15-Site Randomized Trial of Coordinated Care in Medicare FFS

Randall Brown, Ph.D., Deborah Peikes, Ph.D., Arnold Chen, M.D., M.Sc., and Jennifer Schore, M.S., M.S.W.

Medicare beneficiaries in fee-for-service (FFS) who had chronic illnesses and volunteered to participate in 15 care coordination programs were randomized to treatment or control status. Nurses provided patient education (mostly by telephone) to improve adherence and ability to communicate with physicians. Patients were contacted an average of two times per month. The findings after 2 years are not encouraging. Few programs improved patient behaviors, health, or quality of care. The treatment group had significantly fewer hospitalizations in only one program; no program reduced gross or net expenditures. However, effects may be observed when 4 years of followup are available and sample sizes increase.

INTRODUCTION

Chronic illnesses, such as heart disease and diabetes, pose a significant expense to the Medicare Program and a major detriment to beneficiaries' quality of life. 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., 2007). Furthermore, beneficiaries often have multiple chronic illnesses, which compounds the cost and complexity of their care. The 12 percent with three or more of these eight chronic health problems accounted for one-third of all Medicare spending. Coordinating the care these patients require is difficult, because Medicare beneficiaries with one or more of the eight illnesses saw an average of 17 different FFS providers per year during 2002-2005 (Chen et al., 2007), the median patient with coronary artery disease saw 10 different physicians during a year, and there is often no one physician responsible for a beneficiary's care (Pham et al., 2007). Furthermore, the care that Medicare beneficiaries receive for chronic illnesses is often uneven and of poor quality (Asch et al., 2006; Leatherman and McCarthy, 2005; Jencks, Huff, and Cuerdon, 2003).

Despite the costs and complexity of providing effective chronic care, studies have suggested that many acute health problems, and the resulting monetary and social costs, can be prevented if (1) patients are provided with medical care that is consistent with recommended standards (Institute of Medicine, 2001; Shojania et al., 2004); (2) patients adhere to recommended diet, medication, exercise, and self-care regimens (Bodenheimer et al., 2002); and (3) providers communicate better with each other and their patients (Coleman and Berenson, 2004; Stille et al., 2005). A number of small pilot programs designed to improve patients' adherence to treatment regimens and physicians' adherence to professional guidelines have improved outcomes and reduced health care utilization for patients with heart disease (Mattke, Seid, and Ma, 2007; Clark et al., 2005; McAlister et al., 2004). This potential has led many health

The authors are with Mathematica Policy Research, Inc. (MPR). The statements expressed in this article are those of the authors and do not necessarily reflect the views or policies of MPR, or the Centers for Medicare & Medicaid Services (CMS).

maintenance organizations and indemnity insurers to develop their own programs or contract with care coordination (more often called disease management [DM]) providers for such programs (Sidorov et al., 2002; Villagra and Ahmed, 2004 for evidence of the effectiveness of DM for diabetic patients in a managed care setting). However, credible evidence from large-scale studies on the effectiveness of care coordination is not yet available, and the literature shows mixed effects on health outcomes and cost (Mattke, Seid, and Ma, 2007; Gravelle et al., 2007; Smith et al., 2005; Goetzel et al., 2005; DeBusk et al., 2004; Galbreath et al., 2004; U.S. Congressional Budget Office, 2004).

The congressionally mandated Medicare Coordinated Care Demonstration (MCCD) is among the first random assignment multisite studies of care coordination. It tests specifically whether care coordination and DM can lower costs and improve patient outcomes and well being for Medicare FFS beneficiaries with chronic illnesses.

In early 2002, CMS announced the selection of 15 demonstration programs for the MCCD in a competitive awards process under which each was allowed to define, within broad boundaries, its own intervention and target population. Each program began enrolling patients between April and September 2002 and was authorized to operate for 4 years. Eleven of these programs later requested, and were granted, 2-year extensions. Beneficiaries who agreed to participate were randomly assigned by the evaluator, MPR, to either the treatment group, which received the intervention, or the control group, which did not. Both groups continued to have traditional Medicare coverage and were free to access FFS providers in the usual manner. CMS paid each program a negotiated monthly payment for care coordination of $50 to $444 per treatment group beneficiary per month, with a mean of $196.

The 15 programs differed widely in how they implemented their care coordination interventions with patients and providers.[1] All of the programs conducted assessments of patients' needs and condition and developed patient care plans. All but one of the MCCD programs provided patient education to improve adherence to medication, diet, exercise, and self-care regimens. Most of the education consisted of nurses providing factual information; a few also used behavior change models like the transtheoretical approach (Prochaska and DiClemente, 1983) or techniques like motivational interviewing (Emons and Rollnick, 2005). Almost all of the programs used standard curricula and had processes for assessing the effectiveness of the education, ranging from reviewing clinical indicators to assessing patients' self-reported behavior and responses to questions about their knowledge.

Most programs sought to improve communication between patients and providers by training patients to communicate more effectively, and sent physicians regular written reports on patients. Only four programs focused on improving provider practice, in part to minimize the burden on physicians. However, six programs did expect program participants' primary physician to participate in the care coordinators' care planning for patients, and nine programs paid the physician for telephone or in-person meetings or review of program reports. Five of the 9 programs paid the physicians a per capita fee, typically $20 to $30 per month per patient. The programs devoted relatively little attention

[1] Information on the interventions is drawn from two rounds of telephone calls at about months 3 and 36 after startup, an inperson site visit 9 months after startup, and a management information system the authors designed for the demonstration (Brown et al., 2007).

to increasing patients' access to needed support services such as home-delivered meals, transportation, or to coordinating care across providers and settings.

The intensity of interventions varied. Care coordinators' caseloads for programs ranged between 36 and 86 for 11 of the 15 programs; the other 4 had average caregiver caseloads over 100 (Table 1). Because the program was voluntary, care coordinators were able to contact virtually all patients for initial assessments (in person, for 10 of the programs) and later to monitor their well-being and progress. Most programs contacted patients 1 to 2.5 times per month on average, but three contacted patients more frequently (4 to 8 times per month). Most contacts were by telephone; however seven programs provided over one-quarter of contacts in person. The care coordinators (rather than the patients) initiated about 90 percent or more of the contacts in most programs. Three programs used home telemonitoring devices for all patients to transmit patients' weights, other clinical indicators, and symptom reports to their care coordinators daily, and another three programs used such devices for selected patients.

Study Population

Medicare beneficiaries were eligible to volunteer for the study if they were in FFS (traditional) Medicare, had one of the chronic conditions targeted by the program, and lived in the program's catchment area. Ten programs required that the beneficiary have a hospitalization for the target condition in the 12 months (or less) prior to enrollment (although lags between programs' identification of such patients and patient enrollment sometimes led to longer gaps). Each program also defined its own exclusion criteria, with a few programs excluding beneficiaries under age 65 or with end stage renal disease (ESRD), among others. By design, enrollees were not included in the research sample if (1) they were members of the same household as research sample members (to avoid contamination such applicants were automatically assigned the same intervention status as their household member, but these second members were not considered part of the research sample), (2) the programs could not provide correct Medicare health insurance claims numbers that were needed to obtain claims data (very few cases), or (3) they did not meet CMS' three demonstration-wide requirements during one or more months of the followup period (having both Parts A and B coverage, having Medicare as the primary payer, and being in FFS at the start of the followup period).

