Contract No.: 500-95-0047 (09)
MPR Reference No.: 8756-420

MATHEMATICA
Policy Research, Inc.

Second Report to Congress on the Evaluation of the Medicare Coordinated Care Demonstration

December 15, 2006

Randall Brown
Deborah Peikes
Arnold Chen
Judy Ng
Jennifer Schore
Clara Soh

Submitted to:

Centers for Medicare & Medicaid Services
Office of Strategic Planning
C3-20-17
7500 Security Boulevard
Baltimore, MD 21244

Project Officer:
Carol Magee

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Director:
Randall Brown

Table of Contents

Page

Care for beneficiaries with chronic illnesses, such as heart disease and diabetes, is a major expense to the Medicare program, and a major detriment to beneficiaries' quality of life. For example, just under one-half of all beneficiaries in 1997 were treated for one or more of eight categories of chronic illnesses, and they accounted for three-fourths of all Medicare spending in 1998 (Brown et al. 2004) Furthermore, beneficiaries often have multiple chronic illnesses, which compounds the cost and complexity of their care. The 12 percent with three or more of the eight chronic health problems accounted for one-third of all Medicare spending. Coordinating the care these patients require is difficult, because patients with chronic illnesses see an average of 11 different physicians per year (Anderson 2002). Despite these alarming statistics, many of the acute health problems caused by chronic illnesses can be prevented if (1) patients are provided with medical care that is consistent with recommended standards; (2) patients adhere to recommended diet, medication, exercise, and self-care regimens; and (3) providers communicate better with each other and with patients. A number of small pilot programs designed to improve patients' adherence to treatment regimens and physicians' adherence to professional guidelines have been found to be effective in improving patient outcomes and reducing costs (see reviews by Chen et al. 2000; Wagner et al. 2001). This potential has led many health maintenance organizations and indemnity insurers to develop their own programs or to contract with disease management or case management providers for such programs (see Villagra and Ahmed 2004 for evidence of the effectiveness of disease management for diabetic patients in a managed care setting). However, the Medicare fee-for-service program does not cover such services.

Section 4016 of the "Balanced Budget Act of 1997" requires the Secretary to conduct a Medicare Coordinated Care Demonstration (MCCD) which tests whether case management and disease management programs can lower costs and improve patient outcomes and well-being in the Medicare fee-for-service population. In January 2002, the Centers for Medicare & Medicaid Services (CMS) selected 15 demonstration programs in a competitive awards process, under which each was allowed to define its own intervention and target population, within broad parameters. Each program began enrolling patients between April and September of that year and was authorized to operate for 4 years. Beneficiaries who agree to participate are randomly assigned by the evaluator, Mathematica Policy Research, Inc., to either the treatment group, which received the intervention, or the control group. Both groups continued to obtain their traditional Medicare–covered services from fee-for-service providers in the usual manner.

This report to Congress synthesizes findings from the first 2 years of the demonstration programs' operations, focusing on program impacts over the first year after enrollment for beneficiaries who enrolled during the first year, and over the first 25 months of operations for all enrollees. Findings presented include program-specific estimates of impacts on (1) survey-based measures of patients' health status, knowledge, behavior, satisfaction with their health care, quality of care, and quality of life; and (2) claims-based measures of patients' Medicare service use and expenditures, and the quality of care received. The report links differences across programs in these impacts to differences in the interventions and the target populations in order to draw inferences about "what works" and "for whom." This synthesis of findings draws on an earlier report to Congress that described the types of programs and beneficiaries participating in the demonstrations, the interventions the programs have implemented, and how well patients and physicians like the programs (Brown et al. 2004). This report updates that information and adds analyses of Medicare service use and expenditures and a scoring methodology developed specifically for this evaluation to rate the quality of each program's intervention on several dimensions.

The findings in brief indicate that patients and physicians were generally very satisfied with the program, but few programs had statistically detectable effects on patients' behavior or use of Medicare services. Treating only statistically significant treatment-control differences as evidence of program effects, the results show:

- Few effects on beneficiaries' overall satisfaction with care
- An increase in the percentage of beneficiaries reporting they received health education
- No clear effects on patients' adherence or self-care
- Favorable effects for only two programs each on: the quality of preventive care, the number of preventable hospitalizations, and patients' well-being
- A small but statistically significant reduction (about 2 percentage points) across all programs combined in the proportion of patients hospitalized during the year after enrollment
- Reduced number of hospitalizations for only 1 of the 15 programs over the first 25 months of program operations
- No reduction in expenditures for Medicare Part A and B services for any program

Despite the absence of statistically significant treatment-control differences in Medicare expenditures for traditional services, it is possible that some of the programs are cost neutral to date. This could be true because the large variation in Medicare expenditures and the small number of beneficiaries enrolled in some programs make it difficult to draw definitive conclusions—for nine programs, treatment-control differences over the first 25 months of operations are not statistically different from zero, but they are also not significantly different from the average fee paid to the programs. Based on the patterns of differences in hospitalizations, Medicare Part A and B expenditures, and total Medicare expenditures including the care coordination fees, six of the programs are not cost neutral, four probably are not, and five may be cost neutral, over their first 25 months of operations.

The results presented here are not the final word on the programs' impacts—changing ingrained behaviors of physicians and patients and improving communications among non-integrated fee-for-service providers are all difficult tasks to achieve. Furthermore, even if achieved, such improvements in the processes of care may not yield statistically discernable improvements in patients' well-being or reductions in Medicare costs over the first 2 years of program operations. Thus, the estimates presented here may differ from those that will be observed over the full 4 years of operations. Nonetheless, this report provides (to our knowledge) the largest single random assignment study to date of disease management/case management programs, and only the second evaluation ever conducted of such programs in a Medicare fee-for-service setting (the first was Schore et al. 1999).

A. WHAT TYPES OF PROGRAMS AND BENEFICIARIES ARE PARTICIPATING?

The 15 MCCD programs were selected from 58 proposals responding to CMS's solicitation. Programs' hosts had to have experience operating a disease management or case management program that had been shown to reduce hospitalizations or costs in some population or setting. CMS took this approach to maximize the potential for showing, in a time-limited demonstration, that successful care coordination programs used in other settings (typically managed care) could be implemented in a Medicare fee-for-service environment. Each program is offered only to patients living in its catchment area and meeting its approved eligibility criteria—typically, having a particular chronic illness. (Some programs further restrict enrollment to patients who have had a hospitalization during the year or 2 years preceding enrollment.)

