

# Value-Based Purchasing Strategies

---

Joseph P. Newhouse

Harvard University

December 17, 2007

# Outline of Talk

---

- Some caveats about me
- Lots of evidence that providers respond to reimbursement – for better or for worse
- Not a lot of evidence on how best to do:
  - ◆ Pay for performance
  - ◆ Tiered networks
  - ◆ Disease management

# I'm An Economist: We're Said to Know the Price of Everything and the Value of Nothing

---



# I'm Also An Academic: Sometimes We Say the Obvious

---



# And Sometimes We Lose the Forest for the Trees



# Providers Respond to Levels and Methods of Reimbursement

---

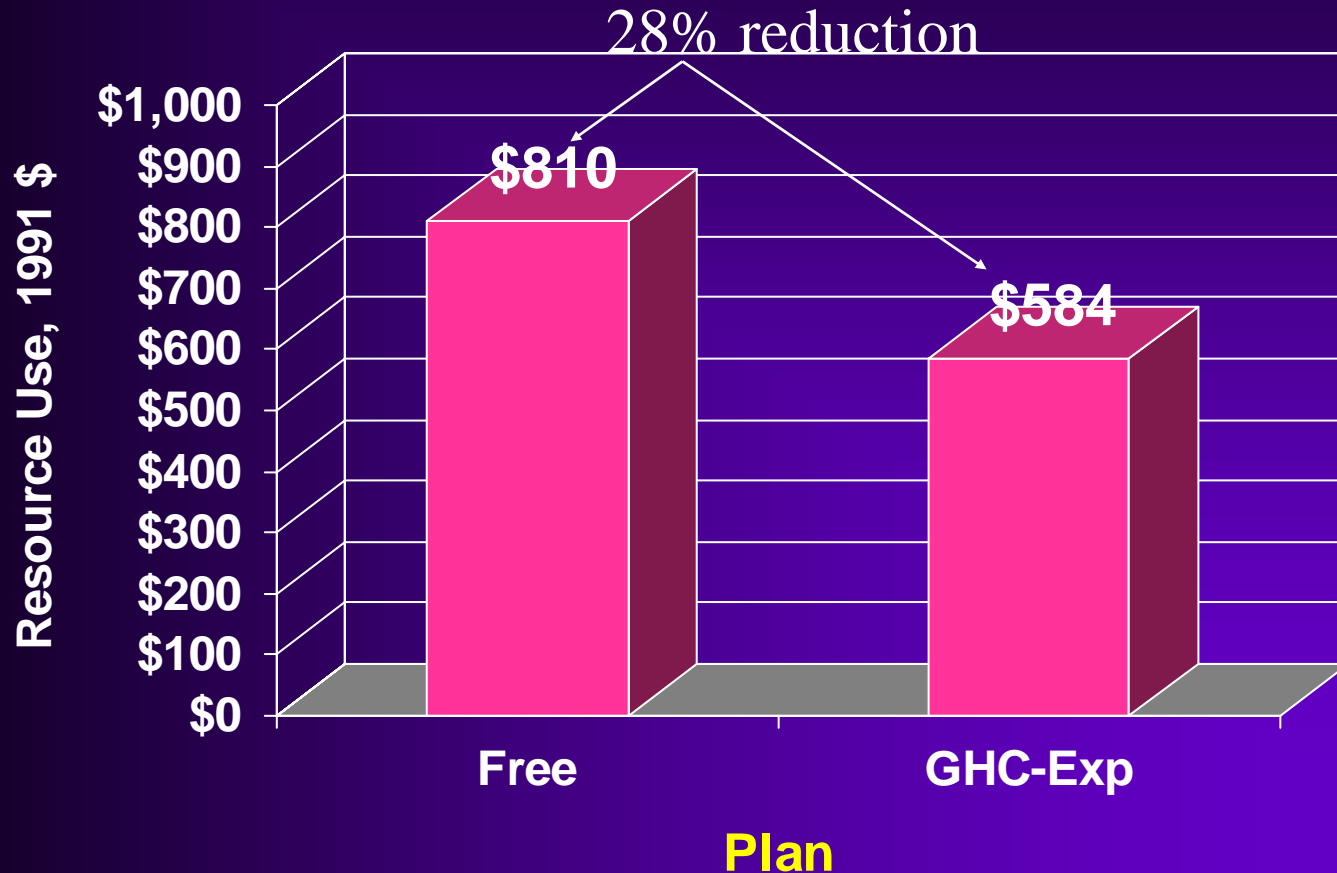
- This is where we have a lot of evidence that shows if the amount or the type of reimbursement changes:
  - ◆ Health plans respond
  - ◆ Hospitals respond
  - ◆ Physicians respond