In each site, eligible applicants to the program were randomly assigned to the treatment or control group, in a 1:1 ratio, at the time they volunteered for the program and signed the patient consent form. The sequence of assignments was generated by randomly selecting 4-digit "strings" of treatment-control assignments, excluding strings with all treatments or all controls, to minimize the likelihood that runs of more than 6 consecutive treatment or control group assignments were made.[2] The sequence was generated by an MPR statistician and neither the process nor the strings were revealed to anyone. Program operators' intake staff recruited patients for the study, and submitted their identifying information through a Web site developed by MPR. The software checked cases to ensure they or a household member had not been previously enrolled, ascertained that the required information was

[2] The strings included 14 of the 16 possible sequences, (e.g., TTCC, TCCT, TCCC, etc.), excluding only TTTT and CCCC. Thus, the maximum number of consecutive controls (or treatments) was six.

Table 1

Selected Features of the Medicare Coordinated Care Demonstration Programs

Program	CC Must Be BSN or MSN Prepared	Typical CC Caseload	Percentage of Patients with Monitoring Contacts	Mean Number of Contacts per Month	Percentage of Contacts In-Person	Initial Assessment Routinely in Person	Home Telemonitor Used[1]	Education Based on Behavior Change Model[2]	Physicians Routinely Expected to Participate in Care Planning	Program Payment to Physicians
Avera	—	86	93.2	8.2	1.6	✓	✓	—	✓	$30 pppm
Carle	—	155	98.6	1.4	31.4	—	—	✓	✓	For Meetings with CCs
CenVaNet	—	75	94.7	1.4	18.1	✓	—	✓	—	—
Charlestown	—	60	99.0	2.3	31.9	—	—	—	✓	$26 pppm
CorSolutions	—	145	100.0	2.6	3.7	✓[3]	—	✓	—	For Telephone Conferences with CCs
Georgetown University	✓	36	98.0	5.9	14.1	✓	✓	—	✓	For Inperson Conferences with CCs
Health Quality Partners	—	106	99.5	2.2	41.6	✓[4]	—	✓	—	—
Hospice of the Valley	—	40	100.0	2.5	37.1	✓	—	✓	—	—
Jewish Home and Hospital	✓	66	85.3	2.5	40.2	✓	—	—	—	$28 pppm
Medical Care Development		70	86.6	1.5	29.4	✓	—	✓	—	$20 pppm
Mercy	✓	50	99.6	1.4	69.2	✓	—	— —	✓	
QMed	—	150	98.9	1.2	7.6	—	—	—	—	For Review of Program Reports
Quality Oncology	—	40	100.0	NA[5]	0.0	—	—	NA[6]	—	For Provision of Medical Records
University of Maryland	✓	71	100.0	3.9	6.5	✓	✓	NA[7]	—	$100 pppm
Washington University	—	70	98.3	1.2	4.7	—	—	✓ —	✓	

[1] QMed periodically tested its patients with an ambulatory ischemia monitor. CenVaNet, Jewish Home and Hospital, and Mercy used home telemonitors for a minority of patients.

[2] Behavior change and readiness-to-change models became more popular during the later years of the demonstration. Many of the programs with ✓'s did not initially include patient educator training in these methods, but introduced it later.

[3] CorSolutions initially contracted with local home health agencies to conduct part of the initial assessment, but discontinued this practice later in the demonstration.

[4] Health Quality Partners routinely assessed only its high-risk patients in person.

[5] Quality Oncology reported that its care coordinators were not recording all their patient contacts; therefore, this figure is not presented.

[6] Quality Oncology targeted cancer patients. Their education is shorter term and focuses on recognition of adverse treatment effects. Thus, behavior change is not relevant to program teaching.

[7] University of Maryland did not provide patient education; its intervention was the provision of home telemonitoring for patients with congestive heart failure.

NOTES: CC is care coordinator. BSN is baccalaureate degree in nursing. MSN is masters degree in nursing. pppm is per patient per month. NA is not available.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

included and met certain validity checks, and returned the random assignment result within 30 seconds after submission. In the five sites that requested it, randomization was performed separately by strata defined by a severity of illness assessment provided by the programs. After random assignment, eligible applicants were notified of their treatment or control group status, and the programs' staff began work with the treatment group only.

The mix of sociodemographic characteristics and chronic conditions of enrollees (measured over the 24 months immediately preceding their enrollment in the demonstration) varied substantially across programs. Compared with all Medicare beneficiaries, enrollees were more highly educated and had higher incomes (Brown et al., 2007), and were less likely to be under age 65, or enrolled in Medicaid (Table 2). The most common conditions the study sample had been treated for in the 2 years before enrollment were coronary artery disease (CAD) (66 percent), congestive heart failure (CHF) (54 percent), and diabetes (41 percent). The proportion originally eligible for Medicare due to disabilities or having ESRD ranged from 1 to 40 percent. Most of the programs enrolled high-cost patients: pre-enrollment Medicare expenditures averaged more than $2,000 per month during the year before enrollment for participants in seven programs, but less than $600 per month for three programs; the average for Medicare beneficiaries in FFS nationwide was $552 per month in 2003 (Centers for Medicare & Medicaid Services, 2006).

Data

Data on hospital use and expenditures were obtained from the Medicare Standard Analytic File. The Medicare National Claims History File provided data on all other services used. Patient characteristics and eligibility for Medicare were taken from the Medicare enrollment database. A patient survey conducted by MPR roughly 10 months after randomization provided data on patient behavior, health outcomes, and satisfaction with health care. The amount CMS paid to the programs for the care coordination intervention for any given treatment group patient was obtained from Medicare claims files with special G-codes designated for the demonstration.

Followup Period

To measure the effects on hospitalizations, Medicare expenditures, and quality of care, we compared outcomes for the treatment and control groups in each program. Outcome measures were constructed for two time periods for which the samples overlapped but differed. Treatment-control differences in quality-of-care measures were estimated by comparing outcomes during the 12 months following the month of random assignment for all beneficiaries randomized during the program's first year of operations. Effects on hospital use and total expenditures per eligible month were estimated over the first 25 calendar months of program operations, using all sample members who were enrolled in the program through the first 25 months, and calculated over all eligible patient-months in that time period.

Sample size for the 1-year followup exceeded 1,000 for four programs, but was less than 120 for three programs. Only six programs had at least 600 sample members, the minimum needed to have 80 percent power to detect effects of 20 percent or more on number of hospitalizations or on binary survey or claims variables with a mean of 0.50. For the 25-month analysis, sample sizes were substantially larger with 11 programs having at least

Table 2

Baseline Characteristics of the Medicare Coordinated Care Demonstration Randomized Through Month 25

