Use of Data Mining Methods in Developing a Prospective Payment System for Inpatient Rehabilitation

Greg Ridgeway and Dan Relles

RAND

Santa Monica, CA

History

Balanced Budget Act of 1997 and BBRA of 1999

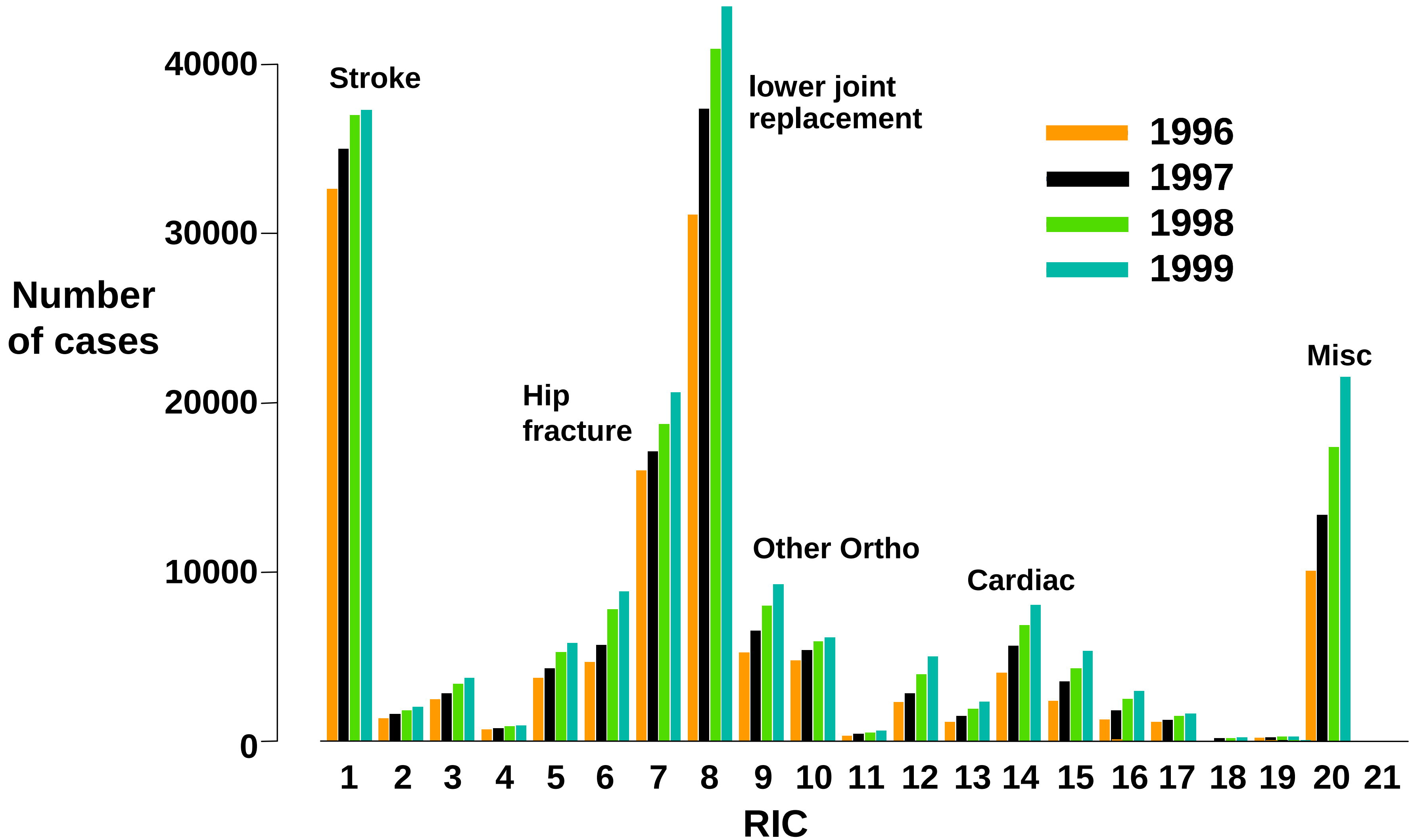
- **Centers for Medicare & Medicaid Services (CMS) must implement a Prospective Payment System for inpatient rehabilitation.**
- **Cases should be classified based on impairment, age, function, comorbidity, and ‘other factors deemed appropriate’**

Medicare data from 1996-1999

- **hospital reported costs**
- **patient disease and functional status data**
- **hospital level data**

We modeled the cost of rehabilitation

Patients Seek Rehabilitation for an Assortment of Impairments



Model

The basic form of the prospective payment system is

$$\text{payment}_{ij} = M \times F_j \times w(\text{age}_i, \text{motor}_i, \text{cognitive}_i) \times c_i \times a_i$$

