Case Study:
The California P4P Program Journey Toward Efficiency Measurement

Dolores Yanagihara, MPH
P4P Program Director
Integrated Healthcare Association

National P4P Summit

March 10, 2009

"Efficiency Measurement: The Pot of Gold At the End of the Rainbow?"

Part II

Overview

- Background
- Current Status
- Findings
- Next Steps

California P4P Participants

Health Plans:

- Aetna
- Anthem Blue Cross
- Blue Shield of CA
- Western Health Advantage
- CIGNA
- Health Net
- Kaiser*
- PacifiCare/United

Medical Group and IPAs:

- 230 groups
- 35,000 physicians

11 million commercial HMO members

** Kaiser participates in the public reporting only*

The Push for Efficiency Measurement

- Demand by purchasers and health plans that cost be included in the P4P equation

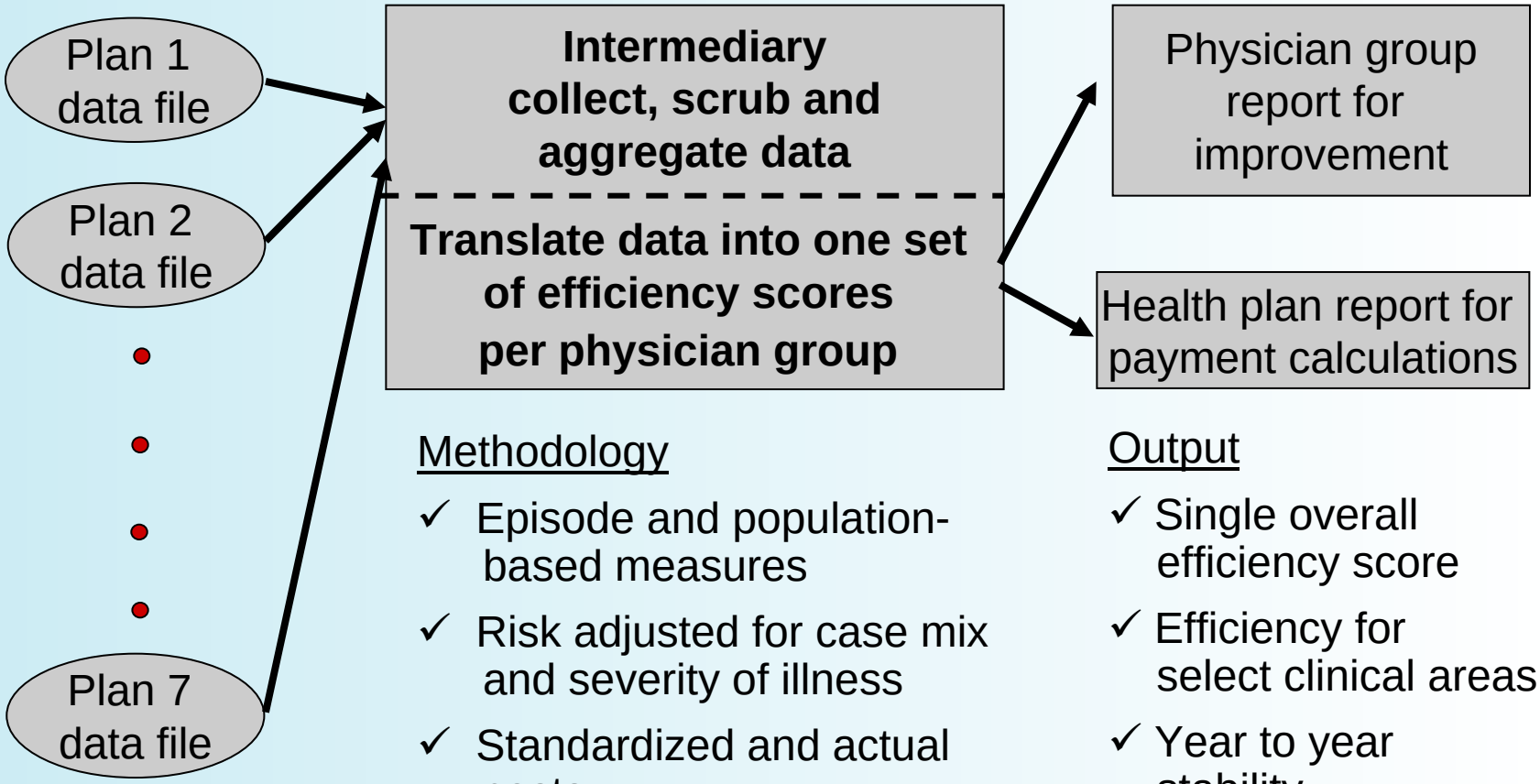
  Quality + Cost = Value

- Opportunity for common approach to health plan and physician group cost/risk sharing
- Demonstrate the value of the delegated, coordinated model of care