In return for providing the care coordination intervention described in its CMS-approved operational protocol, each program receives a negotiated monthly payment for each beneficiary who chooses to enroll and is randomized to the treatment group. Payments to the programs range from $50 per enrollee per month for low-risk patients with one or more of several chronic illnesses in one program to $437 per month for the first 9 months for all patients with congestive heart failure (CHF) enrolled in another program. The negotiated rates were based on the programs' estimates of the cost of their interventions; however, to increase the likelihood that each program would generate net savings to CMS, the rates also were tied to the projected costs of the programs' proposed target populations. If a 20-percent savings in these projected Medicare costs would not be enough to offset the cost of the intervention, either a program restricted the proposed target population to higher-risk cases (such as beneficiaries with a recent hospitalization) or CMS reduced the proposed program payment to meet this constraint. Five programs had monthly fees exceeding $300; six had fees below $175.

The evaluation's first report to Congress on the MCCD showed that the 15 selected programs varied widely in their organizational structures, target populations, and interventions, and that they had varied levels of success in recruiting patients (Brown et al. 2004). The participating organizations include five commercial disease management vendors, three hospitals, three academic medical centers, an integrated delivery system, a hospice, a long-term care facility, and a retirement community (see Table 1). The programs operate in 16 states (mostly in the northeast and Midwest) and in the District of Columbia; five serve beneficiaries living in sparsely populated rural areas. The programs also vary widely in the numbers and types of chronic conditions they target, with six programs targeting only a single condition, three serving patients with less-specific problems (for example, high-risk patients identified from administrative data by an algorithm), and the six other programs falling between these two extremes. Ten programs required that a patient have a hospitalization for the target condition in the year (or less) prior to enrollment.

TABLE 1

CARE COORDINATION PROGRAMS PARTICIPATING IN THE EVALUATION

Host Organization (Average Monthly Fee)	Organization Type	Service Area	Targeted Diagnoses	Number Ever Enrolled After 12 and 24 Months	Medicare Expenditures per Month During Year Before Enrollment
Programs Starting in April 2002					
Carle Foundation ($159)	Integrated delivery system	Rural counties in east central Illinois and west central Indiana	Heart conditions, diabetes, chronic lung disease	2,283 2,642	$521
CenVaNet ($80)	Care coordination provider	Richmond, Virginia	Heart conditions, diabetes, chronic lung disease, cerebrovascular disease	1,074 1,305	$953
Charlestown Retirement Community ($244)	Retirement community	3 retirement communities in the Baltimore area	Heart conditions, diabetes, COPD	430 802	$1,159
Health Quality Partners ($108)	Care coordination provider	Eastern Pennsylvania (rural)	Heart conditions, diabetes, asthma, moderate to severe hyperlipidemia or hypertension	498 1,140	$414
Medical Care Development ($297)	Hospital consortium	Rural areas of Maine	Heart conditions	393 876	$1,718
Mercy Medical Center/North Iowa ($257)	Hospital	Rural areas of Iowa	CHF, chronic lung disease, liver disease, stroke, vascular disease, renal failure	627 865	$1,315
Programs Starting in June 2002					
Avera Research Institute/Avera McKennan Hospital and University Health Center ($316)	Hospital	Rural counties in Iowa, Minnesota, Nebraska, and South Dakota	CHF	318 624	$1,615
CorSolutions ($444)	Care coordination provider	Harris County (Houston), Texas	CHF	671 2,162	$2,644
Georgetown University Medical School ($320)	Academic institution	Washington, DC, and parts of Maryland and Virginia	CHF	108 199	$2,530
Jewish Home and Hospital Lifecare System ($317)	Long-term care provider	Manhattan, New York City	Heart conditions, diabetes, chronic lung disease, cancer, liver disease, stroke or other cerebrovascular disease, psychotic disorder, major depressive or anxiety disorder, Alzheimer's disease or other cognitive impairment	543 766	$1,450
University of Maryland Medical School ($350)	Academic institution	Baltimore	CHF	58 137	$3,299
Programs Starting in July Through September 2002					
Hospice of the Valley ($224)	Hospice	Maricopa County, Arizona (greater Phoenix)	CHF, COPD, cancer, neurological conditions	470 814	$2,174
QMed ($96)	Care coordination provider	2 counties in northern California	CAD	1,404 1,454	$507
Washington University School of Medicine ($173)	Academic institution with care coordination provider	St. Louis, Missouri	High-risk patients who are clinically unstable , targeted through proprietary algorithm	1,425 2,038	$2,263
Quality Oncology, Inc. ($140)	Care coordination provider	Broward County, Florida (Miami)	Cancer	63 141	$2,885

CAD = coronary artery disease; CHF = congestive heart failure; COPD = chronic obstructive pulmonary disease.

The mix of patients enrolled varied across programs on some characteristics, but on others the programs were quite similar.

- The most common primary conditions of program patients were CHF (29 percent of enrollees), coronary artery disease (CAD) (24 percent), and diabetes (13 percent)
- Four programs drew a high proportion of beneficiaries who were older than age 85, and one program targeted and enrolled a high proportion of younger beneficiaries with disabilities
- Compared with all Medicare beneficiaries, the programs' patients generally were substantially more highly educated and had higher incomes
- Most programs enrolled relatively few black or Hispanic patients, few patients younger than age 65, and few patients who also were enrolled in Medicaid

Many of the programs had unexpected difficulty enrolling the target number of patients, with only four exceeding the first-year target of 686 patients that was set by Mathematica Policy Research, Inc. as being the minimum necessary for the evaluation (although two others had over 600; see Table 1). Several programs enrolled less than one-half their targeted number of patients for the first year, citing initial overestimates of the number of eligible patients from their referral sources, physicians' failure to encourage their patients to enroll, high patient refusal rates, and limited care coordinator time to both recruit patients and serve those already enrolled. The programs that were most successful in enrolling patients were those that had a close relationship with physicians before the demonstration started and those with access to databases (such as clinic or hospital records) to identify potentially eligible patients. By the end of the second year, 12 of the 15 programs had over 600 enrollees.