Program	Number	Age		Race	Diagnosis[1]						Medicare Buy-In	ESRD or Originally Eligible Due to Disability	Medical Use	
		≤64	85+	White, Non-Hispanic	CAD	CHF	Stroke	Diabetes	Cancer	Dementia			Number of Hospitalizations	Monthly Expenditures
							Percent							
Avera														
Treatment	292	0.0	21.6	98.6	82.5	99.0	24.3	46.2	23.6	4.5	7.9	11.6	2.18	$1,697
Control	291	0.0	18.9	97.9	76.3	98.6	24.7	39.2	25.4	3.8	8.9	12.4	2.32	1,662
Difference	—	0.0	2.7	0.7	6.2*	0.3	-0.4	7.1*	-1.8	0.7	-1.1	-0.7	-0.14	35.2
Carle														
Treatment	1,178	1.6	11.1	97.5	54.8	32.1	15.9	41.8	23.3	6.0	5.3	8.7	0.48	559
Control	1,161	1.5	12.2	96.0	50.7	26.5	15.6	40.6	19.6	5.8	5.1	9.0	0.46	537
Difference	—	0.2	-1.1	1.4*	4.2**	5.6***	0.3	1.2	3.7**	0.3	0.2	-0.2	0.02	23
CenVaNet														
Treatment	616	0.0	12.7	84.3	73.5	52.4	27.9	55.0	28.9	5.5	5.4	8.1	0.78	946
Control	611	0.0	12.1	83.1	70.7	49.1	29.3	56.0	27.7	5.7	4.9	8.7	0.71	823
Difference	—	0.0	0.5	1.1	2.8	3.3	-1.4	-0.9	1.2	-0.2	0.5	-0.6	0.07	123
Charlestown														
Treatment	370	0.0	45.4	98.9	77.0	46.8	35.7	26.5	33.2	12.4	0.0	1.4	0.82	1,057
Control	369	0.0	42.8	98.1	58.0	42.6	36.0	27.6	33.3	8.1	0.0	3.8	0.84	1,103
Difference	—	0.0	2.6	0.8	19.0***	4.2	-0.4	-1.2	-0.1	4.3*	0.0	-2.4**	-0.02	-46
CorSolutions														
Treatment	1,159	14.2	12.4	66.1	87.1	99.3	42.4	55.1	18.8	12.2	28.0	25.6	2.35	2,779
Control	869	14.3	12.7	63.9	85.9	97.2	43.0	56.9	18.9	14.5	27.9	27.0	2.57	2,943
Difference	—	0.0	-0.2	2.2	1.2	2.1***	-0.7	-1.8	-0.1	-2.3	0.1	-1.4	-0.21**	-164
Georgetown														
Treatment	95	0.0	13.7	35.8	84.2	100.0	32.6	56.8	24.2	17.9	17.9	12.6	2.06	2,265
Control	95	0.0	15.8	42.1	82.1	97.9	35.8	51.6	28.4	13.7	24.2	13.7	3.22	3,164
Difference	—	0.0	-2.1	-6.3	2.1	2.1	-3.2	5.3	-4.2	4.2	-6.3	-1.1	-1.17***	-899**
Health Quality Partners														
Treatment	499	0.0	8.0	99.2	40.9	14.6	19.2	24.3	24.5	1.6	1.4	4.4	0.32	495
Control	493	0.0	7.3	98.4	36.9	11.8	15.0	23.5	25.2	1.8	1.4	2.0	0.36	502
Difference	—	0.0	0.7	0.8	4.0	2.9	4.2*	0.7	-0.7	-0.2	0.0	2.4**	-0.04	-7
Hospice of the Valley														
Treatment	370	0.0	27.6	96.0	63.8	60.0	40.0	33.8	30.8	26.0	16.0	10.3	1.81	2,286
Control	358	0.0	22.9	96.4	65.4	56.7	38.3	29.9	32.7	23.7	17.3	13.7	1.80	2,126
Difference	—	0.0	4.7	-0.4	-1.6	3.3	1.7	3.9	-1.9	2.2	-1.4	-3.4	0.00	161

Refer to footnotes at the end of the table.

Table 2—Continued

Baseline Characteristics of the Medicare Coordinated Care Demonstration Randomized Through Month 25

Program	Number	Age		Race	Diagnosis[1]						Medicare Buy-In	ESRD or Originally Eligible Due to Disability	Medical Use	
		≤64	85+	White, Non-Hispanic	CAD	CHF	Stroke	Diabetes	Cancer	Dementia			Number of Hospitalizations	Monthly Expenditures
							Percent							
Jewish Home and Hospital														
Treatment	352	0.3	36.1	57.4	55.7	45.5	33.8	38.9	29.6	37.8	29.8	11.4	0.83	$1,542
Control	347	0.0	36.6	57.6	47.3	30.0	26.2	34.6	29.4	35.7	27.1	7.2	0.80	1,378
Difference	—	0.3	-0.5	-0.3	8.4**	15.5***	7.6**	4.3	0.2	2.1	2.7	4.2*	0.03	164
Medical Care Development														
Treatment	411	6.8	10.7	99.0	91.2	70.6	22.6	47.2	19.2	3.9	18.7	17.5	2.04	2,014
Control	407	5.4	11.6	99.3	91.9	68.6	22.4	47.2	23.3	4.2	22.1	18.9	2.08	2,066
Difference	—	1.4	-0.8	-0.2	-0.7	2.0	0.3	0.0	-4.1	-0.3	-3.4	-1.4	-0.04	-53
Mercy														
Treatment	420	4.1	16.7	99.5	67.6	65.0	31.4	32.4	25.2	8.6	11.4	18.3	1.36	1,365
Control	422	3.8	18.0	99.8	69.0	63.5	31.0	37.0	26.8	8.8	11.4	15.9	1.40	1,335
Difference	—	0.3	-1.3	-0.2	-1.3	1.5	0.4	-4.6	-1.5	-0.2	0.1	2.5	-0.05	29
QMed														
Treatment	651	6.6	4.6	88.9	46.7	16.7	14.9	26.1	21.7	1.2	11.2	18.0	0.30	565
Control	642	6.7	5.8	90.7	45.6	18.2	16.7	26.2	19.8	2.0	10.6	13.7	0.30	528
Difference	—	-0.1	-1.2	-1.7	1.1	-1.5	-1.9	-0.1	1.9	-0.8	0.6	4.3**	0.00	37
Quality Oncology														
Treatment	65	6.2	13.9	84.6	49.2	16.9	20.0	23.1	100.0	7.7	9.2	15.4	1.02	2,894
Control	63	9.5	9.5	85.7	44.4	17.5	14.3	33.3	95.2	4.8	14.3	17.5	0.99	2,686
Difference	—	-3.4	4.3	-1.1	4.8	-0.5	5.7	-10.3	4.8*	2.9	-5.1	-2.1	0.03	208
University of Maryland														
Treatment	66	12.1	6.1	56.1	83.3	98.5	31.8	47.0	12.1	4.6	18.2	24.2	2.67	3,080
Control	59	15.3	6.8	61.0	74.6	89.8	25.4	40.7	11.9	11.9	11.9	30.5	2.34	3,286
Difference	—	-3.1	-0.7	-5.0	8.8	8.7**	6.4	6.3	0.3	-7.3	6.3	-6.3	0.33	-206
Washington University														
Treatment	968	27.0	10.3	61.0	58.6	44.8	28.0	43.4	40.1	13.5	20.8	40.4	1.79	2,251
Control	964	27.6	8.4	63.3	57.9	43.1	26.0	46.4	36.6	12.0	19.4	42.5	1.83	2,262
Difference	—	-0.6	1.9	-2.3	0.7	1.8	2.0	-3.0	3.5	1.5	1.4	-2.1	-0.04	-12

Refer to footnotes at the end of the table.

