

Contract No.: 100314  
MPR Reference No.: 6175-400

**MATHEMATICA**  
Policy Research, Inc.

**Evaluation of the  
Medicaid Value Program:  
Health Supports for  
Consumers with Chronic  
Conditions**

*Executive Summary*

*August 14, 2007*

*Dominick Esposito  
Erin Fries Taylor  
Kristin Andrews  
Marsha Gold*

Submitted to:

Center for Health Care Strategies  
200 American Metro Boulevard  
Suite 119  
Hamilton, NJ 08619

Project Officer:  
Melanie Bella

Submitted by:

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

Project Director:  
Erin Fries Taylor

## EXECUTIVE SUMMARY

The Center for Health Care Strategies' (CHCS) Medicaid Value Program (MVP) sought to test interventions seeking to improve care for adult Medicaid beneficiaries with multiple chronic conditions. The program was funded by a grant from Kaiser Permanente, with additional funding from the Robert Wood Johnson Foundation. This report provides Mathematica Policy Research's (MPR) evaluation of the MVP program and the estimates of program effects produced by the programs themselves. This study was funded by CHCS to identify best practices and lessons for future replication or testing. This report is composed of two parts—a cross-cutting analysis of findings, and case studies for each of the 10 interventions tested through the MVP program.

### BACKGROUND: MVP AND THE EVALUATION

MVP sought to build knowledge about effective interventions for Medicaid beneficiaries with multiple chronic conditions. MVP grantees were selected through a competitive process. The solicitation was directed to state Medicaid agencies and the organizations with whom they contract to deliver care. Applicants had relative flexibility to define their target populations and intervention strategies as long as they were focused on clients who each had multiple chronic conditions. An independent review panel reviewed applications to provide feedback on their relevance, strength, and the likelihood that each applicant could implement the intervention within the time and with the resources available. The evaluation team provided feedback to the panel on each applicant's evaluability.

Of the organizations submitting proposals, 10 were ultimately selected. Each team received \$50,000 to help offset its costs but was expected to otherwise self-finance its effort. Each "innovation team" was expected to participate in periodic meetings, work with CHCS (and MPR) on implementation and evaluation design, and share information on its efforts and data on their process and outcome measures. The original timeline of about 17 months (September 2005 to January 2007) was extended another six months to compensate for start-up delays and to allow more time for the interventions to generate effects.

The evaluation sought answers to four basic questions:

1. What interventions did MVP grantees implement and what were they trying to achieve with these interventions?
2. To what extent were MVP grantees successful in implementing their interventions and what factors facilitated or impeded this?
3. Did the interventions achieve the outcomes or impacts sought? If not, why? And if so, how? What factors could have made the intervention more successful?
4. How generalizable is the MVP experience? That is, what was learned about the various models as well as their replicability and utility?

Given the availability of resources, the evaluation relied on grantee-submitted information to assess intervention processes and outcomes, complemented by periodic telephone calls and two rounds of formal interviews. Each round consisted of as many as four or five interviews per team to learn more about the experience and how to interpret the data.

To support the program and evaluation, MPR worked with grantees to identify the “logic model” for each of their interventions and used it to define a small number of process and outcome measures that would be tracked over time, preferably for the intervention and a suitable comparison population. MPR helped CHCS develop a template to structure reporting requirements that captured this and other important information. While this structure could not ensure that a rigorous evaluation would be possible, it provided good information on each intervention, some perspectives on its potential, and guidance on priorities for the future.

## **GRANTEES’ INTERVENTIONS WERE DIVERSE**

The 10 MVP teams all sought to improve care for Medicaid beneficiaries, but they did so in a variety of ways and focused on different populations. Table 1 summarizes the interventions tested throughout MVP. Key features of the interventions can be summarized as follows.

- ***Target Population.*** Target populations varied, with four grantees targeting patients with diabetes and comorbidities, three focusing on mental health and substance abuse care, and two grantees focusing more generally on those at high risk for adverse events and clients with high overall costs (and multiple chronic medical conditions). The remaining grantee was more methodologically focused on comparative assessment of health risk screening tools to support systems redesign.
- ***Intervention Focus.*** Of the nine care-focused programs, seven targeted their interventions on patients, all but one of them using a case management and coordination model to improve patient care. The exception augmented a pre-existing disease management program with in-person patient education. Two grantees targeted their intervention on providers, in the hopes of improving the quality of patient care.
- ***Duration.*** Only two interventions were of very short duration (less than 12 months); the rest had reporting periods of 12 months or more, with an average of 15 months. Two interventions had at least a year of operational experience prior to the start of MVP.

