

Contract No.: 100213
MPR Reference No.: 6099-500

MATHEMATICA
Policy Research, Inc.

**Evaluation of the Local
Initiative Rewarding
Results Collaborative
Demonstrations:**

Final Report

October 10, 2006

*Suzanne Felt-Lisk
Gilbert Gimm
Stephanie Peterson*

Submitted to:

Center for Health Care Strategies
P.O. Box 3469
Princeton, NJ 08543-3469

Submitted by:

Mathematica Policy Research, Inc.
600 Maryland Ave. S.W., Suite 550
Washington, DC 20024-2512
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Officer:

Joanie Rothstein

Project Director:

Suzanne Felt-Lisk

This page has been intentionally left blank for double-sided copying.

ACKNOWLEDGMENTS

This report would not have been possible without the helpful cooperation over several years from executives and staff within each of the participating health plans. Individuals who were key to the demonstration and to articulating the experience of the demonstration include but are not limited to: Gary Melton, Rob Jones, Sylvia Carlisle, Lily Quiette, Dori Lange, Christine Gerbo, Dennis Collins, Sam Tobin, and Maureen Hanlon. We also extend heartfelt thanks to the health plan data analysts and data supervisors who worked hard to ensure the data met the needs of the evaluation, including but not limited to: Joseph Sampson, Anna Kolodziej, Fanglin Meng, Phinney Ahn, Robert Hill, Matt Snyderman, Nick Montesa, and Yashodara Rao. We would like to thank Kathy Kim and Beau Carter (formerly of IHA), Ellie Payne, Nikki Highsmith, Elizabeth Cobb, and Joanie Rothstein of CHCS, for providing helpful comments as the project officers for the evaluation during various periods during the Collaborative.

Jim Verdier (MPR) and Glen Mays (formerly of MPR) were key internal advisors to the project team. Glen also contributed substantially to the evaluation design. Katherine Bencio tirelessly programmed the analyses presented here. Lisa Chimento and Nancy Beronja of The Lewin Group consulted with the project team on encounter data issues. Fabrice Smieliauskas served as the project's analyst within MPR for the first two years, collecting data from the plans and conducting many of the first round physician interviews.

Specific to this report, we thank Myles Maxfield and Jim Verdier for their helpful comments on an earlier draft. Karen Rosenthal edited the report, and Donna Dorsey produced it.

This page has been intentionally left blank for double-sided copying.

CONTENTS

Chapter	Page
EXECUTIVE SUMMARY	xi
I DEMONSTRATION OVERVIEW	1
A. VARIATION IN DESIGN	1
B. MOTIVATION AND FINANCIAL ABILITY TO BEGIN INCENTIVES	5
C. LIRR INCENTIVES IN CONTEXT	6
1. Incentives are Accompanied by Other Quality Supports	6
2. LIRR Incentives are Part of a Larger Set of Financial Incentives	6
3. Competition Was a Motivating Force, but the Collaborative Plans Were Not Shadowing Other Plans' Programs	7
II EVALUATION METHODS	9
A. OVERVIEW OF EVALUATION DESIGN	9
1. Implementation/Process Analysis	10
2. Impact Analysis	11
III FINDINGS OF THE IMPLEMENTATION ANALYSIS	17
A. OVERVIEW	17
B. ELIGIBILITY FOR AND DISTRIBUTION OF INCENTIVES THROUGH DECEMBER 2005	18

Chapter	Page
III <i>(continued)</i>	
C. IMPLEMENTATION CHALLENGES	19
1. The Plan Perspective—Provider Incentives	19
2. The Plan Perspective—Encounter Data Incentive	22
3. The Plan Perspective—Member Incentives	23
4. The Provider Perspective	23
D. SUSTAINABILITY	26
IV TRENDS IN OUTCOMES AND IMPACTS OF THE INCENTIVES	27
A. TRENDS IN OUTCOMES	28
1. HEDIS Outcomes for Well-Baby Visits and Adolescent Well-Visits	28
2. Trends in Provider Performance by Provider Specialty, Practice Type, and Patient Volume Among Providers Eligible for an Incentive	30
3. Encounter Data Reporting	31
B. IMPACTS OF THE INCENTIVES	32
1. Well-Baby Visit Incentives	33
2. Adolescent Well-Visit Incentives	35
3. Encounter Data Reporting Incentive	37
4. Performance Trend by Size of Payout	38
V HAS PRACTICE CHANGED? PROVIDERS’ COMMENTS ON THEIR RESPONSE TO THE INCENTIVES	41
A. LEVEL OF AWARENESS OF INCENTIVES	41
B. ACTIONS PROVIDERS HAVE TAKEN TO IMPROVE	42
C. CHALLENGES TO IMPROVEMENT	43
1. Well-Visit Rates	43
D. COMMENTS LINKING IMPLEMENTATION PROBLEMS TO PROVIDER RESPONSE	45

Chapter	Page
VI	SYNTHESIS OF QUALITATIVE AND QUANTITATIVE ANALYSIS OF THE WELL-BABY VISIT AND ENCOUNTER DATA INCENTIVES, PLAN BY PLAN 47
A.	PLAN-BY-PLAN SYNTHESIS OF WELL-BABY INCENTIVES RESULTS 47
1.	Plan D 48
2.	Plan B 50
3.	Plan A 50
4.	Plan C 52
5.	Plan E 53
6.	Plan F 53
7.	Plan G 53
B.	PLAN-SPECIFIC SYNTHESIS OF ENCOUNTER DATA RESULTS 54
1.	Plan C 54
VII	CONCLUSIONS 55
A.	PROVIDER INCENTIVES FOR WELL-BABY VISITS 55
1.	Facilitating Factors 55
2.	Barriers to Greater Improvement 56
B.	ADOLESCENT WELL-VISIT INCENTIVES 57
C.	ENCOUNTER DATA REPORTING INCENTIVES 57
D.	FUTURE RESEARCH 58
1.	Research Questions Directly Suggested by This Evaluation 58
2.	Overall Need for Studies of Pay-for-Performance in Medicaid 59

This page has been intentionally left blank for double-sided copying.

T A B L E S

Table		Page
I.1	INCENTIVE PROGRAMS UNDER LIRR.....	3
1.2	NUMBER OF PHYSICIANS AND MEMBERS POTENTIALLY TOUCHED BY THE WELL-BABY VISIT INCENTIVE PROGRAMS.....	5
I.3	QUALITY SUPPORTS OTHER THAN FINANCIAL INCENTIVES.....	7
II.1	OVERVIEW OF METHOD FOR QUANTITATIVE ANALYSIS OF LIRR OUTCOMES AND IMPACT	14
III.1	IMPLEMENTATION PROBLEMS OR BARRIERS	20
IV.1	HEDIS TRENDS FOR WELL-BABY VISITS AND ADOLESCENT WELL VISITS MEASURES, PARTICIPATING PLANS, 2002-2005.....	29
IV.2	PERFORMANCE OF PROVIDERS ELIGIBLE FOR INCENTIVES, BY PRACTICE CHARACTERISTICS AND BASELINE PERFORMANCE LEVEL	30
IV.3	ENCOUNTER DATA REPORTING RATES AND TIMELINESS, PROVIDERS ELIGIBLE FOR AN INCENTIVE.....	32
IV.4	PERCENT OF HEDIS-ELIGIBLE MEMBERS WITH SIX OR MORE WELL VISITS BY 15 MONTHS OF AGE, 2002-2005.....	34
IV.5	CHANGE IN WELL-BABY PROVIDERS VISIT PERFORMANCE 2003-2005, ELIGIBLE VS. INELIGIBLE FOR INCENTIVES (PROVIDERS WITH AT LEAST 10 HEDIS-ELIGIBLE BABIES)	35
IV.6	PERCENT OF HEDIS-ELIGIBLE MEMBERS AGES 12-21 WITH ANNUAL WELL VISIT	36
IV.7	ENCOUNTER DATA REPORTING OUTCOMES FOR GROUPS ELIGIBLE AND NOT ELIGIBLE FOR THE INCENTIVE.....	38

Table	Page
IV.8 CHANGE IN WELL-BABY VISIT PERFORMANCE, 2003-2005, BY SIZE OF 2003 INCENTIVE PAYOUT (PROVIDERS ELIGIBLE FOR INCENTIVES WITH AT LEAST 10 HEDIS-ELIGIBLE BABIES)	39
IV.9 MEDIAN PERCENT OF PREVENTIVE SERVICE ENCOUNTERS SUBMITTED WITHIN 90 DAYS	39
V. 1 EXAMPLES OF COMMENTS FROM PROVIDERS WHO RESPONDED TO THE INCENTIVES	43
V.2 EXAMPLES OF COMMENTS FROM PROVIDERS WHO DID NOT RESPOND TO THE INCENTIVES	44
V.3 COMMENTS FROM PROVIDERS WHO REPORTED BECOMING DISCOURAGED DUE TO PROGRAM ISSUES	45
VI.1 HEDIS DATA: PERCENT OF HEDIS-ELIGIBLE MEMBERS WITH SIX OR MORE WELL VISITS BY 15 MONTHS OF AGE, 2002-2005	48
VI.2 CHANGE IN WELL-BABY VISIT PERFORMANCE, 2003-2005, BY HEALTH PLAN (PROVIDERS WITH AT LEAST 10 HEDIS-ELIGIBLE BABIES)	49

EXECUTIVE SUMMARY

Work to revise payment policies in a way that would better support the delivery of high quality care has gained momentum in recent years after the Institute of Medicine's landmark report, *Crossing the Quality Chasm* (2001), cited it as a necessary change. Although a handful of commercial managed care companies had established physician performance incentives by 2000, such incentives were quite scarce in Medicaid, and no systematic studies had examined their value. This report examines the first known collaborative effort to establish financial incentives within Medicaid among multiple plans with the same objective.

Specifically, the Local Initiative Rewarding Results (LIRR) Collaborative was established in Fall 2002 to implement incentives designed to improve care for low-income children enrolled in California's Medi-Cal and Healthy Families programs.¹ The collaborative, which included seven "local initiative" low-income-focused health plans in California, was part of the national Rewarding Results demonstration program sponsored by the Robert Wood Johnson Foundation and the California HealthCare Foundation. The demonstration was administered by the Center for Health Care Strategies, and Mathematica Policy Research (MPR) led the evaluation. The demonstration ran through December 2005.

This report provides the final results of the evaluation. The report is based on several major data sources. First, we obtained plan-level HEDIS® data for care provided during 2002-2005 as well as national figures from NCQA for 2002-2004 and California figures obtained from the state for 2003-2004.² Second, plans supplied administrative (encounter) data aggregated to the provider level on provider type and specialty, well-baby visit and well-adolescent performance, and payout for 2003 through 2005. Third, the plans submitted aggregate data three times during the demonstration on the following: the number of physicians and members eligible for incentives, changes in the activities to promote quality that complement the incentives, aggregate payout information, and problems encountered in implementation.

¹ These are the state's Medicaid managed care and SCHIP programs, respectively.

² HEDIS is a registered trademark of NCQA.

Fourth, we conducted two waves of telephone interviews with stakeholders, including key plan executives and staff, and two to six providers in each plan, usually physicians. The interviews were conducted approximately a year after the start of the demonstration, then again in its last quarter. Finally, we drew on notes from steering committee meetings and related follow-up calls with plans that occurred periodically throughout the demonstration.

The evaluation includes both qualitative and quantitative analysis of these data sources, and draws on both in its conclusions. We use “difference-in-differences” analysis to evaluate the impact of the demonstration to the extent possible with available data, but there are several important limitations. We lack a true control group—that is, a set of plans or providers similar to those with incentives except for the presence of incentives. We use several comparison groups, described in Chapter II. While comparison of plans’ HEDIS data to national and California non-LIRR plan data are the strongest comparisons in the study, even then we cannot rule out that there may be differences between the LIRR plans and the typical California or national Medicaid plan that confound the comparison and thus could provide misleading results.

Due to sample size issues, we were not able to use multivariate analysis of the provider-level data to help sort out provider performance improvement associated with incentives versus associated with other provider practice characteristics.

INCENTIVE DESIGN AND IMPLEMENTATION

The incentives under the LIRR Collaborative were designed to improve HEDIS rates for well-baby visits by 15 months, well-adolescent visits, and/or the completeness of encounter data submission. However, the Collaborative plans’ incentive designs varied on key dimensions including whether they addressed all three of these goals, the mix of provider and member incentives, the performance requirement to receive an incentive payment, the mechanism for demonstrating performance, and other features (Table 1). In most cases, the financial incentives were one part of a larger set of incentives being implemented at the plans. Most of the plans began providing other supports to improve quality along with the incentives during the demonstration period. In particular, five plans began giving providers feedback on their performance, and four began mailing periodic reminders to members needing care. Also, three provided training and technical assistance on encounter data submission, and two began consulting with low-performing providers. The plans that implemented the incentives varied from a large plan with over 2,000 primary care providers serving infants and over 4,000 serving adolescents to a small plan with 72 physicians serving infants and 151 serving adolescents. Two of the five plans paid fee-for-service for well care prior to the incentives while in three payment for well visits was wholly paid through providers’ capitated rates.

Five of the original eight plans in the LIRR Collaborative implemented new incentives; four began between July and December 2003, and one began in July 2004. Two of the five plans substantially strengthened their incentives since that time, and another that had targeted incentives to only a few high-volume providers initially expanded the incentives to a few more. The three plans that did not implement the incentives made that decision for very different reasons. One reportedly withdrew from of the Collaborative because it felt the evaluation data requirements were too burdensome. A second announced a major incentive aimed at improving quality, only to find itself in a major financial crisis, making it necessary to cancel the planned

incentive distribution. The third plan decided instead to try re-advertising its existing fee-for-service payments for well-care, along with educating providers about coding and the plan's flexible policies on well-visit timing.

Table 1. Number of Plans Newly Introducing Each Type of Incentive Design

	Well-Baby	Well-Adolescent	Encounter data
Shift risk pool criteria toward quality	1	0	0
Bonus payment for HEDIS-related measures	4	1*	1
Member gift certificates	1	0	0
Member movie ticket(s)	0	2	0

*One additional plan implemented an incentive with several providers that provided \$20 for each adolescent they saw for an annual visit.

Through December 2005, almost \$5.5 million in incentives was paid out to providers under the Collaborative. The plans generally aimed to use the incentives to influence well-child care for the vast majority of babies and adolescents in their networks. A total of over 80,000 children whose care could potentially be improved are served by physicians who were eligible for incentives in 2004. In addition to the provider incentives, about 145,000 adolescents (in two plans) and 56,000 parents of infants (in one plan) were eligible to receive movie tickets or a gift certificate (respectively) if they/their babies received timely well-child visits.

OUTCOMES AND IMPACT OF THE DEMONSTRATION

The trend in LIRR plans' outcomes during the demonstration was overall quite positive. For example, two-year average HEDIS scores for the well-baby visit measure improved by 7.5 to 27 percentage points between the pre- and post-demonstration periods (2002/2003 vs. 2004/2005). Two-year average adolescent well-visit scores also improved, between 2.5 and 10 percentage points. The timeliness of encounter data submissions for preventive service encounters improved from 72 percent of submissions within 90 days in quarters 1 through 4 to 83 percent in quarters 5 through 9 at the one plan that implemented an encounter data incentive.

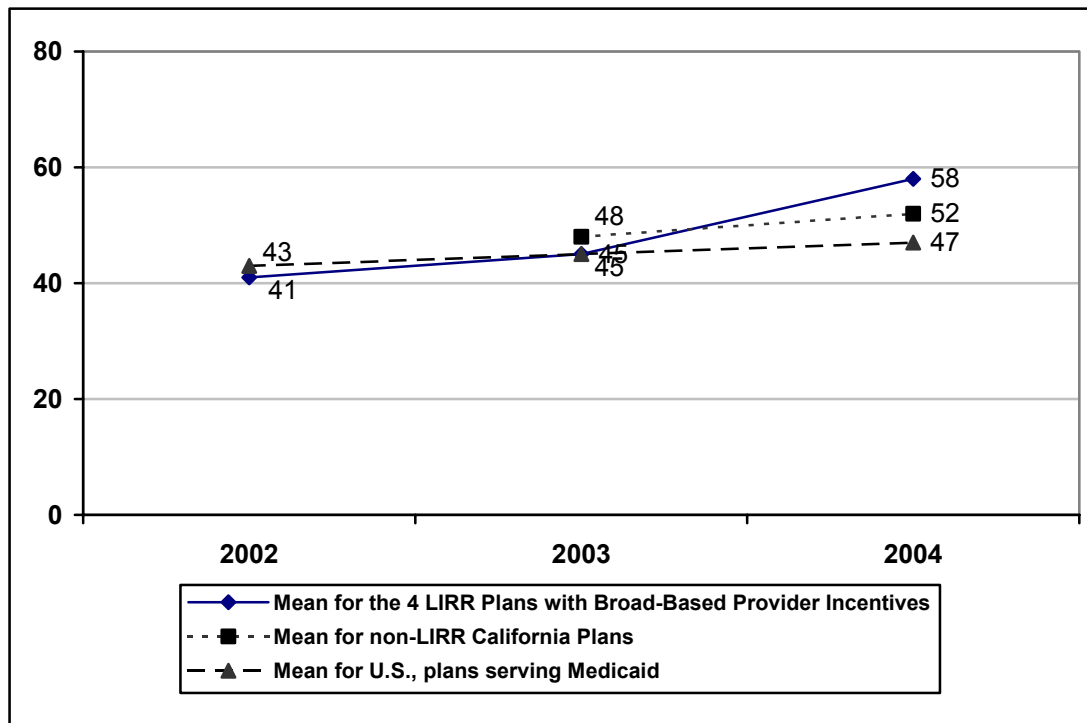
However, the rate of preventive service reporting actually declined for that plan.