Most of the programs succeeded in enrolling patients with serious chronic illnesses, but a few programs enrolled relatively healthy patients. Preenrollment Medicare expenditures averaged more than $2,000 per month during the year preceding enrollment for first-year participants in six programs, but less than $600 per month for three other programs (average Medicare expenditures for noninstitutionalized beneficiaries nationally was $505 per month in 2002). The programs with low-cost enrollees are likely to have difficulty achieving large enough savings to offset the cost of their interventions. In one-half (eight) of the programs, enrolled patients had an average of one or more hospitalizations per year during the 2 years before enrollment. (Three of these programs averaged two hospitalizations per patient per year.) In 13 of the programs, the enrolled patients had higher costs than did diagnostically eligible nonparticipants in the same geographic area during that year. However, the two programs whose enrollees had the lowest preenrollment Medicare costs (about $500 per month) enrolled patients with preenrollment costs and admission rates that were lower than those of eligible nonparticipants. The program with the greatest preenrollment discrepancy between participants and nonparticipants enrolled sizable numbers of beneficiaries it identified as eligible through chart reviews, but many of these enrollees did not meet diagnostic eligibility criteria according to claims data examined here, for the year prior to enrollment.

B. WHAT INTERVENTIONS ARE THE PROGRAMS DELIVERING, AND HOW ARE THEY DOING IT?

The 15 programs differed widely in both how they implemented their care coordination interventions with patients and their involvement with patients' physicians and other providers. Information about the interventions came from interviews with program staff as well as data recorded by care managers on their contacts with patients. Interviews were conducted at three points: by telephone at 3 months after enrollment began; during in-person visits 6 months after the telephone interviews; and by telephone again, roughly 3 years after startup.

The programs differed in their relative emphasis on four major vehicles for achieving better outcomes for patients: improving patients' adherence to treatment and self-care regimens, improving coordination and communication among providers, improving physician practice, and increasing access to support services. All but 1 of the 15 programs stressed patient education to improve adherence and coordination, but most devoted less attention to convincing physicians to change their practices or to improving access to support services.

The programs varied greatly in their approach to care coordination. They differed on the mode and intensity of contacts, staff credentials, ratio of staff to patients, method of monitoring, patient education methods, and approaches to improving communications between physicians and patients and among physicians.

All but two programs required all their care coordinators to be registered nurses, but caseload size varied widely. Thirteen programs required care coordinators also to have specific experience with cardiac, geriatric, medical-surgical, or community nursing. Caseload sizes ranged from a low of 36 patients per care coordinator to a high of 200.

All programs began care coordination with assessments of patients' needs and condition, after which they developed patient care plans. Of the 15 programs, 12 conducted at least part of their assessment in person, even though most of their invention was conducted over the telephone. Ten programs initiated their assessments within 3 weeks after enrollment on average. Only one program (Jewish Home and Hospital) took longer than 6 weeks on average to begin its assessment. The assessments culminated in care plans to fill the gaps in the patients' knowledge and treatment. These plans were developed collaboratively with patients and, when appropriate, with the patients' families.

Most (12) of the programs contacted patients one to three times per month on average (mostly by telephone), but 2 had more frequent contacts. Six of the programs averaged 1.2 to 1.5 contacts per patient per month during the first year after enrollment; another six averaged between 2.2 and 2.9. Avera, however, contacted patients over 8 times per month on average. The great majority of contacts were by telephone, except in Mercy, whose care coordinators conducted over two-thirds of their contacts in person. Patients initiated about 10 percent or less of the contacts in most programs.

Six programs used home telemonitoring devices, although three of these did so on a very limited basis. Electronic devices transmitted patients' weights, other clinical indicators, and symptom reports to their care coordinators on a daily basis. A seventh program provided ambulatory ischemia monitoring. In addition, 13 programs required care coordinators to contact all of their patients at least monthly by telephone or in person.

All but one program provided patient education; almost all used standard curricula and had processes for assessing the effectiveness of the education. Over 85 percent of enrollees in the 14 programs featuring patient education (only the University of Maryland did not) received contacts for educational purposes during their first year in the program. The educational materials were part of electronic databases for some of the programs, and some assessed patients to identify specific learning barriers. Programs assessed effectiveness by reviewing clinical indicators or home monitoring data for evidence of improving health or relied on patients' self-reported behavior changes or responses to questions about their knowledge.

Most programs sought to improve communication between patients and providers by training patients, and they sent physicians regular written reports on patients. Some programs taught patients to take prepared lists of questions to their office visits, while others gave them schedules of tests they should be receiving. While most programs communicated with patients' physicians via written reports, one held formal conferences with participating physicians and one had its quality manager visit physicians to discuss adherence to evidence-based practice, using data obtained from ambulatory ischemia monitoring and physicians' medical records to make their point. Five programs had care coordinators practice in the same location as physicians, enhancing the coordinators' ability to communicate face-to-face with them. Seven programs arranged to have hospitals notify care coordinators when the hospital admitted program enrollees or had their care coordinators review hospital or emergency room (ER) admission lists.

Programs sought to minimize the burden on patients' physicians. Only four of the programs listed improvement of provider practice as one of their approaches for improving patient health. They did so mostly by providing recommendations for specific patients when treatment plans deviated from evidence-based guidelines. One program, however, provided education about such guidelines and offered physicians incentives to participate. Some programs used opinion leaders or advisory boards to encourage physicians' active participation, paying either a monthly stipend per patient ($20 to $30 typically) or paying for participation in meetings or for delivery of medical records.

The programs devoted relatively little attention to increasing patients' access to needed support services. All but one program provided such assistance, such as referring patients to transportation services or home-delivered meals, but only five ever did so for more than one-half their patients.

Programs varied widely in the sophistication of their electronic systems to manage data on patients and program activities. Thirteen programs used these systems to support their work with patients. Among those, 11 generated reports from those systems reminding coordinators about when to contact patients, and 12 used the systems to provide reports on patients' clinical indicators and outcomes.

While information on *what* programs are doing as their interventions and how they do it can be useful for understanding why some programs are more effective than others, it may be more important to know how heavily they focus on particular dimensions of care coordination and how well designed the interventions are on these dimensions. On the surface, many of the programs in this demonstration appear to implement quite similar interventions, yet in-depth discussions with the programs reveal a number of important differences in the intensity of their intended efforts to provide patient education or service arrangement or other possible components of their intervention. To address this issue, the evaluation developed a scoring algorithm for rating each program's interventions on 10 separate domains:

- Program Staffing
- Initial Assessment
- Patient Education
- Improving Communication and Coordination
- Improving Provider Practice
- Service and Resource Arrangement
- Information Technology and Electronic Records
- Ongoing Monitoring
- Quality Management and Outcome Measurement

These ratings were developed independently of the survey and claims data on program outcomes, and without regard to data on contacts supplied by the programs. Researchers scoring the programs relied solely on the information collected during in-person and telephone discussions with the programs about their intervention; estimates of program impacts were not shared with scorers until after they had completed their ratings. Scores were normalized to range from 0 (intervention did not address this domain) to 100 (intervention was extremely well-designed on this domain).