## **GRANTEES SUCCEEDED IN IMPLEMENTING THEIR INTERVENTIONS THOUGH NOT NECESSARILY AS RAPIDLY AS THEY HOPED**

Grantees generally were able to implement the interventions they sought and create the partnerships needed to support those interventions, though in some cases refinements were made.

TABLE 1

## BASIC CHARACTERISTICS OF MVP GRANTEES AND THEIR PILOT INTERVENTIONS

Grantee	Intervention Description	Focus	Target Patient Population	Approximate Size of Patient Population	Study Design	Start Date (Mos. Operational)
CareOregon	Complex case management	Patient based	Costliest patients (top 3 to 5 percent)	330 intervention and 600 comparison	Comparison group	October 2005 <sup>a</sup> (12)
Comprehensive Neuroscience	Health utilization summaries	Provider based	Patients with schizophrenia (summaries sent to their primary care providers)	3,000 patients (2,271 treatment and 729 controls)	Randomly assigned treatment and control groups	May 2005 (17/9)
DC Medical Assistance Adm.	Medical house call program	Patient based	Elderly patients with chronic medical conditions in home setting	85 patients served and 650 comparison	Comparison group	January 2004 <sup>b</sup> (27)
Johns Hopkins HealthCare	Integrated case management	Patient based	Patients with multiple chronic conditions and substance abuse issues	100 intervention and 100 comparison <sup>c</sup>	Comparison group	October 2005 (16)
Managed Health Services	Comparison of two health risk assessment tools	System redesign	SSI clients enrolled in managed care program from April 2005 to November 2005	3,000 (2,800 with at least one assessment completed)	Not applicable	April 2005 (13)
Memorial Healthcare System	Health navigator (case management)	Patient based	Patients with 2 or more conditions, including at least asthma, CHF, diabetes, or hypertension	120 treatment and 40 controls <sup>c</sup>	Randomly assigned treatment and control groups	February 2006 (15)
McKesson Health Solutions	Group diabetes education sessions moderated by health educators	Patient based	Diabetic patients (including those with cardiovascular disease) in state using the McKesson disease management program	28 treatments completed all four sessions and 70 controls <sup>d</sup>	Randomly assigned treatment and control groups	Sessions in April and August 2006 (5/3)
Partnership Health Plan	Preventing Heart Attacks and Strokes Everyday (PHASE)	Provider based	Diabetics with hypertension, cardiovascular disease, or depression	225 intervention and 1,650 comparison <sup>c</sup>	Comparison group clinics	January 2006 (15)
University of California, San Diego	Depression care manager (IMPACT)	Patient based	Diabetic patients with depression participating in a diabetes disease management program	100 patients in intervention group	Three intervention clinics (no comparison group)	July 2006 (10)
Washington State DSHS	Integration of primary, MH/SA, and LTC care with a care coordination team	Patient based	Aged, blind, and disabled Medicaid patients (many of whom have mental health or substance abuse issues)	Average monthly caseload of 2,400, 15,000 comparison	Comparison group	January 2005 (18)

Note: All sample sizes are as reported by grantees as of April 2007, when grantees last reported measures to CHCS.

<sup>a</sup>The first care coordination team was formed as of this date; the number of patients is as of November 2006.

<sup>b</sup>The program began in 1999, but the DC Medical Assistance Administration chose to evaluate it beginning in January 2004. About 500 clients make up the intervention group.

<sup>c</sup>These programs began with more patients but have lost some to disenrollment over time. See case studies in Part 2 of this report for more information.

<sup>d</sup>McKesson randomly assign about 80 patients to the treatment group, but two-thirds decided to not participate in the intervention.

CHF=Congestive Heart Failure; DSHS=Dept. of Social and Health Services; SSI=Supplemental Security Income; MH/SA=mental health/substance abuse; LTC=long-term care.