Were the positive outcomes due to the incentives? We examined this in two ways. First, we compared performance across plans to several comparison groups. Second, for the well-baby visit incentive and the encounter data incentive, we assessed whether the cross-plan finding seemed to logically hold true for each individual plan, given our knowledge of the plan's program.

Cross-Plan Analysis

Provider incentives for well-baby visits may have helped plans improve their HEDIS scores, as plans and providers with broad-based provider incentives improved performance substantially more than the relevant comparison groups (Figure 1, and also see Table IV.4 in the body of the report).

Figure 1. Percent of HEDIS-Eligible Members with 6 or More Well Visits by 15 Months of Age, 2002-2004



We found less evidence that new incentives for adolescent well-care helped plans improve their HEDIS scores. That is, much more than for the well-baby visit analysis, the favorable or unfavorable comparison depended upon the exact plan, timeframe, and comparison group for the analysis. On the positive side, the LIRR plans that implemented incentives all showed changes in HEDIS scores well above the change in the national median change for Medicaid plans during 2002-2004. However, within the LIRR plans, the plan with the greatest improvement in the group implemented no new incentive until July 2005, and the HEDIS scores of several plans that implemented new incentives declined between 2004 and 2005.

The encounter data incentive may have helped the plan that implemented it stem a tide of declining encounter data, with the median rate of reported preventive service encounters per 1,000 dropping by 10.5 percent in eligible groups compared to 43.7 percent in non-eligible groups. Of note, the trend in reporting of services targeted by the incentive (preventive services) mirrored that of non-targeted services. A skeptic might point to this as suggesting that unobserved characteristics of the group of eligibles may be driving the difference between

eligible and non-eligible groups. Alternatively, this same change across services could mean that the technical assistance provided by the plan helped to resolve general reporting problems that affected reporting of all services, and/or could mean groups respond in a more generalized manner to incentives like this rather than specifically changing reporting for targeted relative to non-targeted services. Timeliness of submissions improved equally for both the groups eligible and those not eligible for the incentive.

Cross-Plan Analysis—Combining the Qualitative and Quantitative Data

Provider Incentives for Well-Baby Visits

The totality of evidence from the demonstration suggests that provider incentives can be a useful tool to help improve the percentage of babies that receive the recommended number of well-baby visits, but that it is not easy to achieve a large effect with them in the Medicaid context. Provider incentives are feasible to implement in a variety of types of plans serving Medicaid including small and large plans, plans with a wide variety of provider network arrangements, and plans with and without prior experience with incentives.

Although as a group plans that implemented provider incentives for well-baby care improved more than several comparison groups, the qualitative and quantitative analysis together suggests four of the five plans that implemented the incentives were not able to achieve large effects from the incentive. Due to data constraints we cannot confirm that a large improvement in HEDIS well-baby visit scores in the fifth plan was caused by the incentives.

Adolescent Well-Visit Incentives

The demonstration plans as a group improved substantially on their HEDIS well-adolescent visit scores, however there is less evidence than for well-baby visits that the incentives helped them do so. For example, the two LIRR plans with the most dramatic rises in well-adolescent scores were a plan whose incentive consisted of a \$20 per teen payment to several providers for well-adolescent visits, but that also conducted a major direct outreach campaign to adolescents, and a second plan that did not implement any new incentives until very late in the period (July 2005), but that continued its fee-for-service payment above capitation for adolescent well-visits. The one plan with the new broad-based provider incentives (\$50 per adolescent well-visit) did not improve significantly. Although the plans with new movie ticket incentives improved some, they did not reach the national mean by 2004 and did not improve any further in 2005. Further, the movie ticket redemption rates were low.

We cannot conclude from this that adolescent well-visit incentives do not work. The demonstration did not provide as rich a laboratory for studying the effects of adolescent well-visit incentives as it did for well-baby incentives, since only one plan implemented a broad-based provider incentive and two implemented new movie-ticket incentives. There may be other ways not yet known to these plans to improve the take-up rate of movie tickets, making them potentially more effective than seen here.

Encounter Data Reporting

In the one plan that implemented an encounter data incentive, the incentive may have helped the plan stem a tide of declining encounter data. However, results remain ambiguous, in part because the trend in reporting of service targeted by the incentive mirrored that of non-targeted services.

Other useful qualitative information about encounter data reporting was also uncovered as part of the evaluation:

- The rates of administrative reporting of well-baby visits rose dramatically in Plan G, a plan that did not implement a new incentive, but that used a physician champion to help re-advertise its pre-existing fee-for-service payments over capitation for well visits, along with discussing the flexible policies of the plan and articulating the importance to the plan of timely well-baby visits.
- After reviewing the plan-by-plan stories, we believe that for incentives to be designed effectively, they need to consider providers' starting point with respect to reporting of encounter data. That is, if the incentives are to be calculated based on administrative data, then the amount of behavioral change in reporting required to trigger an incentive payment should be considered as dollar amounts and thresholds for the incentive are set. This follows from qualitatively comparing this aspect of the plans' experiences between the two plans with broad-based incentives where the incentives more likely worked (Plan D and Plan A) to the two where the case for the incentives working is harder to make (Plan B and Plan C).

Plan-by-Plan Analysis of Likelihood of Impact from Well-Baby Visit Incentives

Below we highlight key findings for the five plans that implemented the well-baby visit incentives. The "stories" of the two LIRR plans that did not implement a well-baby incentive are provided in Chapter VI of the report. We first highlight the plan with the most impressive gains, then return to alphabetical order.

Plan D. This plan showed the most impressive improvement in HEDIS scores during 2002-2005, beginning low in 2002 at 33 percent and rising 24 percentage points, 11 percentage points, and 8 percentage points so that in 2005 its score was fully 74 percent above the 2004 national mean for Medicaid plans. The plan's success may have been due to implementing new incentives by building onto pre-existing fee-for-service incentives, which provided payment for well-baby care on top of capitation since 1997. This plan also may have largely avoided communication gaps around its incentive program that plagued several of the other plans.

Plan A. This plan's HEDIS two-year average HEDIS score improved by 9.5 percentage points between 2002/2003 and 2004/2005, rising to equal the national mean by 2004 and further increasing in 2005. Its improvement was far more than the national or California plan averages between 2003 and 2004, and continued to significantly improve through 2005. Despite the improvement, it remained below the other California plans' 2004 average in both 2004 and 2005. One can argue either for or against the likelihood that the incentives caused the

improvement. On the plus side, the plan created physician awareness of the incentives by mandating their attendance at a (conveniently scheduled) meeting to describe the program, as a condition of physician participation in the plan's sizable risk pool. The dollars at stake are substantial. On the minus side, uncertainty around how the program works and the indirect relationship between payout and performance in this program may have dampened enthusiasm for the program among physicians.

Plan B. This plan's HEDIS score also rose impressively during the period, however, the logic that the incentives caused this to occur does not hold when administrative data are examined. The incentive to providers is based on how many of their babies had six well visits appear in the plan's encounter database. If the incentive had caused the improvement in HEDIS scores, we would have expected physicians to report more well-visit encounters. This improvement should be visible by either a substantial improvement in the percentage of babies with four or more visits (a low bar) in the administrative data, or we would have expected that figure to be high and remain high through the period, indicating that provider had already been used to reporting most visits and were improving their provision of care and/or reporting on the fifth and sixth visits. Neither of these was true. The percentage of babies with four or more visits from administrative data was low in 2003 (29 percent) and did not consistently improve. This indicates that the physicians as a group were not good about reporting well-visits at the start, probably because they receive capitated payment their payment does not depend on reporting, and they did not improve in the face of the incentive. Since HEDIS scores reflect the combination of administrative data and documented well-visits from the medical charts, it appears that the boost to HEDIS scores must have been coming from an improvement in care (or care documentation) reflected in the medical charts but unrelated to the incentive.

Plan C. This plan's two-year average HEDIS score rose 7.5 percentage points between 2002/2003 and 2004/2005, and its improvement trend was well above the national average for Medicaid plans. However, its HEDIS score did not further improve between 2004 and 2005, remaining below the national and California plans average in both years. Several other factors described in Chapter VI, including mixed experience with communication with providers, and administrative difficulties with the program, leave us uncomfortable about suggesting whether the evidence on balance is for or against the incentives having had any effect. If they did have an impact, it does not appear to be a large one.

Plan E. Unlike the four plans above, Plan E narrowly targeted its well-baby incentive to a few high-volume providers for the plan. We were less able to assess trends for this plan because of missing HEDIS data in 2002 and the small number of providers with at least 10 HEDIS-eligible babies in the plan's administrative database. Although the plan's HEDIS score improved dramatically between 2004 and 2005 (after having dropped the year prior), our discussions with relevant providers and the plan together suggested that the incentives were likely not the cause of the plan's 2005 improvement. The plan believes that more effective on-site technical assistance to sites eligible for incentives helped make a difference, as well as direct outreach to patients for two high-volume clinics. However, the magnitude of change suggests that provides who did not receive this assistance also improved, for unknown reasons.

FACILITATING FACTORS AND BARRIERS

Facilitating Factors

At least three factors may have facilitated better results from the incentives based on the plan-by-plan analysis described above.

- Good communication between the plan and providers in general and about the incentives.
- Asking for a behavioral change from physicians that is commensurate with the dollars at stake. At the plan in which incentives may have contributed to a large improvement in its HEDIS score, providers were already in the habit of reporting well-baby visits on a regular basis due to the plan's fee-for-service payment policy, so less change was needed on their part than in other plans report the additional visits necessary to receive the new incentive.
- Total amount of payout dollars available to the eligible practices.

The idea that where there were larger payouts this may have fostered greater improvement is also supported by our cross-plan analysis finding that providers who received larger payouts in the early period improved more in the subsequent period for both the well-baby and the encounter data incentives. However, there may have been other factors that explained that pattern as well as high payouts or instead of them. For example, providers in the high payout group may be by nature motivated to perform well, or may be better able to improve performance given their office structure and resources, which could lead them to receive higher payouts at the start and to be more receptive to incentives.

Barriers to Greater Improvement

Across the plans, barriers to greater improvement in the demonstration, while not able to be assessed quantitatively, may have included:

- Relatively weak plan communication with providers about the incentives, with a couple of exceptions as noted above
- Population characteristics and office factors that constrained the extent to which providers who wanted to improve could do so:
 - Population factors were associated with the low socioeconomic status of the population targeted for improvement, and included parents' focus on economic subsistence, which makes timely well-visit care for their children seem less important; lack of transportation; and mobility of the population and the associated likelihood of inaccurate contact information.

-
- Office factors included clinics with a heavy walk-in population, for whom scheduling appointments is not a routine part of their operation; lack of enough staff to support outreach along with not enough money to support expanding staff capacity; and system issues such as lack of technology or underuse of technology.

Several program features in specific plans may also have inhibited greater improvement:

- In one plan, an overly ambitious threshold for providing an incentive, given the starting point of very low administrative data reporting by providers (the incentive was only paid if all six well-visits were found in the data, requiring a large change in practice).
- In another, a low dollar incentive per baby (\$50 for the six visits), a barrier which may have been aggravated by the fact that the typical provider in this plan served very few babies of the relevant age group, and providers' rates of reporting of well-visits at the start were quite low.
- In one plan, offering the incentive to only a handful of high-volume providers. This by definition limited the potential of the incentive to measurably improve the plan's score; the incentive would have had to have had a dramatic effect on all targeted providers, an unlikely scenario even under the best of circumstances.

CONCLUSION

The LIRR demonstration plans were pioneers in implementing financial incentives in the Medicaid context. The evaluation provides some evidence for the effectiveness of the financial incentives the plans implemented for well-baby care, although we believe that any *large* effect from the incentives was limited to one plan (at most). The greatest contribution of the demonstration and evaluation may be that it has generated lessons learned and hypotheses for the next generation of plan incentive programs and research.

Specifically, financial incentives need to be implemented as a program as well as a payment policy. As such, they require plans to effectively communicate to providers about them, to grapple with administrative and system issues in implementation, to ask for provider behavior changes that are reasonable given the starting point and the dollars at stake, and to use the incentives in a context that supports improvement, for example, offering providers monthly contact lists of enrollees due for a well-visit. Supports and realistic expectations for provider change given dollar amounts involved may be particularly important in the Medicaid context.

This page has been intentionally left blank for double-sided copying.

CHAPTER I

DEMONSTRATION OVERVIEW

The goal of the LIRR Collaborative in California was to improve health care for low-income children enrolled in Medi-Cal and Healthy Families by offering financial and non-financial incentives to providers and enrollees. The plans implemented incentives designed to encourage well-baby visits, well-adolescent visits, and improved provider submission of encounter data to plans.

The LIRR Collaborative demonstration was part of the national Rewarding Results demonstration program sponsored by The Robert Wood Johnson Foundation and the California HealthCare Foundation. The Collaborative was the only Rewarding Results grantee to focus on improving care for the Medicaid population. While collaborative activities were funded by the grant, the financial incentives established by the plans were all funded by the plans' own operating budgets. The Integrated Healthcare Association (IHA) administered the demonstration from Fall 2002 to February 2004. The Center for Health Care Strategies administered the demonstration from February 2004 through December 2005.

This chapter describes the quality incentive designs that the plans chose under the demonstration and explains their origin and context. Chapter II introduces the evaluation design and describes the data sources used in subsequent chapters. Chapter III assesses the implementation of the incentives, including the challenges and successes. Chapter IV shows the trends in key outcomes for the demonstration and analyzes the impact of the incentives across the demonstration as a whole. Chapter V shares providers' comments on their response to the incentives, based on the two rounds of provider interviews. Chapter VI synthesizes the qualitative and quantitative analysis of the well-baby and encounter data incentives, plan by plan. Finally, Chapter VII presents our conclusions.

A. VARIATION IN DESIGN

The sets of incentives that were planned and implemented under the LIRR Collaborative were all designed to improve HEDIS rates for well-baby visits by 15 months, raise HEDIS rates for well-adolescent visits, and/or improve completeness of encounter data submission.

However, incentive designs chosen by the plans have varied among the plans on several key dimensions:

- Whether they address one, two, or all three of the above-mentioned goals of the demonstration
- The mix of provider and/or member incentives used to achieve their goals
- The performance requirement to receive an incentive payment, and the relationship between achievement and amount of reward
- Whether incentives were paid in lump sums annually or semi-annually or paid on an ongoing basis (as relevant documentation is submitted)
- Whether plans paid based on administrative data (such that providers qualify automatically) or whether the plan required submission of documentation for each case

Table I.1 provides more specifics about the five plans' incentives that were implemented under the Collaborative, and Table I.2 shows the number of physicians and members potentially touched by the well-baby visits incentive program in each plan. In 2004, a total of 2,393 physicians were eligible for well-baby visit incentives as were almost 62,000 babies. In addition, about 145,000 adolescents (in two plans) and 56,000 parents of infants (in one plan) were eligible to receive movie tickets or a gift certificate (respectively) if they/their babies received timely well-child visits.

Although there were originally eight plans in the LIRR Collaborative, one formally withdrew during 2003. Two more did not implement any new incentives (discussed more in Chapter III). Since one of these two did design its incentives and in fact announced them, we include discussion of that plan where relevant. However, most discussion of implementation and impact (Chapters III and IV) is centered around the five plans that implemented some incentives.

Because none of the plans had seen evidence that any particular type of design worked better than others, their choice of designs were based in large part on their history and structure and their theories about what might work best. For example, Plan A and Plan F already distributed risk pool dollars annually; therefore, shifting the criteria for risk pool distribution to incorporate performance appeared the most logical form for their incentives. Plan C by its structure is an organization that is far-removed from primary care providers, since it works through "plan partners" who in turn often pay capitated rates to large medical groups who in turn pay individual providers. However, Plan C believed its well-baby visit incentives would have to be paid directly to individual providers, and that the combination of provider incentives together with member incentives could be a powerful one.

Table I.1. Incentive Programs Under LIRR

Plan	Well-Baby Care		Well-Adolescent Care		Encounter Data		Other Plan Incentives
	Type Of Incentive	Payout/ Distribution By December 2005	Type Of Incentive	Payout/ Distribution By December 2005	Type Of Incentive	Payout/ Distribution By December 2005	
Plan A	Annual risk pool distribution to providers influenced by performance on well-baby visits	\$1,517,700	One movie ticket for completing a well-visit; annual letter to member to create awareness	1,339 movie tickets were distributed	None		Mammograms, chlamydia screening, well-child visits ages 3-6, eye exams for people with diabetes
Plan B	Bonus of \$100 per baby who gets 6 or more visits by 15 months. Automatic to providers based on administrative data, paid semi-annually. Increased to \$200 in July 2004. Financed from risk pool dollars	\$103,900	Bonus of \$50 per adolescent who gets their well-visits. Automatic to providers based on administrative data, paid semi-annually. Financed from risk pool dollars	\$581,100			Mammograms, cervical cancer screening, chlamydia screening, providing BMI data
Plan C	Bonus of \$25 for each set of 3 well-baby visits (total possible = \$50 per child). Automatic to providers based on administrative data, annually.	\$707,600	Two movie tickets for adolescents completing a well visit; annual letter to member's parent and members and list to providers.	3,309 sets of movie tickets were distributed	Tiered pmpm payment to participating physician groups ("PPGs") for meeting encounter data volume thresholds. Paid semi-annually but \$ calculation based on monthly volume. Up to \$.32 pmpm.	\$2,055,300	

Table I.1 (continued)

Plan	Well-Baby Care		Well-Adolescent Care		Encounter Data		Other Plan Incentives
	Type Of Incentive	Payout/ Distribution By December 2005	Type Of Incentive	Payout/ Distribution By December 2005	Type Of Incentive	Payout/ Distribution By December 2005	
Plan D	\$100 bonus per child receiving 5 th well-child visit by 13 months and Polio, Hep B and Hib vaccines. Another \$100 for both another well-child visit by 15 months and relevant vaccines. Paid to providers as relevant documentation is submitted.	\$434,800 ^a	Offered movie tickets for a few months in 2003 but discontinued due to very low response. Began a financial incentive for IPAs in July 2005 (not studied)		None		Immunizations by age 2, ongoing bonus payments for each well-child and well-adolescent visit, pap tests, chlyaimdia screening, prenatal services, post partum, diabetes
Plan E	Tailored assistance to small number of high volume providers. Financial component: \$50 per child if 4 or 5 visits by 15 months and \$150 if 6 or more.	\$30,000	Adolescent visit incentives to three offices in 2004 and 16 in 2005 \$20 per teen seen	\$34,150	None		Ongoing movie tickets to adolescents, immunizations well-child ages 3-6, prenatal care

^aIn addition to this new incentive added under the LIRR, Plan D has had ongoing fee-for-service incentive payments (above capitation) that totaled \$1.9 million for well-baby visits in 2004.