Principles: Efficiency Measurement in P4P

- Collaborative development/adoption
- Aggregation across plans
- Alignment with national measures when feasible
- Thorough testing and analysis prior to implementation
- Transparent methodology
- Risk adjustment to support fairness
- Rigorous approach for validity and reliability
- Actionable results to support efficiency improvement

Framework: Efficiency Measurement in P4P



Methodology

- ✓ Episode and population-based measures
- ✓ Risk adjusted for case mix and severity of illness
- ✓ Standardized and actual costs

Output

- ✓ Single overall efficiency score
- ✓ Efficiency for select clinical areas
- ✓ Year to year stability

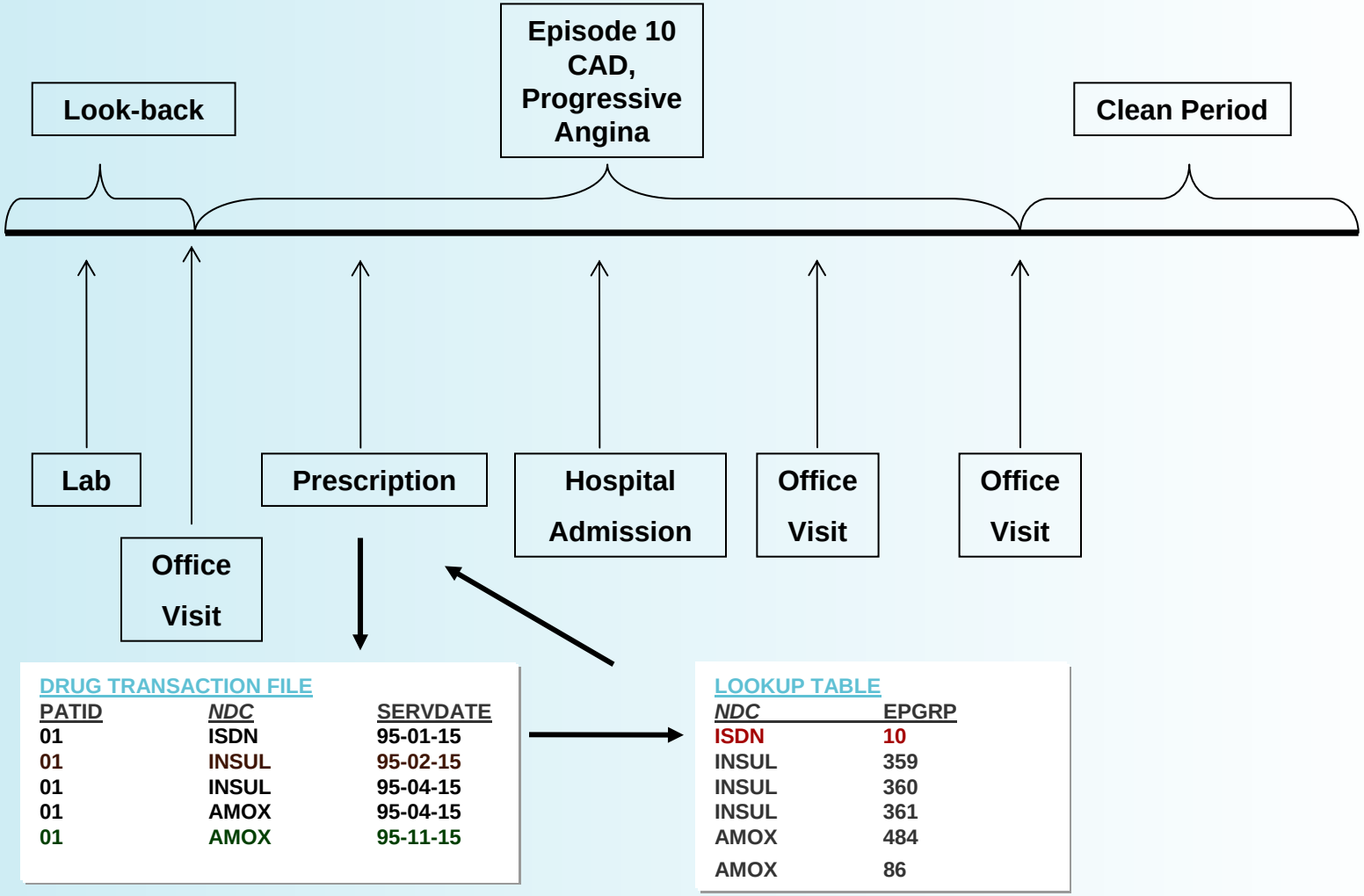
Deciding on an Initial Approach

- Considered standardizing currently used resource use measures (admits/1000, etc.) as interim measures
- Rejected – stakeholders anxious to get to sophisticated efficiency measures ASAP and didn't want to spend resources on standardizing what was already being done

