Cost and Outcomes of Medicare Reimbursement for HMO Preventive Services

Donald L. Patrick, Ph.D., M.S.P.H., David Grembowski, Ph.D., Mary Durham, Ph.D.,
Shirley A.A. Beresford, Ph.D., Paula Diehr, Ph.D., Jenifer Ehreth, Ph.D., Julia Hecht, Ph.D., M.P.H.,
Joe Picciano, and William Beery, M.P.H.

Medicare beneficiaries enrolled in a health maintenance organization (HMO) were randomized to a preventive services benefit package for 2 years or to usual care. At 24- and 48-month followups, the treatment group had completed more advance directives, participated in more exercise, and consumed less dietary fat than the control group. Unexpectedly, more deaths occurred in the treatment group. Surviving treatment-group enrollees reported higher satisfaction with health, less decline in self-rated health status, and fewer depressive symptoms than surviving control participants. Despite these changes, the intervention did not yield lower cost per quality-adjusted life year in this historically prevention-oriented HMO.

INTRODUCTION

Many preventive care programs in the U.S. are not designed specifically to meet the needs of older adults, although persons 65 years or over will make up 20 percent of the total population by the year 2030, compared with about 12 percent currently (Day, 1996). As a group, older adults are more likely than younger persons to have multiple, chronic, and disabling impairments and illnesses: One-half of seniors 65 years or over have a disability (McNeil, 1997). Smoking, body-mass index, and exercise patterns in midlife and late adulthood appear to have an impact on later disability, and disability can be postponed for many low-risk persons (Vita et al., 1998). Some evidence indicates that older adults can benefit just as much from health habit changes, such as smoking cessation, as middle-aged adults (Omenn et al., 1990). Evidence of a sustained trend toward improvement in health status among older adults, however, is lacking (McKusick, 1999).

An increasing proportion of public health care dollars goes to older adults. The rationale for prevention and health-promotion programs often includes the assumption that prevention is less costly than caring or curing (Somers, 1984). The cost of preventive measures such as immunizations, for example, can be considered slight in comparison to the cost of treating infectious diseases once they are contracted. Prevention and health-promotion programs for older adults, however, do not follow such simple logic. Preventing such conditions as hypertension among older adults or preventing deterioration in physical and mental functioning may involve strategies wherein the total costs are as large as any savings that might accrue from their introduction. Given the complexity of the issue, it is therefore not surprising that the

cost-effectiveness of disease-prevention and health-promotion programs is receiving increased attention (Gold et al., 1996).

Some educational interventions can potentially reduce the expense and frequency of use of conventional health care services (Roccella and Lenfant, 1992). Insufficient empirical evidence exists, however, on the costs and effects of interventions that are multifaceted and tied together as a package of services. Cost-effectiveness evaluations of the impact of specific senior health-promotion programs on health behavior have shown conflicting results (Abt Associates, 1995; Burton et al., 1995; Burton, German, and Shapiro, 1997; Lave et al., 1996, Morrissey et al., 1995). The incorporation of preventive programs for older adults in well-established managed care systems has also not been examined for effects on use and cost of care, although an increasing proportion of seniors are receiving services in managed care organizations, and health maintenance is a central theme of managed care.