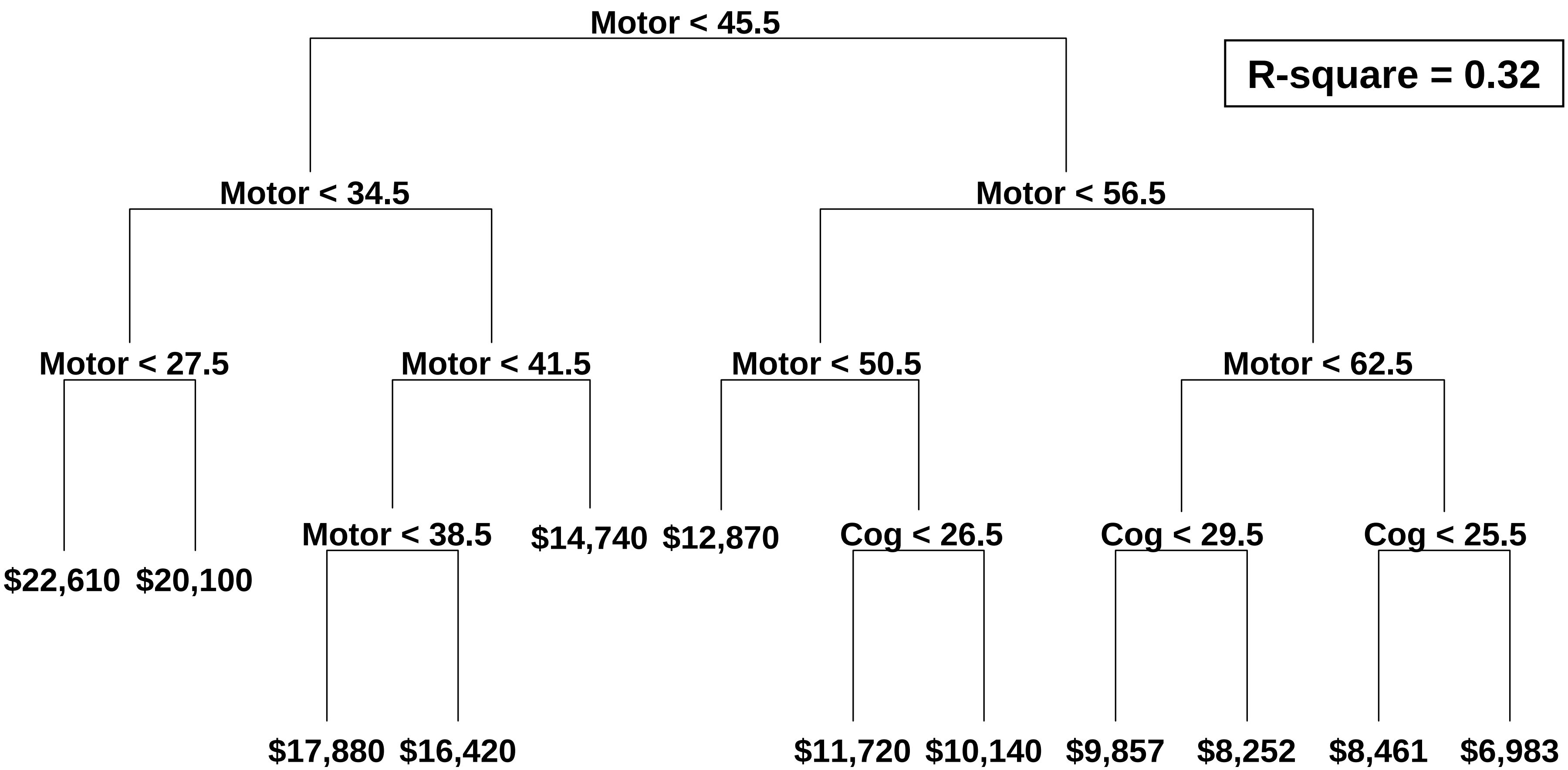
- w **is the main focus of this discussion**
- M **is a fixed budget normalizing constant**
- F_j **is a facility level adjustment**
- c_i **is an adjustment for comorbidities**
- a_i **is an adjustment for "transfer"**
- **Outlier payments will be added for a very small percentage of patients**

Dataset Contents

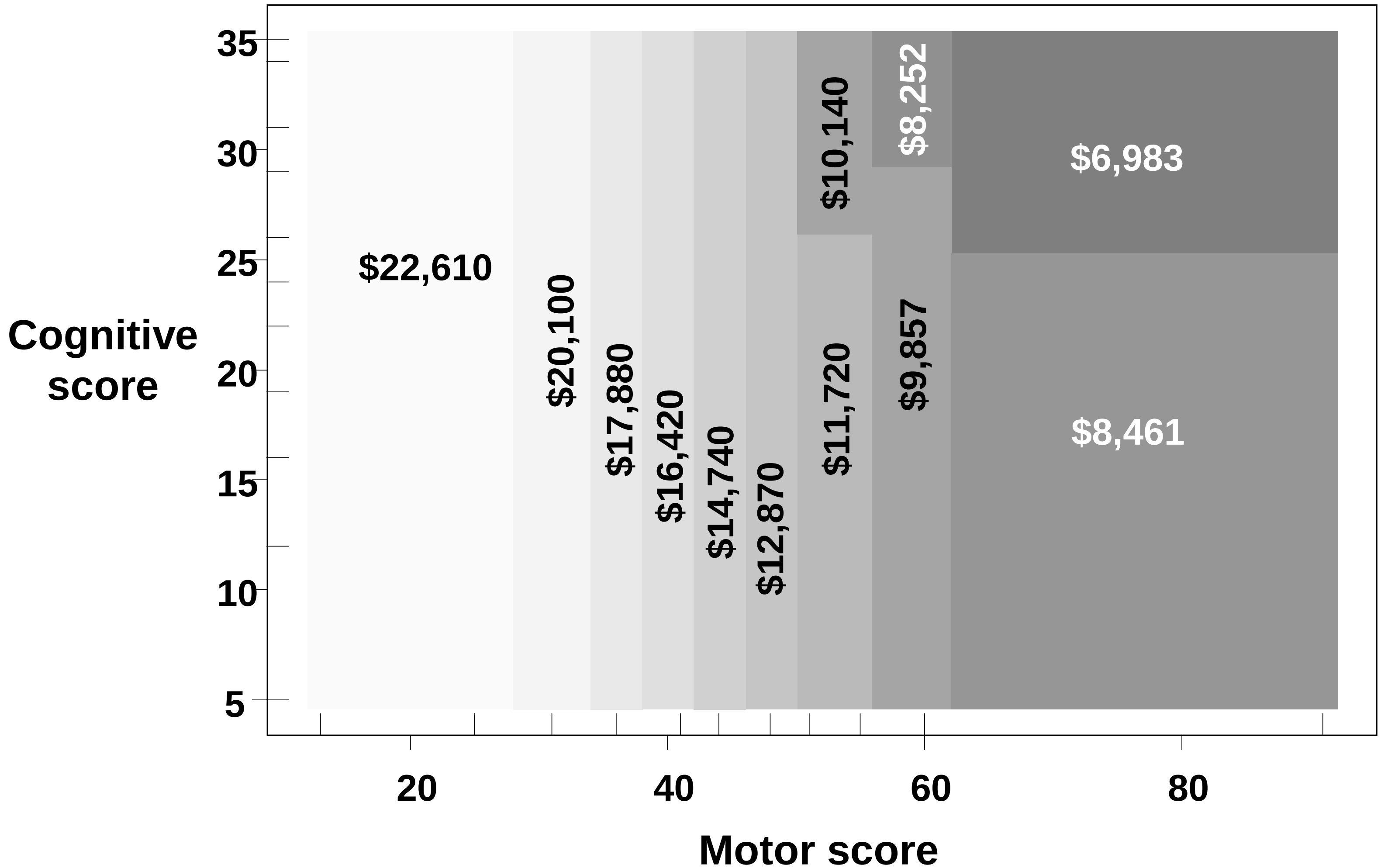
- **Case selection**
  - Cases discharged to the community
  - Eliminate statistical outliers (currently refining)
- **Patient characteristics at admission**
  - Impairment code (maps into 1 of 21 RICs)
  - Age
  - Functional independence measure (FIM)
    - Cognitive FIM components (5)
    - Motor FIM components (12)
- **Measure of resource use**: wage-adjusted cost