Efficiency Measures

1. Generic Prescribing
2. Population-Based
   - Overall Group Efficiency
   - Standardized and actual costs
   - DCG and geographic risk adjustment
3. Episode-Based
   - Overall Group Efficiency
   - Efficiency by Clinical Area
   - Standardized costs only
   - MEG, Disease Staging, and DCG risk adjustment

Episode Construction



DRUG TRANSACTION FILE

PATID	*NDC*	SERVDATE
01	ISDN	95-01-15
01	INSUL	95-02-15
01	INSUL	95-04-15
01	AMOX	95-04-15
01	AMOX	95-11-15

LOOKUP TABLE

NDC	EPGRP
ISDN	10
INSUL	359
INSUL	360
INSUL	361
AMOX	484
AMOX	86

Methodological Considerations

- Use internal benchmarks to calculate "expected"
  - Based on the average risk adjusted cost across all 7 health plans
- 12 month measurement period, unless otherwise indicated through testing
- Outlier methodologies to eliminate 1% of highest and lowest cost episodes
- Clinical exclusions to be determined (e.g. transplants)

CA Advantages for Efficiency Measurement

- Unit of measure – Physician group vs. individual physician measurement makes attribution more reliable
- Large sample size – Aggregation of plan data allows for adequate sample size
- Consistent benefit package – HMO/POS member population provides relatively consistent benefits
- Stakeholder trust – Relatively good

Getting Data

- Sign Business Associate Agreements
- Address antitrust concerns
  - Opinion from legal counsel
  - Guidelines for acceptable reporting
- Overcome confidentiality clauses in contracts
  - Obtain Consent to Disclosure Agreements
    - Physician Groups
    - Hospitals
- Obtain useable data from health plans
  - Multiple data submissions needed