Programs varied widely on each of these domains, especially Quality Management and Outcome Measurement, for which scores ranged from 5 to 91, and Improving Provider Practice, which ranged from 0 to 77. Scores varied less widely across programs on the Problem Identification and the Initial Assessment domains. Average scores were highest for the Initial Assessment and the Monitoring domains, and lowest on average for Improving Provider Practice, reflecting the lesser attention given to this area by most of the programs.

While individual programs often scored extremely well on some domains and poorly on others (at times because a particular domain was not part of its intervention), a few programs had high scores on several domains and others had consistently low ratings across most of the domains. Carle was scored in the top quintile of programs (the 3 highest) on 6 of the 10 domains, and Mercy and Quality Oncology each had 4 scores in the top quintile. The Jewish Home and Hospital and the University of Maryland scored in the bottom quintile on nine and seven of the domains, respectively. Yet both of these programs scored in the top quintile on one domain each. The importance of these rankings is not to identify those programs that do particularly well or poorly across measures, but to determine whether having strong designs in certain domains is consistently associated with having favorable impacts on Medicare costs or the quality of care.

C. HOW DO PATIENTS AND PHYSICIANS LIKE AND RESPOND TO THE PROGRAMS?

Survey data collected on patients in the 12 programs with over 300 enrollees by the end of their first year and on enrollees' physicians in all 15 programs suggest that the programs are popular with both patients and physicians. The patient surveys generally were conducted 7 to 12 months after patients enrolled. Physicians were surveyed in two waves, once about 12 to 15 months after the program in which their patients were enrolled began operations, and a second wave about 18 to 21 months after program startup.

About two-thirds of treatment group patients on average across programs were aware of the program; 15 percent of control group members also reported receiving some care management. Most treatment group members were aware they were receiving care coordination although the percentage varied widely across programs, ranging from only 30 percent in QMed to 81 percent in Mercy saying "yes" when asked, "During the past 12 months, did someone like a nurse, social worker or geriatric nurse help arrange or coordinate your health care?" Thus, the programs generally were successful in establishing a relationship with the patients. However, 3 to 28 percent of the control group also answered "yes" to this question, suggesting that the interventions are not the only source of professional care coordination assistance available in the programs' service areas. Among those saying they received this type of assistance, those in the treatment group generally reported higher levels of satisfaction with the help received. Nonetheless, the fact that about one-third of treatment group members did not report receiving care coordination and that some control group members reported they did receive such assistance makes it more difficult for the programs to demonstrate a significant impact on the treatment group.

Treatment group patients were generally very satisfied with the care coordination they received. Coordinators were rated on four different dimensions—support and monitoring, knowledge and ability to get answers, ability to explain adherence to recommended self-care, and help arranging services—each with three or four specific indicators. About one-third to one-half of the patients surveyed rated their coordinators as excellent on the 14 indicators examined, and most of the rest rated them as "very good." Very few patients (less than 10 percent in nearly all instances) rated the programs as only fair or poor on any of the measures. Care coordinators received especially high marks on indicators of the emotional support and monitoring they offered, especially their "caring attitude," with over 60 percent of the patients on average giving their programs an excellent rating. Patients also rated programs highly on staying in touch (over one-half rating it excellent, on average). Patients gave somewhat lower, but still quite positive, ratings on average for programs' including them and their families in decisions, and for helping them cope with their illness and avoid complications.

Patients were somewhat less impressed with the help they received from programs in arranging appointments or services. Across most programs, about 35 to 40 percent of the patients gave an excellent rating. Exceptions include Carle's higher ratings, and two programs that received markedly lower ratings (the same two programs with low marks on support and monitoring). Substantial minorities of patients (10 to 24 percent) gave the programs a fair or poor rating on this domain. These less favorable ratings are likely to be due to most programs' focusing their attention more on monitoring and education than on arranging services.

Patients had high praise for the care coordinators' knowledge. Over one-half the patients on average rated care coordinators' knowledge as excellent, and only two programs had less than 43 percent giving an excellent rating. About 40 to 43 percent of patients on average rated their programs' care coordinators as excellent on their ability to explain symptoms or get physicians to answer questions or help them to identify early warning symptoms; these rates were similar across most programs.

Finally, a modest proportion of patients gave excellent ratings to care coordinators' ability to explain recommended diet, medication, and exercise regimens. Of all the measures, patients were least likely to give coordinators very high marks on their ability to explain exercise regimens (although few patients rated the programs as fair or poor). The somewhat less enthusiastic ratings on these measures may be due to care coordinators' focusing their education efforts less intensely on exercise than on other patient behaviors.

Overall, a consistent pattern emerges from these numerous patient ratings of the care coordination interventions, with Health Quality Partners consistently receiving notably higher marks than other programs. These high patient ratings were consistent with the evaluation's scoring results, in which Health Quality Partners had the highest score of the 15 programs on patient education and ranked among the top on monitoring as well. Carle and Avera also were rated highly on some patient survey measures, especially those related to providing emotional support and service arrangement. Avera's high ratings on explaining early warning signs is consistent with the scoring algorithm's strong ranking of this program (third highest) on patient monitoring and its use of home telemonitoring, which likely generated follow-up conversations between care coordinators and patients about heart failure and symptoms. Carle's high ratings from patients on getting answers from physicians is consistent with its top score among all programs on improving communications and coordination among providers and the relatively close relationship its program staff had with their patients' physicians. Carle's patients' high ratings of the program on service arrangement is also consistent with Carle having one of the top three scores on service arrangement in the scoring algorithm.

Most of the programs received high ratings from their patients' physicians on most dimensions, although there were clear differences across the dimensions and across programs. Physicians were asked to rate the programs on numerous factors, including their effects on the physician's practice (medical practice, time and paperwork burden, and financial impact if any), patients' education and behavior, service arrangements for patients, care coordination, physicians' relationship with patients, and patient outcomes and behavior. Physicians were also asked to rate care coordinators' clinical competence.

Program physicians widely agreed that the programs made things easier overall for patients and did a good job of monitoring and followup, but they were not always as positive about the usefulness of program reports (42 percent responded these were "very useful") or about the programs' effects on other aspects of their practice. Table 2 illustrates the wide range of responses across programs. Similar wide variation across measures and programs occurred in each of the other categories.

