times D \times N \times RRS \times B$	120B	90B

Actual Vs. Predicted Annual PCAL Dollars for 436 Mid-Size PCPs



Example: Overusing Antibiotics of Concern

- Doctors may over-prescribe **high-cost** or **high-risk** antibiotics as identified by NCQA
- Doctors say: “patients are sicker,” or have conditions that warrant using these “antibiotics of concern”
- Hard to develop clinical guidelines for their appropriate use
- Idea: Use risk adjustment to compare observed use to expected levels, adjusting for patient illness burden

NCQA “Antibiotics of Concern”

Table ABX-C: All Other Antibiotics by NCQA Drug Class			
Description	Prescription		
Absorbable sulfonamide	• sulfadiazine	• sulfasalazine	
	• sulfamethoxazole-	• sulfisoxazole	
Aminoglycoside	• amikacin • gentamicin	• kanamycin • streptomycin	• tobramycin
Cephalosporin (first generation)	• cefadroxil	• cephalexin	
	• cefazolin	• cephradine	
Lincosamide (other than clindamycin)	• lincomycin		
Macrolide (other than azithromycin and clarithromycin)	• erythromycin	• erythromycin lactobionate	
	• erythromycin esto	• erythromycin stearate	
	• erythromycin ethy	• erythromycin-sulfisoxazole	
Penicillin (other than amoxicillin/ clavulanate)	• ampicillin	• penicillin G potassium	
	• ampicillin-sulbacta	• penicillin G procaine	
	• amoxicillin	• penicillin G sodium	
	• carbenicillin	• penicillin V potassium	
	• dicloxacillin	• piperacillin	
	• nafcillin	• piperacillin-tazobactam	
	• oxacillin	• ticarcillin	
	• penicillin G benzat	• ticarcillin-clavulanate	
Tetracyclines	• doxycycline	• minocycline	• tetracycline
Miscellaneous antibiotics	• daptomycin	• nitrofurantoin macrocrystals	
	• fosfomicin	• rifampin	
	• metronidazole	• trimethoprim	
	• nitrofurantoin		

Antibiotic Scripts of Concern

- Verisk Health model (HCCs + age-sex + interactions) predicting number of filled ABX-B prescriptions

	Estimation sample: MarketScan	Validation sample: Pharmetrics
N	17,419,585	2,714,279
Mean	0.45	0.34
SD	1.01	.76
R-Squared	26.5%	25.5%

Lessons: Antibiotics of Concern (ABX)

Findings:

- Only about half of all AB prescribing is for ABX (571 prescriptions per 1000 vs. 1061)
- However, AB prescribing exhibits nearly 3 times the relative variability (CV = 4.72 vs. 1.59).
- Patient-level factors explain about 30% of the individual-level variability in either measure and 95% of variability across patient panels.

Conclusion:

Risk-adjusted ABX is an effective bonus measure because it targets an activity that we want doctors to change, and panels can be adequately adjusted for patient-level differences.

Bonus Payments Discussion

- Risk-adjusted predictions provide credible, empirically-based benchmarks
- Risk models allow incentives to focus on provider-associated, rather than case-mix-based, variations in outcomes
- Practices differ in how much each performance measure matters; bonus calculations should use practice-specific weighted sums of individual measures
- Careful thought and empirical testing needed to guide measure development and decide how individual measures will contribute to the bonus

Conclusions

- This is work-in-progress
- We are partnering with several organizations implementing these models in small and medium-sized pilots
- Preliminary findings are encouraging
- A lot of important work remains to be done

**Thank you for your time and attention
I look forward to our discussions**