Table 2—Continued

Baseline Characteristics of the Medicare Coordinated Care Demonstration Randomized Through Month 25

Program	Number	Age		Race	Diagnosis[1]						Medicare Buy-In	ESRD or Originally Eligible Due to Disability	Medical Use	
		≤64	85+	White, Non-Hispanic	CAD	CHF	Stroke	Diabetes	Cancer	Dementia			Number of Hospitalizations	Monthly Expenditures
							Percent							
All Programs														
Treatment	7,512	7.3	15.0	83.3	66.1	53.9	27.6	41.3	26.8	10.1	13.9	17.1	1.29	$1,544
Control	7,151	7.0	14.7	83.9	62.2	48.9	26.3	40.9	26.3	9.7	13.3	16.9	1.30	1,497
Difference	—	0.2	0.3	-0.6	3.9***	4.9***	1.3*	0.4	0.5	0.4	0.6	0.2	-0.01	47
Medicare Overall	42.3mn	14.4	11.1	84.6	40.2[2]	40.2[2]	NA	12.0	16.9[3]	5.0[4]	18.0	15.2	NA	552

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

[1] Medical conditions treated during the 2 years before randomization, as reported in Medicare claims data.

[2] Data available only for Medicare beneficiaries living in the community, with heart disease, which includes both CAD and CHF; included for comparison purposes only.

[3] Excludes skin cancer.

[4] Includes only beneficiaries with Alzheimer's disease.

NOTES: CAD is coronary artery disease. CHF is congestive heart failure. ESRD is end-stage renal disease. NA is not available.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

600 cases. However, only three of the programs had 80 percent power to detect impacts on expenditures of 20 percent or larger (requiring roughly 1,400 sample members, 700 in each group), given the substantially greater coefficient of variation for expenditures (1.5) than for hospitalizations (1.0).

Medicare expenditures and service use are measured only over those months when the sample member met (for at least one day of the month) the basic eligibility requirements for the demonstration. The evaluation began measuring Medicare expenditures and service use in the first full month after random assignment. Observations are weighted to reflect the number of months the patient was eligible for the study over the time period examined.

STATISTICAL ANALYSIS

An intent-to-treat design was used. All beneficiaries who were randomly assigned were included in the analyses. The nature and intensity of intervention received varied substantially across programs, and across sample members within any program, depending on their interest and assessed needs.

Regression models were used to estimate impacts on hospitalizations and costs. The regressions controlled for age; sex; whether the beneficiary had been treated for CHF during the 2 years before randomization (in programs that did not exclusively target CHF); the number of the following conditions the patient had been treated for during the 2 years before randomization: CAD, CHF, stroke, diabetes, cancer, chronic obstructive pulmonary disease, dementia, peripheral vascular disease, ESRD, depression, and asthma; the annualized number of hospital admissions in the previous year; and total Medicare Parts A and B expenditures per month in the prior year.[3] The survey data were analyzed by comparing the unadjusted means of the treatment and control groups.

Only main effects were estimated at the site level, as sample sizes were not adequate for analysis of subgroup effects. All of the analyses conducted were prespecified in a research design report prepared for the study (available at http://www.mathematica-mpr.com/publications/pdfs/researchdesign.pdf). To address the problem of multiple test bias, given the large number of outcome measures examined for quality of care, we grouped outcomes by domain and did not attribute treatment-control differences in any domain to the effects of the program unless the number of statistically significant findings in that domain was markedly greater than what might be expected to occur by chance.

RESULTS

Patient Knowledge and Behavior

Despite a heavy focus on patient education, only five programs had significant treatment-control differences on any of the eight knowledge or behavior measures examined (Table 2). Only one program had significant favorable differences for two of the measures (exercising regularly, and trying to cut down on drinking). For some measures, this was due in part to the high adherence rate among the control group leaving little room for improvement (e.g., 90 percent for adherence to medications [Brown et al., 2007]).

[3] Various other specifications, including log transformations of expenditures, were also examined; none led to substantively different conclusions. CHF was explicitly controlled for because it was the chronic condition most often targeted, and because costs are substantially higher for patients with CHF than for most other chronic conditions.

Table 3

Effects of the Medicare Coordinated Care Demonstration on Patient's Quality of Care During First Year After Enrollment

Outcome	Number of Programs: With Data[1]	With Moderate Improvements[2]	With Large Improvements[2]	Programs with Impacts
Knowledge and Behavior[3]				
Understands Diet	12	1	0	CenVaNet
Follows Healthful Diet	12	0	1	Washington University
Understands Exercise	12	1	0	Medical Care Development
Exercises Regularly	12	1	0	Health Quality Partners (HQP)
Misses Doses of Medication	12	0	0	—
Visits Physician with List of Questions	12	0	0	—
Tried to Quit Smoking (Smokers Only)	12	0	1	QMed
Tried to Cut Down on Drinking (Drinkers Only)	12	0	1	HQP
Preventative Care				
All Patients				
Flu Vaccine[3]	12	1	0	CenVaNet
Pneumonia Vaccine[3]	12	1	0	HQP
Colon Cancer Screening[4, 5]	14	0	0	—
Screening Mammography[4,7]	14	0	1	HQP
Diabetes Patients[4]				
Diabetes Education	14	0	0	—
Eye Examination	14	0	0	—
Cholesterol or Lipid Test	14	1	1	Carle, HQP
Hemoglobin A1c Test	14	1	0	Carle
Urine Test for Protein	14	0	1	Carle
Congestive Heart Failure (CHF) Patients[4,7]				
LV Function Test	14	0	0	—
Coronary Artery Disease (CAD) Patients[5]				
Cholesterol or Lipid Test	14	1	1	Carle, HQP
Preventable Hospitalizations[4]	14	2	1	CenVaNet, Georgetown, Hospice of the Valley
Health Status and Well-Being[3]				
Emotional Distress	12	3	0	CorSolutions, HQP, Mercy
Depression	12	0	0	—
Poor Sleep	12	2	0	Avera, Jewish Home and Hospital
Pain	12	2	0	Hospice of the Valley, QMed
Effect of Primary Condition on Life	12	1	0	CenVaNet
Physical Health Summary Score	12	1	0	CorSolutions
Mental Health Summary Score	12	0	0	—

[1] Measures for which 12 sites have data were obtained from the patient survey. The claims-based measures excluded quality oncology because the program's focus on beneficiaries with cancer makes measures of general preventative care and preventative care for diabetes, CHF, and CAD irrelevant for the program.

[2] Moderate=a statistically significant treatment-control difference (p<= 0.10) that favors the treatment group and is less than 10 percentage points and less than one-half the control group proportion (pc) or its complement (1-pc). Large=a statistically significant treatment-control difference (p<= 0.10) that favors the treatment group and is more than 10 percentage points or at least one-half the control group proportion [pc] or its complement (1-pc).

[3] Sample sizes for the survey ranged from 395 to 684 per site across the 12 sites in which surveys were conducted. The survey sample was evenly split between treatment and control groups; response rates (from 84.9 to 97.6 percent) were similar for the treatment and control groups in each site.

[4] Sample sizes for the claims-based measures ranged from 55 to 2,042 per site across 14 sites.

[5] Colon cancer screening is fecal occult blood testing, screening colonoscopy, sigmoidoscopy, or barium enema.

[6] Screening mammography is only assessed for females.