In this article, we report on the results of a cost and outcomes evaluation of a demonstration health-promotion program entitled "A Healthy Future," implemented in a well-established staff-model HMO. The package of services comprising the program was designed to be feasible to implement in managed care organizations throughout the U.S. The evaluation was designed to test the cost-effectiveness of the program at two points in time. We expected that the introduction of the preventive-services package would, in the short-term (at 24 months), increase use of preventive services and decrease use of non-preventive ambulatory care services without significant cost savings, because of the introduction of a new benefit. In the long term (at 48 months), we expected an effective intervention to significantly reduce the rate of hospitalization, use of non-preventive ambulatory care services, and total cost of care, while slowing the rate of expected health-status decline among treatment participants. Thus, at 48 months, we anticipated a comparative gain in quality-adjusted life-years between treatment and control groups. Secondary outcomes were considered to be behavioral changes.
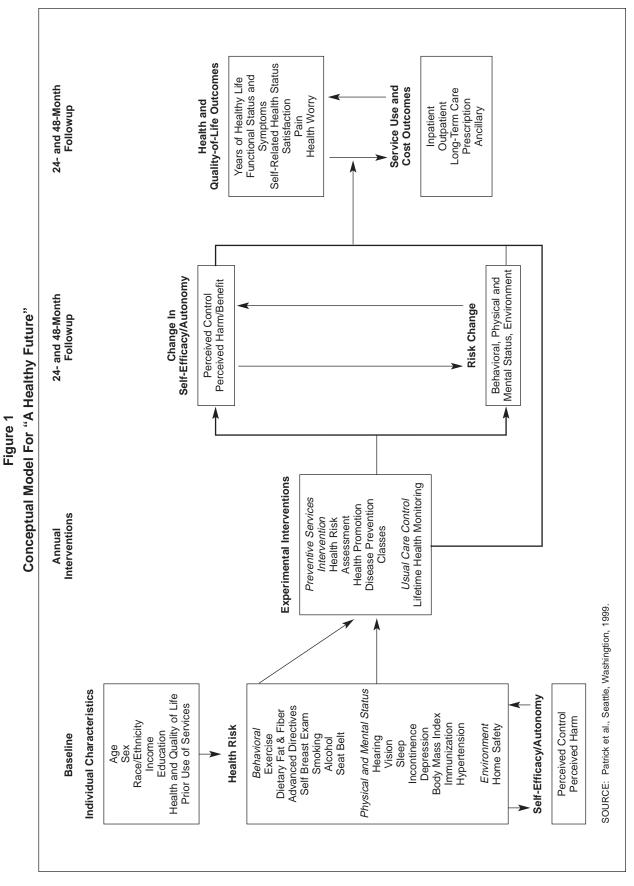
METHODS

Setting

Group Health Cooperative (GHC) of Puget Sound, an HMO in Seattle, Washington, in association with the University of Washington, was one of six sites in a HCFA-sponsored trial to determine whether Medicare payment for preventive services delivered to seniors results in better health and fewer doctor and hospital visits.

Conceptual Model

The conceptual model for "A Healthy Future" posited that exposure to health care interventions would increase self-efficacy in performing preventive behaviors and would reduce health risks, either through beneficial changes in health behavior or through immunizations and screening services, thereby improving health status and reducing costs and utilization in the long run (Figure 1). We applied the closely related principles of self-efficacy and autonomy in designing and implementing the interventions and conducting the evaluation (Grembowski et al., 1993). The theory of self-efficacy, that control over one's behavior can be a determinant of health, is a central concept of Bandura's (1977a,b) social learning theory and can be used to explain health behavior change and maintenance. Based on this theory, raising the level of self-efficacy of older adults for engaging in

Figure 1

Conceptual Model For "A Healthy Future"

health-promoting behaviors and maintaining health should lead to the adoption of beneficial changes in health behavior. Autonomy is a global concept related to the notions of control or sense of control over one's environment and life. We used this concept in relation to decisionmaking among older adults. Promoting autonomy in decisionmaking, particularly about health care, should lead to greater participation in decisions concerning life-sustaining treatment and potentially lead to lower cost of care during the final months of life.

Enrollees

Medicare enrollees in four GHC medical centers ($N = 5,012$) were invited to participate. Enrollees were drawn from the list of seniors who had been enrolled in GHC for at least 1 year and were receiving services from GHC physicians who agreed to participate in the trial. Participating physicians screened potential enrollee lists and excluded from the 5,720 enrollees 329 seniors who were severely cognitively impaired or known to have terminal illnesses that would likely result in death before the project was concluded. We attempted to exclude as few senior enrollees as possible, because this trial was designed to investigate the cost and effectiveness of a benefit potentially to be made available to all eligible Medicare enrollees.

Procedures and Measures

Each of the participants completed a baseline mail questionnaire and a 30-minute telephone interview. Dillman's (1978) mail survey methods were used to design an easy-to-read questionnaire that minimized response burden and encouraged participation. The mail questionnaire contained the senior health-risk assessment items used for both the evaluation component and intervention plans for treatment participants, information on self-efficacy for behavioral change, measures of functional status and perceived health, preferences for life-sustaining treatment, and out-of-plan use of services. The questions ranged over such topics as exercise, dietary fat and fiber, advance directives, breast self-examinations, smoking, alcohol and seat belt use, hearing and vision status, sleep, incontinence, depression, body mass index, immunization, hypertension, medications, and home safety. (Definitions of these measures may be found in the Technical Note).