Table 1.2. Number of Physicians and Members Potentially Touched by the Well-Baby Visit Incentive Programs

Plans Implementing Incentives Under LIRR	Number of Physicians Offering Well Visits to Medi-Cal Members Ages 0-15 months (2004)		Number of Medi-Cal Members Ages 0-15 months, December 31, 2004 ^a	
	Total	Eligible for Incentive	Total	Enrolled with a Provider Eligible for Incentive
Plan A	72	72	4,841	4,841
Plan B	368	368	3,112	3,112
Plan C	2,050	1,506	37,813	36,787
Plan D	449	428	15,656	15,656
Plan E	175	19	1,763	649
Total		2,393		61,045

^aAn additional 733 babies were eligible across the plans through the Healthy Families program, representing about one percent of the babies potentially touched by the programs.

The plans financed their incentives with either “new” money or reallocated money, which could affect provider perception and response. Three used what was essentially “new money,” that is, administrative dollars freed up from lower costs relative to revenue and/or excess reserves. Two reshuffled existing dollars—one by changing risk pool distribution criteria and the other by revamping its existing incentives program to better align with HEDIS.

B. MOTIVATION AND FINANCIAL ABILITY TO BEGIN INCENTIVES

All five plans sought to improve their HEDIS scores through the incentives because of their importance: plans’ standing relative to other plans in the state and the nation on various HEDIS measures is a matter of professional pride; their performance is important when seeking NCQA accreditation; plans believed that state-level incentives related to HEDIS measures were under development (this did occur as discussed below); and plans are non-profit, committed to quality care for low-income populations by their mission. The particular measures for the demonstration were selected primarily because of the low scores across the group of interested plans at the time the demonstration was being planned.

The potential for direct financial savings from improving well-baby and well-adolescent visits was not a motivating factor—and little discussed—probably because no relevant data were known to assess such potential, and because quality improvement, not cost reduction, was the demonstration’s goal. However, because of the structure of Medi-Cal managed care in the counties in which the Collaborative plans operate, beneficiaries are in one of only two plans.

Therefore, the plan is likely to incur costs from medical or behavioral health problems that could have been prevented or minimized through timely well-baby or well-adolescent visits, which likely contributed to some plan executives’ interest. Some plans also believe that their

non-profit status played a key role in moving forward, along with their commitment to pass as much money through to providers as possible to help them continue to serve the plan's Medi-Cal patients.

The demonstration reinforced the fact that health plan financial stability is a critical factor enabling incentives when one of the LIRR Collaborative members canceled its risk share distribution—through which it was to implement its incentives—as one of a series of major steps to try to regain a financial foothold after a crisis unrelated to the Collaborative.³ Other participating plans have not thus far experienced any substantial financial problems during the demonstration.

C. LIRR INCENTIVES IN CONTEXT

1. Incentives Are Accompanied by Other Quality Supports

Other quality supports that could improve the same HEDIS rates targeted by the financial incentives were established by the plans before and during the demonstration. Table I.3 shows that two supports, performance feedback to providers on the targeted measures and reminders mailed to members, were often established during the demonstration period.⁴ Consulting with low-performing providers began in two plans. Plans were generally already helping to link members to providers based on primary language spoken and/or distance/travel time prior to the demonstration. One plan also spearheaded a major effort to use its own staff to call the parents of all of its adolescent members who were due for a well-care visit, explaining the importance and encouraging them to make the appointment. That same plan hired a staff member to support outreach for well-baby visits and immunizations in several clinics that could not accept financial incentives.

2. LIRR Incentives Are Part of a Larger Set of Financial Incentives

LIRR incentives are most often one piece of a larger set of financial incentives that the participating plans are using to help improve their HEDIS rates. Specifically, some plans give incentives for mammograms (2 plans), chlamydia screening (2), well-child visits ages 3-6 (2), immunizations (2), prenatal care (1), cervical cancer screening (1), and eye exams for people with diabetes (1).

Several plans also had financial incentives that targeted the same HEDIS measures as the demonstration, but they were not counted for evaluation purposes as LIRR incentives because they pre-dated the demonstration. For example, one plan had an ongoing incentive of movie tickets for adolescents who submitted documentation of a well-care visit, and three plans pay a

³ A “risk share distribution” occurs on a periodic basis in some plans, usually annually, to distribute back dollars to providers that are withheld by the plan from capitated payments up-front in order to incentivize providers along certain dimensions. Traditionally risk share distributions have been based on utilization measures and/or volume of members, however, quality incentives can also be built into the formula.

⁴ Mailed reminders to members include an annual letter to adolescent members and/or their parents, and reminder letters to parents of infants.

fee-for-service amount to providers that supplements capitation for well-child visits (including well-baby and/or well-adolescent visits).

Table I.3. Quality Supports Other Than Financial Incentives

	Number of Plans (of 6 submitting the information)	
	Began Activity During Demonstration Period	Already Had Activity in Place
Performance Feedback to Providers		
Well-Baby Visits	5	0
Well-Adolescent Visits	4	0
Consultation With Low-Performing Providers		
Well-Baby Visits	2	0
Well-Adolescent Visits	2	0
Mailed Reminders to Members		
Well-Baby Visits	4	1
Well-Adolescent Visits	4	1
Reminders to Providers of Members Needing Care		
Well-Baby Visits	1	1
Well-Adolescent Visits	2	0
Generic Provider Mailings on Preventive Care Topics		
	2	1
Training and Technical Assistance on Encounter Data Submission		
	3	1
Periodic Reports to Providers on Encounter Data Volume		
	1	2
Linking Patients to Providers		
Well-Baby Visits	0	6
Well-Adolescent Visits	0	5
Other	2	1

Source: Plan tracking tools submitted to MPR for 2003 and 2004.

3. Competition Was a Motivating Force, but the Collaborative Plans Were Not Shadowing Other Plans' Programs

In evaluating the demonstration, it is important to document the Collaborative's market context and how it may have affected the design, implementation, and response. For example, in theory, three types of possibilities could have affected the demonstration implementation and outcomes.

First, did competing plans provide a major motivating force for implementing incentives in the participating plans? If so, understanding these dynamics could be important in making incentive programs appealing to other plans nationally. In fact, the plans' desire to improve their HEDIS scores relates to their market in the sense that they hope to demonstrate they are equal to or better than their competitor. However, plans were not responding to specific motivating actions of competitors. Several believed other plans in the market had some incentives in place that applied to providers for commercial and/or Medi-Cal populations, but they were not clear about the details and said they faced no competitive pressure (e.g., pressure from providers) to implement incentives. Several noted they were the dominant force in their Medi-Cal market area—the market leader not a follower. Thus, we do not find any specific unique market factors that motivated the demonstration.

Second, are other plans in the market (whether they compete for Medi-Cal business or not) implementing incentives that are similar to or different from the Collaborative's incentives? This may be important if the other incentives in the market interact in a combined or competing way with the LIRR incentives on LIRR providers who contract with multiple plans. Because we interviewed only a handful of providers in each plan, who tended to be high-volume providers with closer-than-average ties to the plan, we cannot rule this out as an issue. However, our interviews did not find this problem. Only one provider we interviewed reported having multiple plan contracts each with different incentives, but he did not view the differences or related paperwork as a problem, and noted that most focus on immunization. Health plans reported not hearing anything from their providers about other plans' incentives, although one plan believed its providers were becoming more vocal with other plans about its incentives.

Third, in what ways did the market dynamics facilitate or impede extending the LIRR collaboration to include other plans in ways that maximize the incentive's effect on providers? There were early discussions at the LIRR Collaborative Steering Committee meetings of attempting to collaborate with competitors on the incentives in order to have a larger impact on providers because each plan accounts for only a small share of their business. During the discussion, it was clear that some plans were open to this possibility, while others had such an acute rivalry with their competitors that the idea was viewed as not worth considering. In the end, it was considered infeasible for the LIRR Collaborative to take on extended collaborations under the demonstration.

CHAPTER II

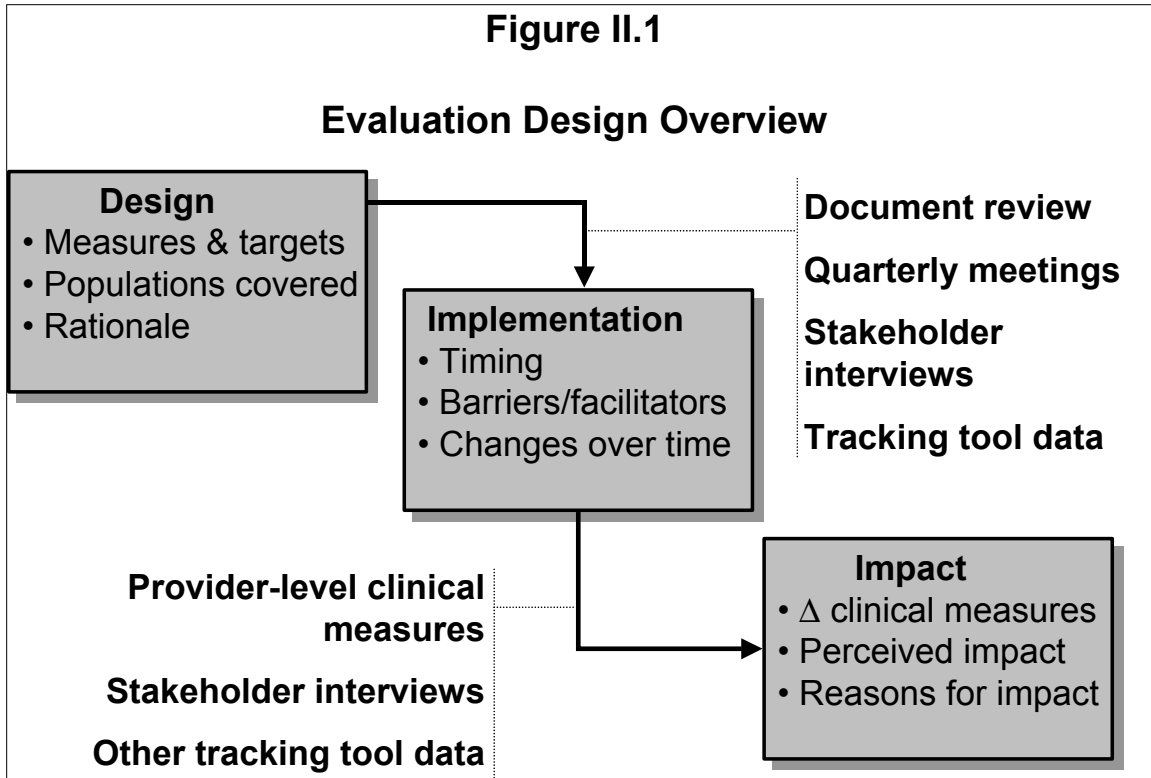
EVALUATION METHODS

In February 2003, the California HealthCare Foundation named Mathematica Policy Research as the prime contractor to evaluate the LIRR Collaborative demonstration through December 2005. The study evaluated the implementation and impact of the quality incentives newly established by seven Local Initiative plans serving the Medicaid and Healthy Families populations in California.

A. OVERVIEW OF EVALUATION DESIGN

The major goals of the evaluation were to evaluate the impact of the various incentive approaches on well-baby and well-adolescent care, and to identify the behavioral, organizational and operational factors that helped determine impact.

To measure the impact of the incentives on well-baby and well-adolescent care, we (1) analyzed administrative (claims or encounter) data submitted by providers to the health plans, (2) compared the plans' HEDIS (Health Plan Employer Data and Information Set) scores over time to those of other Medicaid/SCHIP plans in California and nationally, and (3) conducted interviews with plans and providers. To identify the factors that helped determine impact, we used interviews and document reviews analyzed together with the quantitative data. This mixed-method design minimized the evaluation's vulnerability to the kinds of data misinterpretations that typically threaten the validity of non-experimental studies that rely exclusively on quantitative evaluation methods. It also enabled us to provide health plans, providers, Medicaid agencies, foundations, and others interested in implementing rewards strategies with information on the factors that can contribute in important ways to the success or failure of performance incentive initiatives. Figure I.1 provides a schematic overview of the evaluation design, which is described below. The design includes two main components: an implementation/process analysis and an impact analysis.



1. Implementation/Process Analysis

The implementation/process analysis documented the achievements of the LIRR health plans, individually and collectively, in launching and sustaining their incentive systems, and identified the factors that facilitated or inhibited these achievements. When used together with the quantitative estimates of impact, the implementation/process analysis provided insights on why the various incentive systems did or did not produce their intended results. The implementation/process analysis had three main features:

- ***Collection and review of project-related descriptions and other documentation.*** In order to fully understand the implementation decisions and actions taken by Collaborative members at the outset and throughout the project, we obtained and reviewed quarterly reports and project descriptions generated by the plans for CHCS as well as meeting minutes generated by the Collaborative in the course of its work.
- ***Tracking of progress in developing and implementing incentive systems.*** We collected qualitative information on each plan's implementation progress from an electronic tracking tool that plans completed annually. This tool captured basic information on participating providers, enrollees, and incentives. It also provided information on changes to the incentives as well as changes to the environment in which the incentives are being implemented (state policy changes, marketplace changes, and the like). We obtained additional information through follow-up

telephone contacts with health plan staff annually to understand key changes to the performance measures and incentives that are reported, and by attending key meetings of the Collaborative, including quarterly steering committee meetings (often in person) and other relevant meetings (by conference call).

- ***Interviews with plans, providers, the Center for Health Care Strategies, and state Department of Health Services (DHS) officials using semi-structured telephone interviews.*** We conducted two waves of interviews with key Collaborative stakeholders during the course of implementation: wave one was conducted from October 2004 to March 2005, and wave two in fall 2005. These interviews collected qualitative information on implementation successes and setbacks and the perceived impact and value of the incentive systems. In each wave, we aimed to conduct interviews with six providers per plan and with key people at each plan (including the medical director, the CEO, and key staff working on the incentives effort). Providers were asked for their views on the incentive systems, and to explain any changes in practice that they believe affected their scores.⁵ Interviews included both higher and lower-performing providers, and focused particularly on larger-volume providers. Two of the six providers requested for each plan were requested to be “leadership” physicians, that is, physicians likely to have relatively more broad-based knowledge about incentive design and implementation than the typical physician. In practice, the “leadership” physicians generally had relatively close ties to the plan but their understanding of the incentives and providers’ reactions to it were not as different from other physicians as we had hoped. In the second wave, we interviewed a mix of physicians interviewed in the prior round as well as newly identified physicians. Plans provided the contact information for the providers.

2. Impact Analysis

The evaluation also assessed the Collaborative’s incentive systems impact in terms of their perceived impact, changes over time in HEDIS scores and other plan-level measures, and changes in provider-level performance.

Perception of Impact. As described above, we collected qualitative information on the perceived impact of each health plan’s rewards systems through two waves of semi-structured telephone interviews with health plans, physicians, and other key stakeholders. Physicians’ comments on their own practice changes are considered a useful though not statistically representative source of information for the analysis (Chapter VI).

Changes Over Time in HEDIS Scores. We used information on plan-level HEDIS scores to compare these measures descriptively over time (2002 to 2005) to determine whether the direction and magnitude of change supports the hypothesis that the Collaborative’s incentive

⁵ Providers interviewed are generally practicing physicians; in one case a nurse practitioner joined a physician on an interview; in a few cases we interviewed the office manager who handles submissions for the incentive; and in two cases we interviewed the heads of provider groups that receive encounter data incentives.

systems are improving plan-level performance. We compared the changes in HEDIS scores observed among the seven Local Initiative plans that implemented incentives to the changes in scores observed among other Medicaid-serving plans in California and nationally over the same time period.

The Impact of Incentives on Provider-Level Performance. Plans also provided an annual Provider Level Data Submission for 2002-2005. For every primary care provider in a plan’s network, this submission supplies: provider demographics, clinical performance measures, encounter data measures, incentive payment measures, and utilization measures. To assess the likely impact of the various types of incentives, we analyzed the patterns of change in performance as they relate to the key characteristics of the providers.

“Difference-in-difference” analysis is used to evaluate the incentives’ impact. In other words, we examined whether the incentives established by the seven plans have different effects on provider performance over time relative to available comparison groups. Table I.1 lists the research questions and summarizes the method we used for comparative analysis of demonstration outcomes and impact.