[7] Enrollees were defined as having diabetes, CHF, or CAD if they had a Medicare claim with such a diagnosis in the 2 years prior to enrollment; diagnosis categories are not mutually exclusive.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

Quality of Preventive Care

Only 2 of the 15 programs showed compelling evidence of effects on quality of preventive care indicators during enrollees' first year in the program (Table 3). Carle Clinic and Health Quality Partners each appear to have improved 4 of the 11 measures of general and disease-specific preventive care. Carle had moderate to large treatment-control differences in testing for cholesterol, hemoglobin A1C, and urine protein in beneficiaries with diabetes, and testing for cholesterol in beneficiaries with CAD. Health Quality Partners' treatment group had significantly higher rates of pneumonia vaccination, screening mammography, and cholesterol testing in both diabetes and CAD patients. The treatment groups in CenVaNet, Georgetown, and Hospice of the Valley had significantly lower rates of potentially preventable hospitalizations.

Health Status

Eight of the programs each had one or two statistically significant differences favoring the treatment group, among the seven outcome measures related to patients' health status and quality of life (Figure 1). All of these differences were modest in size. None of the programs had statistically significant treatment-control differences in mortality (Brown et al., 2007).

Looking across the various indicators of quality of care, we see little evidence that the programs individually or as a group had marked effects (Figure 1). Only Health Quality Partners had consistently favorable effects on substantially more quality indicators than would be expected by chance (7 of the 27 measures).

Medicare Service Use

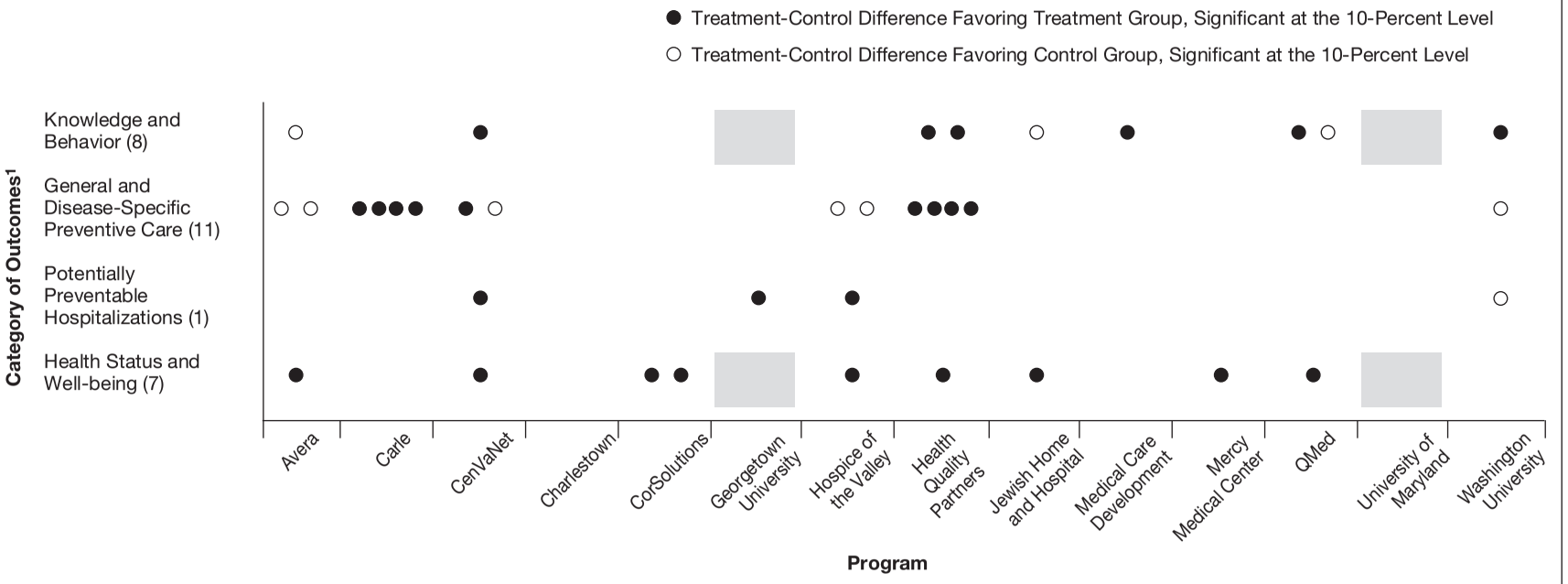
Overall, combining the 15 programs (Table 4), the treatment group experienced 4 percent fewer hospitalizations than the control group during the first 25 months of operations, but the modest difference was not statistically significant ($p = 0.145$). The difference was statistically significant for only 1 of the 15 programs, Mercy, where the average number of hospitalizations for the treatment group was 27 percent lower than that for the control group ($p = 0.003$).

Medicare Expenditures

Looking at the 15 programs combined, there was no effect on monthly Medicare expenditures over the 25-month period, even before considering the care coordination fees (Table 5). Mercy's treatment group's 27 percent fewer hospitalizations resulted in 13 percent ($154) lower monthly Medicare expenditures relative to the control group over the first 25 calendar months, and the p-value (0.105) was just above the 10-percent significance level for a two-tailed test. The difference, however, is not enough to offset Mercy's average effective care coordination fees of $245 per month over this time period. Some other programs had lower expenditures for the treatment than control group, but none of these were statistically significant. One program, Charlestown, had average monthly Medicare expenditures that were 21 percent ($212) higher for the treatment group. Analyses conducted using the logarithm of expenditures as the dependent variable (to account for the right-skewed distribution of costs per month) improved the statistical precision, making both of these sites' estimates significantly different from zero at the 0.01 level. None of the other programs' estimated effects were

Figure 1

Treatment-Control Differences on Quality of Care Among First Year Enrollees in the Medicare Coordinated Care Demonstration



[1] Number of measures is shown in parentheses.

NOTES: Sample sizes for the survey ranged from 395 to 684 per site across the 12 sites in which they were conducted (the 3 sites with fewer than 125 enrollees were excluded from the survey). The survey sample was evenly split between treatment and control groups, and response rates (ranging from 84.9 to 97.6) were very similar for both groups in each site. Sample sizes for the claims-based measures ranged from 55 to 2,042 per site across 14 sites (Quality oncology was not included because of its focus on cancer and the inapplicability of most of the measures in this figure). The Georgetown and University of Maryland programs did not have sufficient numbers of enrollees to be included in the patient survey, so the survey-based measures are shaded for these two programs.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

Table 4

Enrollees' Average Annualized Number of Hospital Admissions Per Year Through First 25 Months of Program Operations

	Sample Size		Average Annualized Number of Hospital Admissions				
Program	Treatment Group	Control Group	Treatment Group	Control Group	Treatment-Control Difference	Percent Difference	*p*-Value
Avera	292	291	1.51	1.45	0.06	4	0.728
Carle	1,178	1,161	0.52	0.54	-0.02	-4	0.538
CenVaNet	616	611	0.74	0.70	0.03	4	0.636
Charlestown	370	369	0.79	0.69	0.09	14	0.236
CorSolutions	1,159	869	1.80	1.89	-0.09	-5	0.395
Georgetown	95	95	1.64	1.86	-0.22	-12	0.487
Health Quality Partners	499	493	0.37	0.41	-0.04	-10	0.505
Hospice of the Valley	370	358	1.25	1.46	-0.21	-14	0.127
Jewish Home and Hospital	352	347	0.88	0.88	0.00	0	0.992
Medical Care Development	411	407	1.39	1.38	0.01	1	0.959
Mercy	420	422	0.73	1.01	-0.27	-27	0.003
QMed	651	642	0.37	0.39	-0.02	-4	0.740
Quality Oncology	65	63	1.18	1.43	-0.25	-18	0.510
University of Maryland	66	59	2.33	2.36	-0.03	-1	0.950
Washington University	968	964	1.42	1.34	0.08	6	0.381
Overall	7,512	7,151	0.91	0.95	-0.04	-4	0.145