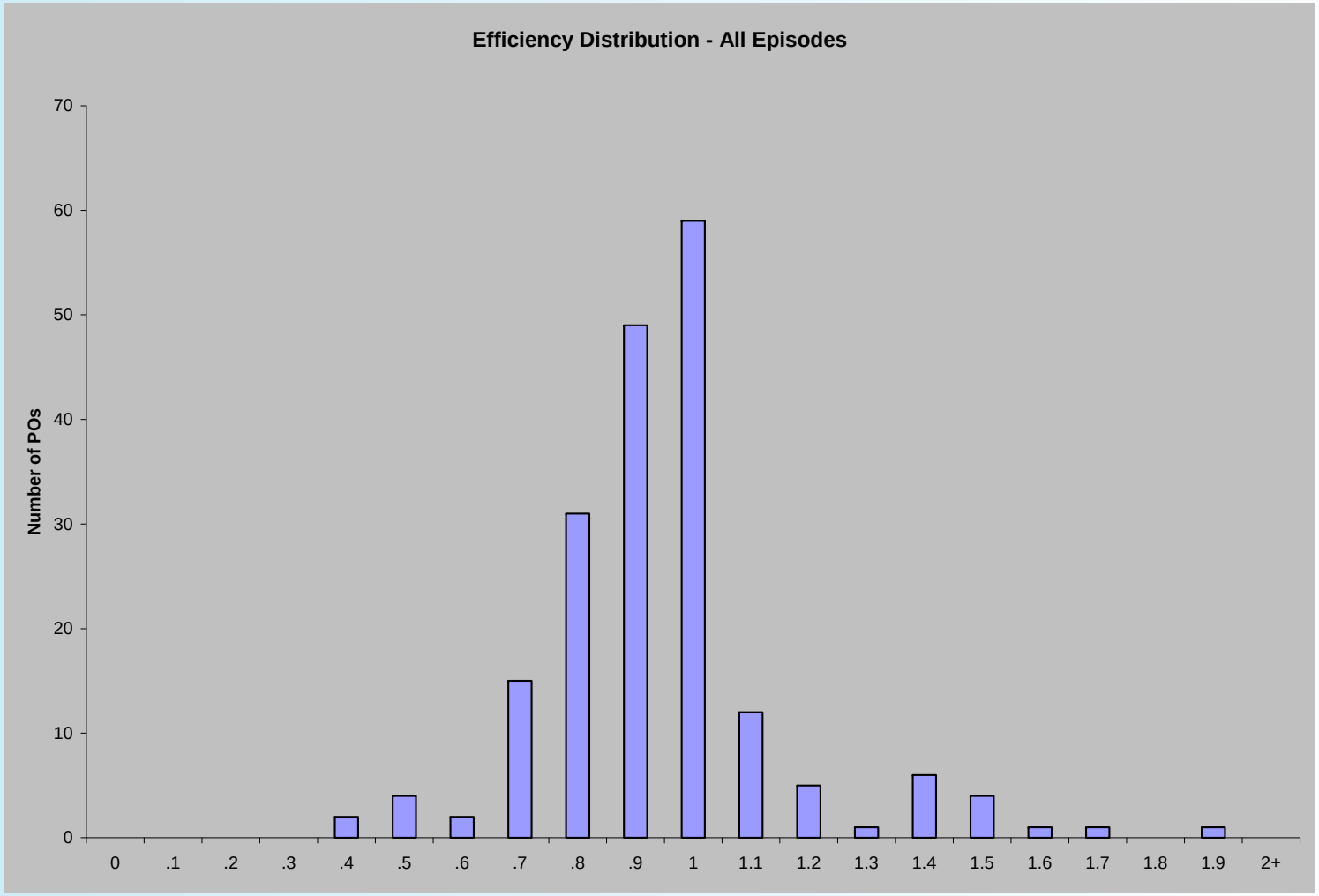
Development Timeline

November 2005 – July 2006	RFP process for vendor selection; Thomson Reuters Healthcare selected
October 2006	Established multi-stakeholder Technical Efficiency Committee
March 2008	BAAs signed and data received from all plans
July 2008	Round 1 testing complete
September – October 2008	Data quality meetings with health plans
January 2009	Round 2 testing complete