# Health Plans: RAND Experiment Results from 30 Years Ago

---

- Still the only randomized trial of a group and staff model HMO, Group Health Co-operative of Puget Sound (GHC)
- Showed major reduction in \$, hospital use
  - ◆ Two groups, one randomized to free care at GHC, one with free care in the fee-for-service system in Seattle

# HMO Resource Use Much Less



Hospital admissions down 39%; visit rates similar

Predicted means in 1991 dollars. t's vs GHC-E: Free 3.2, GHCC 0.78. Values include out-of-plan use.



# Outcomes

---

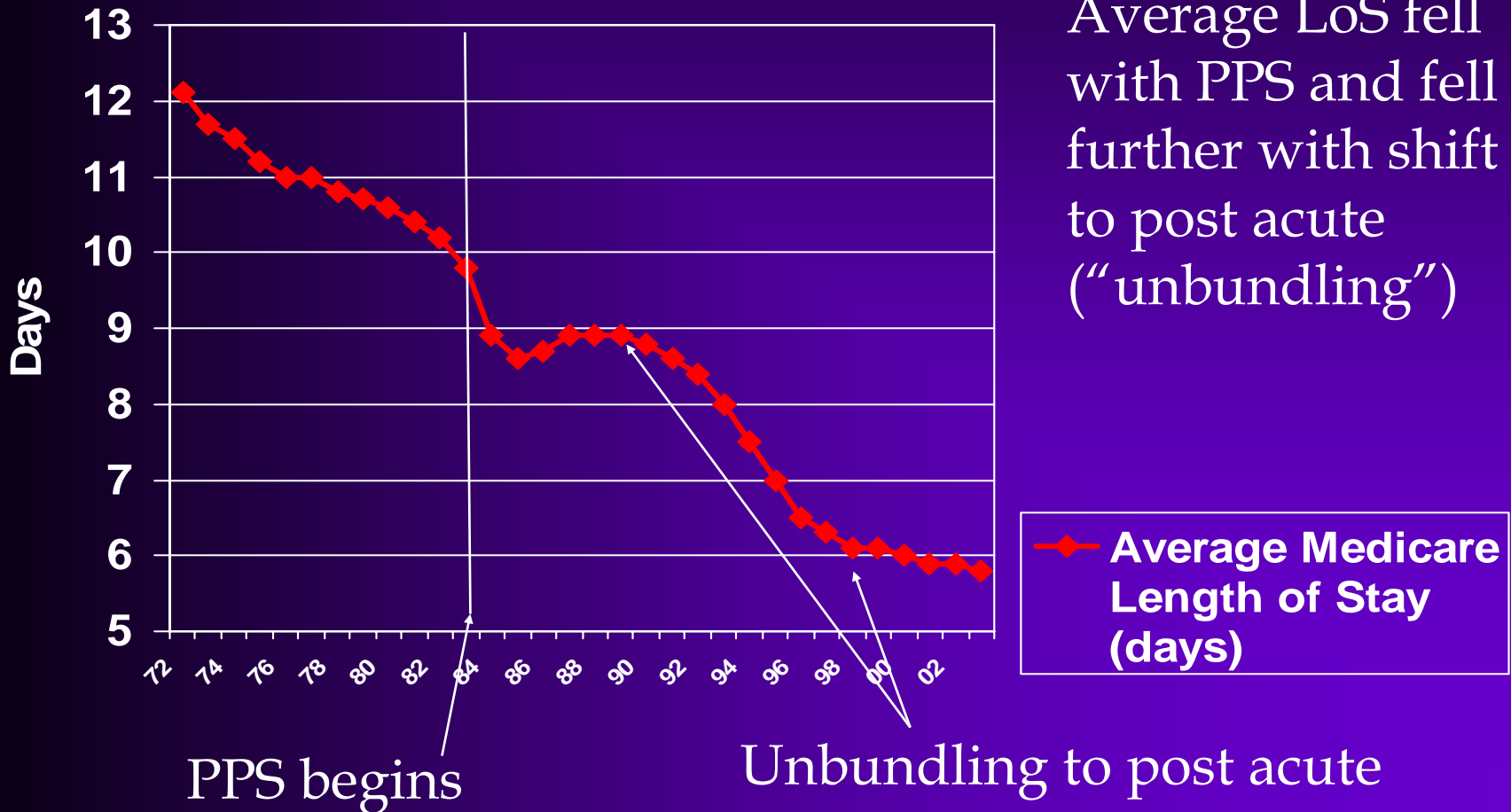
- Measures of self-rated health (like SF-36), physical functioning, mental health, health habits (like smoking), and many physiologic measures (like blood pressure) showed little or no difference between the two groups