NOTES: Regression adjusted. Observations are weighted by the number of months in the followup period that the sample member meets Centers for Medicare & Medicaid Services' eligibility requirements: being in fee-for-service, having both Parts A and B coverage, and having Medicare as the primary payer.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

Table 5

Enrollees' Average Monthly Medicare Parts A and B Expenditures Without Care Coordination Fees Through First 25 Months of Program Operations

	Sample Size		Average Monthly Medicare Part A and B Expenditures Without Care Coordinators Fees				
Program	Treatment Group	Control Group	Treatment Group	Control Group	Treatment-Control Difference	Percent Difference	*p*-Value
Avera	292	291	$1,401	$1,470	-70	-5	0.641
Carle	1,178	1,161	691	699	-7	-1	0.861
CenVaNet	616	611	895	847	48	6	0.477
Charlestown	370	369	1,216	1,004	212	21	0.058
CorSolutions	1,159	869	2,494	2,700	-206	-8	0.229
Georgetown	95	95	2,082	2,358	-276	-12	0.534
Health Quality Partners	499	493	609	608	1	0	0.989
Hospice of the Valley	370	358	2,058	2,061	-2	0	0.990
Jewish Home and Hospital	352	347	1,707	1,815	-108	-6	0.606
Medical Care Development	411	407	1,531	1,569	-39	-2	0.820
Mercy	420	422	1,039	1,193	-154	-13	0.105
QMed	651	642	606	686	-80	-12	0.349
Quality Oncology	65	63	4,178	4,280	-101	-2	0.882
University of Maryland	66	59	3,178	3,178	0	0	1.000
Washington University	968	964	1,962	1,893	68	4	0.558
Overall	7,512	7,151	$1,283	$1,314	-31	-2	0.368

NOTES: Regression adjusted. Observations are weighted by the number of months in the followup period that the sample member meets Centers for Medicare & Medicaid Services' eligibility requirements: being in fee-for-service, having both Parts A and B coverage, and having Medicare as the primary payer.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

significantly different from zero at the 0.05 level.

Cost Neutrality

The evaluation also assessed whether the programs were cost neutral; that is, whether the costs of delivering care coordination were covered by reductions in traditional Medicare expenditures (Table 6). Overall, total costs, including the care coordination fees, increased by 11 percent ($p<0.001$). Six of the programs had costs that were significantly higher for the treatment group. Despite the absence of statistically significant treatment-control reductions in Medicare expenditures for traditional services, it is possible that some of the remaining nine 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 these 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 amount of savings needed to cover the average fee paid to the programs for providing care coordination. To draw inferences about these nine programs, we examined the patterns of differences in hospitalizations, traditional Medicare expenditures, and total Medicare expenditures including the care coordination fees.

Four programs are probably not cost neutral, because they did not reduce hospitalizations, which account for the largest share of costs. The other five programs (Table 6) could conceivably be cost neutral over their first 25 months of operations. All but QMed had relatively large treatment-control differences in hospitalizations of between 12 and 27 percent, and in QMed's case, the modest (4 percent) difference may be enough to cover their low care coordination fees. Thus, these five programs may actually be generating savings in traditional expenditures that are sufficient to offset the program fees, even though two of the programs have larger estimated losses than the programs classified as probably not cost neutral. However, the estimates are too imprecise at this time for the evaluation to conclude that there are such savings, or that any such savings are large enough to cover the average fee paid for care coordination.

CONCLUSIONS

Over the first 2 years of program operations, most of the demonstration programs did not achieve their objectives of improving care and reducing hospitalizations and costs. While the available sample size at this stage did not provide sufficient power to detect modest size effects on costs for most programs, only five of the programs could possibly be viewed as cost neutral to date, and none showed evidence of actual cost savings. The lack of effects on hospitalizations (for which smaller effects were detectable due to the smaller variance) together with the absence of effects on patient self-care and adherence, despite high engagement rates, reinforces the conclusion that only a few of the programs could have been cost neutral. Even though 10 of the 15 programs had negative treatment control differences in Part A and B expenditures, the differences for 7 of the 10 programs was 8 percent or less of the control group mean and only one program (with very few cases) had estimated Part A and B savings large enough to offset the program fees. Thus, the findings are not encouraging overall, despite the programs

Table 6

Cost Neutrality Through Month 25 of Program Operation for Enrollees During the First 25 Months

Program	Average Care Coordination Fee per Month of Followup[1]	Treatment-Control Differences in Medicare Expenditures per Month, Including Care Coordination Fee				
			80 Percent Confidence Interval			
		Difference	Lower Bound	Upper Bound	Percent of Control Mean	*p*-Value
			Not Cost Neutral			
Charlestown	$233	$445	$303	$588	44.4	0.000
Washington University	166	234	84	383	12.3	0.045
			Probably Not Cost Neutral			
University of Maryland	321	321	-692	1,334	10.1	0.685
Avera	271	201	11	391	13.7	0.175
Carle	152	145	90	199	20.7	0.001
CenVaNet	72	120	34	207	14.2	0.076
Health Quality Partners	105	106	5	206	17.4	0.179
			Possibly Cost Neutral			
Hospice of the Valley	190	188	-35	412	9.1	0.280
Jewish Home and Hospital	260	152	-116	420	8.4	0.468
Medical Care Development	180	141	-76	359	9.0	0.406
CorSolutions	315	109	-110	328	4.0	0.525
Mercy	250	96	-26	217	8.0	0.312
Georgetown	296	20	-546	587	0.9	0.963
QMed	88	8	-102	118	1.2	0.924
Quality Oncology	81	-20	-894	854	-0.5	0.976
Overall	196	144	99	188	11.3	0.000

[1] The amount paid to a program as recorded in the Medicare claims data differs from the program's approved per member per month fee for active patients because some patients disenrolled from the programs, but were retained in the research sample.

NOTE: Estimates obtained from weighted least squares regression, with weights equal to number of months observed.

SOURCE: Brown, R., Peikes, D., Chen, A., and Schore, J., Mathematica Policy Research, Inc., 2008.

having prior experience operating care coordination programs in other settings that they considered successful.

With 15 programs targeting diverse conditions and patients and employing different interventions, it is difficult to ascertain the intervention features responsible for the few programs with favorable impacts observed. It may be that the clinical integration of the physicians in the Carle and Health Quality Partners programs played a role in their improvements in process measures of quality. Closer monitoring of patients' status may have helped lower rates of potentially preventable hospitalizations in the CenVaNet, Georgetown, and Hospice of the Valley programs. Of note, the Mercy program (the only one with statistically significant reductions in total hospital use) had by far the highest proportion among all programs of contacts conducted in person (two-thirds), and appeared to excel at identifying problems and planning care, delivering patient education, and improving communication and coordination among patients and physicians (Brown et al., 2007). While Mercy's hospital impact was the only statistically significant estimate among the 15 programs and might therefore be due to chance, the large magnitude and low p-value (0.003) argue for this being a true impact. Five of the programs had treatment group hospitalization rates 10 percent or more below control group rates, but only one program had a treatment group rate exceeding the control group rate by more than 6 percent. This pattern suggests that a subset of programs may have truly reduced hospitalizations, even though there may not be enough precision for each of the individual estimates to be statistically significant.