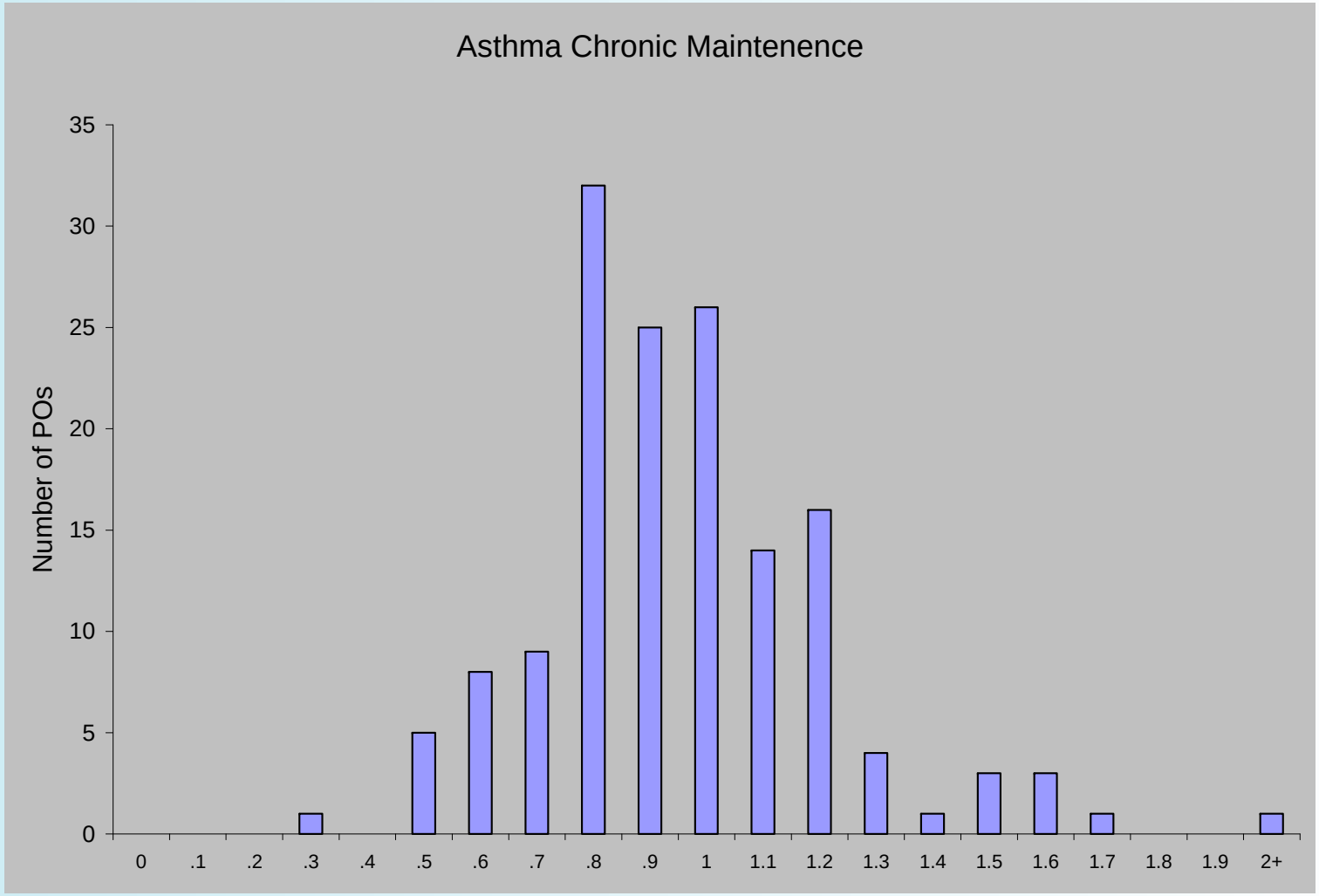
Findings: High Level

- Collected and aggregated data from 6 health plans
  - Numerous data fixes needed to standardize data across plans
- Produced overall population and episode-based efficiency results
  - Reasonable and normally distributed
- Drilled down to single episode groups and service categories
  - Greater granularity of drill down = more data gaps/inconsistencies identified

Episode-Based Overall Efficiency



Efficiency for Asthma Episode Group



Conclusion after Testing Round 1

- Current state of our dataset does not support comprehensive efficiency measurement using episodes of care
  - PO-specific results will not yet be disseminated
- Specific data issues have been identified and can be acted on
- It's too early to determine whether the data can be sufficiently improved

Data / Methodology Enhancements

- Shared organization-specific data quality reports with plans and POs
  - Increased understanding of content of files
  - Identified cause of discrepancies
  - Identified more complete sources of data
- Modified Facility Outpatient Standardized Pricing Approach
  - Collapsed ASC into Outpatient Facility
  - Removed “trivial” cases
  - Priced all remaining cases using APC system

Findings: Data Quality

- Variation in Place of Service coding on facility claims
  - Affects assignment of standardized pricing
- Inconsistent availability of procedure codes on facility claims
  - Affects outpatient standardized pricing
- Varying availability of diagnosis codes
  - Affects inpatient standardized costs and risk adjustment
- Overall efficiency score is strongly correlated with the Hospital Outpatient efficiency score
  - True driver of efficiency or data issues?

Place of Service on Facility Claims

Facility - Place of Service	Plan A	Plan B	Plan C	Plan D	Plan E	Plan F
Office	0%	0%	0%	22%	0%	0%
Inpatient	22%	25%	43%	19%	25%	30%
Outpatient	44%	60%	45%	32%	52%	54%
ASC	6%	4%	4%	0.1%	5%	4%
ER	9%	4%	4%	23%	11%	8%
Pharmacy	20%	0%	0%	0%	0%	0%
SNF	0.3%	0.4%	1%	0.5%	0.1%	0.6%
ESRD	0%	5%	2%	0%	0.1%	4%