CART creates patient classification

A 12-Node Tree for Stroke



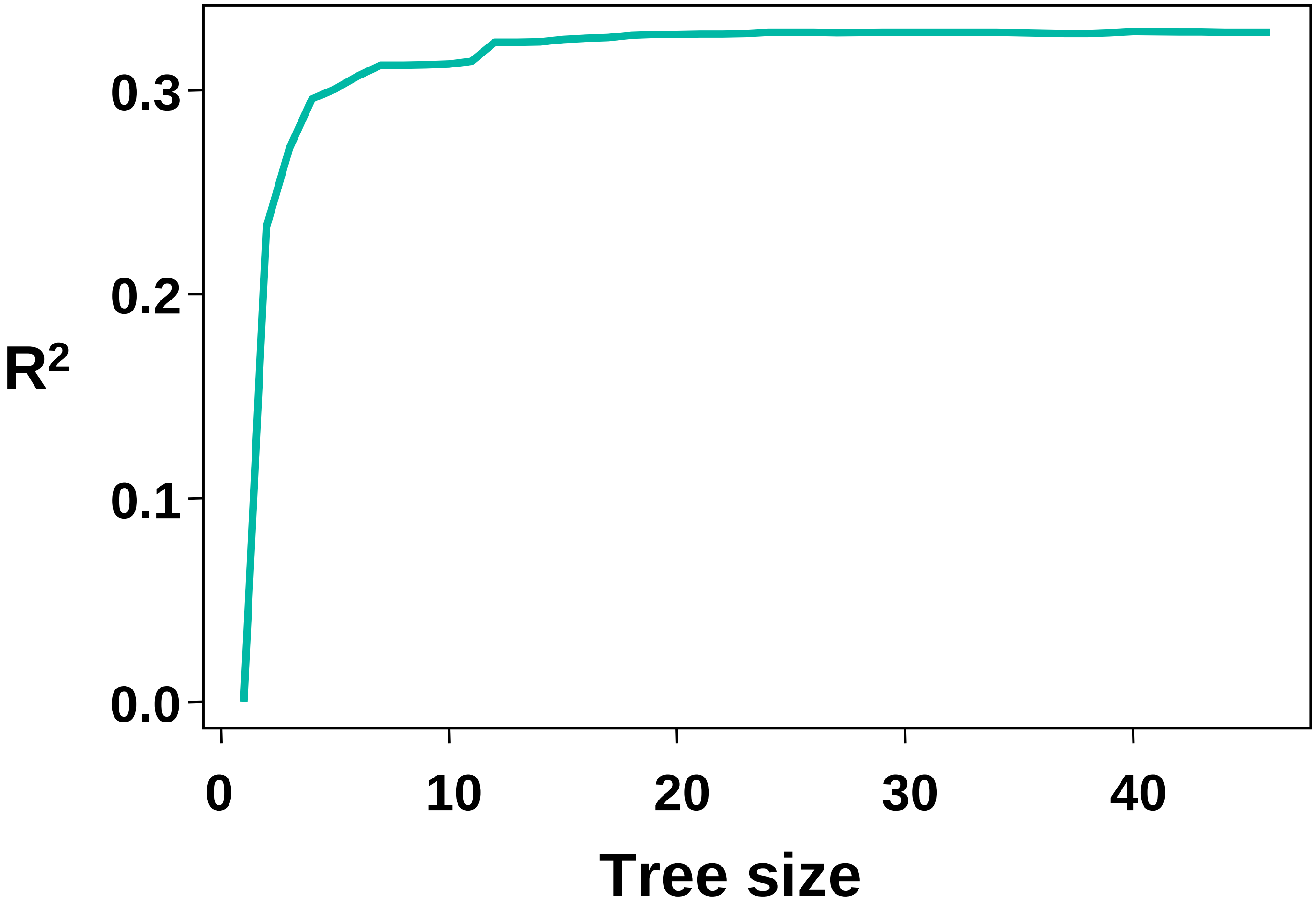
CART Costs



Cognitive score
35
30
25
20
15
10
5
$22,610
$20,100
$17,880
$16,420
$14,740
$12,870
$10,140
$11,720
$8,252
$9,857
$6,983
$8,461
20
40
60
80
Motor score

Gain in R^2 for Larger Trees Tends to Be Minimal

Stroke, 1997



Stopping Rule Dramatically Affects Size of Tree

Fit Year	Max	1 SE
96	218	95
97	244	97
98	333	123
99	325	126
96-97	398	142
98-99	483	180

Choosing the Size of the Tree Balances Policy and Statistics Goals

	Small trees	Large trees
Payment formula	Simple	Complex
Case management	Simple	Complex
Capacity to fit	Low	High
Variance	Low	High

We made further restrictions on monotonicity and eliminating splits with cost estimates that practically did not differ.