# My Conclusion

---

- A capitated, integrated medical care system used the hospital much less and saved \$, but
  - ◆ Since 1980 the age-adjusted hospital admission rate in the US has fallen 32 percent so these savings may no longer be available
  - ◆ Not clear that such organizations can expand
    - Local history of HCHP; Kaiser market share relatively static; minor part of Medicare
    - How to enter and attract capital?

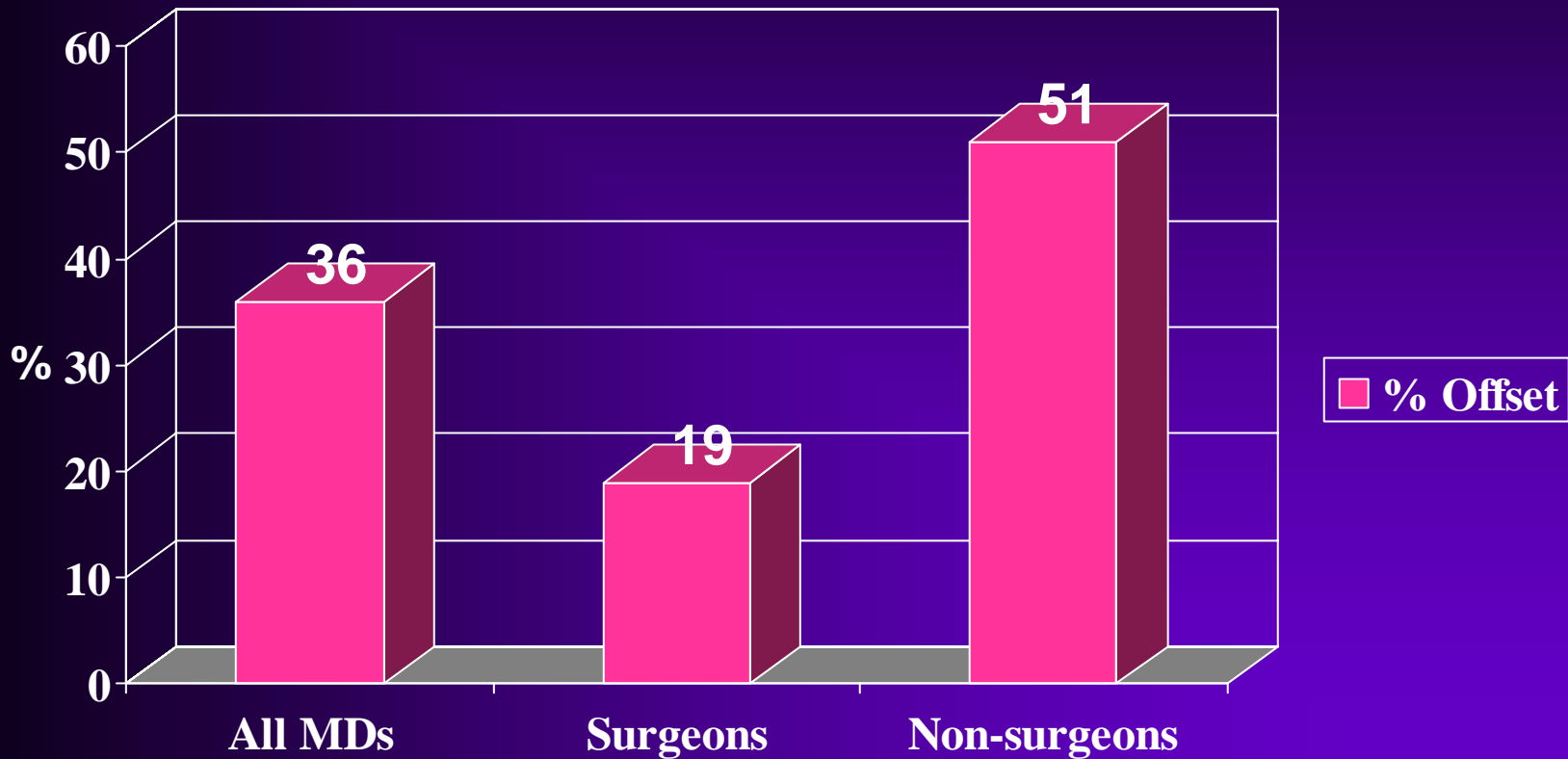
# Hospitals Respond to PPS



# Physicians Respond to the Level and Basis of Payment

- Next slide shows “offset” effect, or how much MDs increase Medicare “volume” if Medicare fees cut
- Other findings on basis of payment:
  - ◆ UK expanded GP capitation to cover hospitalization costs; elective admissions fell
  - ◆ Salaried residents (vs FFS) saw fewer patients
  - ◆ Danish MDs with partial capitation, partial FFS saw more patients than with full capitation