Procedure Codes on Facility Claims

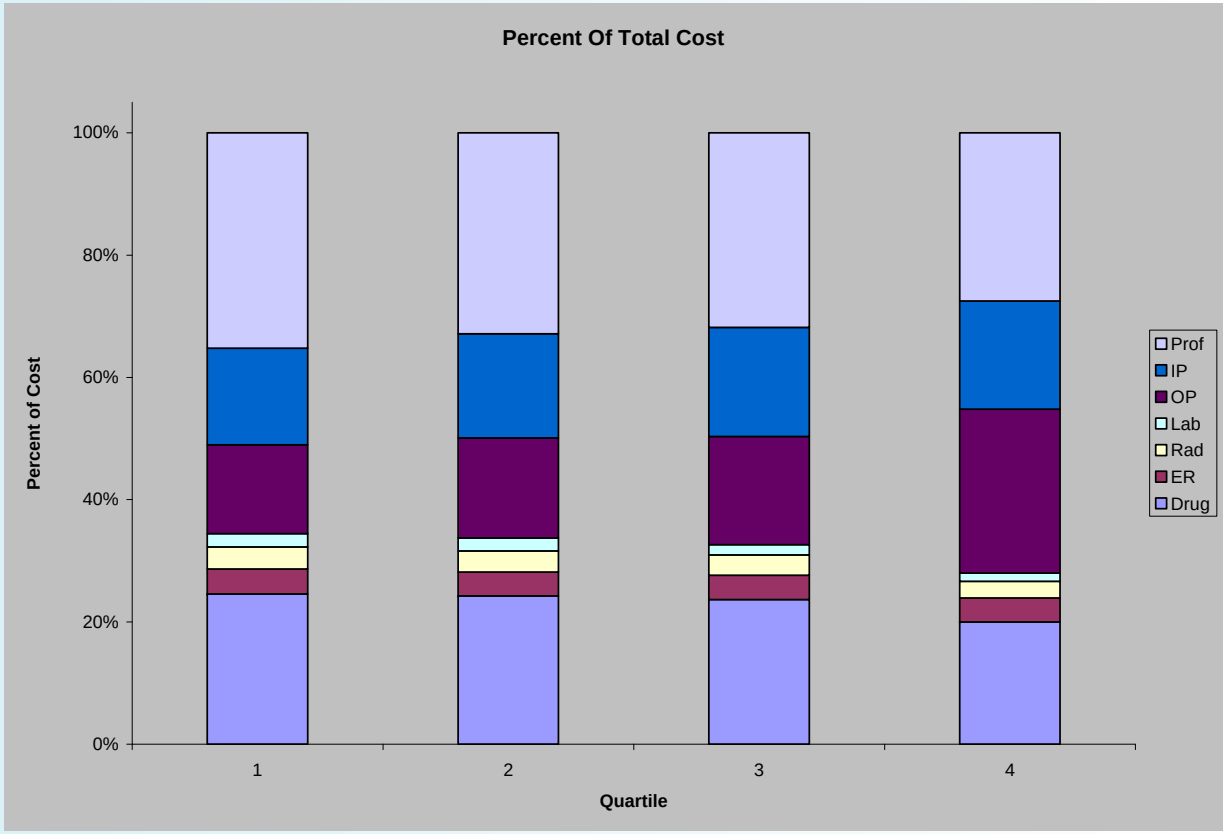
Facility Proc	Plan A	Plan B	Plan C	Plan D	Plan E	Plan F
Percent Coded	61%	29%	TBD	12%	52%	30%

Diagnosis Codes

Professional	Plan A	Plan B	Plan C	Plan D	Plan E	Plan F
Dx1	100%	100%	100%	100%	100%	100%
Dx2	9%	33%			42%	36%
Dx3	4%	14%			14%	15%
Facility	**Plan A**	**Plan B**	**Plan C**	**Plan D**	**Plan E**	**Plan F**
Dx1	100%	100%	100%	100%	100%	100%
Dx2	100%	92%	70%	13%	60%	73%
Dx3	73%	64%	47%	8%	37%	47%
Dx4			31%	5%	23%	23%
Dx5			21%	2%	14%	13%
Dx6			14%	1%	9%	9%
Dx7			10%	1%	7%	6%
Dx8			7%		5%	5%
Dx9			5%		4%	3%