TABLE 2

PHYSICIANS' SATISFACTION WITH CARE COORDINATION
(Percentages)

	Mean	Minimum	Maximum
Program Reports on Patients Very Useful	42	0	91
Medical Practices a Little or a Lot Better on:			
Reducing problems with polypharmacy	56	11	81
Reducing telephone time	55	4	88
Making things easier for staff	56	22	86
Making care more evidence-based	49	20	95
Making it easier overall to care for patients	75	33	100
Monitoring and Followup Very Good/Excellent	71	38	100

Note: The mean is the average across the 15 programs.

The patients' primary physicians in general were pleased with the program overall. Across the 15 programs, on average, a majority (67 percent) of physicians felt that the program increased patients' overall quality of care, and 80 percent said they would recommend the program to patients and colleagues (about 60 percent said they would "definitely" recommend the program and the remainder said they would "probably" recommend the program).

There were some major variations across programs in physicians' ratings. For example, 95 percent of physicians in Charlestown found the program improved patients' quality of care and would definitely recommend the program to others, while only 11 percent of physicians in Quality Oncology were as impressed on either measure. Charlestown consistently received higher ratings from its patients' physicians than did other programs, while three programs (CenVaNet, QMed, and Quality Oncology) consistently received lower ratings from their physicians than did the other programs.

In general, physician satisfaction ratings corresponded with scoring algorithm results based on discussions with program staff and physicians. For example, Carle's and Charlestown's physicians, who consistently rated their programs more highly than did the others on physician practice effects, also scored in the top quintile for improving provider practice (Carle being the top scorer in this category). Mercy's program received higher physician ratings than the other programs on perceived service arrangement and care coordination effects, consistent with its scoring in the top two quintiles for the categories of service and resource arranging, and for improving communication and coordination. Similarly, at the other end of the spectrum, QMed and Quality Oncology had overall physician satisfaction ratings across all categories that were consistently lower than the cross-program average by more than 1 standard deviation, which coincides with their scoring algorithm ratings that place them in the bottom two quintiles.

Few significant differences were observed between treatment and control group members on satisfaction with the process of care. Despite the generally favorable rating that treatment group patients and physicians gave to most of the programs' care coordination efforts, the treatment group did not consistently report higher satisfaction than control group members with indicators of the quality of the health care they received from the various providers they saw. The indicators include ratings of the degree of choice in treatment that patients feel they have, the extent to which providers keep in touch with each other, the explanations received from specialists, explanations of side effects, explanations of treatments, explanations of tests, and the quickness of receiving test results. The treatment group members were significantly more likely than the control group members to report feeling they had a choice in the treatment of their condition in only 1 of the 12 programs included in the survey (Avera). Differences favoring the treatment group occurred most often for providers keeping in touch (5 of the 12 programs). Treatment group members in four of the programs also gave more favorable ratings than the corresponding control group on explanation of treatments. Satisfaction with explanations of side effects and explanations from specialists were significantly greater for the treatment than the control group for only two and three of the programs, respectively, and with explanation of tests for only one. None of the programs had impacts on the timeliness with which test results were delivered, according to the treatment-control differences.

A few programs appeared to have more impact than others on patients' satisfaction with their overall care. Avera's and Mercy's treatment groups each gave significantly higher ratings than their control groups on three of the six measures. Three programs had significant differences on two of the six measures, and three programs had significant effects on one of the measures. The four other programs included in the survey had no discernable effect on patients' satisfaction with care.

D. HOW DO THE PROGRAMS AFFECT ADHERENCE AND QUALITY OF CARE?

The care coordination programs were expected to improve patients' adherence to recommendations and their quality of care, which, in turn, was expected to lead to improvements in patients' health and well-being. The evaluation compared the treatment and control groups' receipt of health education, knowledge and behavior about self-care, quality of care, and health status and well-being to determine whether the programs had the intended effects. Measures of preventive care and preventable hospitalizations over the year after enrollment were constructed from Medicare claims data for all first-year enrollees enrolled in 14 programs. (The measures were not appropriate for Quality Oncology, which targeted patients with cancer.) The analysis also draws on the patient survey responses to examine receipt of education, knowledge, behavior, adherence, receipt of care, and functioning. Table 3 summarizes the results.

Overall, the programs appeared to have no consistent discernible effect across numerous measures of behaviors and outcomes except receipt of health education. While there were isolated treatment-control differences for a few outcomes for a few programs, there was no pattern suggesting that the programs, as a group or individually, had true effects in any area besides receipt of health education. Favorable effects were observed for 1 or 2 measures of health status and well-being (out of the 9 examined) for 8 of the 12 programs.

TABLE 3

TREATMENT-CONTROL DIFFERENCES ON QUALITY OF CARE, AMONG FIRST YEAR ENROLLEES

Category of Outcomes (Number of Measures)	AVE	CAR	CEN	CCI	COR	GEO	HOS	HQP	JHH	MCD	MER	QMD	UMD	WSH
Health Education (5)	●●	●●●●	●●●●●		●●		●	●●●●●	○	●●	●●●●●	●		●○
Knowledge and Behavior (8)	○		●					●●	○	●		●○		●
Service Arrangement and Unmet Needs (7)				●	●●		●	●			○	○		
General and Disease-Specific Preventive Care (12)	○○	●●●●	●○				○○	●●●●						○
General and Disease-Specific Potentially Preventable Hospitalizations (8)	○		●●○			●●	●○	●		○				○○
Functional Status (9)	○	○○○		○	○				○	●	●	●		○
Health Status and Well-being (8)	●		●	○	●●●		●●	●	●	○	●	●		

Source: Treatment-control differences from patient survey and Medicare claims data. The Georgetown and University of Maryland programs did not have sufficient numbers of enrollees to be included in the patient survey, and so the survey-based measures are shaded for these two programs. The measures summarized in this table were not appropriate for the Quality Oncology program, which focused on cancer patients, and so it is not shown.

Note: ● = Treatment-control difference favoring the treatment group, significant at the 10-percent level.
○ = Treatment-control difference favoring the control group, significant at the 10-percent level.

<u>Abbreviations for MCCD Programs (Columns)</u>
AVE = Avera
CAR = Carle
CEN = CenVaNet
CCI = Charlestown
COR = CorSolutions
GEO = Georgetown University
HOS = Hospice of the Valley
HQP = Health Quality Partners
JHH = Jewish Home and Hospital
MCD = Medical Care Development
MER = Mercy
QMD = QMed
UMD = University of Maryland
WSH = Washington University

The large effects on health education did not lead to effects on self-reported knowledge, adherence, or health-related behaviors. The treatment groups in all but 1 of the 12 surveyed programs were significantly more likely than their corresponding control groups to report having received education on health behaviors. The most common effects were on receipt of education about diet and exercise, followed by the receipt of health educational materials, education on recognizing when to seek urgent care, and education on the importance of medication adherence.