Description of Recommended FRGs

RIC		Number of Nodes	Components present
1	Stroke	14	M, C, A
2	Brain injury – traumatic	5	M, C
3	Brain injury – nontraumatic	4	M
4	Spinal cord – traumatic	4	M
5	Spinal cord – nontraumatic	5	M, C
6	Neurological	4	M
7	Orthopedic – Hip fracture	5	M
8	Orthopedic – Replacement of lower extremity joint	6	M, C
9	Orthopedic – Other	4	M
10	Amputation – lower extremity	5	M
11	Amputation – other	3	M
12	Arthritis – Osteoarthritis	5	M, C
13	Arthritis – Rheumatoid, other arthritis	4	M
14	Cardiac	4	M
15	Pulmonary	4	M
16	Pain Syndrome	2	M
17	Major multiple trauma, no brain or spinal cord injury	3	M
18	Major multiple trauma, with brain or spinal cord injury	4	M, C
19	Guillain-Barre	3	M
20	Miscellaneous	5	M, A
21	Burns	2	M
	Total	95	

Problems with Only Considering CART

- Although the FRGs achieve an R^2 of about 0.35, we want to know if that is far from the best achievable
- The number of nodes can be very large and it is difficult to decide when to stop
- We derived our FRGs with a restrictive functional form. We want to know how well they perform

The Computational Experiment

Method ×	Index	× Fit year	× Evaluation Year
	Standard FIM Motor and Cognitive Scores	1996	
CART	Standard Scores – transfer to tub	1997	1996
OLS	Decompose Motor into ADLs and mobility (w/o tub transfer)	1998	1997
GAM		1999	1998
MART	Decompose Motor into transfer (w/o tub transfer), locomotion, sphincter, and self care	1996-7	1999
		1998-9	
	18 FIM Components		

Development of a Gold Standard

- We considered two more flexible, state-of-the-art regression methods and compared their predictive performance to CART's
  - Generalized additive models (GAM)
  - Multivariate adaptive regression trees (MART)
- Methods
- Empirical results

Generalized Additive Model (GAM)

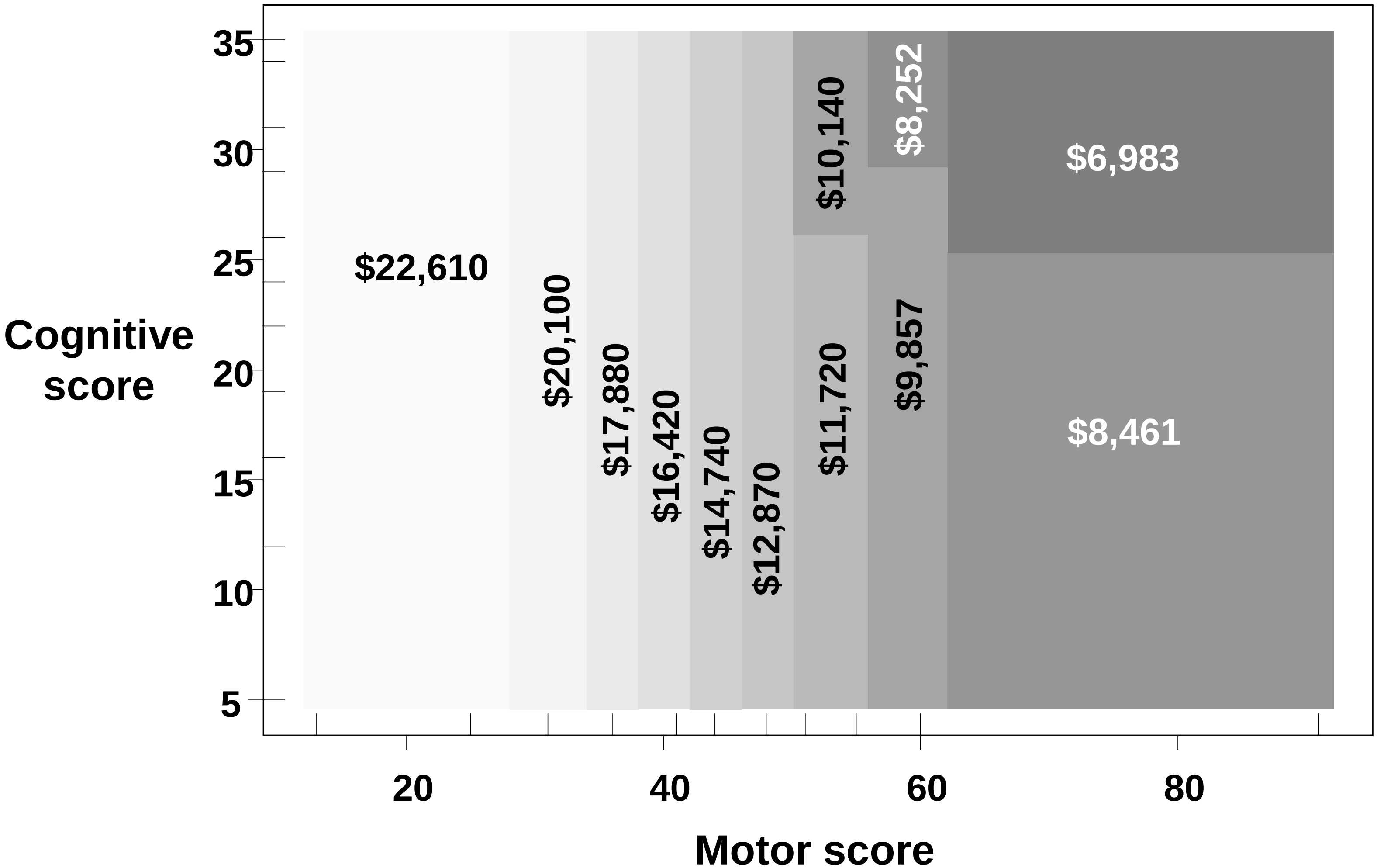
- **CART allows for large jumps in cost curves**
- **We really believe that patients with similar motor scores should have similar costs**
- **Assumes that the log(cost) is the sum of smooth functions of the predictors**

$$\log(\text{cost}) = f_1(\text{age}) + f_2(\text{motor}) + f_3(\text{cognitive}) + \varepsilon$$

- **GAM is designed to find smooth *f*'s that maximize R^2**
- **GAM uses no interaction terms**