# Medicare Offset Estimates



**36%** of a cut in fees is offset by higher volume

# A Schematic of Payment

---



Capitation

Episode,  
incl. post  
acute

Hospital Stay

Hospital Day,  
Bundled Visit  
e.g., post-op

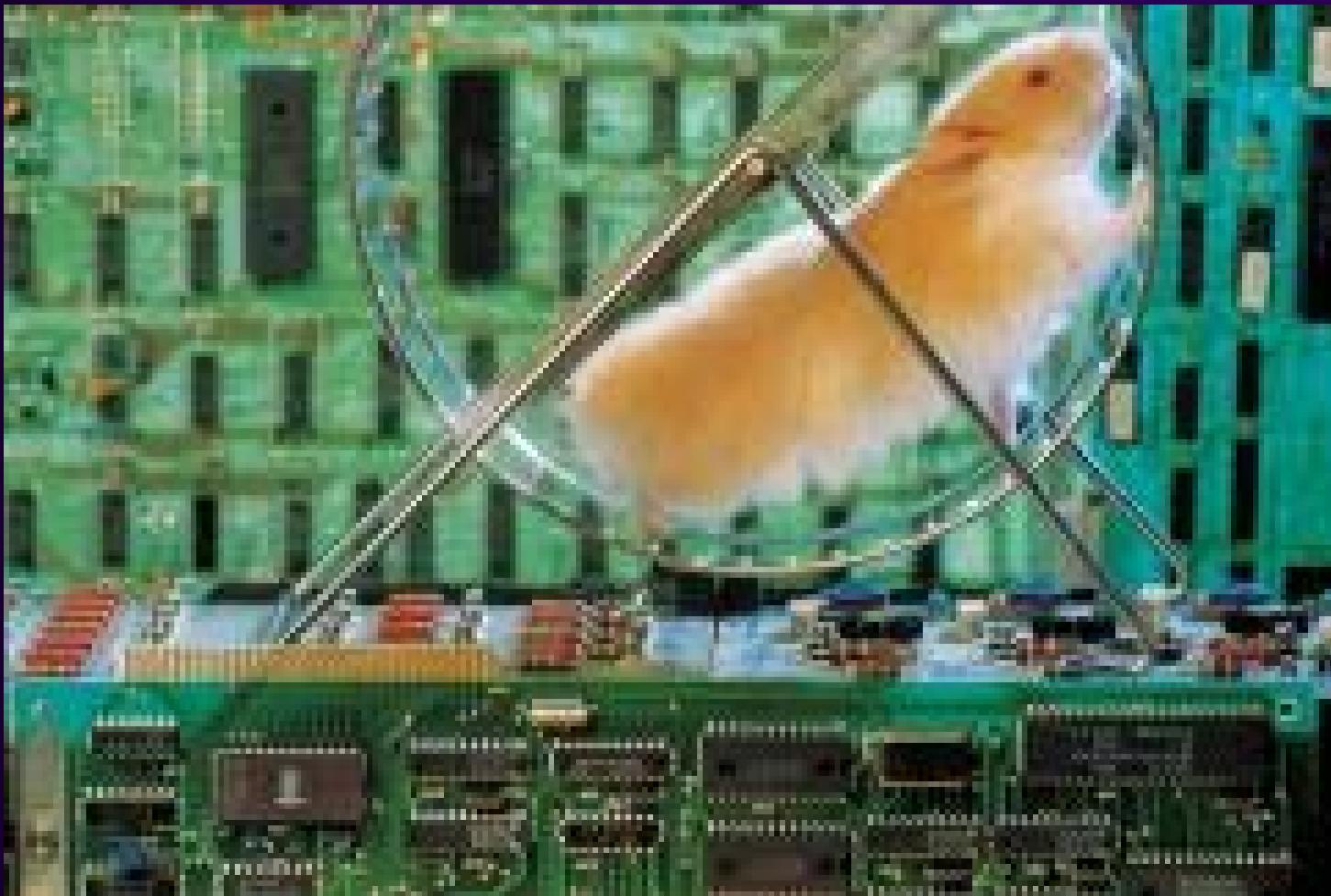
FFS

Moving left to right weakens the incentive for:

*Efficient Production*, but also for *Selection of Good Risks*;  
*Underservice*, but greater incentive for *Overservice*  
*Unbundling Services*

# The Physician Who Is Paid Using Fee-for-Service

---