Start-up delays were common among the grantees. Grantees varied in the size of the intervention group they aimed for from the start, with two substantially larger than the others. The small size of the target populations for many interventions reflects a combination of inherently small numbers of people with certain complex conditions, limited resources of some grantees, and the challenges associated with recruitment for some of the interventions (such as problems with contact information and lower than expected disease prevalence).

## **LEADERSHIP COMMITMENT AND OTHER FACTORS ARE CRITICAL FOR IMPLEMENTATION TO SUCCEED**

The evaluation identified at least five factors that were important across grantees in influencing their success at implementation. First, and consistent with many other studies, nearly all the grantees said that strong leadership commitment from the top of their organization was very important. Second, grantees were most successful at implementation in environments where conditions were favorable—that is, where there were no competing priorities or constraints that limited the attention to (and sometimes the resources for) the intervention. Third, staff, patient, and provider buy-in is critical; staff and patient buy-in is essential in patient-based interventions and provider support essential if changing provider behavior is the focus. Fourth, support and leadership by the Medicaid agency is critical for many grantees to open doors because the agency has authority over program policy and operations; for some, however, equivalent leadership by organizations given major authority by the state can substitute for Medicaid support. Fifth, the ability to standardize the intervention early on, with highly-specified intervention activities and protocol documentation, made it much easier to communicate what was needed and avoid later delays or confusion among those who implement the interventions.

## **GRANTEES FOUND IT EASIER TO IMPLEMENT THE INTERVENTIONS THAN TO GENERATE EVIDENCE OF THEIR EFFECTS ON OUTCOMES**

Each grantee succeeded in implementing its intervention as intended (though perhaps not at the intended scale or speed). However, grantees found it easier to implement changes to their interventions than to design them so that intervention outcomes could be rigorously evaluated. Such an evaluation requires that implementation be strong, solid measures of process and outcome be reported, appropriate comparison data be available for similar populations not subject to the intervention, and intervention scale be sufficiently large that program effects of meaningful magnitude can be detected if they exist.

Through their work with CHCS and MPR, all MVP grantees developed and reported some data on process and/or outcomes for the population in which they intervened. Grantee reporting periods ranged from fewer than 6 months (UCSD) to 27 (DCMAA); the average reporting period was about 15 months and 8 of 10 grantees reported data for 12 months or more. However, individual participants may have participated in interventions for shorter periods of time since many of the interventions had rolling enrollment.

Given the objectives of MVP, understanding what the interventions may yield in terms of improved care for Medicaid beneficiaries with multiple chronic conditions was an important

question for analysis. Whether this question could be answered depends on: (1) the clarity of the intervention (can it be described operationally) and whether it was implemented; and (2) the rigor with which it is possible to determine whether the change had positive effects on outcomes.

To support our analysis of outcomes, we examined each project to assess it against these two criteria (see Table 2). The projects generally were stronger on the first criteria than the second. While most grantees had at least “medium” strength in terms of the clarity of their intervention, definition of the target population, and consistency with available evidence of good practice, only two had a sufficiently well-defined comparison group design, sample size, and patient participation rate (where applicable) to support a rigorous assessment of impacts (Washington State, CNS).<sup>1</sup> While this is a major limitation to our overall assessment of MVP, reported findings on the intervention process for other grantees suggest some innovative and potentially promising programs were successfully implemented.

---

<sup>1</sup> The strongest analyses of outcomes (an “impact study”) include an assessment of intervention-comparison differences with appropriate statistical tests. Only two grantees provided tests for all their outcome measures (CNS, McKesson) and a third (Hopkins) did so for one measure. Most grantees had neither the organizational capacity to conduct these tests nor adequate person-level data. However, for grantees that had large sample sizes and plentiful data, we could make some educated guesses as to the promise of interventions based on the reported measures and what we learned about the interventions during the evaluation.