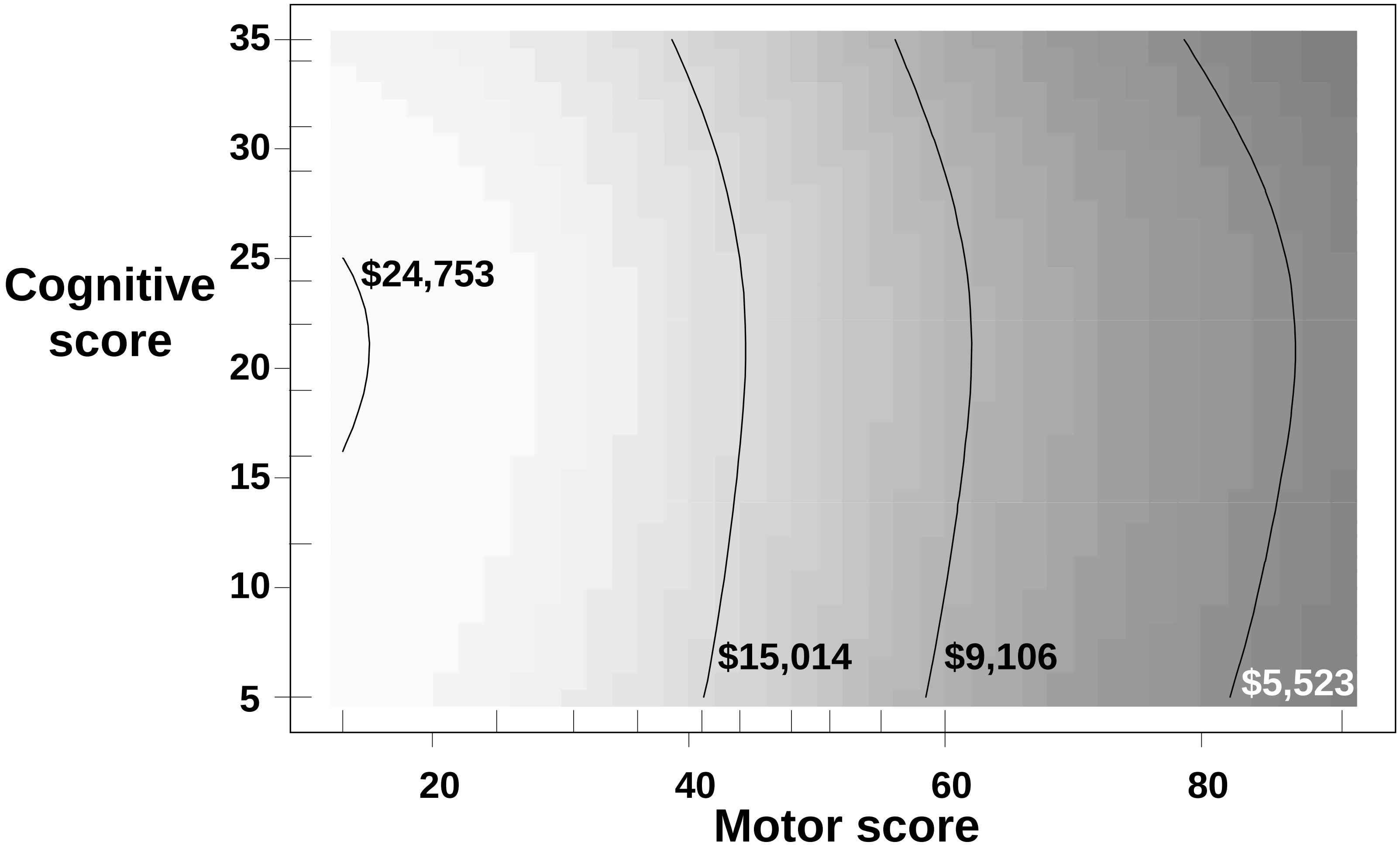
Reprise: CART Costs

Motor and Cognitive, Stroke, 1998-9



GAM Costs

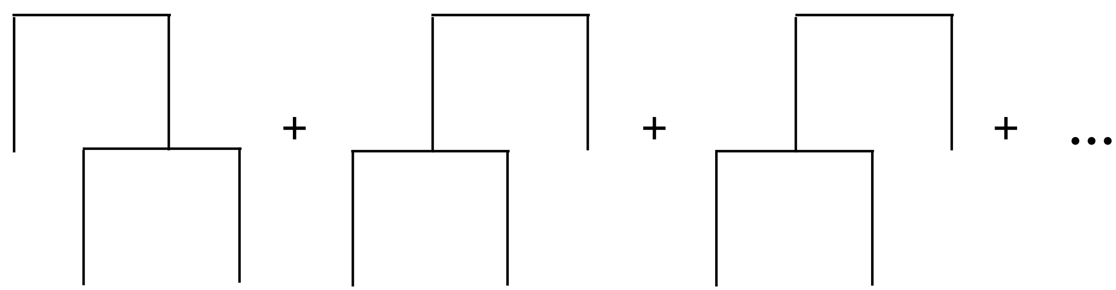
Motor and Cognitive, Stroke, 1998-9



Multivariate Adaptive Regression Trees (MART)