We did not have available a true control group—providers with no incentives who are similar in other ways to providers with the incentives. The following limitations apply and are taken into consideration as we draw conclusions in Chapter VII:

- Plan-level HEDIS benchmarks for non-LIRR plans in California and for the U.S. are the best comparison groups available to the study, however, one can argue that the average CA Medicaid plan or the average U.S. Medicaid plan may not be similar in important ways to the average LIRR plan. If that were the case, interpreting a difference in the trends between the groups as an impact could be misleading. Because of the lack of literature on the characteristics of plans that are important to success with incentives, we cannot either support or refute this type of charge by comparing relevant characteristics.
- Providers who are not eligible for incentives within the LIRR plans (across the plans) are a third comparison group used in the analysis. However, one of the LIRR plans that did not implement incentives comprises a large percentage of the non-eligible provider groups and thus dominates the results for the non-eligible group. If there was something importantly atypical about this plan, the comparison could be misleading. The atypical characteristic that we are aware of with this plan is that it was in financial crisis during the demonstration period. However, during this time it continued to pay primary care providers fee-for-service payment above capitation for well visits, as it always had.
- Primary provider groups within the same plan who were not eligible for the encounter data incentive are a fourth comparison group, used in the encounter data analysis. While we know the non-eligible primary provider groups have fewer plan members than those eligible for the incentive, the plan is not aware of any

other differences in the types of groups that make up the set of non-eligible providers.

We do not adjust the data for differences in patient characteristics such as age, gender, or health condition. We do not believe that statistical adjustments at this level of detail were either necessary or feasible because of the relative homogeneity of the pediatric Medi-Cal populations served by the local initiative plans and providers as well as the difficulty and cost of obtaining reliable and standardized data on these characteristics. Moreover, the difference-in-difference methodology we use implicitly controls for differences in the patient populations that do not change over time. Our analysis of the provider-level data assumes that in general providers' patient populations remain relatively stable over the period of our study (2003-05); we do not adjust impact estimates for changes in patient characteristics across providers. We did not have the data to test whether such an adjustment is advisable, although we do not have any information to suggest it is necessary.

Table II.1. Overview of Method For Quantitative Analysis of LIRR Outcomes and Impact

Research Question	Data Source(s)	Unit of Analysis	Measure of Interest	Expected/Hypothesized Result	Comparison Group(s) Available
Well-Baby Care					
Did the timeliness of well-baby care improve in participating plans? How did they improve relative to others?	LIRR plans' HEDIS scores, 2002-2005 NCQA data on national mean HEDIS scores State of California data on HEDIS scores for California plans not part of the LIRR	Individual LIRR plan, and groups of plans (e.g., all those that implemented broad-based provider incentives)	Percentage of babies who meet HEDIS continuous enrollment criteria and had at least six well-visits by age 15 months	Improvement in plan performance relative to comparison groups	National mean HEDIS score for Medicaid plans, 2002-2004 Mean HEDIS score for non-LIRR Medicaid plans in CA, 2003-2004 HEDIS scores of LIRR plans that did not implement new well-baby visit incentives, 2003-2005
How did performance change for incentive-eligible providers versus providers not eligible for an incentive?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Percentage of babies who meet HEDIS continuous enrollment criteria and had at least four well-visits by age 15 months	Greater improvement by providers eligible for the incentive relative to those not eligible	Providers not eligible for a well-baby incentive within LIRR plans who did and did not implement an incentive
How did performance change for incentive-eligible providers that at baseline were high-performing, versus those that were low-performing?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Same as above (at least four well visits)	Greater improvement by low-performers relative to high performers (cite JAMA article)	None—can review only performance strata within eligible group of providers
Did performance change differently for pediatricians versus family practice and general practitioners?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Same as above (at least four well visits)	Exploratory analysis	None—can review only specialty type within eligible group of providers
Did performance change differently among individuals and small practices versus medical groups?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Same as above (at least four well visits)	Exploratory analysis	None—can review only practice type within eligible group of providers

Table II.1 (continued)

Research Question	Data Source(s)	Unit of Analysis	Measure of Interest	Expected/Hypothesized Result	Comparison Group(s) Available
Did performance change differently for higher-volume providers versus lower-volume providers?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Same as above (at least four well visits)	Exploratory analysis. Higher volume providers would usually be expected to improve more than lower-volume providers, but the set of panel providers excludes providers with <10 HEDIS-eligible babies, so truly low-volume providers are already excluded from the analysis.	None—can review only specialty type within eligible group of providers
Did providers who received larger incentive checks in 2003 improve more than others?	Administrative (encounter) data submitted by the plans for the evaluation, aggregated to the provider level	Groups of providers across the plans	Same as above (at least four well visits)	Greater improvement by providers who received bigger checks.	Providers who were eligible but received less than the median but more than 0. Providers who were eligible but received nothing.
Adolescent Well-Visits					
Did the percentage of adolescents receiving an annual well visit improve in participating plans? How did they improve relative to others?	Participating plans' HEDIS scores, 2002-2005	Individual LIRR plan	Percentage of children ages 12-21 meeting HEDIS continuous enrollment criteria who had an annual well visit	Improvement in plan performance for plans that implemented a provider incentive or member incentive relative to comparison groups	National mean HEDIS score for Medicaid plans, 2002-2004 Mean HEDIS score for non-LIRR Medicaid plans in CA, 2003-2004 HEDIS scores of LIRR plans that did not implement new adolescent well-visit incentives, 2003-2005
Encounter Data					
Did the completeness of encounter data reporting for preventive visits improve among groups eligible for the incentive? Did they improve relative to groups not eligible?	Administrative data from the one relevant plan, containing counts of preventive service and other encounters, member months, and count of preventive service encounters submitted within 90 days, by provider group, for program quarters 1-9	Primary provider group (defined by the plan, includes multiple medical groups and practices)	Preventive service encounters per 1,000 member months	Improved completeness of reporting by groups eligible relative to groups not eligible for the incentive	Primary provider groups not eligible, within the same plan

Table II.1 (continued)

Research Question	Data Source(s)	Unit of Analysis	Measure of Interest	Expected/Hypothesized Result	Comparison Group(s) Available
Did the completeness of encounter data reporting for preventive services targeted by the incentive improve more than reporting for other services among eligible groups?	Administrative data from the one relevant plan (see just above)	Primary provider group	Preventive service encounters per 1,000 member months	Improved completeness of reporting for services targeted by the incentive relative to services not targeted	Rates of reporting of services not targeted by the incentive, within the same plan
Did the groups who were the poorest reporters initially improve more than others?	Administrative data from the one relevant plan (see just above)	Primary provider group	Preventive service encounters per 1,000 member months	The 25th percentile point for the completeness measure will improve more than the median among eligible groups	Change in the median is compared with change in the 25th percentile point
Did the timeliness of encounter data reporting for preventive services improve among eligible groups? How did timeliness change relative to groups not eligible?	Administrative data from the one relevant plan (see just above)	Primary provider group	Percent of preventive service encounters submitted within 90 days	Improved timeliness of reporting among groups eligible relative to those not eligible for the incentive	Timeliness of reporting for non-eligible groups within the same plan
Did groups who received larger checks in quarters 1-4 improve preventive services reporting more than others in quarters 5-9	Administrative data from the one relevant plan (see just above)	Primary provider group	Preventive service encounters per 1,000 member months	Groups who receive larger checks in quarters 1-4 will improve more than others	Groups eligible for the incentive who were paid less than the median

CHAPTER III

FINDINGS OF THE IMPLEMENTATION ANALYSIS

A. OVERVIEW

As described in Chapter I, five of the original eight plans in the LIRR Collaborative have implemented new incentives. Most began their incentives during July to December 2003; one began in July 2004. Since then, plans have made the following major changes:

- One plan substantially strengthened its well-baby visit incentive by announcing a new requirement in mid-2004 that to receive *any* risk pool dollars for contract year 2005 (paid in 2006), providers must be above the 2004 median performance on the well-baby visit measure [Plan A]. The plan found that its initial incentives design—a point system that gave some credit to providers for each 15-month-old with four visits and more credit for five or six visits—resulted in poor performers who served a high volume of patients at the four-visit level who received as much as or more than high performers who served fewer. With the revised design, poor performers who do not improve will get no risk pool distribution. The minimum threshold, calculated from administrative data only, is not a high hurdle—only about 15 percent.
- Another plan (Plan E), which began by offering its incentives to only three high-volume providers, expanded the offering to several more.
- A third (Plan B) increased its bonus amount in July 2004 from \$100 per baby for six well-baby visits by 15 months to \$200 per baby for six visits.

The three plans that did not implement the incentives had very different reasons for not doing so. One withdrew from the Collaborative reportedly because it felt the evaluation data requirements were too burdensome. The second worked hard to overcome historical political barriers to implementing the incentives, and in fact announced a major shift in its method for allocating risk pool dollars, but they had to cancel the risk pool distribution entirely as one of many steps to recover financial viability after a crisis. The third plan decided not to implement any new incentives largely because it thought it might be able to accomplish the same thing by re-advertising its ongoing fee-for-service payment for well-child care that supplements capitation, along with a campaign to get doctors to incorporate the components of well-visits

into sick visits. The plan's analysis shows most of its adolescents see a plan primary care provider for at least one sick visit during the year. This plan also mentioned that providers on its board resist "tampering with" reimbursement.

In the course of the demonstration's implementation, both plans and the demonstration itself experienced turnover of key staff. Several plans experienced turnover of medical directors who had been key champions for the incentives and made the decision to participate in the LIRR Collaborative. While the turnover was unrelated to the incentives, it created challenges in terms of continuity. The loss of the Plan C executive who had headed the Collaborative's encounter data workgroup may have stalled momentum among some plans that were initially interested in improving their encounter data through incentives.⁶ The demonstration itself began under the administration of the Integrated Healthcare Association (IHA). In February 2004, demonstration administration shifted from IHA to the Center for Health Care Strategies.

B. ELIGIBILITY FOR AND DISTRIBUTION OF INCENTIVES THROUGH DECEMBER 2005

When they implemented incentives, the plans generally aimed to influence well-child care for the vast majority of babies and adolescents in their networks. A total of almost 150,000 children whose care could potentially be improved are served by physicians who were eligible for incentives in 2004 (61,800 babies ages 0-15 months and 88,000 adolescents).⁷ In terms of provider eligibility, the vast majority of physicians who serve babies in the plans' networks (at least 73 percent) were eligible for the well-baby incentives. The only exception was one plan that took a far more targeted approach, offering the incentives to only seven high-volume providers by the end of 2004. In all, 2,431 physicians were eligible for the well-baby incentives (across all five plans),⁸ and 923 physicians were eligible for well-adolescent visit incentives (in two plans) in 2004.

Incentives paid out to providers under the Collaborative through December 2005 total about \$5.5 million:⁹

- \$2.8 million for well-baby visit incentives, across five plans
- \$615,250 for well-adolescent visit incentives, in two plans
- \$2.1 million for encounter data incentives, in one plan

⁶ In May 2003, the oral report from the encounter data subcommittee of the Collaborative indicated that encounter data incentives would be paid quarterly by all plans but one; however, only two plans actually implemented an encounter data incentive.

⁷ Nearly all the babies were Medi-Cal members (99 percent) rather than Healthy Families members, and 84 percent of the adolescents were Medi-Cal members.

⁸ Sixty-two percent of the eligible physicians were located in one plan (Plan C).

⁹ This indicates the amount paid out by December 2005, but covers services for various periods consistent with plans' varying incentive payment strategies.

Several of the plans indicated the payout figures were under their allotted budgets for the incentives in 2004 and hoped the payout would increase in the coming months to reflect more widespread participation and better performance. Payouts generally did increase in 2005 relative to 2003/2004. One plan increased its well-baby incentive payout from \$181,000 in the first year to \$526,600 for 2005, while another increased from \$143,300 to \$286,500. When available, the number of physicians participating in the incentive program is a meaningful indicator of physician involvement. This calculation was possible for the two plans that required physician offices to submit documentation to receive an incentive. For the one of these plans, the average three-month participation rate was 55 percent for the well-baby visits (of 428 eligible providers) and 53 percent for the well-adolescent visits (of 537 eligible providers). In a second plan, with targeted eligibility, three of seven eligible provider sites participated (fewer than 19 physicians).

In addition to the provider incentives, a few plans provided member incentives. About 145,000 adolescents (in two plans) and 56,000 parents of infants (in one plan) were eligible to receive movie tickets or a gift certificate (respectively) in 2004 if they/their babies received timely well-child visits. The number of these groups who actually participated and thus received the items in 2004 was much smaller: 3,200 adolescents and parents of 1,800 infants. The percentage of eligible adolescents who turned in a form to receive the offered movie tickets may be termed the “take-up rate.” Take-up rates for the movie tickets were low in the first year and did not increase substantially in the second year of the program. For example, one plan’s take-up rate for its movie tickets was .8 percent for adolescents in the first year and 1.8 percent in the second year. The other plan reported a movie ticket take-up rate of 4.7 percent for year 1, and only 2 percent in year 2. A third LIRR plan did not implement any *new* member incentives during the period, but a long-standing incentive of movie tickets for adolescents is considered an important ongoing initiative by the plan and take-up rates in 2004 were at 13.5 percent, suggesting the two plans newer to this type of initiative could potentially improve take-up over time to at least that level.

C. IMPLEMENTATION CHALLENGES

1. The Plan Perspective—Provider Incentives

We found that implementation challenges arose to some extent at almost every step of implementation except design. Overall, plans found that successfully implementing incentives took considerable effort, and more than they anticipated. Table III.1 shows the problems plans identified through the tracking tools they submitted to the evaluation. The text below explains the issues identified in our first wave of plan interviews.

a. Design and Approval to Proceed

Adopting a Unilateral Approach to Design Kept the Process Simple. The plans did not struggle much with incentive designs, which were typically decided on a fairly unilateral basis, that is, without much provider input. Although a unilateral approach is not recommended by other qualitative research on incentives that has come out since the demonstration began (Verdier et al. 2004), we do not see evidence to date that the unilateral approach to initiating the

Table III.1. Implementation Problems or Barriers

Problem or Barrier	Number of Plans		
	2003	2004	2005
Number of Plans Responding to the Question	6	4	3
Lack of Awareness or Understanding of Incentives			
By Physicians	3	3	3
By Members	3	0	1
Negative Physician Reaction to Incentives	1	0	1
Problems Obtaining Claims/Encounter Data Needed for Performance Measures	3	1	1

Source: Plan tracking tools submitted to MPR.

incentives mattered much for these particular incentives, since the incentives were typically supported by the physicians we interviewed and the plans reported no negative feedback from providers. Note that these particular incentives relate to recommended access rather than to a particular clinical guideline. However, one plan convened focus groups of providers and members, and chose movie tickets for its adolescent well-care visit member incentive on the basis of the focus group comments. Members had ranked movie tickets as a desired item second only to concert tickets.

Plans Faced Only Isolated Resistance to Approving the Incentives. Health plan boards typically were presented with the proposed designs and readily approved them, except in one plan where the providers who sit on the board strongly preferred no changes to the reimbursement arrangements. The plan's CEO then became involved in persuading the relevant individuals—with eventual success. A key factor in changing their position was reviewing with them the actual data and improvements of greatest concern. Also, since the plan that does not contract with individual providers, but wanted to pay individual providers, they needed to negotiate this point with the organizations with which they contract. There was some resistance by one of these organizations, which wanted to receive the money directly, but the organization relented.

b. Roll-Out

Gaining Physician Attention Was Difficult. The plans found it difficult to gain the attention of physicians in their network, and it appears that they had mixed success in gaining attention during roll-out, although we do not have statistically representative data to judge awareness. In 2004, three of four plans answering our question about implementation problems cited lack of physician awareness or understanding as a problem. The providers we interviewed who were closely connected with the plans (such as those on the board or quality committee)

were typically aware in a timely manner, but several of the community physicians we interviewed were either unaware of the incentives or had learned of them only three to six months before our call, even though the programs had been rolled out about a year earlier. The plan that greatly strengthened its incentives in the second year made the greatest effort to ensure awareness at that time (summer 2004), mandating that all physicians attend one of three sessions to learn about the incentives as a condition of participating in the risk pool in 2005. Other plans typically sent letters, provider bulletins, or a more stylized page of information; some also included mention in their provider newsletter and some held meetings that were attended by a relatively small fraction of the network physicians. For plans that require provider offices to submit documentation, training of provider office staff was another important activity—both initially and on an ongoing basis—because office staff change positions frequently.

Administrative Challenges at Start-Up Were Significant for Three Plans. The administrative aspects of roll-out were challenging for three of the plans for different reasons. Plan C faced a challenge because the plan does not contract with individual physicians, yet the well-baby incentives were to be paid to individuals in a large network. Therefore, it had to create a new database with tax ID numbers and ensure current contact information for the full set of eligible providers. Plan D had to reorganize people and establish a process to handle incoming documentation, since it required providers to submit chart documentation of the visits and immunizations, which required substantial effort from several staff. Later the plan was able to shift to allow online submission, which reduced the effort. Plan E was implementing incentives as a component of a wholly revised strategy for interacting with and influencing providers; therefore major personnel and process changes were made along with implementation, though not solely due to the Collaborative’s incentives.

c. Performance Feedback and Payout

Plans Faced System and Data Challenges. Two of the three plans that used administrative data to calculate providers’ performance and payout amounts encountered data and system challenges, and one of the plans processing provider-submitted data cited programming issues as a problem with initial implementation. One plan has allowed providers to review their data for accuracy before finalizing the payout. The plan acknowledges that sometimes the data are wrong, and providers are allowed to send in records to support corrections. Correcting data does add to the cost of administering the incentives (and checking data adds effort on the provider side as well). In the other plan with system issues, the first payout was delayed by several months due to a major systems change at the plan from one data warehouse to another. The change affected access to the data and absorbed the time of personnel who were needed to check the data prior to payout.

Both plans and physicians mentioned miscoding of well-care visits as an issue that contributed to data inaccuracies and produced lower-than-accurate scores. To help with the problem, one of the plans (Plan A) held a forum for billing and administrative personnel to instruct providers on properly using codes to get “hits” for the incentive.