TABLE 2

RATINGS OF GRANTEEES' INTERVENTION DESIGNS, IMPLEMENTATION, AND IMPACT ANALYSES

Intervention	Patient-based Interventions							Provider-based Interventions			System Redesign
	CareOregon	DCMAA	Hopkins	McKesson	Memorial	UCSD	Washington State DSHS	CNS	Partnership	MHS	
Design Implementation	Low	Medium	Medium	Medium	Medium	High	High	Medium	High	N.A.	
Impacts Analysis	Medium	Medium	Medium	Medium	Medium	Low	High	Medium	Medium	N.A.	
Research Design	Low	Low	Low	Low	Low	Low	High	High	Low	Low	
Impacts on Outcomes	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Yes	No	Unknown	N.A.	

N.A. = not applicable.

## **OUTCOMES FOR THE TWO MOST RIGOROUSLY DEFINED EVALUATIONS SHOW POSITIVE RESULTS FOR ONE BUT NOT THE OTHER**

**Washington State’s Medicaid Integration Partnership** focused on better coordination of primary care, mental health, substance abuse, and long-term care for categorically needy aged, blind and disabled beneficiaries. Under the intervention, these services (previously provided separately) were integrated under one contract with a single health plan (Molina Healthcare of Washington), on a phased basis, including health risk assessment, monitoring of patient symptoms, provider education, and coordination of services, which is particularly intense for those with extensive needs. All eligible beneficiaries were automatically enrolled though they had the option to opt out. The intervention appears to have slowed the rate of inpatient admissions and mental health hospital days among enrollees, improved client satisfaction with some aspects of care delivery (for example, shorter wait times for routine care appointments), and improved care coordination for clients. While the details of the intervention would need to be adapted to each state organizational context, the approach appears relevant to other states. Further, the focus on integration addresses an important area of long-standing interest and provides evidence that care could potentially be improved by centralizing attention to diverse components of care that are often independently provided.

**Comprehensive Neuroscience’s (CNS) Medical Risk Management Project** attempted to improve the quality of care for a large number of people with a low-cost intervention that distributed information to primary care providers on the services that their schizophrenic patients used in the prior year. Because they had a strong and well-implemented design (randomly assigned treatment and control groups), a rigorous impact evaluation could be conducted, indicating no detectable effects on outcomes. The project team experienced a variety of operational problems which probably contributed to the absence of effects (for example, delays in tracking patients and providers, patients without a medical home, limitations in communication with providers); importantly, the team worked hard to address these limitations as they arose which may ultimately influence the scope of the intervention and lead to more promising outcomes. The results suggest that providing information to providers on the care used by their patients is not effective alone and CNS intends to use this insight to strengthen the intervention in the future. This project illustrates the importance of having a valid comparison group design and highlights the caution with which promising trends in the less rigorously defined MVP interventions should be interpreted. Nearly all outcomes were lower during the intervention period compared with the baseline period for both the treatment and control groups. Without a rigorous research design, one might confuse these trends as impacts when, in reality, there were no differences among the two randomly assigned groups.

## **WHILE OUTCOMES CANNOT BE ASSESSED, THE OTHER INTERVENTIONS ALSO GENERATED IMPORTANT INSIGHTS ON CHANGING CARE PROCESSES**

- The Johns Hopkins intervention aimed to use case management within a managed care plan and better communications across sectors of the system to improve care coordination for adult Medicaid beneficiaries with a history of substance abuse and high health care costs, with a focus on improved access to services. Results suggest that use of such services may have increased in the intervention group relative to the



comparison group, though there were design limitations. Since the intervention sought to affect access to these services, it is regrettable that the context (unavoidably small numbers of eligible patients) did not allow a more rigorous test of impacts on process and outcome measures.