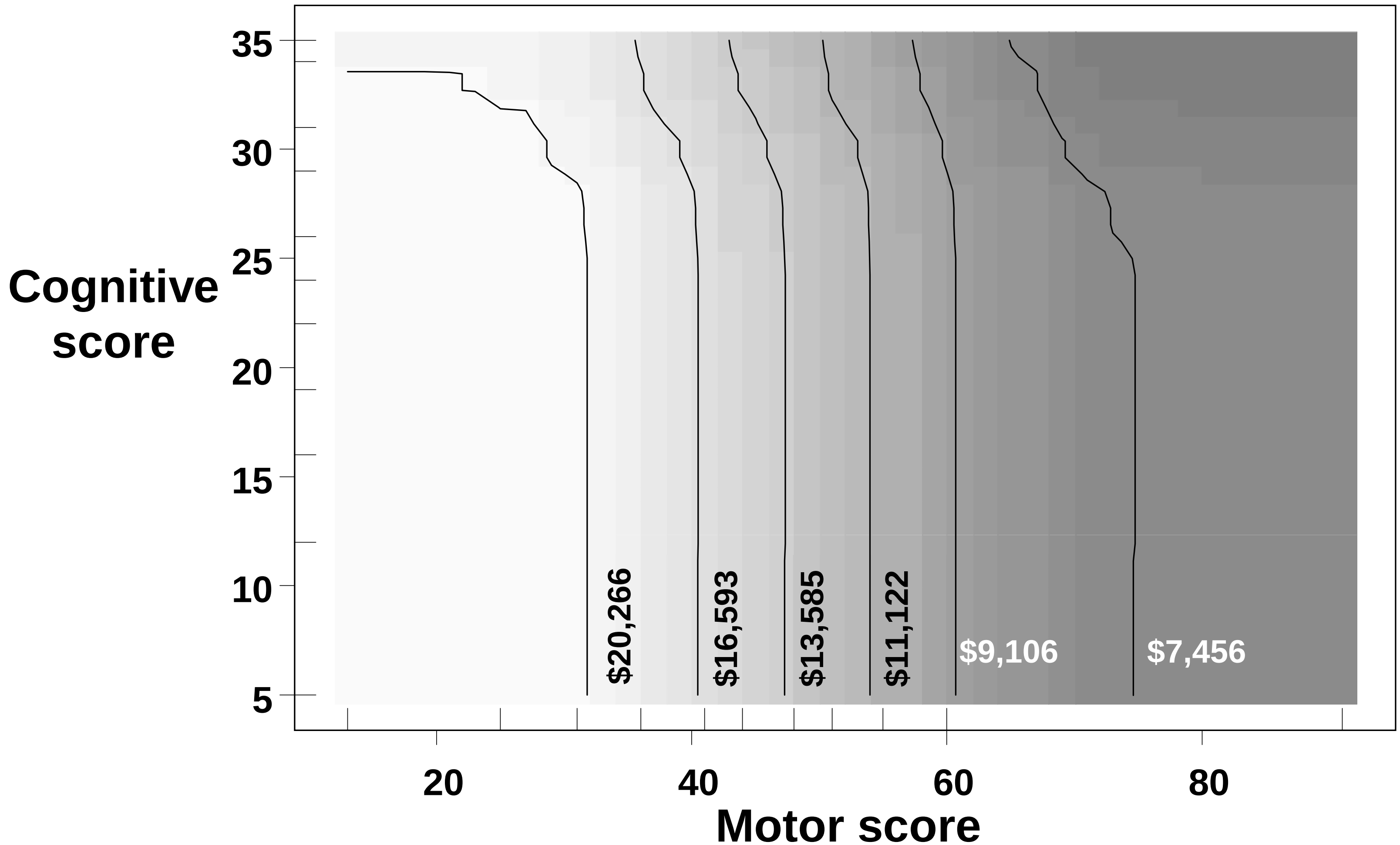
- **Like GAM, MART can find non-linear relationships**
- **It can also find interaction effects in the predictor variables**
- **MART fits an initial, simple CART model, then iteratively fits the residuals with additional CART models**
- **The sum of many CART trees can model complex, non-linear relationships between cost and the predictor variables**

$f(\text{age,motor,cog}) =$



MART Costs

Motor and Cognitive, Stroke, 1998-9



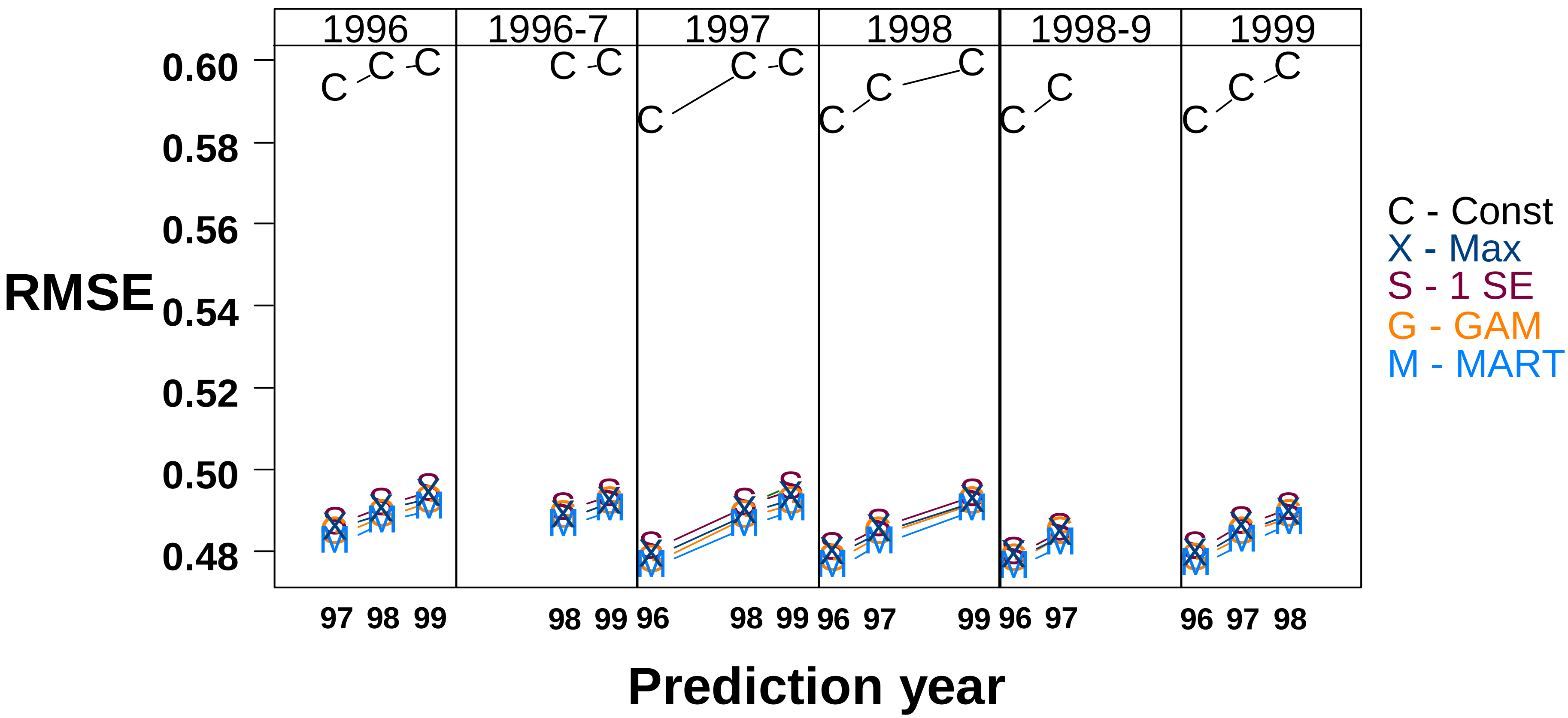
Aggregate Performance of the Various Methods – R^2