After the questionnaires and signed consent forms were returned, subjects were randomized to treatment or control group. A followup telephone interview was conducted to complete an interviewer-administered health-status questionnaire and to provide missing data detected in the mail questionnaires. The telephone interview contained an older adult version of the Quality of Well-Being Scale (Kaplan, Bush, and Berry, 1976) developed for use in this project as the major health-status endpoint.

After randomization, enrollees assigned to the treatment group ($n = 1,282$) were invited to take advantage of the benefit package and services for 2 years. Control-group participants were given usual care, including access to usual preventive services as described in this article. Both groups were assessed in followup surveys conducted 24 months and 48 months following randomization. Twenty-four-month and 48-month mail questionnaires and telephone interview surveys were conducted in a manner nearly identical to the initial surveys, with minor question additions and deletions.

Cost assessment was based on data obtained from the GHC utilization and cost database. Utilization and costs from charge data were calculated for five different

categories of care: (1) inpatient services, (2) outpatient services, (3) long-term care services, (4) prescription services, and (5) ancillary services. The cost of services provided by non-GHC facilities was obtained from GHC billing data sets. All cost/charges were obtained for each participant enrolled during the month for any observation period. The capitation provided for treatment-group enrollees was not added to total costs; however, all utilization for treatment-group enrollees, including the preventive-services package, was included. Yearly aggregates were calculated for each participant, and costs were applied for each category for each patient and summarized for the participants in order to compare treatment- and control-group arms of the trial 1 year prior to clinic implementation and 24 and 48 months after clinic implementation.

Preventive Services Package

The package of preventive services offered to participants in the treatment group consisted of four distinct components: (1) health-risk assessment, (2) health-promotion visit, (3) disease-prevention visit, and (4) followup classes. A computer program was developed to determine treatment participants' health risks using information collected in the baseline mail questionnaire and a telephone followup interview based on state-of-the-art risk-assessment techniques provided by the Centers for Disease Control and Prevention. After the health-risk assessment was conducted, participants in the treatment group were invited to a 90-minute health-promotion visit with trained nurses in the medical center where the participants usually received care. At this visit, the senior enrollees' health-risk appraisals were reviewed, positive behaviors rein-forced, and referrals made to interventions for appropriate risk areas that the seniors were interested in addressing.

Three universal interventions were given to all participants: (1) counseling intended to improve exercise behavior; (2) counseling to promote a diet low in fat and high in fiber; and (3) counseling to complete advance directives, including a living will and/or durable power of attorney for health care. Written materials were made available, classes were offered, and a recommended health-promotion plan, including recommended physician followup actions, was discussed with each participant. Following the health-promotion visit, classes were held with particular emphasis on exercise and planning ahead (promoting advance directives, housing plans, and long-term care insurance).

The disease-prevention component consisted of a 1-hour visit with the nurses and physicians who were seen regularly by the enrollees. The physicians conducted histories and physical exams, reviewed the seniors' health promotion plans, and encouraged participation in health-promotion interventions prescribed during the health-promotion visit. The nurse followup included further teaching in specific areas of risk and dealt with any remaining concerns of the enrollees. Primary care followup was conducted principally by registered nurses. In the second year of the project, both the health-promotion and disease-prevention components were offered to seniors, with additional work on continuing areas of health risk. The only classes offered in the second year of the intervention were group-exercise sessions. GHC received a capitation rate of $186.03 per year for the preventive-services package, as well as $20.00 for each baseline health-risk assessment conducted for treatment-group participants, from HCFA. All providers who

participated in the demonstration were trained by the intervention team from the Center for Health Promotion at GHC using standardized training materials and sessions.

Control-group participants were offered the usual package of benefits from GHC that traditionally had an emphasis on prevention in primary care for senior enrollees. This benefit package included preventive services when requested by control participants or ordered by the physician. No reminders or periodic reviews were conducted for control participants to ensure they received preventive services as part of their usual care.

Study of Participation

A substudy was conducted 3 years after the invitation to examine factors influencing participation and health status among participants and non-participants in the demonstration. Automated records were examined to investigate demographic and utilization differences. Of the 2,454 non-participants, 900 people were selected through sampling stratified by age. Data on the Quality of Well-Being Scale were obtained for 680 (75 percent) of sampled non-participants.