# The Physician Who Is Paid Using Capitation





# My Personal Views

---

- Capitation may be too strong an incentive for underservice and selection; fee-for-service may be too costly
- I have advocated a mixed system for both health plans and physicians, e.g., part payment with capitation, part with FFS
  - ◆ Current proposal to pay a PCP a monthly fee for a medical home and using discounted fee-for-service approximates this

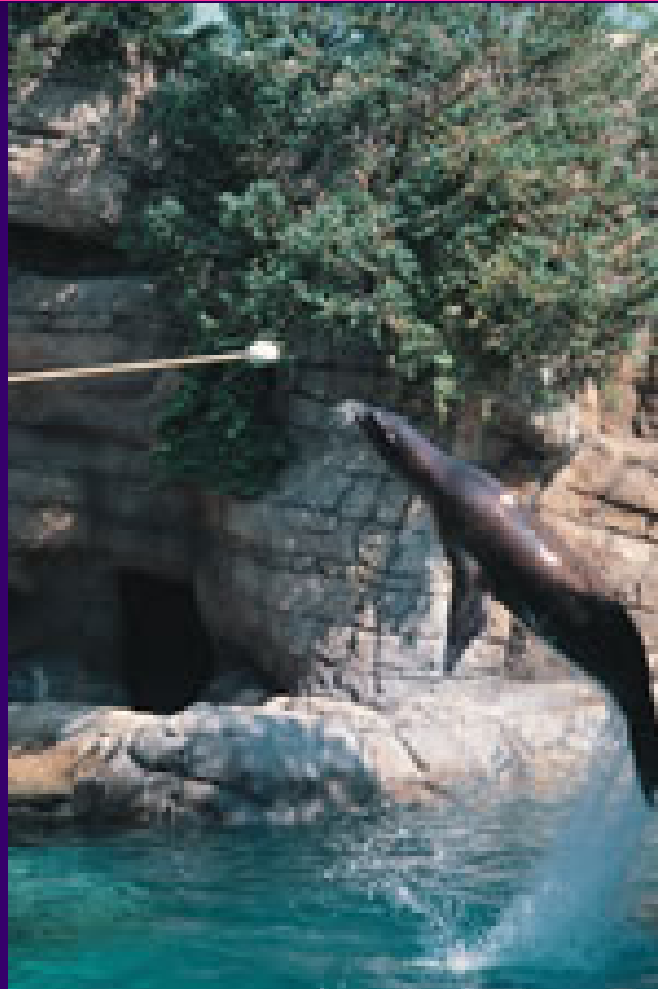
# Many Current Initiatives but Little Evidence of Good Outcomes