Fit Year	Evaluation Year	Const	CART Max	CART 1 SE rule	GAM	MART
96	97	0.16	0.35	0.34	0.36	0.36
	98	0.15	0.33	0.33	0.35	0.35
	99	0.15	0.32	0.32	0.33	0.33
97	96	0.17	0.36	0.35	0.37	0.37
	98	0.15	0.34	0.33	0.35	0.35
	99	0.15	0.32	0.32	0.34	0.34
98	96	0.17	0.36	0.35	0.37	0.37
	97	0.16	0.35	0.34	0.36	0.36
	99	0.15	0.33	0.32	0.34	0.34
99	96	0.17	0.36	0.35	0.37	0.37
	97	0.16	0.35	0.34	0.36	0.36
	98	0.15	0.34	0.33	0.35	0.35
96-97	98	0.15	0.34	0.33	0.35	0.35
	99	0.15	0.33	0.32	0.34	0.34
98-99	96	0.17	0.36	0.36	0.37	0.37
	97	0.16	0.35	0.35	0.36	0.36

Aggregate Performance of the Various Methods – RMSE

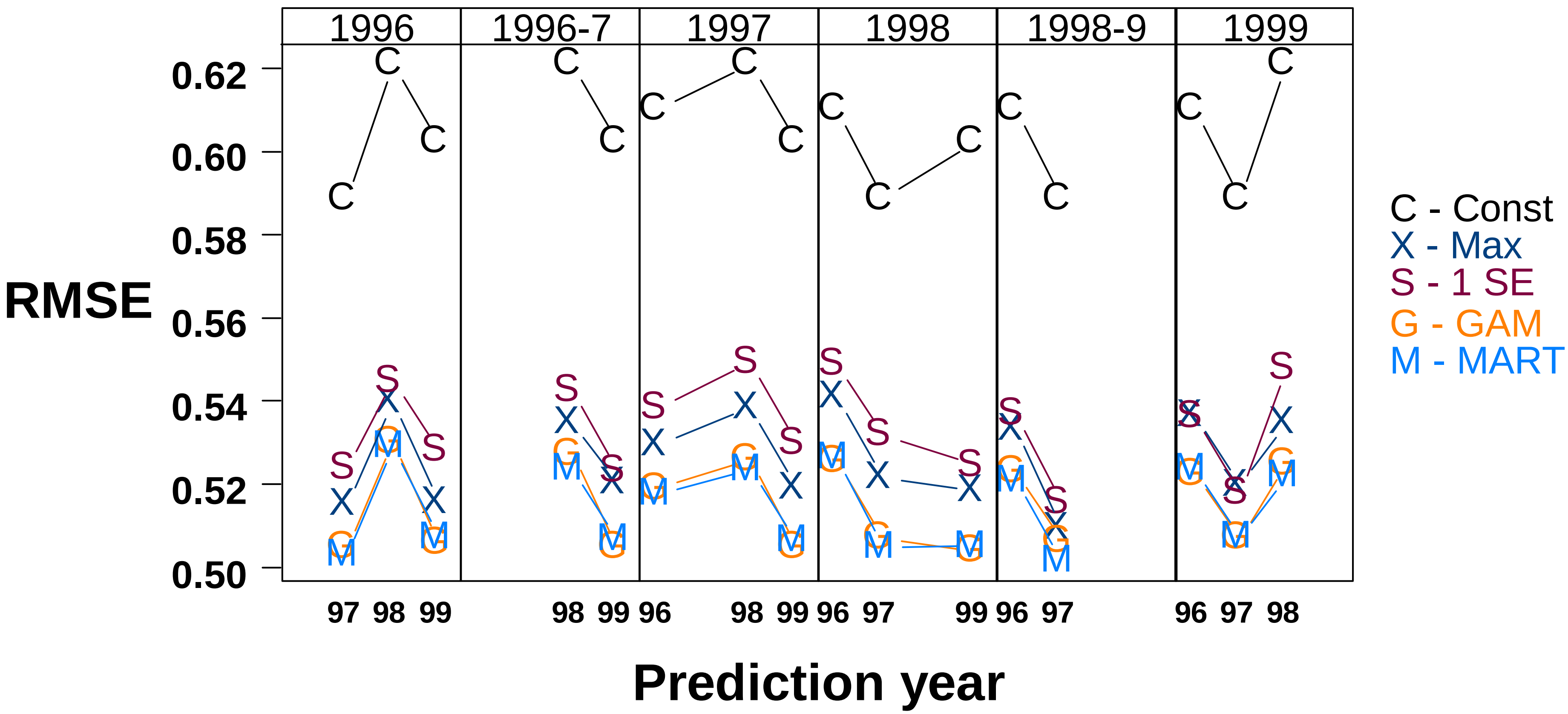
Fit Year	Evaluation Year	Const	CART		GAM	MART
			Max	1 SE rule		
96	97	.541	0.477	0.480	0.473	0.473
	98	.545	0.483	0.486	0.479	0.479
	99	.546	0.486	0.489	0.482	0.482
97	96	.535	0.471	0.473	0.467	0.467
	98	.545	0.482	0.485	0.479	0.478
	99	.546	0.486	0.488	0.482	0.482
98	96	.535	0.471	0.473	0.468	0.468
	97	.541	0.477	0.479	0.474	0.473
	99	.546	0.484	0.486	0.481	0.481
99	96	.535	0.472	0.474	0.468	0.468
	97	.541	0.477	0.479	0.474	0.473
	98	.545	0.481	0.483	0.479	0.478
96-97	98	.545	0.482	0.483	0.479	0.478
	99	.546	0.485	0.486	0.482	0.481
98-99	96	.535	0.470	0.471	0.468	0.467
	97	.541	0.475	0.477	0.473	0.473