Two plans reported having to continue to perform the payout function of the incentive program manually in 2005, despite earlier efforts intended to automate the process. As a result,

in one of the plans payouts that were originally intended to occur twice per year occurred annually instead. Both are expecting that the system issues will be worked out within the next year to allow automated payout.

Ensuring the Incentives Reach Their Targets Was an Unexpected Difficulty. For two plans in particular, getting the incentives to the targeted individuals has been challenging in certain instances. Specifically, both plans have experienced difficulty with high-volume providers in their networks that employ physicians whose employment and/or contractual arrangements prohibit the employed individuals from receiving financial incentives. Sometimes this meant the money went into the clinic's general operating fund rather than the individual provider/staff. In other cases the problem has seemed unsolvable and the provider is deemed ineligible for the incentives. In one case where a financial incentive was not feasible, the plan made available a plan-employed staff member to conduct outreach to parents of babies who needed visits and/or immunizations.

While getting the payout to the targeted individual can be a problem, some of the plans target payment to a contracted entity, which is often an organization rather than an individual. Plans expressed uncertainty about whether any money is passed down to the relevant individuals when a contracted organization is paid, rather than an individual, and in fact, they assumed that it often was not passed down.

2. The Plan Perspective—Encounter Data Incentive

Plan C found (and providers confirmed) considerable challenge in effectively implementing the encounter data incentive in a plan with complex arrangements for care delivery.¹⁰ The first challenge is that only one entity can be the explicit target for the incentive, yet several organizations and individuals and their systems must all work smoothly together to permit a complete, accurate encounter dataset at Plan C. When only one of the relevant entities receives the reward for improvement, it may be less likely that the entire process receives the necessary attention. In this case, the Participating Physicians Group (PPG) receives the incentive. The PPG is a large provider organization that may include multiple medical groups and/or IPAs, which in turn include multiple physician offices. Some pass down the incentive to smaller entities within the PPG, but some do not.

The second major challenge is that due to the multiple system interfaces as well as the human component required to produce complete, accurate data, it is challenging even to diagnose the origin of a data problem so that it can be addressed. In June 2004, Plan C held a summit to (1) discuss strategies to improve encounter data submission, (2) identify ways to help more medical groups meet the encounter data benchmarks, and (3) hear best practices from groups/IPAs that have been successful in earning the incentives. Over 40 individuals representing plan partners, medical groups, and IPAs joined plan staff for the day.¹¹ Technical assistance, free-of-charge, from Plan C staff and its encounter data consultant was offered to the groups/IPAs to assist them in better success with the incentive program.

¹⁰ The challenges are likely to be lessened in a plan that contracts directly with individual providers.

¹¹ Summits were also held in 2003 and 2005.

The third challenge is simply the limited amount of money available from the plan for the incentive, especially when the solution to the problem may be considerably more expensive. One PPG explained that it has its own internal incentive for its IPAs to improve encounter data, with one-third of a bonus pool distributed on that basis. The payout was \$250,000 last year, in contrast to two payments it received from Plan C of \$10-15,000. This PPG notes that it is the top-performing IPA on encounter data, in an environment where medical groups tend to do better since they can exert more direct influence on their physicians. However, despite its efforts, it is only in the mid-range of performance on this measure among Plan C PPGs. Plan staff noted hearing from some providers that it costs them much more to submit an encounter data form than the top incentive offered (\$.32 pmpm).

3. The Plan Perspective—Member Incentives

The Need to Obtain State Approval Delayed One Program. The need for state approval of the member incentive program delayed implementation in one plan by about six months.¹² The plan had to overcome initial state concerns to ensure the cash value of the gifts to an individual would not exceed \$50 per year, and that the movie tickets were not at all health-related.

Automating the Process Proved Challenging. Working to automate the process of handling the cards that are sent in to claim the incentives has been a challenge. It has required regular meetings and brainstorming among several different departments within the plan including claims, information systems, provider operations, health services. (The cards are sent to the adolescents and must be signed by the doctor providing the well-care visit in order for the member to redeem the movie ticket.)

One Plan Found It Needed to Counter Fraud. An initial spate of fraudulent cards was a problem at one plan, which reported that about 20 percent of the cards initially sent to the plan claiming movie tickets were either signed by the doctor without a well-visit, or the doctor's signature was forged by the member. The plan continues to be vigilant and investigates and follows up on patterns that may constitute fraud.

4. The Provider Perspective

Physicians we interviewed in the five plans that implemented incentives raised two common themes regarding implementation.

Clear Communication About the Incentives Is Critical. Providers we interviewed in both waves frequently told us that better communication from the plans about the incentives would be helpful. In one plan providers received a single check and aggregate data across several performance measures, so the providers did not have the ability to distinguish their performance on well-baby or well-adolescent visits from their performance on other measures, or to see how much of their check was attributable to performance on the various component

¹² A second plan that implemented a movie tickets incentive for adolescents earlier than this plan did not report any delay due to state approval.

measures. In other plans, providers were not sure if the information they had received contained such detail, however, most felt that such information would be useful. A few felt that this information was not necessary because they were too busy to read such information. In other plans, the communication issues were more mixed. Some providers suggested that it was difficult to understand how the incentives work or who the money went to, and one provider received a notice about the incentives that indicated she would be individually assessed and rewarded, only to find she could not receive the money as an individual because of her employment arrangement. Those interviewed in the second wave tended to want the plans to communicate more on how the incentive program works, and/or detail on earnings, including how a provider could receive the maximum amount of the incentive.

Physicians affiliated with one plan provide an exception to the overall pattern of sentiment that communication could be improved. The physicians were complimentary about the plan's communication on several dimensions:

- Communication in general was said to be fine with this plan (in contrast to its competitor).
- One provider remembered specifically discussing his/her scores with the plan, and found the plan's explanation around the measures very straightforward and helpful.
- Another said the feedback he/she receives on performance is "simple, quick to read, but gives you the essential information—how much you got, how much you could have gotten, and what you need to do to make up the difference."
- A physician from a large group explained that because they were quite large (over 50 physicians) they didn't have to go to the plan to hear about the incentives; rather, the plan sent a staff member to them to explain.
- One of the physicians commented that the plan has roundtable discussions every month or so which is great. "They explain changes and what is going on within [plan name] and they ask if any of us (physicians) have any concerns, which we discuss frankly." The physician commented that this makes them feel a part of what they are doing.

Providers Appreciate When Other Supports for Outreach Accompany the Financial Incentives. Specifically, some of the plans provided support in the form of lists of children who were about to turn 15 months old, and one provided pre-addressed reminder cards that the office could stamp and send. Some of the practices we spoke with used and appreciated these tools. At the same time, several physicians believed the plans could be doing more: for example, several believed that the plan had better address information or change-of-address information for their patients than they did, and that information should be shared routinely to facilitate outreach. Some of the physicians suggested the plan should do the outreach centrally, although it was not clear that they were suggesting eliminating the provider incentives. In addition, some providers in the second wave of interviews stressed that if they are going to receive additional

support that the information should be accurate. Some physicians felt that lists that were sent to them or the data being used to determine if they qualified for the incentive was not always accurate. For example, one physician commented that the patient list sent to him was “roughly 50 percent” inaccurate. Many others stated that at least 5 percent of the contact information the plan included was inaccurate.

Other Physician Comments on Design and Implementation. At least one physician commented as follows on each of several other design and implementation features:

- The clinical appropriateness of the structure of one of the incentive measures in one plan was questioned. Specifically, providers receive a \$100 bonus for a patient if the patient has received a well-child visit between 12 and 15 months of age with full IPV, Hep B, Hib and DtaP (immunizations). There should be six months between the third and fourth DtaP shot for it to be maximally clinically effective. Often, the third shot is not given exactly on time, leaving the clinician with a choice between forgoing the incentive or ordering the DTaP earlier than is ideal. The plan made a change in 2005 to accommodate this concern.
- Because it is impossible to get some patients to come in for their visits, the plan should reward provider effort (e.g., reward for three attempts to schedule a patient) rather than actual visit timeliness.
- To enable coherent medical group management, the medical group should at a minimum receive a copy of the information on performance and amount of bonus being provided to its employed providers.
- For the particular measures in the Collaborative, the money should go to the group rather than the individual provider, since it is outreach by administrative staff that will contribute most to improving the visit rates.
- Member incentives (where not present) would be very helpful accompaniments to the provider incentives and could help overcome the difficulties providers have in getting people in for appointments.
- For the adolescent well-visit movie-ticket incentive, the plan should consider allowing the physician to submit a note authorizing movie tickets. The provider could then improve the effectiveness of its outreach by guaranteeing the adolescent would get movie tickets even if he/she did not have the card the plan sent on his/her birthday. This provider said that adolescents often lose the cards between their birthday and the sometimes much later appointment.
- However two providers in the second wave of interviews did not like the idea of having to fill out or keep the forms at the physician office because it created additional paperwork that they had to manage.

- The plan should not change the design of the incentive program each year. It has made it too difficult to follow.
- A few providers recommended other incentives either in addition to well-baby and well-adolescent or to replace them: One provider commented that there should be incentives set up for chronic conditions (such as diabetes and asthma). Another provider stated that an incentive should be given for those providers that regularly update their data systems to lower the level of clinical errors. Another provider thought that incentives for attending meetings was important. Finally, another provider thought that his plan should add a newborn incentive for exams within the first weeks after leaving the hospital.

D. SUSTAINABILITY

All the plans that implemented incentives are continuing with incentive programs, which they expect to refine over time to make them more effective. In particular, the plans are anxious for more information about whether incentives work and which incentive strategies bring better results; they expect this information to come from the LIRR evaluation.

In addition, state Medicaid policy changes have affected plans' thinking about their future incentives. By the time of our plan interviews in fall 2004, plans were aware that the state Medicaid agency planned to begin using HEDIS measures as a major factor in auto-assignment of Medi-Cal members, though the details had not yet been finalized. On the positive side, plans cite the state's new policy as an important reason to continue with their incentive programs, although not specifically the well-baby visit incentive, since that is not included as one of the state's selected measures. One plan stated, however, that it viewed the well-baby visit incentives as key to helping the plan to achieve its immunization goals, which are a measure included in the state's new auto-assignment policy. Plans cite the downside of the policy as inhibiting their willingness to work collaboratively with competing plans for quality improvement, since the plan in each county with the highest HEDIS scores will be rewarded with additional membership.

CHAPTER IV

TRENDS IN OUTCOMES AND IMPACTS OF THE INCENTIVES

This chapter first displays the trends in outcomes for well-baby visits, well-adolescent visits, and rates of encounter data reporting for the plans that implemented new incentives. Next, the chapter draws on comparative analysis to the extent possible to examine the likely impact of the incentives. The final section examines whether provider groups with larger payouts early in the program improved more than others. Key findings include:

- We find generally positive trends in outcomes during 2002-2005. For plans that implemented new incentives:
 - Two-year average HEDIS scores for the well-baby visit measure improved by 7.5 to 27 percentage points between the pre- and post-demonstration periods.
 - Two-year average HEDIS scores for the well-adolescent visit measure improved by 2.5 to 10 percentage points between the pre- and post-demonstration periods.
 - The timeliness of encounter data submissions for preventive service encounters improved from 72 percent received within 90 days in quarters 1 to 4 to 83 percent in quarters 5 to 9. However, the rate of preventive service encounter reporting actually declined.
- Provider incentives for well-baby visits may have helped plans improve their HEDIS scores, as plans and providers with broad-based provider incentives improved performance substantially more than the several relevant comparison groups.
- We found less evidence that new incentives for adolescent well-care helped plans improve their HEDIS scores. That is, much more than for the well-baby visit

analysis, the favorable or unfavorable comparison depended upon the exact plan, timeframe, and comparison group in the analysis.

- The encounter data incentive may have helped one plan stem a tide of declining encounter. However, results remain ambiguous, in part because the trend in reporting of services targeted by the incentive mirrored that of non-targeted services.
- Groups with larger payouts in the early period improved more than others in a subsequent period on the well-baby and encounter data outcomes.

A. TRENDS IN OUTCOMES

1. HEDIS Outcomes for Well-Baby Visits and Adolescent Well-Visits

The percentage of 15-month-old babies who received the recommended number of well visits (six) improved substantially over time in all the plans that implemented provider incentives targeted to this measure (Table IV.1). In particular, the far right column on Table IV.1 summarizes the positive change between the two years that are mostly prior to the incentives (since the incentives began no earlier than July 2003) versus the two years after the incentives were implemented. The plans' scores improved by 7.5 to 27 percentage points. Note that the plans that undertook the incentives varied widely in their scores prior to the demonstration, and that variation remained after implementing their incentives. In other words, if one ranked the four plans with all four years of data based on their pre-demonstration scores, none of the plans had changed their rank order by the end of the demonstration.

The percentage of adolescents who received an annual well-visit also improved in the plans that implemented incentives. In this case, the improvements were more modest than for well-baby visits, with two-year averages improving between 2.5 and 10 percentage points.

Outcomes measured by HEDIS data include both administrative data and medical chart reviews, using a standard methodology. They are thus relatively unaffected by changes in administrative data that may have occurred during the demonstration period,¹³ and offer the best assessment of whether care actually changed plan-wide between the period before and after the demonstration.

¹³ Some plans argue that it is easier to get credit for a well-visit through administrative data, where components of the visit are not examined, than it is through chart review, where to give credit for a well-visit the components of the well-visit must be in place, giving plans with good administrative data systems an edge in their scores.

Table IV.1. HEDIS Trends for Well-Baby Visits and Adolescent Well Visits Measures, Participating Plans, 2002-2005

	Pre- implemen- tation 2002	Year of implemen- tation 2003 ^a	Post- Implementation		Percentage Point Change 2002-2004	Percentage Point Change 2003-2004	Percentage Point Change 2003-2005	Two-Year Average 2002/2003	Two-Year Average 2004/2005	Percentage Point Change In Two-Year Average
			2004	2005						
1. Percent of HEDIS-Eligible Members with 6 or More Well-Baby Visits by 15 Months of Age										
Plans with Broad-Based Provider Incentives										
Plan A	42	37	47	51	5	10	14	39.5	49.0	9.5
Plan B	49	39	61	54	12	22	15	44.0	57.5	13.5
Plan C	33	40	44	44	11	4	4	36.5	44.0	7.5
Plan D	39	63	74	82	35	11	19	51.0	78.0	27.0
Plan with Narrowly Targeted Provider Incentive										
Plan E	NA	61	56	73	NA	-5	12	NA	64.5	NA
2. Percent of HEDIS-Eligible Members Age 12-21 with Annual Well-Visits										
Plans with Provider Incentive										
Plan B (Broad-Based)	30	38	38	35	8	0	-3	34.0	36.5	2.5
Plan E (Narrow, & Major Outreach)	36	38	45	49	9	7	11	37.0	47.0	10.0
Plans with New Movie Ticket Incentives										
Plan C	23	38	37	37	14	-1	-1	30.5	37.0	6.5
Plan A	28	26	37	25	9	11	-1	27.0	31.0	4.0

^aFour of the five plans that implemented incentives did so during July-September 2003. One plan (PLAN D) implemented its incentive in July 2004.

Table IV.2. Performance of Providers Eligible For Incentives, by Practice Characteristics and Baseline Performance Level

	Mean Performance of Providers, Percent of Babies Turning 15 Months with 4+ Well-Baby Visits				
	2003	2004	2005	Change in Performance 2003 vs. 2004	Change in Performance 2003 vs. 2005
Type of Providers with at least 10 HEDIS-Eligible Babies					
Specialty					
Pediatrics (n=168)	49.9	55.8	51.9	5.9	2.0
Family or General Practice (n=44)	52.1	54.3	52.6	2.2	0.5
Type of Practice					
Individual or small practices (n=202)	54.8	58.7	56.0	3.9	1.2
Medical groups (n=57)	41.7	44.4	43.3	2.7	1.6
Patient Volume (Number of LIRR Plan- Enrolled HEDIS-Eligible Babies Turning 15 Months)					
Lower volume (10-19) (n=92)	37.3	46.5	40.8	9.2	3.5
Middle (20-29) (n=19)	47.1	48.9	42.5	1.8	-4.6
High Volume (30+) (n=172)	62.4	63.4	63.2	1.0	0.8
Performance Tier					
Low (n=74)	18.0	42.6	42.6	24.6	24.6
Middle (n=108)	55.2	60.9	56.5	5.7	1.3
High (n=101)	76.8	63.4	63.2	-13.4	-13.6

2. Trends in Provider Performance by Provider Specialty, Practice Type, and Patient Volume Among Providers Eligible for an Incentive

Plans' administrative data show the pediatricians and family and general practitioners who had at least 10 HEDIS-eligible babies turning 15 months old performed about the same on average during 2003, and both specialties as a group improved slightly during 2004/2005 (Table IV.2). Note that for this analysis, we calculate the percentage of babies with at least four visits by 15 months of age, rather than the percentage with at least six visits as we did for the HEDIS analysis above. We made this change because plan data systems very often are unable to identify one to two visits for the baby while the baby is still under the mother's ID number.

Individual or small group practices (fewer than three physicians) with at least 10 HEDIS-eligible babies turning 15 months performed a bit better than medical groups on average in 2003, and both groups improved roughly similarly during 2004/2005.

Higher-volume providers in the analysis did not improve more than others, in fact it was the lower-volume group that improved the most. This is not surprising because the analysis only includes the approximately 15 percent of providers who had the highest volumes of

HEDIS-eligible babies. The common wisdom that performance will improve more with incentives among higher-volume providers may well still hold, with higher-volume providers being defined as *all* those who had at least 10 HEDIS-eligible babies turning 15 months enrolled with the plan. (We cannot test this by analyzing the performance trend of those with fewer than 10 HEDIS-eligible babies turning 15 months because their performance figures are not at all stable over time—a random change in just one or two babies would have a dramatic effect on their scores.) The higher mean performance of the highest-volume providers throughout the period is notable, and may suggest an interesting area for follow-up research, to learn more about who they are and how high volume correlates with other practice characteristics associated with strong performance.