Analysis

The analysis of data to test the short-term (24-month) and long-term (48-month) hypotheses were in the first instance descriptive, comparing patterns of self-efficacy and behavior change, service use, and health status between the treatment and control groups. Bivariate statistical tests were performed to identify significant differences between groups. Student t-tests were calculated to determine whether treatment and control participants had significantly different health risks, health outcomes, and costs. One challenge occurred in accounting for deaths in the treatment and control groups. Analysis of survivors only favors the group with more deaths because more of the sickest enrollees are removed from the complete-case analysis. We were able to address this problem only in the analysis of changes in the Quality of Well-Being because death has a value in this scale. In the analyses of mortality risk, multiple logistic regression analyses were used to adjust for baseline variables comparing the treatment group with the control group. Some analyses were repeated using Cox proportional hazards analysis.

RESULTS

Participation

Fifty-one percent ($n = 2,558$) of all eligible enrollees gave their consent to participate, with participation rates varying between 50 and 62 percent across the four medical centers (Durham et al., 1991). The substudy of participants indicated that non-participants tended to be older, live farther from clinics, and visit their physicians less frequently than participants. Participation was higher for enrollees whose spouses also were invited to participate as opposed to those who either had no spouse or whose spouse was not invited. Higher health status measured by the Quality of Well-Being Scale was associated positively with participation, although those with the very best health status were significantly less likely to participate, indicating a u-shaped curve between health status and participation.

About 90 percent of the treatment group had health-promotion and disease-prevention visits in the first year of the intervention. Visits were about 7 percentage points lower in the second year of the intervention. About 78 percent ($n = 1,005$) of the treatment group received all four visits (two

visits, health-promotion and disease-prevention, per year), and 9 percent (*n* = 119) had no visits. In contrast, only 24 percent (*n* = 302) of the treatment group attended one or more classes. Even though the exercise and planning-ahead classes were offered to virtually all treatment-group enrollees, only a minority attended.

Between 83 and 88 percent of treatment-group enrollees received the three universal interventions (exercise, nutrition, and planning ahead). Exposure to other preventive services ranged between 17 percent (seat belts) and 76 percent (home safety). Much of this variation is a result of treatment enrollees' health risks. As expected, at-risk treatment-group enrollees were more likely to receive a preventive service for a particular health risk than not-at-risk participants (except for the exercise and home-safety interventions). Among at-risk treatment-group enrollees across areas, exposure to the intervention ranged between 45 and 89 percent.

At the 2-year followup survey, treatment and control enrollees indicated whether they had ever talked to their doctors, nurses, or pharmacists about topics in the health-promotion visits. The percentage of participants who report discussing these topics was much higher in the treatment group than in the control group, verifying that differential exposure to preventive services occurred between the two groups, as intended. About 89 percent of treatment participants had a second health-promotion visit. Of these, the percentage of seniors reporting attempts to reduce their health risk varied greatly, ranging between 5 and 76 percent across areas.

Baseline Differences

Because enrollees were randomized into the treatment and control groups, we expected no differences between these groups on the key variables in the evaluation. Key differences were found in a few measures (Table 1). Compared with control enrollees at baseline, the treatment group reported worse self-rated health, more stress in the past year, a higher percentage of participants with hypertension, and more immunizations. Although a small but significant difference was obtained when self-rated health was measured as a continuous variable, no significant difference was detected in the percentage of treatment and control participants rating their health as fair or poor (18.3 and 16.2 percent, respectively; chi-square $p = 0.158$).

Attitudinal and Behavioral Outcomes

At the 24-month followup (Table 2), treatment participants had significant health-risk improvements in 9 out of 17 areas: (1) alcohol, (2) breast cancer screening, (3) exercise, (4) hypertension, (5) medication awareness, (6) nutrition, (7) planning ahead, (8) use of seat belts, and (9) smoking. Improved health risks were also observed at the 48-month followup in these areas, except medication awareness ($p = 0.07$). In contrast, health risk declined significantly in two areas, hearing and home safety, at both followups. At the 48-month followup, no significant differences in self-efficacy beliefs were observed between the treatment and control participants in all areas. Trend test *z*-value and chi-square significance tests indicate that treatment-group enrollees improved significantly more than control participants in three areas: (1) physical activity, (2) completion of advance directives, and (3) the proportion receiving flu shots. In addition, a *t*-test comparison indicated that the treatment group reduced dietary fat more than the control group, based on percentage of calories from