CART Captures Nearly All Explainable Variation for Stroke



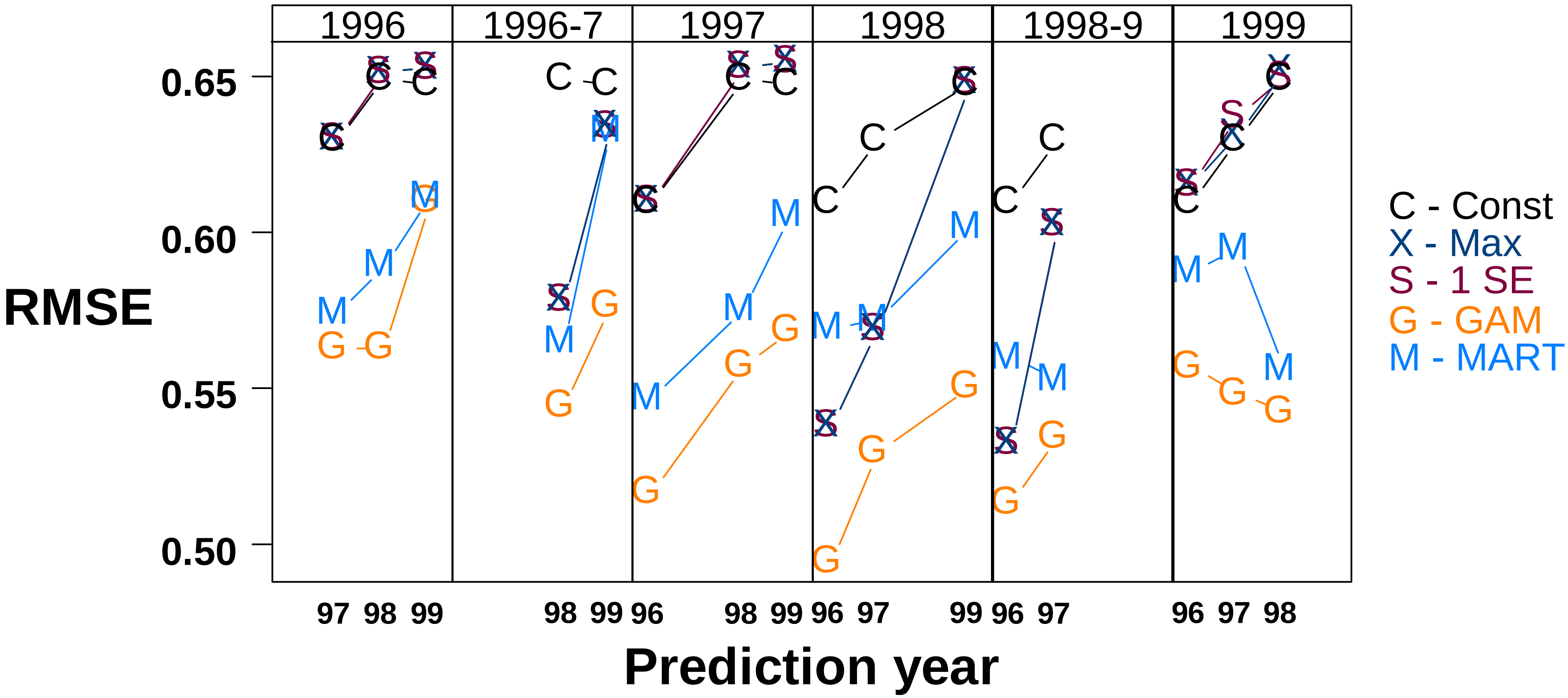
N **= 32687, 35026, 37012, 37340**

CART Captures Most of the Explainable Variation for Traumatic Brain Injury



N = 1383, 1629, 1871, 2053

CART Does Not Do Well on the Burns RIC



N **= 70, 103, 111, 102**

Simulation to Evaluate Potential CART Hospital-Level Distortions

- **Used CART and MART to develop wage-adjusted payment formulas**
- **Aggregated up to hospital-year level**
- **Compared CART vs. MART annual hospital payments**

$$\text{ratio}_j = \frac{\sum_i M \times F_j \times CART(\text{age}_i, \text{motor}_i, \text{cognitive}_i)}{\sum_i M \times F_j \times MART(\text{age}_i, \text{motor}_i, \text{cognitive}_i)}$$

CART and MART Would Pay Hospitals Similarly

Hospital Payment Ratio	Percent of Hospitals (Case Weighted)			
	1996	1997	1998	1999
90	0.0	0.0	0.0	0.0
94	0.1	0.0	0.0	0.0
95	0.0	0.0	0.2	0.1
96	0.4	0.3	0.3	1.1
97	2.5	2.5	2.4	2.1
98	11.6	9.8	11.9	8.0
99	21.6	25.8	21.8	24.6
100	28.9	28.4	30.9	29.7
101	22.3	22.0	21.5	24.3
102	9.7	8.3	8.3	7.0
103	1.9	2.7	2.2	2.4
104	0.8	0.2	0.5	0.6
105	0.1	0.1	0.1	0.2
106	0.2	0.0	0.0	0.0
107	0.0	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0

94%

Conclusions from Looking at Gold Standard Models

- MART and GAM do about the same, always better than CART
- Overall, CART RMSEs are about .005 higher than MART or GAM, so error bands expand by 2%
- CART RMSEs are within 90% of the RMSE distance between RIC average prediction and MART or GAM. CART is fairly close in performance to the ideal model
- Hospital level payments are almost the same under either CART or MART payments
- Gold standard models offer a useful perspective on the performance of a case classification system