---

- Pay-for-Performance (P4P)
- Tiering networks
- Disease Management
- And what evidence there is says to me that we haven't yet figured out how to do it

# Pay for Performance (P4P)

---



# Pay-for-Performance

---

- Seems logical (Mark McClellan: “you get what you pay for”), but many issues in how to do it, for example:
  - ◆ How many \$ are at risk and how should payment be structured?

# PacifiCare Experience in California

---

- Plan paid bonus of ~5% of cap rate if MD was above the prior year 75<sup>th</sup> percentile of all MDs on a measure, so most of the \$ went to MDs already above the threshold
- Hence expensive per unit of improvement
- Typical PacifiCare market share was 15%, so financial incentive to MD was small

# One Modest Positive Result

**Table 1.** Improvement in Clinical Quality Scores for Quality Incentive Program (QIP) Measures\*

	Pre-QIP, %	Post-QIP, %†	Difference (Post – Pre), % (SE)	P Value
<b>Cervical cancer screening</b>				
California (n = 134)	39.2	44.5	5.3 (1.6)	<.001
Pacific Northwest (n = 33)	55.4	57.1	1.7 (0.9)	.03
Difference	-16.2	-12.6	3.6 (1.8)	.02
<b>Mammography</b>				
California (n = 134)	66.1	68.0	1.9 (1.1)	.04
Pacific Northwest (n = 32)	72.4	72.6	0.2 (1.1)	.43
Difference	-6.3	-4.6	1.7 (1.5)	.13
<b>Hemoglobin A<sub>1c</sub> testing</b>				
California (n = 134)	62.0	64.1	2.1 (1.0)	.02
Pacific Northwest (n = 31)	80.0	82.1	2.1 (3.3)	.20
Difference	-18.0	-18.0	0.0 (3.5)	.50

\*Data from authors' analysis of PacifiCare physician group performance reports for 2001-2004. Predicted values obtained from generalized estimating equation models of performance.

†For the purposes of this analysis, we defined the post-QIP period as beginning with the data reported for the first quarter of 2003.