Outpatient Hospital and Overall Efficiency

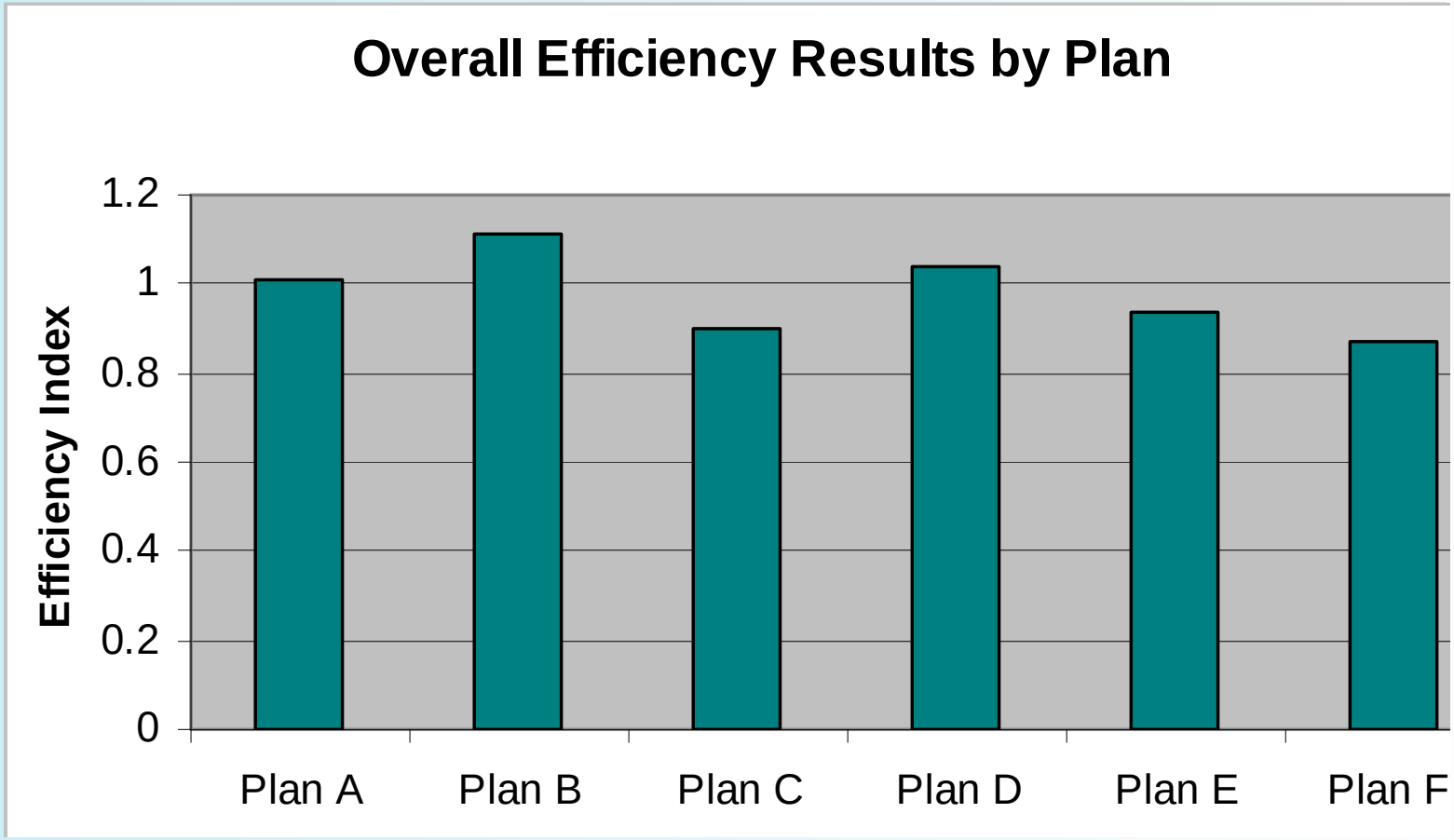
Service Category	Correlation
OP Hospital	0.772
Drug	0.443
Prof	0.302
IP Facility	0.278
ER	0.204
Rad	0.151
Lab	0.037