- McKesson’s project added an intensive in-person group educational component to standard disease management for aged, blind, or disabled Medicaid clients with diabetes. The results, especially in Oregon, suggest that group educational sessions might have promise to increase patient self-efficacy and hemoglobin A1c testing beyond that of standard disease management. Scale, however, appears to be an issue in this intervention, as McKesson reached far fewer patients than it intended. Any other organization seeking to replicate this intervention should study the reasons for low enrollment carefully because reaching a larger share of potentially eligible people is likely crucial to generating meaningful effects on patient outcomes.
- DC’s medical house call program aims to provide a medical home to people who otherwise cannot physically travel to a physician’s office. A Medicaid waiver option for elderly and disabled clients, the program coordinates care for chronically ill individuals who prefer to remain at home. The program targeted an important high-cost population in an innovative way. Those in the intervention had care patterns consistent with what one would desire—higher use of personal care assistants, durable medical equipment, and medications as well as fewer nursing home admissions and nursing home days. However, the comparison group used to estimate program impacts was not a strong one and the program only collected data during the intervention period. These are serious methodological weaknesses that limit what can be learned about outcomes. However the intervention appears an interesting one that could have promise, so it could warrant more rigorous testing and study in other locations.
- Memorial’s health navigator intervention added a social worker to its existing disease management program to help patients understand the health and non-health services available to them. The health navigator’s role was to conduct patient home visits, complete assessments, and develop care plans. The health navigator completed assessments with all patients she visited and completed a care plan with a high proportion of them. Treatment group members had nearly twice as many contacts with either the health navigator or their primary disease managers compared with control group members. All these process measures are considered, by Memorial, as prerequisites for improving longer-term outcomes. One of Memorial’s early challenges included defining a clear role for the health navigator and integrating her with existing disease management staff. Standardization of these roles is critical for successful replication.
- CareOregon provided team-based case management to patients with various chronic medical conditions with the intent of varying the intensity of the intervention based on client needs to maximize impact on utilization and costs. For example, some clients could be referred to mental health services and others linked to community resources. Setting standards for such a flexible intervention is difficult. While the intervention was not standardized at the outset of MVP, the project team made great

strides over the course of the intervention to define roles for intervention staff and standardize protocols of care. CareOregon found that clearly defined staff roles and protocols for staff improved delivery of the intervention. Because the intervention changed over time and also was not paired to a similar comparison population, it is not possible to gauge the potential of the intervention to generate the savings it hoped.

- Partnership's provider-based intervention aimed to improve patient quality of care for patients with diabetes and other comorbidities. Partnership made a conscious decision to work with specific practices with which it has long-standing arrangements and to give these practices flexibility to make changes as they saw fit. Partnership found that involving a team from each office promoted ownership and helped office staff better understand the intervention; however the design did not generate sufficiently detailed information on the intervention or credible estimates of its effects. Partnership also had a parallel program for diabetes that was patient-focused. Their experience helped generate insight on the importance of coordinating intervention practices with the activities of existing interventions to avoid duplication.
- UCSD added a depression treatment program to a diabetes disease management program at three community clinics; both programs have been studied independently, but never together. Regrettably, the project experienced delays in start up related to the need to line up funding and then subsequent problems in implementation related to obtaining funding for care for uninsured patients and operational challenges (including coordination between clinic staff and the depression care manager). They also found lower than expected prevalence of depression in the target population. Despite these factors, once the depression care manager began working with patients the intervention was intensive, suggesting that the intervention could hold promise if it could eventually be scaled and implemented long enough.
- The Managed Health Services project addressed a policy question important to many Medicaid policymakers: Can we identify clients in need of case management services more efficiently than through resource-intensive health risk assessments? After reviewing two different risk assessment tools (one based on patient self-reports and the other on claims data), MHS believes that the claims-based tool coupled with other data offers an opportunity to identify clients in need of case management more efficiently than is possible with self-reported data. However the design of the study limits the confidence in such conclusions. Because the issues addressed are important, it could be valuable to study the question further with a more focused design accounting for how case management decisions are made.

## **MOST GRANTEES HAVE CONTINUED THEIR INTERVENTIONS AFTER MVP FORMALLY ENDED**

In April 2007, most of the grantees (seven of nine) were continuing their interventions even though MVP had formally ended and each of them appeared to have fairly good prospects for longer-term sustainability. An eighth intervention (Hopkins) was not continued per se, but several of its activities were institutionalized into standard program operations. The ninth intervention (CNS) was funded by the state of Missouri to continue in a modified form. As with

implementation, support from top leadership was critical for sustainability. Funding is an important issue for interventions' sustainability, particularly those that hire dedicated staff. The availability of such funding obviously also is influenced by leadership commitment. Most grantees said the business case (return on investment) was important but only two grantees planned to measure it following the completion of MVP. In several cases, grantees viewed the business case as resting less on short-term gains than on long-term impact on cost or on the organization's financial strength.