Four of the programs (Carle, CenVaNet, Health Quality Partners, and Mercy) had favorable treatment-control differences across four or more of the five measures of patient education examined. Despite the treatment group members being more likely to say they had received health education, there were no effects for any of the 12 programs on patients' self-reported adherence to diet, exercise, or taking medications. Only scattered favorable effects were observed on self-reported understanding of healthy behaviors, but these were too sporadic to suggest meaningful effects for all but one or two programs. Across measures, four programs (Carle, CenVaNet, Health Quality Partners, and CorSolutions) had somewhat more favorable treatment-control differences than the other programs.

The programs had no discernable effects on service arrangements or unmet needs. While treatment group members in all 12 programs included in the survey were more likely than control group members to report receiving care coordination services (not included in Table 3), as intended, only 3 programs exhibited significant favorable treatment-control differences on other measures of unmet needs or service arrangements. Furthermore, two programs each had one outcome measure for which the *control* group had significantly better outcomes than the treatment group.

Only two programs appear to have made clear improvements in the quality of preventive care (Carle and Health Quality Partners), or to have reduced the number of preventable hospitalizations (Georgetown and Hospice of the Valley). The treatment groups were more likely than the control groups in Carle and Health Quality Partners to receive vaccination and (for women) screening mammography, and recommended blood and urine tests among beneficiaries with diabetes and coronary disease. Georgetown and Hospice of the Valley had significantly fewer "preventable" hospitalizations per beneficiary overall in their treatment groups than in the respective control groups. (Potentially preventable hospitalizations are inpatient admissions for common, acute medical conditions that, in the consensus of expert clinicians, generally should not progress to requiring inpatient care if treated in a timely fashion with adequate outpatient primary care; see Kozak et al. 2001.)

Only two programs (CorSolutions and Hospice of the Valley) had favorable effects on multiple measures of patient well-being, and these were only for selected measures. The treatment groups in those two programs were significantly more likely to report feeling their condition placed less of a burden on family than were the control groups (both programs), feeling calm and peaceful (in CorSolutions only), and having less pain (in Hospice of the Valley only). However, even these two programs had a favorable effect on only two or three of the eight measures of well-being that were examined. In addition, only three programs had a favorable treatment-control difference on any of the nine survey-based measures of functioning (for example, ability to eat independently), and, for six programs, the treatment group reported significantly *worse* health status on one or more measures. However, it is difficult to conceive of a mechanism by which programs would adversely affect patients' functioning. Furthermore, one should expect about one-half the sites to have one significant negative estimate out of the nine measures used just by chance. Finally, there is no evidence of adverse effects on other health outcomes. Thus, these scattered treatment-control differences showing worse functioning for the treatment group than the control group are interpreted as chance differences, rather than as evidence that six of the programs have caused patients' functioning to decline.

E. HOW DO THE PROGRAMS AFFECT MEDICARE SERVICE USE AND COST?

By improving patient adherence, the timeliness of response to worsening symptoms, or other aspects of the quality of care, care coordination programs are expected to reduce hospitalizations, the key factor in reducing Medicare expenditures for beneficiaries with chronic illnesses. On the one hand, the need for emergency room care and other expensive Medicare services that often follow hospitalizations (such as that provided by skilled nursing facilities and home health agencies) may also be reduced. On the other hand, some types of service use and expenditures could increase if the programs increase patients' visits to physicians for preventive care or to address symptom exacerbations. To measure these effects, the evaluation compared the treatment and control groups in each program on Medicare service utilization and expenditures. The measures were constructed for the year after enrollment for patients enrolled during the first year of program operations, and for all patients during the programs' first 25 months of operations.

Only 1 of the 15 programs (Mercy) showed a statistically significant reduction in hospitalizations, and none of the programs had significantly lower expenditures for Medicare Part A and Part B services. In eight other programs, the treatment group had fewer hospitalizations than controls during the first 25 months of program operations, but the observed differences could not be attributed with confidence to the intervention, rather than to chance. Four of these programs had 10 to 18 percent fewer hospitalizations among treatment group members than among control group members, but none of these differences were statistically significant. Furthermore, there was no difference in Medicare expenditures for two of these programs and the other two had very few enrollees. However, hospitalizations and Medicare expenditures were 14 and 21 percent higher, respectively, for the treatment than the control group in the Charlestown program, the only program for which a statistically significant difference in expenditures was observed. For the first year after enrollment, for all programs combined, the treatment group had 2 percent fewer patients admitted to the hospital, a statistically significant difference, but the differences in both the number of hospitalizations and Medicare expenditures were very small and not significantly different from zero.

The treatment group's significantly (27 percent) lower hospitalization rate in Mercy did not result in a statistically significant difference in Medicare expenditures, although expenditures were 13 percent lower for the treatment group over the 25-month period since startup. Medicare expenditures for Part A and B services were lower for the treatment group than the control group by at least 10 percent in two other programs (Georgetown and QMed), but neither difference was close to being statistically significant.

Cost neutrality cannot be rejected for some of the programs. These results suggest that none of the demonstration programs is cost neutral—that is, none has generated statistically significant evidence of savings in Medicare expenditures that could offset the fees paid to the program. However, that conclusion is less clear when one considers the large variance in the estimates. That is, while the evaluation cannot reject the hypothesis that savings in Medicare Part A and B expenditures are zero, for some programs it also cannot reject the hypothesis that savings are large enough to cover the average fee paid to the programs for care coordination. For six programs (shown in the bottom panel of Table 4), cost neutrality can be rejected statistically—net costs have increased for these programs. For the nine other programs, the evaluation cannot formally reject the hypothesis that total average Medicare expenditures per month for the treatment group, including the care coordination fee, are equal to expenditures for the control group (cost neutrality). However, for the four programs with small (less than 10 percent) treatment-control differences in hospitalization or expenditures on Part A and B services, cost neutrality seems unlikely. Failure to reject the cost neutrality hypothesis in these cases may be due to low statistical power resulting from small sample sizes and high variance of Medicare expenditures. Four other programs, however, have treatment-control differences in hospital admissions of 10 percent or greater, and (in two cases) differences in Part A and B expenditures that are large enough to essentially offset the fees. The difference in hospitalizations is smaller for a fifth program (QMed), but the fee for this program is quite low and is almost fully offset by the treatment group's 12 percent lower Medicare expenditures for traditional services. Thus, these five programs may actually be generating savings in Part A and B expenditures that are sufficient to offset the program fees. However, the estimates are too imprecise at this time for the evaluation to definitively conclude that there are such savings, or that they are large enough to cover the average fee paid for care coordination.