Pacific Northwest Was Control Group

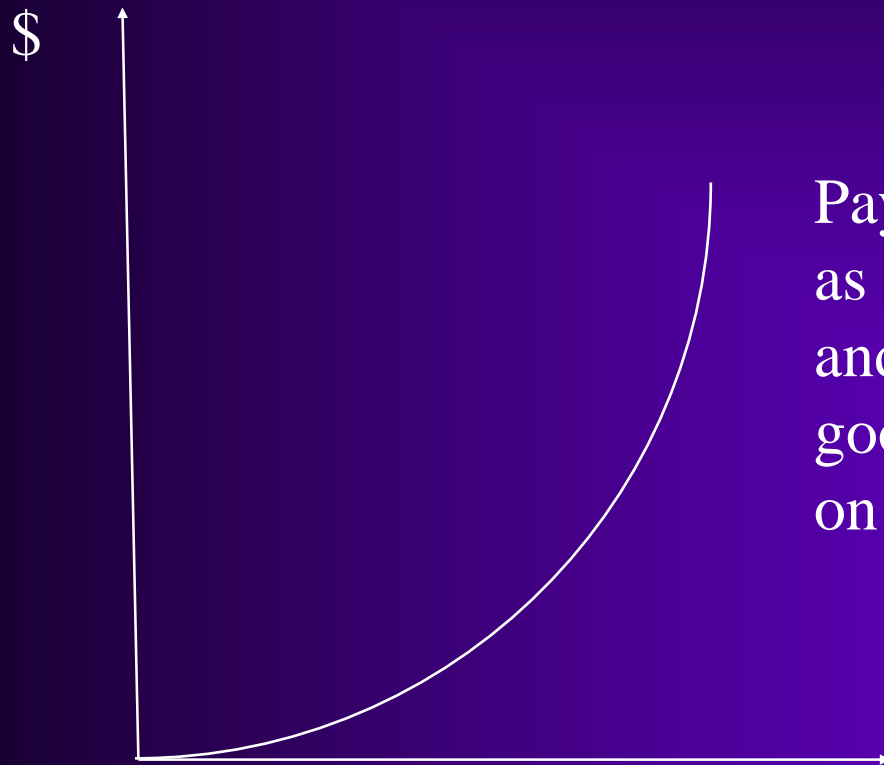
# Other Issues with P4P

---

- Teaching to the test and ignoring unmeasured diseases
- Risk adjustment if based on outcomes
- Expense of auditing

# What I Would Try

---



Pay more at the margin  
as it gets harder to improve  
and if that seems to get  
good results, up the money  
on the table

Performance Measure(s)



# Tiered Networks

---

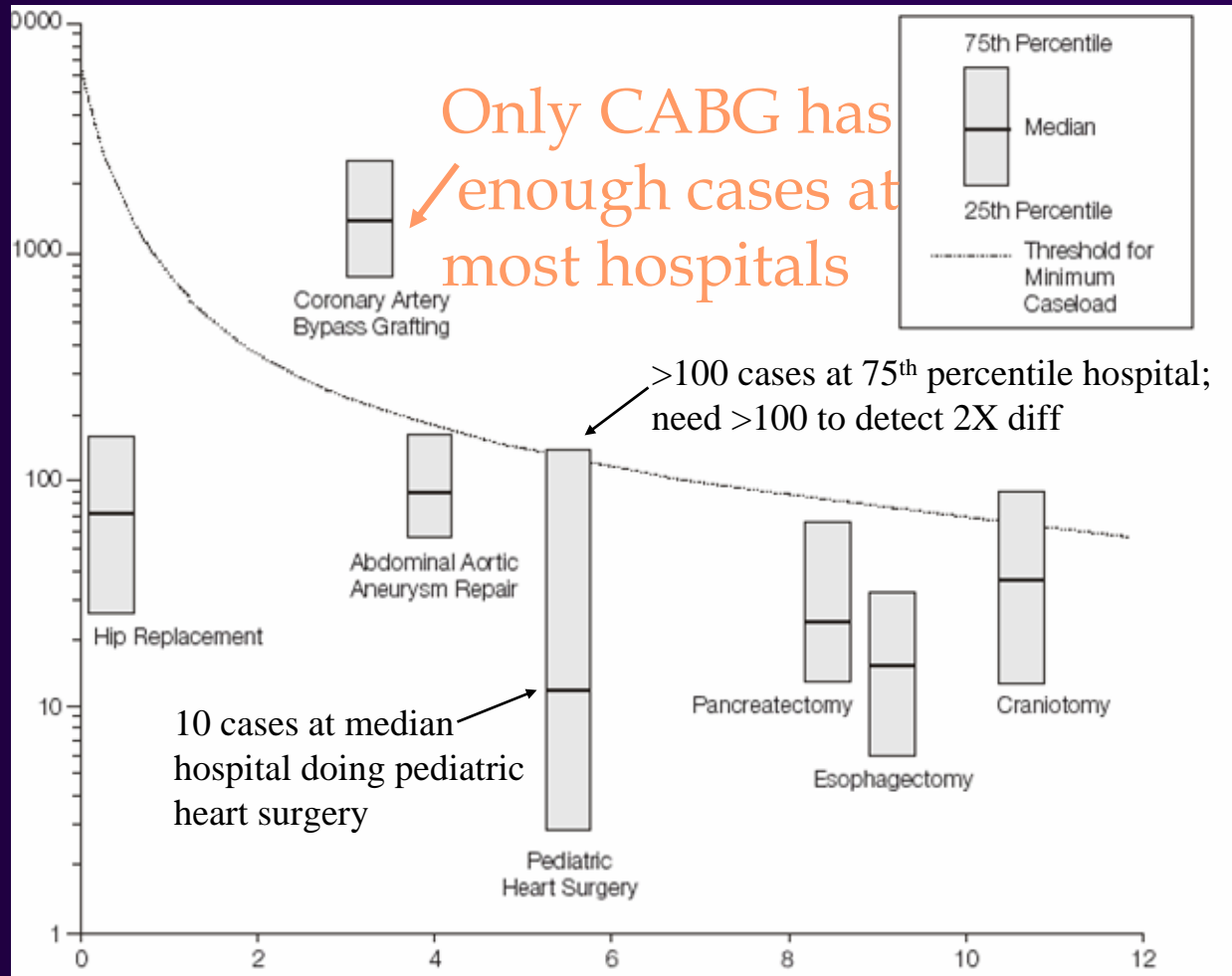
- Major payers other than traditional Medicare now using in several markets
- Idea is to lower cost sharing if enrollee uses certain physicians
  - ◆ Sometimes only specialists, not PCPs
- But how to choose MDs for favored tier?
  - ◆ Criteria? Is enough information available (sample size)? Selection? Teaching to Test?