Our general negative findings are consistent with results from the Medicare Health Support (MHS) program. In MHS, designed to be a population-based version of DM programs in FFS Medicare, commercial DM/coordinated care providers guaranteed savings for all (10,000 or more) Medicare patients with certain severe chronic illnesses in large health care markets (available at: http://www.cms.hhs.gov/CCIP/downloads/EOP_Fact_Sheet_FINAL_012808.pdf).[4] Four of the original nine programs dropped out, and CMS recently announced that none of the remaining five were generating savings in Medicare expenditures large enough to offset program fees. According to CMS, the programs will need to achieve unrealistically large gross savings of 20 to 40 percent in their final year of operations to break even.

The decision by CMS to use a randomized design to properly assess these programs, regardless of how promising they appeared to be in the early 2000s, ensures that the estimates provided here do not suffer from biases inherent in less rigorous approaches to estimating program impacts. A simple pre-post analysis of expenditures for the treatment group—often the research design behind results cited by DM vendors to potential clients—showed large drops in expenditures for the year after enrollment relative to the year before for 10 of the 15 programs. The results from the randomized design shows that these declines are not due to program effects, but rather reflect regression toward the mean. This study also benefits from having good data on the costs of health care (and not just health care utilization) and the costs of providing the interventions, which are essential for the cost benefit analysis; many previous studies have lacked such data (Mattke, Seid, and Ma, 2007). Another strength of the study is that the evaluation collected detailed qualitative and quantitative data documenting that patients had received the intervention, information

[4] A formal report on the findings is not publicly available yet.

again often missing from previous studies. Such data is necessary to avoid making what has been called "Type III errors"—either incorrectly ascribing a lack of impacts to inadequate implementation rather than to deficiencies in the intervention itself, or vice-versa (Carroll et al., 2007; Oakley et al., 2006). Also, selected programs were all required to have prior experience delivering such interventions and at least some evidence of effectiveness, which addresses the common concern that new programs should not be expected to have impacts until they are established. Finally, the followup observed is longer than in many prior studies.

Despite these strengths the study has several limitations, two that make these mid-program findings more ambiguous than we would like and two that limit the policy inferences that can be drawn from the study due to its design. The two limitations that contribute to the uncertainty about the findings are the relatively short followup period, and the modest sample sizes. The followup period, while longer than in most studies, is still relatively short. Our results are limited to an average of just over 14 months of followup, so findings may differ when we examine the full 4-year demonstration period, covering a longer period of exposure and a more mature stage of operations.

The second factor that makes the results ambiguous is that the study is underpowered at this point to detect effects on costs unless they are quite large. Demonstration programs were expected to enroll a minimum of 678 beneficiaries in their first year, a sample size that would be adequate to detect effects of 20 percent on number of hospitalization or on binary survey outcomes with a mean of 0.50 (that is a detectable difference of 10 percent points), assuming a 90-percent response rate. While larger sample sizes would clearly have been preferable, most of the programs were unable to enroll even these modest numbers during their first year. Furthermore, several published studies showed other coordinated care programs with impacts substantially larger than 20 percent. In addition, even if cost impacts of 20 percent can not be detected, the minimum sample size is adequate to detect 20 percent reductions in hospitalizations, well below the rates reported in some programs (Rich et al., 1995; Naylor et al., 1999; Lorig et al., 1999; Chen et al., 2000). As enrollment continues over the next 2 years, sample sizes will continue to grow, leading to greater precision for final results.[5]

The study limitations that affect our ability to draw broader inferences about care coordination and DM are that (1) the demonstration programs did not appear to implement some proven care coordination interventions (and therefore may not be a good test of the true potential of care coordination), and (2) the small size of the programs provides no indication of whether the more effective programs still could be effective at a much larger scale. Naylor et al. (1999) and Rich et al. (1995) have shown in small, single-site randomized trials that an aggressive but time-limited intervention for patients transitioning from hospital to home (a "teachable moment" when patients might be especially receptive to behavior change) can significantly reduce the likelihood of readmission at low cost. Only 2 of the 15 programs in the MCCD program tried to recruit hospitalized patients prior to discharge, and neither implemented a limited term, discharge transition component. Studies such as those by Naylor and Rich also suggest that programs that are not heavily reliant

[5] While the survey inquiries may have led control group members to improve self-care behavior and outcomes, this is highly unlikely, given the difficulty of getting patients to change their behavior.

on the involvement of patients' physicians, like many of the ones tested in the MCCD, do have the potential for effectiveness despite the concerns of some authors (Bodenheimer et al., 2002; Geyman, 2007). Similarly, although randomized trials have shown that fostering patients' self-efficacy through peer-led group sessions can reduce hospitalizations and costs (Lorig et al., 1999; Wheeler, 2003), none of the programs incorporated such features. In some cases, however, programs did base their telephonic interventions on other behavior change models with evidence of effectiveness (Prochaska and DiClemente, 1983).

Finally, this study does not offer guidance about the scalability or optimal design of coordinated care programs. The MHS program was designed to provide an easier-to-administer program, in which a small number of entities would take financial risk for large numbers of chronically ill beneficiaries. On the other hand, the current interest in medical homes designates physician practices as the place where care coordination should occur for beneficiaries—a model and size more consistent with this demonstration. An intermediate model is also being tested under the Care Management for High Cost Beneficiaries Demonstration. While the MCCD study does not shed light on the relative merits of these three designs, if it develops that some of the care coordination sites are cost neutral, and (as appears likely) none of the MHS sites are even close to cost neutrality, it would appear that moderate size units are more likely to be effective than large scale, externally based programs. At this point, however, the evidence for the effectiveness of medical homes and high cost case management is even more limited than the evidence on the MCCD programs.

While some may argue that the interventions may be more effective in commercial or Medicaid populations, or in a Medicare managed care context, we suspect that these settings would engender the same difficulties as encountered in the demonstration. For example, it may be true that younger individuals are more amenable than elderly Medicare beneficiaries to behavior modification, and that Medicaid beneficiaries present a greater opportunity for savings because of high rates of inappropriate and fragmented care. It may also be the case that managed care plan members can benefit from plans' stronger leverage over provider behavior and greater access to timely data on use of services and medications. Nonetheless, the challenges in effecting substantial and lasting changes in patient behavior (for example, weight loss, smoking cessation) and provider behavior and the results presented here suggest that claims of program effectiveness in other populations need to be rigorously tested in randomized studies.

Our generally negative findings, together with those from other recent CMS demonstration and pilot experiences, suggests that DM and care coordination programs may not be the panacea that many payors have hoped for and many vendors proclaimed. Additional research remains to be done, both in this study and in future studies. A few of the MCCD programs show promise of achieving cost neutrality, suggesting that further study of program features is necessary to develop an evidence base for what seems to work best for different types of patients and settings, and what features should qualify a program for Medicare reimbursement if evidence of cost savings is demonstrated over the longer followup period. More definitive results will come from data on the full 4 years of program operations.

ACKNOWLEDGMENTS

The authors thank Carol Magee, Cynthia Mason, and Renee Mentnech for their insightful comments on our report from which the findings in this article were drawn. We also thank Amy Zambrowski, Carol Razafindrakoto, and Licia Gaber Baylis for their help constructing the data sets and running the statistical software to produce the estimates for the article.

REFERENCES

Asch, S.M., Kerr, E.A., Keesey, J., et al.: Who Is at Greatest Risk for Receiving Poor-Quality Health Care? *New England Journal of Medicine* 354(11):1147-1156, March 16, 2006.