F. SYNTHESIZING THE FINDINGS: WHAT WORKS, AND WHAT DOESN'T?

Given that few of the programs have shown convincing evidence to date of reducing beneficiaries' need for hospitalizations and saving money or of improving the quality of care received, there is relatively little assessment that can be done yet of "what works." The one program for which there were statistically significant estimates of reductions in hospital use (Mercy Medical Center in Iowa) differed from the other programs in that it had by far the highest proportion of contacts conducted in person (two-thirds), and it excelled at Problem Identification and Care Planning, Patient Education, and Improving Communications and Coordination between patients and physicians. The program also had large impacts on patient education, as judged from the patient survey, and was rated highly by the patients' physicians.

In the evaluation's follow-up discussions with the programs, Mercy's staff attributed the reductions in hospitalizations they achieved primarily to getting patients to see their physicians quickly when symptoms worsened or problems arose. By identifying looming problems before they became severe and convincing patients of the urgency of seeing a physician (or contacting physicians directly on behalf of patients when necessary), Mercy staff felt they were able to prevent the patients' health from deteriorating to the point where a hospital admission would be necessary. They felt this preventive effect typically arose through quickly getting patients on needed medications or different dosages of their current medications.

The four other programs for which the treatment group had 10 to 20 percent fewer hospitalizations than the control group (although these differences were not statistically significant) also scored highly on one or more domains. For example, Georgetown and Health Quality Partners both scored in the top quintile on Initial Assessment. Quality Oncology scored

TABLE 4

TREATMENT-CONTROL DIFFERENCES IN HOSPITALIZATIONS AND MEDICARE EXPENDITURES, WITH AND WITHOUT PROGRAM FEES, OVER THE 1ST 25 MONTHS OF PROGRAM OPERATIONS

(Percentages)

		Monthly Medicare Expenditures	
	Annual Number of Hospitalizations	Without Care Coordination Fees	Including Care Coordination
May Be Cost Neutral			
Mercy Medical Center	***-27***	-13	8
Quality Oncology	-18	-2	0
Hospice of the Valley	-14	0	9
Georgetown University	-12	-12	1
QMed	-4	-12	1
Probably Not Cost Neutral			
CorSolutions	-5	-8	4
University of Maryland	-1	0	10
Jewish Home and Hospital	0	-6	8
Medical Care Development	1	-2	9
Not Cost Neutral			
Health Quality Partners	-10	0	***17***
Carle Foundation	-4	-1	***21***
Avera	4	-5	***14***
CenVaNet	4	6	***14***
Washington University	6	4	***12***
Charlestown	14	***21***	***44***
Overall	**-4**	**-2**	***11***

Note: Bolded italicized numbers denote statistically significant treatment-control differences at the 10-percent level for hospitalizations and Medicare expenditures without fees, and at the 20-percent level for expenditures including care coordination fees. Negative estimates imply that hospitalizations or Medicare expenditures (with or without the fee included) are lower for the treatment group, a favorable outcome. Positive estimates suggest that the treatment group used more services and cost Medicare more than the control group.

in the top quintile on four domains—Staffing, Information Technology, Ongoing Monitoring, and Quality Management.

Programs that seemed to improve preventive care (Carle and Health Quality Partners) also scored well on patient survey indicators and tended to receive high ratings on the scoring algorithm for various aspects of their interventions. Carle scored higher than all other programs on 5 of the 10 indicators. Health Quality Partners scored at the top on patient education. However, neither of these programs generated reductions in Medicare expenditures for traditional services; thus, both significantly increased net costs to Medicare. This lack of reduction in expenditures may be due in part to the fact that patients in these two programs had far lower preenrollment Medicare expenditures than all but one of the other programs.

Programs that exhibited no effects on hospitalizations, costs, or quality-of-care indicators gave a range of reasons why they were unable to reduce the need for hospitalizations. Reasons included the still-short time frame over which the analysis was conducted; the belief that some of their patients were either too debilitated or not sick enough to benefit from their interventions; and the belief that physicians in their service areas had an intractable tendency to send patients to the emergency room, which is more expensive, rather than to find time for office visits when patients exhibited worsening symptoms.

Looking across the characteristics of the five programs most likely to be cost neutral over the first 2 years of operation and the two that appear to have improved the quality of care seems to confirm the finding in Chen et al. (2000) that no single program feature or characteristic seems to be associated with a greater likelihood of program "success." Nor does the absence of a particular feature seem to doom a program to relative failure. However, how *well* programs perform their functions (based on information obtained from program staff and assessed by the evaluator) does appear to be associated with program success.

While no firm conclusions can be drawn as yet about which MCCD programs really are effective (because samples are still relatively small and the follow-up period relatively short), those programs that are most promising to date share few common structural features. Two of the programs with the most success in improving quality (Health Quality Partners and Carle) operate in rural areas, as does Mercy, the sole program with statistically significant effects on the number of hospitalizations. Yet Avera and Medical Care Development also operate in rural areas and show no such promising results to date. Two of the programs with the most favorable expenditure results (Quality Oncology and Georgetown) have fewer than 100 treatment group members—Medical Care Development is the only other program serving fewer than 300 patients. However, the results for these two programs may be due more to the imprecision of the estimates than to the excellence of the interventions. The five other relatively promising programs have substantially more patients. All four programs whose care coordinators have average caseloads of 50 or fewer patients are among the most effective programs, but the three other relatively effective programs have average caseloads in the highest range (over 75 patients). Three of the five programs operated by commercial disease management programs were among the seven promising programs, but the four other promising programs had hospitals, clinics, or academic medical centers as hosts. Other program characteristics examined seem equally unrelated to whether a program was one of the more effective seven.

How well designed programs were on various dimensions appeared to have a somewhat stronger association with performance than did structural characteristics. Strong performance in any particular domain does not appear to be necessary or sufficient for a program to be relatively successful. However, there are some clear patterns of association between how programs scored on the 10 domains examined and the programs' ability to improve quality or generate reasonably favorable expenditure comparisons. The domains most strongly associated with the promising programs are Staffing (the five programs with the highest ratings on staffing were all among the seven most promising programs), Improving Communications and Coordination (five of the six top programs on this domain were promising programs), Patient Education (four of the top five programs were promising), and Quality Management and Outcome Measurement (four of the top five programs were promising). Characteristics decidedly *not* associated with stronger quality or cost performance included Improving Provider Practice, Service and Resource Arranging, Information Technology, and (perhaps surprisingly) Ongoing Monitoring. For each of these characteristics, only one or two of the five top-rated programs were among the seven programs classified as most promising to date.