As a group, providers who were low performers in 2003 improved dramatically by 2005, far more than “middle” performers and in great contrast to the decline in performance among the high performers in 2003.¹⁴ Despite the great gains by the low performers, they did not even come close to “catching up” to the others. The low performers’ average score was 77 percent lower than the high performers’ average and 67 percent lower than the middle performers in 2003, and by 2005 it was 33 percent lower than the high performers and 25 percent lower than the middle performers. The pattern of low scorers improving and high scores declining in scores is consistent with the phenomena known as “regression to the mean,” which holds that across measurement, of all types, very high measure, tend to decrease and very low measures tend to increase, without cause.

Although unlike HEDIS trends, trends in performance based on administrative data are sensitive to changes in reporting as well as changes in care, only one plan that implemented incentives undertook a major effort to improve reporting. We found overall reporting rates declined during the period in that plan (discussed below), so we included it in this analysis. It is possible, however, that one of the reasons for the large increase in performance noted for the lowest-performing plans is that they had not been routinely reporting well-visits and began to do so.

3. Encounter Data Reporting

The median rates of reporting for preventive service encounters declined by about 10 percent between the first four program quarters and the later program quarters (Table IV.3). However, the timely submission of preventive service encounters improved. In the first four program quarters, on average 72 percent of encounters were submitted within 90 days. In the later quarters, this rose to 83 percent.

To see if the poorest reporters of preventive service encounters improved over the period differently from the average, we also examine the trend in the 25th percentile and did find improvement. During program quarters 1 to 4, one-fourth of the eligible groups reported 646 or fewer preventive service encounters per 1,000 member months. This “25th percentile point”

¹⁴ High, middle, and low performance was defined separately for each plan based on the 33rd and 66th percentile scores. When plans equaled the percentile cut-off point, they were all placed in the higher of the two tiers.

was 42 percent lower than the median. This figure rose to 695 per 1,000 for program quarters 5 to 9, so that for the second period the 25th percentile point was 31 percent lower than the median.

Table IV.3. Encounter Data Reporting Rates and Timeliness, Providers Eligible for an Incentive

	Eligible		
	Q1-Q4 (n=42)	Q5-Q9 (n=43)	Percent Change Q1-Q4 to Q5-Q9
Preventive Service Encounters Per 1,000 Member Months			
Median	1,119	1,001	-10.5
25 th Percentile	646	695	7.6
Percent of Preventive Service Encounters Submitted Within 90 Days	72	83	11

B. IMPACTS OF THE INCENTIVES

The positive trends in performance described above may occur for many reasons. The comparative analysis presented below helps us assess whether the trends reflect a general tendency or whether they suggest the incentives worked to improve timely well-baby visits, annual adolescent well-visits, and better encounter data reporting. Ideally, either an evaluation demonstration uses a randomized design, or a comparison group is selected to be just like the demonstration sites, except for the implementation of incentives. This ideal case would ensure that any differences in outcomes between the two groups are extremely likely due to the incentives. However, this is not the ideal case. The following limitations apply to the specific comparison groups used below:

- Plan-level HEDIS benchmarks for non-LIRR plans in California and for the U.S. are the best comparison groups available to the study, however, one can argue that the average CA Medicaid plan or the average U.S. Medicaid plan may not be similar in important ways to the average LIRR plan. If that were the case, interpreting a difference in the trends between the groups as an impact could be misleading. Because of the lack of literature on the characteristics of plans that are important to success with incentives, we cannot either support or refute this type of charge by comparing relevant characteristics.
- Providers who are not eligible for incentives within the LIRR plans (across the plans) are a third comparison group used in the analysis. However, one of the LIRR plans that did not implement incentives comprises a large percentage of the non-eligible provider groups and thus dominates the results for the non-eligible group. If there was something importantly atypical about this plan, the comparison could be misleading. The atypical characteristic that we are aware of

with this plan is that it was in financial crisis during the demonstration period. However, during this time it continued to pay primary care providers fee-for-service payment above capitation for well visits, as it always had.

- Primary provider groups within the same plan who were not eligible for the encounter data incentive are a fourth comparison group used in the encounter data analysis. While we know the non-eligible primary provider groups have fewer plan members than those eligible for the incentive, the plan is not aware of any other differences in the types of groups that make up the set of non-eligible providers.

In sum, while we draw on all the comparison data available to us, the results are suggestive rather than conclusive, both because we do not have a large and “clean” control group, and because the statistical analyses that we had hoped to do (including multivariate analysis with sufficient sample size and a strong comparison group) proved impossible due to small sample sizes.

1. Well-Baby Visit Incentives

This section answers the question: Did the timeliness of well-baby care improve in participating plans and among providers eligible for the incentives relative to comparison groups?

a. Comparing HEDIS Data for Plans with LIRR Incentives to National and California Means

We used two main comparison groups: the national average HEDIS scores for Medicaid plans, and the California HEDIS scores for non-LIRR plans.

Compared to the national average for Medicaid plans, the timeliness of well-baby visits improved much more in plans that implemented broad-based provider incentives under the LIRR Collaborative.¹⁵ These plans, which lagged the national average slightly in 2002, improved to exceed the national average by 21 percent in 2004 (Table IV.4).

The LIRR plans with broad-based provider incentives also improved more compared to other Medicaid plans in California. While 2002 data were not available, in 2003 the LIRR plans slightly lagged the mean for other Medicaid plans in California (6 percent below the mean), but improved to exceed it by 10 percent in 2004.

The LIRR plans with broad-based provider incentives improved more than a third comparison group as well. The LIRR plans that implemented only narrow provider incentives or did not implement any at all scored higher at the start than the LIRR plans that implemented new incentives. However, they improved less during 2003-2004 so that by 2004 the LIRR plans that implemented broad-based incentives had essentially “caught up.”

¹⁵ Information on variances in the samples used to calculate the HEDIS scores was not available to allow for statistical tests of significance.

Table IV.4. Percent of HEDIS-Eligible Members with Six or More Well Visits by 15 Months of Age, 2002-2005

	2002	2003	2004	2005	Percentage-Point Change 2002-2004	Percentage-Point Change 2003-2004
National Mean for Medicaid Plans	43	45	47	NA	4	2
Mean for "Other" CA Medicaid Plans	NA	48	52	NA	NA	4
Plans with Broad-Based Provider Incentives						
Plan A	42	37	47	51	5	10
Plan B	49	39	61	54	12	22
Plan C	33	40	44	44	11	4
Plan D	39	63	74	82	35	11
All (Mean)	41	45	57	58	16	12
Plan With Narrowly Targeted Provider Incentive						
Plan E	NA	61	56	73	NA	-5
LIRR Plans with No New Incentive						
Plan F	NA	61	61	NA	NA	0
Plan G	NA	48	56	53	NA	8
Both (Mean)	NA	55	59	NA	NA	4

b. Comparing LIRR Plans' Provider Groups Eligible vs. Not Eligible for Incentives, Using Administrative Data

Although their gains were small, the providers in our analysis who were eligible for incentives improved more than those not eligible for incentives (Table IV. 5). The providers eligible for the incentive began the period in 2003 with a mean score 33 percent lower than those not eligible for the incentive, and ended the period in 2005 with a mean score 29 percent lower than those not eligible for the incentive.

For most of the provider subgroups of interest, our sample size was not large enough to compare providers eligible versus not eligible for the incentives. However, we could compare the improvement among pediatricians and among individual or small practices who were eligible versus not eligible. In both these subgroups, the groups that were eligible for the incentives improved no more than the groups that were not eligible for the incentives.

Table IV.5. Change in Well-Baby Providers Visit Performance 2003-2005, Eligible vs. Ineligible for Incentives (Providers with at Least 10 HEDIS-Eligible Babies)

Providers	2003	2004	2005	Change in Performance 2003 vs. 2004	Change in Performance 2003 vs. 2005
All Eligible for Incentives (n=283)	53.3	57.1	54.7	3.8	1.4
All Ineligible for Incentives (n=22)	79.5	76.3	77.1	-3.2	-2.4
Pediatricians					
Eligible for incentives (n=168)	49.9	55.8	51.9	5.9	2.0
Ineligible for incentives (n=12)	80.4	85.6	87.0	5.2	6.6
Individual or Small Practices					
Eligible for incentives (n=202)	54.8	58.7	56.0	3.9	1.2
Ineligible for incentives (n=14)	83.1	82.8	87.1	-0.3	4.0

2. Adolescent Well-Visit Incentives

a. Comparing LIRR Plans with Movie Ticket Incentives to National and California Means

The two plans with new movie ticket incentives improved their percentage of adolescents receiving an annual well visit more during 2002-2004 compared to the national mean for Medicaid plans (Table IV.6). They began the period lagging the national average substantially (by about 30 percent), and improved so that by 2004 they lagged by much less—about 8 percent. However, in one of these plans, all the improvement occurred between 2002 and 2003, so that its 2003-2004 change in score was -1 , compared to the national average change of $+3$.

Compared to other California Medicaid plans, the plans with new movie ticket incentives improved more during 2003-2004, but the difference was very slight (5 percent versus 3 percent for other California Medicaid plans).¹⁶ The mean for the other California Medicaid plans is considerably lower than the national mean (16 percent lower in 2003), and shows that California Medicaid plans as a group have had difficulty performing well on this measure, not just the LIRR plans.

The two plans with new movie ticket incentives did not improve more than the three other LIRR plans with no new incentives for adolescent care during the period.

¹⁶The state of California did not require plans to report this HEDIS measure for 2002.

Table IV.6. Percent of HEDIS-Eligible Members Ages 12-21 with Annual Well Visit

	2002	2003	2004	2005	Percentage-Point Change 2002-2004	Percentage-Point Change 2003-2005	Two-Year Average 2002/2003	Two-Year Average 2004/2005	Percentage-Point Change in Two-Year Average
National Mean for Medicaid Plans	37	37	40	NA	3	NA	37.0	NA	NA
Mean for "Other" CA Medicaid Plans	NA	31	34	NA	NA	NA	NA	NA	NA
Plans with Provider Incentive									
Plan B (Broad-Based)	30	38	38	35	8	-3	34.0	36.5	2.5
Plan E (Narrow & Major Outreach)	36	38	45	49	9	11	37.0	47.0	10.0
Plans with New Movie Ticket Incentives									
Plan A	28	26	37	25	9	-1	27.0	31.0	4.0
Plan C	23	38	37	37	14	-1	30.5	37.0	6.5
Both (Mean)	26	32	37	31	12	-1	28.8	34.0	
Plan with No New Incentive 2003-2004									
Plan D ^a	38	44	52	59	14	15	41.0	55.5	14.5
Plan F	36	37	45	45	9	NA	36.4	45.0	8.6
Plan G	NA	34	33	35	NA	1	NA	34.0	NA
All (Mean)	37	41	49	46	12	NA	NA	44.8	NA

^aThis plan implemented a financial incentive for IPAs in its network in July 2005.

b. Comparing LIRR Plans with Broad-Based Provider Incentive for Adolescent Well-Care to National and California Means

The plan with a broad-based provider incentive for well-adolescent visits improved more than the national mean during 2002-2004. However, all of this improvement occurred between 2002 and 2003. Therefore the 2003-2004 improvement for this plan was 0 compared to the national average improvement of 3 percent.

Compared to other California Medicaid plans, the plan with the broad-based provider incentive did not improve as much as the other California Medicaid plans group during 2003-2004 (0 versus 3 percent improvement).

Compared to the two LIRR plans with no new incentives during 2003/2004 for well-adolescent visits that also had data for 2002-2004, the plan with the broad-based provider incentive improved less well or about equal during the period (8 versus 14 and 9 percent, respectively).

c. Comparing an LIRR Plan with Narrow Provider-Based Incentive and Major Direct Outreach to National and California Means

The plan with a narrowly targeted provider incentive (three offices in 2004 and 16 in 2005) along with a major outreach effort improved its performance more than the national mean for Medicaid plans during 2002-2004 (9 vs. 3 percentage points). In 2002 it had lagged the national mean by 3 percent, and improved by 2004 so such that it exceeded the national mean by almost 13 percent. It also improved more than the mean for other California plans during 2003-2004 (7 vs. 3 percentage points), exceeding the 2004 mean for other California plans by 32 percent.

Compared to the two LIRR plans with no new incentives for well-adolescent visits that also had data for 2002-2004, the plan with the narrow provider incentive plus major outreach improved about the same or less (9 vs. 14 and 9 percentage points, respectively).

3. Encounter Data Reporting Incentive

For this analysis we compare the trend in rates of encounter data reporting between providers eligible and those not eligible for the incentive within the plan that implemented the incentive.

The rates of preventive service encounters for providers eligible for the encounter data incentive did not decline as much as they did for the comparison group of providers not eligible for the incentive (Table IV.7). In addition, the 25th percentile point improved for the group eligible for incentives in contrast to a large decline for groups not eligible.

Surprisingly, the trend in encounter data submission rates for preventive services mirrored the trend for other types of encounters not targeted by the incentive. This could mean that the technical assistance provided by the plan helped to resolve general reporting problems that

Table IV.7. Encounter Data Reporting Outcomes for Groups Eligible and Not Eligible for the Incentive

	Eligible			Not Eligible		
	Q1-Q4 (n=62) ^a	Q5-Q9 (n=63) ^a	Percent Change Q1-Q4 to Q5-Q9	Q1-Q4 (n=25) ^a	Q5-Q9 (n=24) ^a	Percent Change Q1-Q4 to Q5-Q9
Preventive Service Encounters per 1,000 Member Months						
Median	1,119	1,001	-10.5	641	361	-43.7
25 th Percentile	646	695	7.6	404	148	-63.4
Other Encounters per 1,000 Member Months						
Median	2,715	2,590	-4.6	2,518	1,195	-52.5
25 th Percentile	1,973	1,937	-1.8	797	253	-68.3
Percent of Preventive Service Encounters Submitted Within 90 Days						
Median	72	83	11	72	83	11

^aThe n's are for the per 1,000 figures. The n's for the percent submitted within 90 days are: Eligible Q1-Q4: 42; Eligible Q5-Q9: 43; Not eligible Q1-Q4: 20; Not eligible Q5-Q9: 19.

affected reporting of all services, and/or could mean groups respond in a more generalized manner to incentives like this rather than specifically changing reporting for targeted relative to non-targeted services. Alternatively, a skeptic might point to this as suggesting that unobserved characteristics of the group of eligibles may be driving the difference between the eligible and non-eligible groups, rather than the incentives driving the difference.

Both eligible and non-eligible groups improved similarly on the percent of preventive service encounters submitted within 90 days.

4. Performance Trend by Size of Payout

For this analysis we compare the well-baby visit outcomes and rates of reporting of encounter data between provider groups who received higher vs. lower incentive amounts early in the demonstration.

For both well-baby visit performance and encounter data reporting, we find that groups who received higher payouts in the first period performed better than comparison groups in the subsequent period (Tables IV.8 and IV.9). Seeing a significant check may have helped catch the groups' attention regarding the incentive program and led to better performance in subsequent years. However, because the high-payout group was relatively high-scoring at the start (especially for the well-baby visit measure), it may be that providers who are by nature motivated to perform well on this measure, or who are better able to perform well given their office structure and resources, are more receptive to incentives, which could cover their costs to further improve. That is, we cannot assume that a more sizable check given to low performers

(e.g., a deposit to catch their attention regarding the program) would result in similar improvement, although the findings do not rule out that this could work.

Table IV.8. Change in Well-Baby Visit Performance, 2003-2005, by Size of 2003 Incentive Payout (Providers Eligible for Incentives with at Least 10 HEDIS-Eligible Babies)

Size of 2003 Incentive Payout	Mean Performance of Providers, Percent of Babies Turning 15 Months with 4+ Well-Baby Visits			Change in Performance 2003 vs. 2004	Change in Performance 2003 vs. 2005
	2003	2004	2005		
Median or Greater (n=29)	58.5	65.2	63.2	6.7	4.7
Less than Median (n=56)	48.0	51.1	48.4	3.1	0.4
No Incentive Payout (n=198)	52.8	55.6	53.1	2.8	0.3

Source: MPR analysis of provider-level administrative claims data from LIRR participant health plans.

Note: Median 2003 incentive payouts were calculated separately for each individual health plan. The 2003 median incentive payout across health plans was \$833 per provider.

Table IV.9. Median Percent of Preventive Service Encounters Submitted Within 90 Days

	Eligible		Percent Change Q1-Q4 to Q5-Q9
	Q1-Q4 (n=42)	Q5-Q9 (n=43)	
Paid \geq 50 th Percentile Payout in Q1-Q4 (n=21)	76	84	8
Paid < 50 th Percentile Payment in Q1-Q4 (n=15)	72	75	3

Source: MPR analysis of group-level administrative data submitted by Plan C.

This page has been intentionally left blank for double-sided copying.

CHAPTER V

HAS PRACTICE CHANGED? PROVIDERS' COMMENTS ON THEIR RESPONSE TO THE INCENTIVES

The two waves of provider interviews for the evaluation suggest that providers' level of awareness about the incentives grew over the course of the demonstration, but clearly the physicians in these plans' networks are still not closely attuned to the details. Many described increasing their outreach to parents of babies, using the monthly lists sent by the plans of babies turning 15 months, and some tied this specifically to the incentive providing resources to enable and reward this increased effort. However, as a group, the physicians we interviewed had relatively close ties to the plan, so they are probably more likely to be acting on the incentive than the average plan physician.