Table 1

Comparison of Baseline Measures for Treatment- and Control-Group Participants

Baseline Measure	Treatment Group (n = 1,282)		Control Group (n = 1,276)		p-value[1]
	Measure	Standard Error	Measure	Standard Error	
Individual Characteristics					
Mean Age	72.7	0.15	72.8	0.15	0.608
Percent Female	60.1		61.8		0.380
Mean Household Income	$22,199	373	$22,450	393	0.643
Mean Years of Completed Education	13.0	0.10	13.0	0.10	0.770
Health Status					
Quality of Well-Being[2]	0.704	—	0.700	—	0.264
Self-Rated Health[3]	3.30	0.30	3.39	0.30	0.023
Perceived Quality of Life[2]	7.86	0.04	7.94	0.04	0.150
Health Worry[4]	3.16	0.08	3.03	0.08	0.272
Stress Past Year[4]	3.52	0.08	3.25	0.07	0.011
Chronic Disease Index[4]	2.63	0.08	2.61	0.08	0.870
Depression[5]	8.68	0.21	8.30	0.20	0.188
Behavioral Health Risk					
Exercise	62.4	—	59.3	—	0.111
Dietary Fat	28.3	—	27.3	—	0.556
Dietary Fiber	67.5	—	67.4	—	0.968
Advance Directives	68.4	—	66.4	—	0.255
Breast Self-Exam	72.5	—	72.9	—	0.868
Smoking	8.4	—	10.0	—	0.160
Alcohol	16.5	—	16.1	—	0.788
Seat Belts	25.7	—	27.2	—	0.380
Physical and Mental Health Risk					
Hearing	12.4	—	12.8	—	0.777
Vision	23.6	—	20.8	—	0.081
Sleep	29.6	—	31.0	—	0.443
Incontinence	34.6	—	33.6	—	0.618
Depression	26.6	—	25.4	—	0.486
Immunizations	36.6	—	32.0	—	0.014
Hypertension	38.9	—	34.0	—	0.011
Medications	65.0	—	64.0	—	0.616
Environmental Health Risk					
Home Safety	22.5	—	21.6	—	0.611
Outcome Expectations					
Overweight	7.37	0.08	7.37	0.07	0.978
Too Much Fat	8.64	0.06	8.61	0.06	0.788
Too Little Exercise	7.17	0.08	7.24	0.08	0.538
Too Much Alcohol	9.35	0.05	9.36	0.05	0.858
Smoking	9.45	0.04	9.51	0.04	0.348
Efficacy Expectations					
Weight	6.47	0.09	6.39	0.09	0.521
Amount of Fat in Diet	6.85	0.09	6.83	0.09	0.870
Exercise	6.49	0.09	6.72	0.09	0.064
Amount of Alcohol	9.48	0.05	9.54	0.05	0.376
Smoking	9.12	0.08	9.04	0.07	0.477

[1] All continuous measures (those with standard errors) were tested using t-tests; all categorical measures were tested using chi-square tests.

[2] Higher values indicate better status.

[3] Excellent = 5; higher values indicate better status.

[4] Higher values indicate worse status.

[5] Measured using the Centers for Epidemiological Studies Depression scale (Radloff, 1977).

SOURCE: Patrick et al., Seattle, Washington, 1999.

Table 2

Percent Change in Health Risk Among Treatment- and Control-Group Participants at 24-Month Followup