# Even at the Hospital Level Sample Size Is a Problem

- Dimick, et al. studied seven procedures: CABG, AAA, pancreatic resection, esophageal resection, pediatric heart surgery, craniotomy, THR
- They asked: What is the minimum sample size necessary to be able to detect a mortality rate double the national average with 5% type I error and 80% power?
  - ◆ Need more cases to detect smaller difference

# Not Enough Cases to Detect Mortality Differences 2X Nat'l Avg

Observed 3 yr Hospital Case Loads



Operative Mortality Rates

JAMA 2004; 292: 849

# Sample Size Problems Are Worse at the Individual MD Level

- Little variation in hospitalization, visit rates among diabetics from MD-specific factors
- MD needs  $> 100$  diabetic patients to have 80% reliability on various measures
  - ◆ 80% reliability means 80% of the variance in visits is from MDs style, 20% from chance
  - ◆ None of the 250 PCPs in the sample had more than 85 diabetic patients, median had 29
  - ◆ But enough for groups of MDs

# Selection

- Patient selection can importantly influence a physician's profile
  - ◆ Hofer: If a below average MD in year 1 replaced worst 5% of his patients with respect to glycemic control with patients at the average; then most would be above average in year 2
    - But replacing the worst 5% only means replacing 1 to 3 patients in a given practice; encouraging one patient to go elsewhere for care doesn't seem hard

# Personal Preferences

---

- I would (for now) limit tiering to specialists and I would include most specialists in the market in the favored tier
  - ◆ More reliable, especially for specialists doing procedures

# Disease Management - 1

- My take on the literature: Not many robust results
- A recent review:

“Overall, disease management does not seem to affect utilization except for a reduction in hospitalization rates among patients with congestive heart failure and an increase in outpatient care and prescription drug use among patients with depression. When the costs of the intervention were appropriately accounted for and subtracted from any savings, there was **no conclusive evidence that disease management leads to a net reduction of direct medical costs.**”

# Disease Management - 2

---

- Of course, there could be better outcomes that are worth the cost
- Results probably depend upon targeting, which is not standardized
  - ◆ Want to work with non-compliers who will change; not clear to me how good targeting is
- Are we getting better on targeting?



# A Final Comment

---

- Hard for private sector payers to be successful with these strategies if Medicare sends different signals; also issues of coordination among private payers
- Hard for Medicare to have many \$ depend on performance for political reasons
  - ◆ Some beginnings, but I am skeptical Medicare will soon put UK type monies on the table