Finally, the characteristics of the patients enrolled appeared to be unrelated to the relative success of the programs to date. Three of the seven promising programs targeted patients with a single disease; the other four targeted multiple diseases. All three of the programs that enrolled patients with average preenrollment Medicare expenditures of under $600 per month were among the top seven performers, but three others of the top performers were among the six programs whose patients had average expenditures in excess of $2,000 per month. None of the other patient characteristics examined (age, education, income, race) appeared to be related to programs' likelihood of success.

The current findings suggest that hiring excellent staff and performing certain key functions well are the most important determinants of the likelihood that a program might successfully improve patient outcomes or save enough in Medicare expenditures to cover the cost of its intervention. The results to date are thus consistent with findings from Chen et al. (2000) that a few factors were common to most successful programs, including hiring well-trained, experienced nurses with at least a baccalaureate degree, but many other factors, such as having sophisticated electronic health records, were not required.

G. LONGER FOLLOWUP AND MORE OBSERVATIONS ARE NEEDED FOR DEFINITIVE FINDINGS

Due to the small sample sizes, the high variability in Medicare costs, and (for some programs) the small amount of savings required to cover the cost of the intervention, there remains uncertainty over whether nine of the programs generate savings, and if so, whether they are large enough to offset the fees. We cannot conclude with confidence that any of the programs *generate savings* in Medicare expenditures on the normal Part A and B services, because none of the estimated treatment-control differences are significantly different from zero. However, we also cannot conclude with confidence that these programs *increase net costs* to CMS. That is, there is a nontrivial possibility that these programs do generate enough savings in Medicare Part A and B expenditures to offset the modest program fees (typically 2 to 13 percent of the Part A and B expenditures), despite the fact that none of the estimates of such savings are statistically significant. The wide confidence intervals around the estimated savings in Part A and B expenditures encompass both zero (implying no effect) and the average fee paid (implying savings large enough to offset the fee). The conservative inference is that the programs were not cost neutral over the first 25 months, but there is a substantial possibility**,** given that the statistical power to detect true net savings in these nine programs ranges from only 11 percent to 77 percent, that such a conclusion is not correct for some of the programs. Furthermore, effects may yet emerge for some programs as the program and patients gain more experience and as any cumulative effects of the interventions on patient and provider behavior begin to be reflected in outcomes.

Although none of the impact estimates available at this time suggest that the demonstration programs are having large effects on patients' behaviors or outcomes, effects on Medicare service use and expenditures might be observed when the full 4 years of data on all patients become available. Physicians have been responding favorably to the programs—an important factor, given the widespread recognition that few care coordination programs are likely to succeed without significant cooperation and reinforcement from patients' physicians (Chen et al. 2000; Schore et al. 1999). Even more important, patients appear to have formed a bond with their care coordinators, and to trust their judgment.

The absence of large effects on the patient adherence measures may be somewhat discouraging for programs, but it does not necessarily imply that the programs are having no effect on patients' behavior. Relative to the control group, patients of several programs reported better access to information and appointments and better communication among their providers. Furthermore, the finding that program patients were not significantly more likely to report eating a healthy diet or exercising regularly may have a positive explanation—it is possible that, as a result of program education, the treatment group had higher standards as to what constitutes "healthy" or "regular." If that is true, their actual adherence may be better than the control group's, but the survey measures reported here may not reflect it. In addition, in many cases, behavioral change takes time; some changes do not occur until patients have experienced an adverse event that makes them recognize the value of adhering to advice from their physicians or care coordinators. Programs report that they expect it to take a few years to observe changes in patients' behavior and the effects of those behaviors on the patients' health and service use. The observed improvements in preventive care in some programs also may not result in lower hospitalizations or costs for a few years. Thus, there is reason to believe that some programs may have effects over the longer run.

The final evaluation report will assess the effectiveness of the demonstration programs by estimating program impacts on Medicare service use, expenditures, and quality of care over the first 4 years of program operations. The report will also describe the features of the program or target populations associated with effectiveness (if any). CMS has extended the end dates by 2 years to 2008, for the 11 demonstration programs that requested extensions. The four other programs will end in 2006 as originally planned. CMS granted the extensions because the Balanced Budget Act of 1997 authorizes CMS to continue any programs that are found to be cost-effective after the demonstration ends. The Act defines cost-effectiveness as either (1) reducing Medicare expenditures, or (2) not increasing Medicare expenditures while increasing the quality of services furnished and beneficiaries' and providers' satisfaction. The new end dates allow 11 of the demonstration programs to continue operating until the final evaluation findings are available. This extension allows any of the programs that the final evaluation report finds to be cost-effective to remain operating, rather than shutting down in 2006 and having to restart later.

REFERENCES

Anderson, Gerard, and Jane Horvath. "Chronic Conditions: Making the Case for Ongoing Care." Baltimore, MD: The Johns Hopkins University, December 2002.

Brown, Randall, Jennifer Schore, Nancy Archibald, Arnold Chen, Deborah Peikes, Karen Sautter, Bridget Lavin, Sherry Aliotta, and Todd Ensor. "Coordinating Care for Medicare Beneficiaries: Early Experiences of 15 Demonstration Programs, Their Patients, and Providers." Report to Congress. Princeton, NJ: Mathematica Policy Research, Inc., May 2004.

Chen, Arnold, Randall Brown, Nancy Archibald, Sherry Aliotta, and Peter Fox. "Best Practices in Coordinated Care." Princeton, NJ: Mathematica Policy Research, Inc., February 29, 2000.

Kozak, Lola Jean, Margaret J. Hall, and Maria F. Owings. "Trends in Avoidable Hospitalizations, 1980–1998." *Health Affairs*, vol. 20, no. 2, March–April 2001, pp. 225–232.

Schore, Jennifer, Randall Brown, and Valerie Cheh. "Case Management for High-Cost Medicare Beneficiaries." *Health Care Financing Review*, vol. 20, no. 4, summer 1999, pp. 87-101.

Villagra, Victor, and Tamim Ahmed. "Effectiveness of a Disease Management Program for Patients with Diabetes." *Health Affairs*, vol. 23, no.4, July–August 2004, pp. 255–266.

Wagner, E.H., B.T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi. "Improving Chronic Illness Care: Translating Evidence into Action." *Health Affairs*, vol. 20, no. 6, 2001, pp. 64–78.