This suggests that either the grantees are sufficiently convinced there is a business case for their interventions going forward despite the lack of empirical evidence, or that the business case is not as important as they report. Most of these interventions do not appear to be very resource-intensive. Organizations may feel that spending such modest sums does not justify the need for rigorous evidence of effectiveness, particularly if it promotes innovation and demonstrates the sponsor's efforts to help patients and improve care or if it generates goodwill among invested staff. Because of the way organizations operate, this could constitute a sufficient business case for leadership at sponsor organizations.

## **MANY INTERVENTIONS APPEAR REPLICABLE BUT MOST REQUIRE FURTHER STUDY TO DETERMINE THE VALUE OF DOING SO**

The replicability of an intervention depends on: (1) the clarity and specificity of intervention activities (do we know what the intervention is in enough detail that another organization could repeat it); and (2) its organizational and environmental context (how unique is its the setting in which the program took place and how applicable is it to other settings). In addition, whether or not it makes sense to replicate an intervention depends on what is known about its value (are there potential benefits to organizations implementing it and to their patients or providers in terms of favorable impacts on quality, patient outcomes or cost in the short- or long-term).

Most grantees thought that their interventions were replicable. We tend to agree. By and large, the interventions appear relatively "generic" efforts that could work in many, though not necessarily all, environments, with some modest tailoring to fit particular organizational features. Most interventions appear to have sufficient documentation to support efforts at replication. However, in a few cases, replication would be difficult because the interventions were not well documented and standardized protocols were not developed.

The more challenging issue involves whether it makes sense to encourage replication. The grantees generally thought that doing so would be valuable even if they were not able to show empirical evidence on outcomes or business returns. Because these are relatively low-cost interventions, there may be organizational returns to doing so, as noted previously. However, MVP was initiated as a vehicle for identifying ways to improve care for adult Medicaid beneficiaries with multiple chronic conditions. The Washington State intervention had relatively strong evidence of effectiveness; the CNS intervention did not. Some others showed promise in terms of potentially improved processes of care but further testing would be required to judge their effects on outcomes.

## **GRANTEES VALUED THE SUPPORT OF MVP AND CHCS**

Grantees valued the support provided by CHCS and the MVP structure as they pursued their interventions. The structure provided by MVP (including the framework for reporting measures and the role of CHCS in keeping grantees on target) was the most valued area of support. Participants also found the meetings useful and the seed money important in allowing them to conduct their interventions and garner internal support. Association with a project like MVP also added prestige to their efforts. They suggested, however, that communication and support between meetings could have been stronger. Grantees with less experience seemed particularly interested in ongoing general support, whereas others focused more on specific areas for which they sought technical support. The majority said Kaiser Permanente sponsorship added to the value they gained from MVP. (Others had no opinion or were not aware of the sponsor). While Kaiser Permanente was less visible to grantees than CHCS, grantees saw Kaiser as opening doors to potential opportunities and lending prestige to the effort.

## **CONCLUSIONS**

MVP was formed to help expand knowledge of ways to improve care for adult Medicaid beneficiaries with multiple chronic conditions. The program succeeded in generating interest among states and health plans in developing such interventions and in building on that interest to select 10 interventions for implementation. MVP also was successful in implementation. Though progress was slower than many grantees initially hoped, each grantee was able to implement its intervention and eight had at least one year of operational experience before MVP ended. In most cases, grantees continued their interventions after the formal program ended. Further, grantees still appeared enthusiastic about their work at the end of the program and positive about the contribution made by CHCS and the MVP program structure to their efforts.

MVP was much less successful in rigorous, empirical testing of the effectiveness of the interventions. The focus on logic models and measures succeeded in generating quantitative measures on a few critical process and outcome measures. However, only two of the interventions had a sufficiently strong comparison group methodology and enough participants to support formal testing of impacts. This outcome is not surprising, given the limited resources CHCS had available to support data collection for rigorous evaluation and the limited resources available to many of the grantees.

Given the impetus behind MVP, one key question remains: What does the program contribute to our understanding on how to improve care for its target population—Medicaid beneficiaries with multiple chronic conditions? We believe the contribution has been positive on several dimensions.