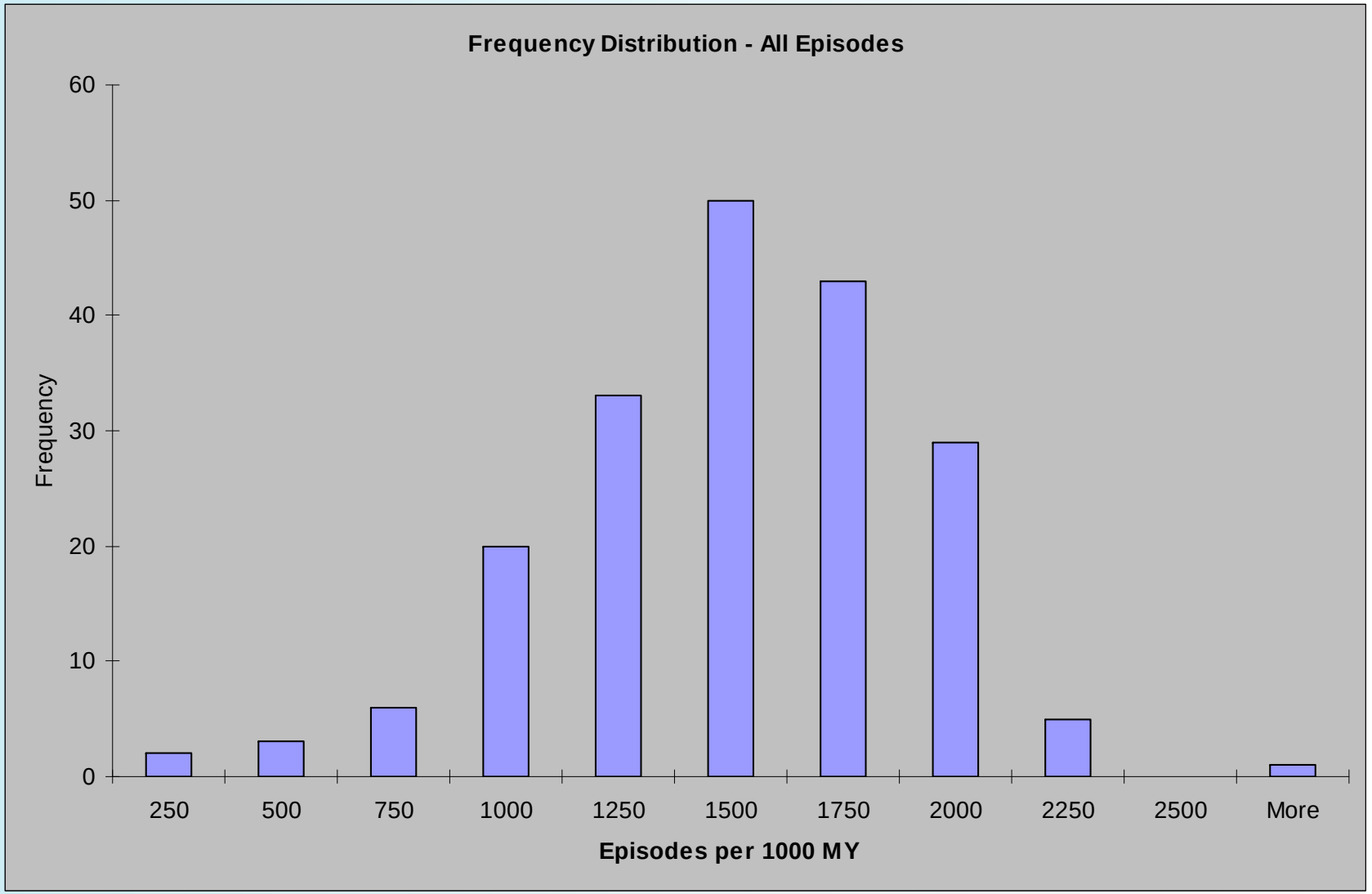
Findings: Episodes of Care

- Variation in overall efficiency across plans
  - True differences or driven by data?
  - Similar finding in MA
- Limited number of high cost episode groups occur frequently enough to produce results for at least 50% of POs
- Episodes that are driven by pharmacy, professional, and lab are the most reliable
- Number of episodes per 1000 member years varies greatly across POs
  - Proxy for data completeness

Plan Level Episode Efficiency



	Episode Type	Percent of Cost	Percent of POs with 30+ Episodes
1	Diabetes Mellitus Type 2 and Hyperglycemic States Maintenance	5.6%	84.9%
2	Renal Failure	5.5%	37.0%
3	Essential Hypertension, Chronic Maintenance	4.5%	88.5%
4	Angina Pectoris, Chronic Maintenance	4.3%	66.7%
5	Neoplasm, Malignant: Breast, Female	3.2%	39.1%
6	Delivery, Vaginal	2.5%	63.5%
7	Osteoarthritis, Except Spine	2.3%	77.6%
8	Asthma, chronic maintenance	2.2%	77.6%
9	Other Arthropathies, Bone and Joint Disorders	2.0%	88.0%
10	Human Immunodeficiency Virus Type I (HIV) Infection	1.7%	15.1%
11	Rheumatoid Arthritis	1.5%	39.6%
12	Neoplasm, Malignant: Colon and Rectum	1.4%	18.8%
13	Delivery, Cesarean Section	1.4%	34.4%
14	Other Inflammations and Infections of Skin and Subcutaneous Tissue	1.2%	90.1%
15	Other Gastrointestinal or Abdominal Symptoms	1.1%	85.9%
16	Complications of Surgical and Medical Care	1.1%	47.9%
17	Multiple Sclerosis	1.0%	15.6%
18	Infections of Skin and Subcutaneous Tissue	1.0%	81.3%
19	Other Ear, Nose and Throat Disorders	1.0%	89.1%


Frequency Distribution - All Episodes
60
50
40
30
20
10
0
Frequency
250
500
750
1000
1250
1500
1750
2000
2250
2500
More
Episodes per 1000 MY

Conclusion after Testing Round 2

- Data does not yet support episode of care based measurement for payment but is now good enough for sharing with POs
- Many episode groups should be discarded because numbers too small for reliability
- Some of remaining episodes may be ready for "prime time"

Current Considerations

- Setting data thresholds for participating in efficiency measurement
- Aggregating to the episode summary group or body system level
- Using episode construct and disease staging to assess appropriateness of high volume, high cost procedures
- Adapting NCQA's Relative Resource Use measures to the physician organization level

Going Full Circle

- Development of episode and population-based measures taking too long
- Need to address affordability of HMO product now
- Standardized currently used appropriate resource use measures for implementation in MY 2009
  - Inpatient acute care discharges PTMY
  - Bed days PTMY
  - Readmissions within 30 days
  - ED Visits PTMY
  - Outpatient surgeries — % done in ASC
  - Generic prescribing

California Pay for Performance

For more information:

www.iha.org

(510) 208-1740

Pay for Performance has been supported by major grants from the California Health Care Foundation