Bodenheimer, T., Lorig, K., Holman, H., et al.: Patient Self-Management of Chronic Disease in Primary Care. *Journal of the American Medical Association* 288(19):2469-2475, November 20, 2002.

Brown, R., Peikes, D., Chen, A., et al.: *The Evaluation of the Medicare Coordinated Care Demonstration: Findings for the First Two Years.* Second Report to Congress. Mathematica Policy Research, Inc. Princeton, NJ. March 21, 2007.

Carroll, C., Patterson, M., Wood, S., et al.: A Conceptual Framework for Implementation Fidelity. *Implementation Science* 40(2):1-9, November 30, 2007.

Centers for Medicare & Medicaid Services: *Health Care Financing Review, Statistical Supplement, 2005*. U.S. Government Printing Office. Washington, DC. July 2006.

Chen, A., Brown, R., Archibald, N., et al.: *Best Practices in Coordinated Care*. Mathematica Policy Research, Inc. Princeton, NJ. March 22, 2000.

Chen, A., Brown, R., Esposito, D., et al.: *Final Report to Congress on the Evaluation of Medicare Disease Management Programs.* Mathematica Policy Research, Inc. Princeton, NJ. July 27, 2007.

Clark, A.M., Hartling, L., Vandermeer, L.B., et al.: Meta-Analysis: Secondary Prevention Programs for Patients with Coronary Artery Disease. *Annals of Internal Medicine* 143(9):659-672, November 1, 2005.

Coleman, E.A. and Berenson, R.A.: Lost in Transition: Challenges and Opportunities for Improving the Quality of Transitional Care. *Annals of Internal Medicine* 141(7):533-536, October 5, 2004.

DeBusk, R.F., Houston Miller, N., Parker, K.M., et al.: Care Management for Low-Risk Patients with Heart Failure: A Randomized, Controlled Trial. *Annals of Internal Medicine* 141(8):606-613 October 19, 2004.

Emons, K.M. and Rollnick, S.: Motivational Interviewing in Heath Care Settings: Opportunities and Limitations. *American Journal of Preventive Medicine* 20(1):68-74, January 2001.

Galbreath, A.D., Krasuski, R.A., Smith, B., et al.: Long-Term Healthcare and Cost Outcomes of Disease Management in a Large, Randomized, Community-Based Population with Heart Failure. *Circulation* 110(23):3518-3526, Winter 2004.

Geyman, J.P.: Disease Management: Panacea, Another False Hope, or Something in Between? *Annals of Family Medicine* 5(3):257-260, May-June, 2007.

Goetzel, R.Z., Ozminkowski, R.J., Villagra, V.G., et al.: Return on Investment in Disease Management: A Review. *Health Care Financing Review* 26(4):1-19, Summer 2005.

Gravelle, H., Dusheiko, M., Sheaff, R., et al.: Impact of Case Management (Evercare) on Frail Elderly Patients: Controlled Before and After Analysis of Quantitative Outcome Data. *British Medical Journal* 334(7583):31-34, January 6, 2007.

Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC. National Academy Press. 2001.

Jencks, S.F., Huff, E.D., and Cuerdon, T.: Change in the Quality of Care Delivered to Medicare Beneficiaries, 1998-1999 to 2000-2001. *Journal of the American Medical Association* 289(3):305-312, January 15, 2003.

Leatherman, S. and McCarthy, D.: *Quality of Health Care for Medicare Beneficiaries: A Chartbook.* The Commonwealth Fund (Vol. 815). May 2005. Internet address: www.cmwf.org/publications/publications_show.htm?doc_id=275195 (Accessed 2008.)

Lorig, K., Sobel, D., Stewart, A., et al.: Evidence Suggesting that a Chronic Disease Self-Management Program Can Improve Health Status While Reducing Hospitalization. *Medical Care* 37(1):5-14, January 1999.

Mattke, S., Seid, M., and Ma, S.: Evidence for the Effect of Disease Management. Is $1 Billion a Year a Good Investment? *American Journal of Managed Care* 13(12): 670-676, December 2007.

McAlister, F.A., Stewart, S., Ferrua, S., et al.: Multidisciplinary Strategies for the Management of Heart Failure Patients at High Risk for Admission: A Systematic Review of Randomized Trials. *Journal of the American College of Cardiologists* 44(4):810-819, August 18, 2004.

Naylor, M.D., Brooten, D., Campbell, R., et al.: Comprehensive Discharge Planning and Home Follow-Up of Hospitalized Elders: A Randomized Clinical Trial. *Journal of the American Medical Association* 281(7):613-620, February 17, 1999.

Oakley, A., Strange, V., Bonell, C., et al.: Process Evaluation in Randomized Controlled Trials of Complex Interventions. *British Medical Journal* 332(7538):413-416, February 18, 2006.

Pham, H.H., Schrag, D., O'Malley, et al.: Care Patterns in Medicare and Their Implications for Pay for Performance. *New England Journal of Medicine* 356(11):1130-1139, March 15, 2007.

Prochaska, J.O. and DiClemente, C.C.: Stages and Processes of Self-Change of Smoking: Toward an Integrative Model of Change. *Journal of Consulting and Clinical Psychology* 51(3):390-395, June 1983.

Rich, M.W., Beckman, V., Wittenberg, C., et al.: Multidisciplinary Intervention to Prevent the Readmissions of Elderly Patients with Congestive Heart Failure. *New England Journal of Medicine* 333(18):1190-1195, November 2, 1995.

Shojania, K.G., McDonald, K.M., Wachter, R.M., et al.: *Closing The Quality Gap: A Critical Analysis of Quality Improvement Strategies, Volume 1—Series Overview and Methodology.* Technical Review 9. Agency for Healthcare Research and Quality. Rockville, MD. 2004.

Sidorov, J., Shull, R., Tomcavage, J., et al.: Does Diabetes Disease Management Save Money and Improve Outcomes? A Report of Simultaneous Short-Term Savings and Quality Improvement Associated with a Health Maintenance Organization-Sponsored Disease Management Program Among Patients Fulfilling Health Employer Data and Information Set Criteria. *Diabetes Care* 25(4):684-689, April 2002.

Smith, B., Forkner, E., Zaslow, B., et al.: Disease Management Produces Limited Quality-of-Life Improvements in Patients with Congestive Heart Failure: Evidence from a Randomized Trial in Community-Dwelling Patients. *American Journal of Managed Care* 11(11):701-713, November 2005.

Stille, C.J., Jerant, A., Bell, D., et al.: Coordinating Care Across Diseases, Settings, and Clinicians: A Key Role for the Generalist in Practice. *Annals of Internal Medicine* 142(8):700-708, April 19, 2005.

U.S. Congressional Budget Office: *An Analysis of the Literature on Disease Management Programs.* U.S. Congress. Congressional Budget Office. Washington, DC. 2004.

Villagra, V. and Ahmed, T.: Effectiveness of a Disease Management Program for Patients with Diabetes. *Health Affairs* 23(4):255-266, July/August 2004.

Wheeler, J.: Can a Disease Self-Management Program Reduce Health Care Costs? The Case of Older Women with Heart Disease. *Medical Care* 41(6):706-715, June 2003.

Reprint Requests: Randall Brown, Ph.D., Mathematica Policy Research, Inc., 600 Alexander Park, Princeton, NJ 08540. E-mail: rbrown@mathematica-mpr.com