While we did not interview enough physicians per plan to draw conclusions about physicians' views on each plan's incentive arrangement, we were interested to see if there were any strong patterns of physician opinion by plan. We found the patterns persisted across the plans rather than being plan-specific. For example, at least one of the second-round interviewed providers in four of the five plans with new incentives complained about inadequate communication from the plan about the incentives. The exception was one of the two plans with the most effective incentives (see Chapter IV), but we were only able to interview two physicians affiliated with that plan on the second round of interviews.

A. LEVEL OF AWARENESS OF INCENTIVES

During the first wave of interviews, all of the physicians we interviewed who were closely affiliated with the plan knew that there were incentives in place, but several of the community physicians without close ties either were unaware or learned of the incentive only a few months prior to our call—about a year after the start of the incentives. Two knew there were incentives, but were unsure which plan offered them. In general, physicians were best able to talk about their experience with and attitudes toward the plan's full incentive program, rather than the specific Collaborative measures on well-baby or well-adolescent care.

By the second wave of interviews, all but one physician we interviewed were aware of the specific Collaborative measures on well-baby or well-adolescent care but still at a very general level. Most did not know at least some of the plan's specific features and many of these physicians commented that they were not too concerned with knowing the specifics. At least one provider stated that he was too busy to learn the details.

Many of the physicians we interviewed were not very attuned to the incentives, beyond basic awareness that they existed. Of the 22 physicians we interviewed in plans with active incentive programs in the first round, half were at least vaguely aware of the incentive amount paid to their practice, and half were not.¹⁷ By the second wave, the vast majority (all but 3 of 18) were at least vaguely aware of the amount paid—that is, they could at least say whether the incentive amount was more or less than 10 percent of their practice income. However, in both waves, many of those who could approximate a payment amount thought in terms of the amount of the check for the full incentive package, not measure by measure.

Of note, one physician we spoke with who received an unusually large incentive check reported that he had received several calls from other physicians asking how he did it (word apparently traveled that he received a large check). He said that he does not reveal the amount of the check, but instead tells them how his office is completely structured around preventive medicine. Office staff will tell a mother that her child needs to come in for a well-visit, then he/she repeats it when the doctor sees the patient. After hearing it enough, the mother brings her child in. Another key factor this provider cites is the high volume of plan members he serves, which also factors into the amount of the payout.

B. ACTIONS PROVIDERS HAVE TAKEN TO IMPROVE

A majority of the providers we interviewed had increased their emphasis on outreach, and said the incentive had highlighted the need to do so (for the first wave of interviews, 11 of the 17 for whom the question applied¹⁸ and for the second wave of interviews, 15 of the 18). Many commented that it was not the potential for money that drove them to these actions, although they supported keeping the incentives and felt there were some providers who would be motivated by money. Note that those we interviewed were generally high-volume providers with relatively close ties to the plan, so they were more likely to be acting on the incentive than the average plan physician. Table IV.1 provides some paraphrased examples from our interview for each type of change.

¹⁷ During wave one of the interviews, of the half (11) that did not know how much they had received from the incentive, 7 said they never saw the amount (often in these cases the check was probably paid to their organization rather than to them personally), and 4 either could not remember (3) or said they saw the amount but did not pay attention/did not care (1). Of the 3 that did not know how much they had received in wave two of the interviews, 2 had never seen the amount and one said he could not remember.

¹⁸ Physicians we spoke with for whom the question did not apply included physicians in the plan that canceled its incentive, and a few physicians who were either new to the plan, received in-kind support but no financial incentive, or an adolescent specialist who did not receive a financial incentive but was able to comment usefully on the member incentive for adolescents.

Table V. 1. Examples of Comments from Providers Who Responded to the Incentives**Better Outreach—More Reminders/Trying Harder**

1. *Me and my office staff responded positively to the incentive. Since we got the list of kids missing visits, the NP in my office goes through the charts, and calls 30 to 40 patients per month. We made this change for all our patients, not just [this plan's] patients—[this plan's] patients are not a large portion of our membership.*
2. *Overall we have changed in response to the incentive in that our staff is very engaged, is now constantly on the lookout for something. When the chart comes in we look for whether they need follow-ups, or what.*
3. *We decided to start calling people to come in to the extent we can, because it's the right thing. I admit the incentive did raise our awareness about the issue, and every penny counts to make this practice run.*
4. *For well-baby visits, we are in the process of trying to start calling them to come in.*
5. *I think I have tried harder to get the patients in as a result of the incentive. Our office staff now send out more postcard reminders to patients. But this is a walk-in clinic, it's very hard to get babies in for six visits.*
6. *Monthly, the front desk staff calls all the patients to come in who are on the plan's list and about to turn 15 months of age. Here the incentive money is paid directly to the employees rather than to the clinic or the physician.*

Taking Advantage of Sick Visits

7. *If I see from my score that I am slacking, I step up my efforts. But only sick people come into this clinic.*
8. *I did not specifically try to increase the frequency of follow-up of patients, but what I would do is if a patient came in for a sick visit I would stretch it into a well-baby visit to do better on the incentives.*

Changing Scheduling Frequency or Process

9. *We are trying to change our scheduling a little bit to get some people in earlier. Some people come in at 15.5 or 16 months and we don't get the bonus for them.*
10. *Both my office staff and I have been taking steps. We have tried to track and follow up with eligible patients more, and schedule more frequent follow-ups.*

C. CHALLENGES TO IMPROVEMENT

Not all physicians responded to the incentives. Table IV.2 provides examples of providers' comments who did not respond.

1. Well-Visit Rates

Many of the physicians we spoke with discussed the difficulties they face in improving well-visit rates for the Medi-Cal population. The two major types of challenges appear to be

Table V.2. Examples of Comments from Providers Who Did Not Respond to the Incentives

No Change
<p>1. <i>I haven't changed my office practice...it's just too difficult to get the patient to come in right on time at the proper time, so I don't try. Some physicians are abusing the system, giving kids immunizations before they are due, just in order to get the bonus.</i></p>
<p>2. <i>We did an experiment with calling 20 to 30 adolescents to see if they'd come in, but many could not be reached and it came to almost nothing. It's discouraging and I don't feel like doing it anymore.</i></p>
<p>3. <i>Our clinic is not as yet doing anything differently because of the incentive program. [The plan] sends out a list of our new enrollees, but we have no one to follow up on it. A lot of the problem is just a lack of manpower for someone to stay on top of these things. [later in the conversation] It's not obvious that we'll individually see a monetary reward [payment goes to group]... it's not the monetary aspect by itself but the link to pushing for quality improvement that would influence our providers.</i></p>
<p>4. <i>It is hard to keep momentum up when you don't even know if it is effective.</i></p>

population factors and office system factors. These factors were cited in both waves of interviews.

Population Factors. The low socioeconomic status of the population targeted for improvement was cited over and over as a major problem in getting the babies in for timely visits. Specifically, the factors mentioned included:

- Parents' focus on economic subsistence, which makes timely well-visit care for their children seem less important
- Lack of transportation
- Mobility, including changing health providers and/or plans, moving from city to city, and the associated likelihood of inaccurate contact information

The difficulties in getting adolescents in for well-visits were also bemoaned by many, although the issues mentioned were less often socioeconomic status and more a lack of willingness/ motivation to come in. Several noted that children involved in sports are often asked to get an annual physical, but those who are not involved in sports—who may be at risk for the most serious problems—are not.

Office Factors. The office factors that challenged improvement included the following:

- Clinics with a heavy walk-in population mentioned this arrangement was not conducive to trying to get people in on schedule
- Lack of enough staff to support outreach (or enough outreach), and not enough money to support expanding staff capacity
- System issues: lack of a computer (though this practice is getting one next year); no system to track patients across the clinic’s 15 doctors to identify those that need follow-up; not taking advantage of technology on hand for follow-up, although “every bit of money would help us improve.”
- Eclectic set of demands: A family practitioner explained that his practice includes many concerns beyond child health, Medi-Cal, and managed care (including minor surgery and OB), so it is difficult for him to be continuously attentive to meeting well-baby and well-adolescent visit goals—“Last month, I focused on well-baby visits for [the plan] and this month I’m focusing more on workers’ comp.”

D. COMMENTS LINKING IMPLEMENTATION PROBLEMS TO PROVIDER RESPONSE

During the second wave of interviews, a few of the 18 providers interviewed expressed discouragement with the plans’ incentive programs. That is, the physicians indicated an initial attempt to respond to the incentive, but said they became discouraged by the way the program worked. One of these concerned a plan with large improvement relative to comparison groups, while the other two concerned the plan whose data clearly showed little effect from the incentives. Table V.3 lists these providers’ comments, which faulted the accuracy of the data used and general rather than specific nature of the data sharing; the feeling that the plan was not taking the program seriously; and communication problems.

Table V.3. List of Comments from Providers Who Reported Becoming Discouraged Due to Program Issues

Became Discouraged
<ol style="list-style-type: none"> 1. <i>We are not sure how we really performed the data are so general. No one is sure how accurate the data are. It is too general to tell. They need to break down the data to let us know what is what...we are starting to get discouraged.</i> 2. <i>[The plan] is not taking it seriously so we have given up and no longer take it seriously. We would take this program seriously if we thought they were.</i> 3. <i>It would be beneficial if [the plan] provided more information on the incentives, what I could receive and what I did receive as well as a comparison to others in the area...[the plan] did not communicate well, they make too many changes so I’ve given up trying to follow them or make any changes to my office practice. I simply collect the check and that is the end of it.</i>

This page has been intentionally left blank for double-sided copying.

CHAPTER VI

SYNTHESIS OF QUALITATIVE AND QUANTITATIVE ANALYSIS OF THE WELL-BABY VISIT AND ENCOUNTER DATA INCENTIVES, PLAN BY PLAN

This chapter discusses each plan’s well-baby visit outcomes in light of its program features and implementation experience. This plan-by-plan synthesis is the basis for identifying patterns across the plans and ensuring that the overall conclusions presented in Chapter VII are consistent with the stories of the individual plans as well as the qualitative and quantitative cross-plan analysis. The plans are discussed beginning with the plan that showed the greatest improvement in its two-year pre-demonstration versus post-demonstration HEDIS score on the well-baby visit measure. We also discuss here the two plans that did not implement a well-baby incentive. At the end of the chapter, we synthesize the plan-specific results for the encounter data reporting incentive. We do not discuss the experience with the adolescent well-visit measure because we believe the relevant material was covered in chapters III and IV in a way that supports our overall conclusions for that measure. For the encounter data incentive, we discuss only the plan that implemented the incentive, because with one exception, the other plans did not make any significant new efforts to improve encounter data reporting, and the exception plan is related to well-baby visit timeliness.

A. PLAN-BY-PLAN SYNTHESIS OF WELL-BABY INCENTIVES RESULTS

The table from Chapter IV showing HEDIS outcomes for the well-baby visit measure is repeated here for the reader’s convenience (IV.1), along with a second table that will be referenced throughout this section. For each plan, Table IV.2 shows the mean performance of providers with at least 10 HEDIS-eligible babies on the percentage of 15-month old babies who received four or more well-visits, based on administrative data only.

Table VI.1. HEDIS Data: Percent of HEDIS-Eligible Members with Six or More Well Visits by 15 Months of Age, 2002-2005

	Pre Impelemen- tation: 2002	Year of Implemen- tation: 2003 ^a	Post- Implementation:		Percentage- Point Change 2002-2004	Percentage- Point Change 2003-2004
			2004	2005		
National Mean for Medicaid Plans	43	45	47	NA	4	2
Mean for "Other" CA Medicaid Plans	NA	48	52	NA	NA	4
Plans with Broad- Based Provider Incentives						
Plan A	42	37	47	51	5	10
Plan B	49	39	61	54	12	22
Plan C	33	40	44	44	11	4
Plan D	39	63	74	82	35	11
All (Mean)	41	45	57	58	16	12
Plan with Narrowly Targeted Provider Incentive						
Plan E	NA	61	56	73	NA	-5
LIRR Plans with No New Incentive						
Plan F	NA	61	61	NA	NA	0
Plan G	NA	48	56	53	NA	8
Both (Mean)	NA	55	59	NA	NA	4

^aFour of the five plans began implementation during July-September 2003. The fifth (PLAN D) began its LIRR incentives in July 2004.

1. Plan D

Among the plans in the study, Plan D had the most impressive improvement during 2002-2005, with its two-year average score on the well-baby visit measure rising by 27 percent between 2002/2003 and 2004/2005. By 2005, its score was fully 74 percent higher than the 2004 national mean for Medicaid plans (Table VI.1). For several years prior to the demonstration, this plan had been paying fee-for-service on top of capitation to providers for well-visits, including well-baby visits. That is, physicians receive a monthly capitated amount for each plan enrollee to cover expected costs for all their care; in addition, they may bill a standard amount for each well visit. The fee-for-service payout is substantial—for example it was \$45 per visit, totaling to almost \$2 million in 2004 fee-for-service payments for well-baby visits alone. The new incentive implemented under the LIRR provided an extra bonus amount of \$100 when a child receives recommended vaccines and five well-visits by 15 months, and then the opportunity for another \$100 if the sixth visit is present along with the recommended vaccines. Total payout under the new bonus nearly doubled from the first to second program

years, but continues to be far lower than continuing payout for the pre-existing fee-for-service incentive.

With this context in mind, we believe the exceptional improvements that we see in this plan from 2002-2005 are not likely due to the new bonus incentive alone. The largest improvement occurred between 2002 and 2003, before the bonus incentive even began in July 2004—the plan’s score improved from 39 percent of 15-month olds receiving their recommended six visits to 63 percent receiving them.

The plan believes that the fee-for-service payments continued to encourage improved timeliness of well-baby care in combination with the bonus incentive during the study period, fueling continued improvements during 2004 and 2005. Our analysis of administrative data for providers with at least 10 HEDIS-eligible babies (including 97 percent of Plan D network primary care providers) found that their mean performance on the measure percentage of babies receiving four or more well-visits by 15 months was high and remained stable throughout the period (at 68.4 to 68.7 percent) (Table VI.2). This pattern is consistent with the idea that the plan’s bonus payment—focused on the fifth and sixth visits—was helping the plan’s score increase during 2004 and 2005.

Table VI.2. Change in Well-Baby Visit Performance, 2003-2005, by Health Plan (Providers with at Least 10 HEDIS-Eligible Babies)

	Mean Performance of Providers, Percent of Babies Turning 15 Months with 4+ Well-Baby Visits				
	2003	2004	2005	Change in Performance 2003 vs. 2004	Change in Performance 2003 vs. 2005
Eligible for Incentive (n=283)					
Plan A (n=24)	70.4	74.6	71.9	4.2	1.5
Plan B (n=41)	29.0	31.4	25.3	2.4	-3.7
Plan C (n=76)	30.7	41.5	37.1	10.8	6.4
Plan D (n=139)	68.4	68.6	68.7	0.2	0.3
Plan E (n=3)	nc	nc	nc	nc	nc
Not Eligible for Incentive (n=43)					
Plan F (n=15)	83.2	76.3	77.1	-6.9	-6.1
Plan E (n=5)	nc	nc	nc	nc	nc
Plan D (n=2)	nc	nc	nc	nc	nc
Plan G (n=21)	27.0	60.7	56.7	33.7	29.7

Source: MPR analysis of provider-level administrative claims data from LIRR participant health plans. One plan (Plan G) was excluded from the analysis due to major changes in data reporting levels between 2003 and 2004 which reflect reporting rather than care.

nc = not calculated separately due to insufficient sample size (n<10) of observations at the individual plan level.

Of note, a plan in the LIRR Collaborative that did not implement any new incentives but which had the same fee-for-service payment for well visits on top of capitation in place (Plan F) had a similar high score in 2003, but did not improve in 2004, in contrast to Plan D which added the bonus program and further improved by 11 percent.

Also worth noting, three of the four providers we spoke with in wave one were clear about how the incentive program worked (more than was the norm overall), and the two providers we were able to speak to in wave two felt communication about the incentives had been good. Thus, we think communications issues that can stifle the potential effect of incentives may not have been as much of an issue in this plan, although the small number of interviews suggests caution on this point.

2. Plan B

Plan B also showed impressive improvements of 13.5 percentage points in the HEDIS well-baby visit measure between 2002/2003 and 2004/2005, however, we are hard-pressed to cite any program features that may explain why this plan would have improved more than the average plan. Provider communications seemed to be an issue for the plan. For example, information on providers' performance was not broken down between the well-baby measure and other measures, so they could not understand how the payout related to each measure. Also, with payout focused solely on the sixth visit using administrative data, payout was relatively low, with only 28 of 189 network primary care providers receiving a payout in 2004. In July 2004 in an attempt to raise the effectiveness of the incentives, the amount of the incentive was doubled to \$200 for the six-visit series. However, the number of providers receiving a payout in the first half of 2005 did not rise.

Without the fee-for-service payment for well-care to encourage reporting, the percentage of 15-month-olds with four or more visits based on administrative data alone was quite low (only 29 percent), and did not consistently improve. Because of this plan's network structure, 95 percent of its providers had 10 or more HEDIS-eligible babies and thus were included in the administrative data analysis, so the analysis fairly represents the plan's overall trend using administrative data. While the mean performance rose just slightly in 2004 it dropped in 2005 to a level that was below the starting point (Table IV.2). Since the incentive payout was based solely on administrative data, the low level of reporting and lack of improvement over time seems to suggest that the plan's HEDIS score improvements may have been due to factors other than its LIRR incentive program.

3. Plan A

This plan's two-year average HEDIS score improved by 9.5 percentage points between 2002/2003 and 2004/2005, rising to equal the national mean for Medicaid plans by 2004 and continuing to increase in 2005. While this plan's incentive program officially began in August 2003, informal feedback indicated that providers were largely unaware of the program during July 2003 and early 2004. The plan's program was designed to take advantage of the plan's annual risk pool payout, switching the criteria for the payout toward quality measures.