First, from a process perspective, MVP demonstrated the value of using logic models and process measures to help grantees be more clear about their interventions and what they hoped to achieve. Even though MVP did not generate solid evidence of effects, the descriptive information supported by this approach will make it easier for others to learn from the MVP experience.

Second, MVP generated evidence suggesting that well-conceived efforts to better integrate care across the range of services (primary care, mental health, substance abuse, and long-term care) required by beneficiaries with multiple chronic conditions, difficult though that may be, have promise. This promise is best reflected in the Washington State Medicaid Integration Partnership but also in the Johns Hopkins care management model. Each of these aimed to modify the way benefits were used and to better integrate care across sectors of services. The interventions also were structured so that financial incentives reinforced the goals of health care services integration.

Third, the findings show that it is not just what the intervention is that matters, but also that the *intensity* of the intervention is likely to be important to improving outcomes for patients with multiple chronic illnesses. This is best illustrated by the challenges CNS faced in generating strong positive effects for what in effect was a relatively low-intensity intervention. However, other grantees also found it challenging to implement their interventions (CareOregon) or to intervene in a way that reflected a sufficient change from standard practice that it was reasonable to expect changes in outcomes (Partnership Health Plan).

Fourth, MVP brings to light what could be some difficult or even insurmountable challenges in building a strong empirical evidence base on ways to improve care for adult Medicaid beneficiaries with multiple chronic illnesses. As MVP grantees found, many relevant subgroups are, by definition, small in number. Further, existing administrative data may not enable sponsors to identify this group reliably. Because costs for these groups tend to be high and numbers small, the power with which interventions can be tested will be constrained inherently by the chance that a single “outlier” patient with a particularly poor and costly outcome may drive the estimates of effects on costs. Utilization-based measures are less sensitive to this constraint but the shift in focus away from resource considerations could make it harder to assess the business case for interventions.

## RECOMMENDATIONS

We believe that these conclusions highlight at least three recommendations for future attention pertaining to improving care for adult Medicaid beneficiaries with multiple chronic conditions.

**First, favor multi-faceted yet well-targeted interventions with sufficient intensity to affect outcomes.** The populations targeted by MVP interventions have complex conditions and multiple needs. These patients interface with the health care system in a variety of ways. CHCS may not want to promote a particular model of care (such as the chronic care model), but it would seem critical to focus on interventions that have the potential to drive change in ways that align processes to reinforce improvements in care and outcomes. Such an orientation seemed to be best reflected in the Washington State intervention and it is intriguing that this program provided the most concrete evidence.

**Second, put greater emphasis on learning and design before testing.** While CHCS scanned the environment prior to implementing MVP, the program was not conceived in a prescriptive fashion and allowed grantees substantial flexibility to develop their own

interventions for testing. To different degrees, each of the grantees found they needed to spend substantial time defining their intervention more clearly before they could proceed. Often, changes in care processes were being implemented for the first time or conceived without benefit from existing experience elsewhere (if it existed). Diversity also limited what grantees could learn from one another or others could learn by examining the collective experience. Given the challenges illustrated by MVP in assessing the effects of interventions, we believe it valuable to spend substantially more time exploring potential interventions for their promise so that efforts and tests could be focused on those that are most promising. Rapid cycle methods are well-suited toward developing testable models, especially if complemented by a rigorous and comprehensive review of existing experience in improving care for adults with chronic illness.

**Third, consider multi-site tests of the most promising interventions and convince funders to invest the resources needed for rigorous evaluation.** Creating change through small-scale interventions that are narrowly focused geographically or defined such that they reach small numbers of people, however sick they are, makes it hard to test interventions. If there are particularly promising interventions, it could be strategically of value to focus resources on bringing these to scale for rigorous testing. For example, for a chronically ill population with average annual hospitalization rate of one per patient, detecting a 15 percent difference in hospitalizations would require a treatment group of 550 or more patients (who participate in the intervention) with a randomly assigned control group of equal size. By standardizing intervention strategy (even with allowable customization by site), one can better pool results to better capture their impact. Beyond the numbers, multi-site tests also add insight on the replicability of an intervention across sites, especially if there is sufficient data to assess effectiveness at the site level as well as across sites.