To improve awareness of the incentives and strengthen their effectiveness, in summer 2004 the plan mandated as a condition of participation in the plan's risk pool that physicians attend one of three scheduled sessions with the plan to explain its incentive program. The message given at the sessions was that well-baby visit performance along with several other targeted HEDIS measures had impacted the providers' risk pool distributions that year, would impact them again this year, and in the following year (payout was scheduled for June 2006); thus, any provider that had not improved up to the median from two years previous would not receive a risk pool check. This scheduled phase-in of the tougher policy meant that in 2005 the plan let providers know who were below the median for the prior year that the full amount of their check is at risk for next year if they do not improve.

The method for calculating the incentive was a formula that gave points for each of the fourth, fifth, and sixth well-baby visits (from administrative data) for the babies who turned 15 months of age. The payout amounts for this plan both in total and per provider paid were large relative to other plans (\$745,000 in 2004, or about \$20,700 per provider), due to the relatively large size of the risk pool that pre-dated the program. However, despite the announcement of the well-baby emphasis, the actual relationship between well-baby visit performance and payout was muted in the first two program years by a formula that rewarded the total number of members served by the providers. Also, providers were accustomed to receiving large checks for the annual risk pool payout, so the large amounts would not be expected to have the same effect as if they had been "new money." The result was a fairly indirect relationship between payout and performance on the well-baby visit measure. The formula is being amended in the coming year to implement a more direct relationship.

Our analysis of provider performance using administrative data is consistent with the incentives having an effect. Because this plan pays its providers on a fee-for-service rather than capitated basis for both well visits and other services, our analysis of provider performance using administrative data showed high and relatively stable performance similar to Plan D, with slight increases in 2004 and 2005 relative to 2003. With this plan the room for improvement would have come mostly on the fifth and sixth recommended visits to fuel the improvement in HEDIS scores.

With this program evolution and the outcomes trend in mind, we could reasonably argue either for or against an effect. The main argument for an effect is as follows:

- The plan's July 2004 mandated meetings with physicians about the incentive program raised awareness among physicians about the importance of well-baby visits to the plan and to their present and future incomes.
- While the specifics of the program may not have been clear to the physicians (then or later), maybe it doesn't matter—as discussed in Chapter V, the physicians we interviewed tend to think about such programs in general terms rather than attend to the details.

- The plan's trial run of its new tougher policy in 2005 may have contributed to higher rates in 2005, by further emphasizing the program and relevance to the group's risk pool payout check, particularly for low-performers.
- The amounts at stake were relatively high, and as discussed in chapter IV, this may encourage us to think the incentive may have been effective.
- While the plan's outcomes did not improve as much as Plan D or Plan B, it improved far more than the national or California plan averages between 2003 and 2004, and continued to improve by a significant margin between 2004 and 2005 (where national and California benchmark data are not yet available).

The main arguments against an effect or a noteworthy effect are:

- It is questionable whether physicians understand the current formula for payout; the uncertainty around how the program works and the indirect relationship between payout and performance may have dampened buy-in for the program among physicians.
- While the plan improved more than average over time, the results are far from spectacular—despite the improvement it remained below the other California plans' 2004 average in both 2004 and 2005.

4. Plan C

This plan's two year average HEDIS score rose 7.5 percentage points between 2002/2003 and 2004/2005, and its improvement trend was well above the national average for Medicaid plans. However, its HEDIS score did not further improve between 2004 and 2005, remaining below the national and California plans average in both years. Plan C's incentive structure was unique among the LIRR plans in two ways. First, it offered the incentives to individual physicians by reaching through several layers of organization in an attempt to generate a direct effect. Second, it combined physician incentives and member incentives in its program, hoping that addressing the issue from both sides would be more effective than either side alone. It gave providers \$25 for each set of three well-baby visits (total payout up to \$50 per child), using administrative data for the calculation, and offered parents a \$25 gift certificate for completing the well-visit series with their child.

Our analysis of provider-level administrative data is consistent with the plan's incentives having an effect. Because, unlike Plan D and Plan A, this plan does not have any fee-for-service payment for well-care, its scores on well-baby visit measures using administrative data are quite low, but they did improve substantially during the study period. The mean performance on the percentage of 15-month old babies with four or more visits rose from 30.7 percent in 2003 to 37.1 percent in 2005 (Table VI.2)—the largest increase of any of the plans.

The plan's experience with communicating about the program and implementing it appears somewhat mixed. Several of the providers we spoke with didn't understand very well how the

program worked or doubted that the administrative data used for payout were accurate given all the organizational layers between them and the plan. Semi-annual payouts were planned, but the cycle during the period was annual, which substantially delayed the first payout due to computer system issues and difficulties in automating the process. In addition, the take-up rate for the well-baby incentive turned out to be disappointingly low (see Chapter II). And dollars at stake for providers were quite low relative to the other plans, with the other plans offering at least the equivalent of \$100 per child, twice the level offered by this plan. The relatively low dollars offered per child were more of an issue when one considers the typically low volume of babies served by the typical provider eligible for the incentive. Only 10 percent of Plan C's primary care provider physicians served 10 or more HEDIS-eligible babies turning 15 months in 2003.

On balance, while the incentives may have had some effect at this plan, we think if they did the effect was not very large.

5. Plan E

Unlike the four plans above, Plan E narrowly targeted its well-baby incentive to a few high-volume providers for the plan. Its HEDIS score on the well-baby visit measure dropped by 5 percent during 2003-2004, although it then improved dramatically in 2005, rising to be only 11 percent lower (9 percent) than the leading plan in 2005, and climbing to far above the California and national averages for 2004.¹⁹ The plan also hired a staff member to explicitly assist the relevant providers in outreach to parents of babies who need well-visits and immunizations.

The very small number of eligible provider groups prohibits us from examining the trend in mean performance among them with administrative data; only eight providers for the plan had at least 10 HEDIS-eligible babies in 2003, though these providers accounted for 31 percent of total member months for babies of that age group, and only three of the eight were eligible in both 2003 and 2004. Our discussions with the relevant providers and the plan together suggested that the incentives were not likely to be the cause of the plan's improvement in 2005.

6. Plan F

Of the two plans that did not implement new incentives during this period, one of these plans, Alameda, was focused on financial survival. Not surprisingly, this plan's performance on the well-baby visit measure did not improve between 2003 and 2004. Fee-for-service payment for well-visit care, begun before the study period and continuing throughout it, may help explain why this plan's well-baby visit rate remained relatively high in 2003 and 2004, despite lack of any further improvement. The plan's rate was 30-36 percent above the national mean in both years.

¹⁹ We could not calculate the two-year average score for 2002/2003 since the plan did not calculate this HEDIS measure in 2002 because it was not required to do so by the state.

7. Plan G

The second plan that did not implement new incentives during the period also had ongoing fee-for-service payment for well-visits, on top of capitation. In summer 2004, plan executives met with physicians affiliated with provider groups that account for 49 percent of the plan's membership. Accompanied by a physician champion known to the physicians (also a plan board member), the plan sought to improve providers' awareness of the availability of fees for well-visits and to remind or teach them to code properly for well-visits. At the meetings, the plan told the physicians they could bill retrospectively for any visits they had not yet submitted for 2004. The meetings were reported to be fairly well-attended, and they were followed up by additional communications through the champion to others who missed the meetings. The plan also clarified its flexible policy on "the periodicity schedule." That is, the state Medicaid fee-for-service program requires specific time intervals between one well-baby visit and the next, in order to pay the provider. According to the plan, physicians did not realize they had more flexibility with the plan to "catch up" members on well-visits and immunizations in ways that do not quite match with the state's periodicity schedule.

The plan's HEDIS scores improved more than the national average and more than the mean for other California plans. Most remarkable, however, was the plan's improvement on its administrative data scores. The mean performance for providers with at least 10 HEDIS-eligible babies turning 15 months in 2003, on the percentage of babies with four or more visits, rose dramatically from 27 percent in 2003 to 61 percent in 2004, then dropped only slightly in 2005 to 57 percent (Table VI.2).

Thus, while it implemented no new incentives, the plan's "re-advertising" of its pre-existing incentives (along with coding instruction) may have contributed to a major improvement in reporting encounter data, along with a more modest but still significant improvement in its HEDIS scores on the well-visit measure between 2003 and 2004. The care improvements either appeared to be short-lived or the high 2004 HEDIS score was partly due to chance. Its HEDIS score dropped back again in 2005, although the administrative data suggested that reporting rates for the higher-volume providers remained high.

B. PLAN-SPECIFIC SYNTHESIS OF ENCOUNTER DATA RESULTS

1. Plan C

As presented in Chapter IV, the rates of reported preventive service encounters unexpectedly dropped for both groups eligible for the encounter data incentive and groups not eligible for it, although the rate dropped less on average for the groups eligible for the incentive. While the money at stake for this incentive was considerably larger than for the well-baby incentive, totaling \$842,500 during the four quarters September 2003 through August 2004, or \$14,500 per paid provider, the context is different. Unlike the well-baby incentive, which was targeted at individual practitioners, the encounter data incentive was aimed at "PPGs"—participating physician groups that tend to be large organizations bound together to contract with the health plan (sometimes including multiple medical groups).

Given favorable results in the comparative analysis, the incentives along with the technical assistance the plan provided may have had a positive effect on reporting. However, a more certain lesson learned from this plan's experience is that identifying and fixing encounter data reporting problems in the context of a complex set of organization relationships, with capitated payment for services as the financial context, is tremendously difficult and not solved by these incentives alone.

This page has been intentionally left blank for double-sided copying.

CHAPTER VII

CONCLUSIONS

The LIRR demonstration plans were pioneers in implementing financial incentives in the Medicaid context. The evaluation provides some evidence for the effectiveness of the financial incentives for well-baby care, although we believe that any *large* effect from the incentives was limited to one plan (at most). The greatest contribution of the demonstration and evaluation may be that it has generated lessons learned and hypotheses for the next generation of plan incentive programs and for research.

A. PROVIDER INCENTIVES FOR WELL-BABY VISITS

The totality of evidence from the demonstration suggests that provider incentives can be a useful tool to help improve the percentage of babies that receive the recommended number of well-baby visits, but that it is not easy to achieve a large effect with them in the Medicaid context. Provider incentives are feasible to implement in a variety of types of plans serving Medicaid including small and large plans, plans with a wide variety of provider network arrangements, and plans with and without prior experience with incentives.

Although as a group plans that implemented provider incentives for well-baby care improved more than several comparison groups, the qualitative and quantitative analysis together suggests four of the five plans that implemented the incentives did not achieve large effects from the incentive. Due to data constraints, we cannot confirm that a large improvement in HEDIS well-baby visit scores in the fifth plan was caused by the incentives.

1. Facilitating Factors

In the plan where incentives may have contributed to a large improvement in its HEDIS score, the new incentives built on top of a pre-existing payment policy that provided a substantial well-visit fee for service on top of capitation for well-baby visits. This pre-existing policy may have enhanced the effect of the incentives in two ways:

- To receive the fee-for-service payments, providers were already in the habit of reporting well-baby visits on a regular basis, so less change was needed on their

part than for providers in plans that were not already reporting well-baby visits very regularly.

- The total amount of money available to practices that is associated with timely well-baby care was substantial when one considers the payment policy plus the new bonuses.

Also, communications with providers about the incentives appeared not to be problematic in this plan as was the case for others.

Two characteristics of a second plan's program when taken together may have helped it achieve some gains from its incentives, although the case for effect from the incentives is less clear than for the previously mentioned plan:

- The use of the plan's risk pool to implement the incentive meant that substantial dollars were at stake
- The plan essentially mandated physician attendance at a meeting to explain the incentive program (to participate in the risk pool), and as part of the process of strengthening the incentive in 2005 showed physicians along with their risk pool check exactly what they could have received had they scored higher.

2. Barriers to Greater Improvement

Across the plans, barriers to greater improvement in the demonstration, while not assessed quantitatively, may have included:

- Relatively weak plan communication with providers about the incentives, with a couple of exceptions as noted above
- Population characteristics and office factors that constrained the extent to which providers who wanted to improve could do so:
 - Population factors were associated with the low socioeconomic status of the population targeted for improvement, and included parents' focus on economic subsistence, which makes timely well-visit care for their children seem less important; lack of transportation; and mobility of the population and the associated likelihood of inaccurate contact information
 - Office factors included clinics with a heavy walk-in population for whom scheduling appointments is not a routine part of their operation; lack of enough staff to support outreach along with insufficient money to support expanding staff capacity; and system issues such as lack of technology or under-use of technology.

-
- Several program features in specific plans may also have inhibited greater improvement:
 - In one plan, an overly ambitious threshold for providing an incentive, given the starting point of very low administrative data reporting by providers (the incentive was only paid if all six well-visits were found in the data, requiring a large change in practice).
 - In another, a low dollar incentive per baby (\$50 for the six visits), a barrier which may have been aggravated by the fact that the typical provider in this plan served very few babies of the relevant age group, and providers' rates of reporting of well-visits at the start were quite low.
 - In one plan, offering the incentive to only a handful of high-volume providers. This by definition limited the potential of the incentive to measurably improve the plan's score; the incentive would have needed to have a dramatic effect on all targeted providers, an unlikely scenario even under the best of circumstances.

B. ADOLESCENT WELL-VISIT INCENTIVES

The demonstration plans as a group improved substantially on their HEDIS well-adolescent visit scores, however there is less evidence than with well-baby visits that the incentives helped them do so. For example, the two LIRR plans with the most dramatic increases in well-adolescent scores were a plan whose incentive consisted only of payment to one high-volume provider for outreach calls to adolescents, but that also conducted a major direct outreach campaign to adolescents, and a second plan that did not implement any new incentives, but continued its fee-for-service payment above capitation for adolescent well-visits. The one plan with the new broad-based provider incentives did not improve significantly. Although the plans with new movie ticket incentives improved some, they did not reach the national mean by 2004 and did not improve any further in 2005. Further, the movie ticket redemption rates were low.

We cannot conclude from this that adolescent well-visit incentives do not work. The demonstration did not provide as rich a laboratory for studying the effects of adolescent well-visit incentives as it did for well-baby incentives, since only one plan implemented a broad-based provider incentive and two implemented new movie-ticket incentives. There may be other ways not yet known to these plans to improve the take-up rate of movie tickets, making them potentially more effective than seen here.

C. ENCOUNTER DATA REPORTING INCENTIVES

A new encounter data incentive was implemented by only one plan. The incentive may have helped the plan stem a tide of declining encounter data. However, results remain ambiguous, in part because the trend in reporting of service targeted by the incentive mirrored that of non-targeted services.

Other useful qualitative information about encounter data reporting was also uncovered as part of the evaluation:

- The rates of administrative reporting of well-baby visits rose dramatically in a plan that did not implement a new incentive, but that used a physician champion to help re-advertise its pre-existing fee-for-service payments over capitation for well visits as well as to discuss the flexible policies of the plan and articulate the importance to the plan of timely well-baby visits.
- After reviewing the plan-by-plan stories, we believe that for incentives to be designed effectively, they need to consider providers' starting point with respect to reporting encounter data. That is, if the incentives are to be calculated based on administrative data, then the amount of behavioral change in reporting required to trigger an incentive payment should be considered as dollar amounts and thresholds for the incentive are set. This follows from qualitatively comparing this aspect of the plans' experiences between the two plans with broad-based incentives where the incentives more likely worked (Plan D and Plan A) to the two where the case for the incentives working is harder to make (Plan B and Plan C).

D. FUTURE RESEARCH

1. Research Questions Directly Suggested by This Evaluation

As we studied experience with and outcomes of the new incentives, the role of encounter data reporting surfaced repeatedly. Plans in the demonstration without fee-for-service payments for well-care had very low rates of reporting. Further, communication around fee-for-service payment for well-visits is easy for providers to grasp and track. Therefore, we wonder, is fee-for-service payment for well-visits a necessary foundation for an effective incentive program aimed at improving timeliness of preventive care?

The two plans with the greatest improvements in their HEDIS scores on their adolescent well-visit measure give us reason to suggest future research in this area. Specifically, these two plans had very different strategies:

- One (Plan E) combined a pre-existing movie ticket incentive with a major direct outreach campaign, and supplemented these efforts with the financial incentive under the demonstration to several providers.
- The other (Plan D) continued its pre-existing fee-for-service payment above capitation for adolescent well visits, and implemented a new financial incentive for IPAs in mid-2005.

Future research might further explore the dynamics of member and provider incentives through a survey of relevant providers in Plan D and a comparison group to identify whether the providers with the fee-for-service incentives are outreaching more to adolescents than others, as well as a survey of parents of teens in Plan E and a comparison group to obtain their

view on the plan's outreach strategy and assess their and their teens' reported attitudes toward well-care relative to others.

Finally, the decline in encounter data reporting at Plan C despite their major attempt to improve it is worrisome without more information. Is this type of decline being experienced in many other plans? Does it reflect reporting system issues that have worsened, a true decline in care provided, for example due to changes in patient acuity, or a change in physicians' willingness to report due to additional pressure on their practice?

2. Overall Need for Studies of Pay-for-Performance in Medicaid

Additional studies of pay-for-performance programs are needed to research their impacts using stronger comparison groups and statistical methodologies not feasible in this study. Many studies are underway through the Rewarding Results project, but we cannot assume that results from commercial plans will hold when applied in the Medicaid context. The population factors in particular frustrate providers attempting outreach with this population, where that is the main activity necessary to improve the quality target. Provider office factors are non-trivial as well, because providers serving many Medicaid patients tend to have fewer resources for process improvement.

In Medicaid in particular, it seems likely that the complements to pay-for-performance for providers that are implemented with it—whether they are direct outreach efforts by the plan, facilitating sharing of best practice information among providers, or member incentives—may be critical to achieving more than a marginal effect. Therefore, future demonstrations and studies in this area may best focus on finding effective quality improvement techniques that boost the power of financial incentives with other types of assistance.