Risk Area	Percent in Both Groups at Risk (Baseline)	Treatment Group (n = 1,211)				Control Group (n = 1,234)				Trend Test	
			No Change				No Change				
		Decline[1]	At Risk	Not at Risk	Improve[2]	Decline[1]	At Risk	Not at Risk	Improve[2]	z-value[3]	p-value[4]
					Percent						
Behavior											
Physical Activity	60.9	13	35	26	27	13	38	28	21	2.45	0.020
Dietary Fat and Fiber[5]	58.8	11	41	29	19	12	42	30	17	0.08	0.398
Advance Directives[6]	66.5	2	33	30	35	4	48	30	18	10.40	0.000
Breast Self-Exam	72.7	11	51	17	21	10	56	17	17	1.50	0.130
Smoking	9.2	1	6	91	2	1	6	90	3	-2.00	0.399
Alcohol	16.3	3	10	81	6	4	9	81	7	-0.06	0.398
Seat Belts	26.4	3	16	71	10	3	16	69	12	-0.08	0.398
Physical and Mental Status											
Hearing	12.6	4	10	84	2	3	10	84	2	-0.10	0.397
Vision	22.2	11	13	67	9	11	12	69	8	-0.01	0.399
Sleep	30.3	12	18	59	11	13	21	57	10	1.63	0.106
Incontinence	34.1	9	23	57	11	9	24	58	9	1.14	0.208
Depression	26.0	12	15	62	11	14	14	62	11	0.04	0.399
Body Mass Index	26.9	5	23	69	3	5	24	67	4	0.69	0.314
Flu Shot	34.3	2	19	62	17	3	19	66	12	2.09	0.045
Medications	65.0	9	51	27	13	11	49	25	15	0.00	0.399
Life Events/Stress	7.1	6	2	86	6	6	1	89	4	0.40	0.368
Environment											
Home Safety[7]	22.0	9	14	71	6	10	16	70	5	1.89	0.067

[1] Decline means participant was not at risk at baseline but was at risk at followup.

[2] Improve means participant was at risk at baseline but was not at risk at followup.

[3] Test.

[4] A 2 x 4 chi-square test for association.

[5] t-test results indicate the intervention group reduced dietary fat more than the control group (p < 0.005).

[6] Decline in advance directives means participants reported having an advance directive at baseline but not at followup.

[7] Only among those in treatment group at baseline.

SOURCE: Patrick et al., Seattle, Washington, 1999.

dietary fat. At the 48-month followup, improvement was maintained only in completion of advance directives and flu shots (data not shown).

Mortality

The mortality rate among the treatment-group participants (5.5 percent) was higher than among control participants (3.3 percent) (Table 3). This was observed at the 24-month followup ($p = 0.006$) and at the 48-month followup, although the difference at 48 months was of borderline significance ($p = 0.062$). When participants are stratified by baseline age, no mortality differences exist between treatment and control participants age 65 to 74 at either followup. Among participants age 75 years or over, however, a significant difference in the mortality rate at 24 months ($p = 0.005$) persisted between the treatment and control participants at 48 months ($p = 0.05$). As the randomization results indicated, differences between treatment- and control-group participants at baseline suggested a less healthy treatment group. More detailed analyses of mortality findings are reported elsewhere (Patrick et al., 1995).

Health Status of Survivors

Treatment-group participants who survived to the 24-month followup improved significantly in satisfaction with health and living (perceived quality of life) and reported higher self-rated health, fewer depressive symptoms, and less health worry than control-group participants (Table 4). For both groups, Quality of Well-Being scores declined by very small amounts during the 24-month period. At the 48-month followup, the treatment group reported higher satisfaction with health and living, higher self-rated health, less depression, and less health worry compared with control participants. The comparisons for differences in depression and health worry were statistically significant, consistent with our hypothesis.

No differences, however, were observed between the treatment and control groups for the Quality of Well-Being score (excluding deaths). This scale allowed us to plot the quality-adjusted survival of participants in the treatment and control groups (Figure 2). These results indicate that adjusting the survival experience with health and function showed an advantage to the control group over the treatment group in health decline. The control group experienced less health decline on the Quality of Well-Being scale than the treatment group. The majority of excess deaths occurred in treatment-group enrollees age 75 or over. The differences observed increased when restricting the analyses to this age group.

Utilization

In the year prior to clinic implementation, no statistically significant differences in utilization and costs existed between the treatment and control groups. As shown in Table 5, total GHC and non-GHC costs did not differ among the treatment and control group at either 24 or 48 months following clinic implementation. Utilization of clinic visits and non-GHC costs for outpatient visits, laboratory tests, and prescriptions were higher among treatment than control participants in the 24 months following clinic implementation. This is not surprising because the averages include the health-promotion and disease-prevention visits provided during the intervention. No other utilization and cost comparisons differed. Total costs did not include the capitation provided by HCFA for treatment-group participants ($186 per treatment-group enrollee per year), indicating that total cost

Table 3
Mortality for Treatment and Control Participants at 24- and 48-Month Followup, by Age and Group

Participant Group	24-Month Followup				48-Month Followup			
	Alive		Dead		Alive		Dead	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
All Participants[1]								
Total	2,445	95.6	113	4.4	2,168	90.7	223	9.3
Treatment	1,211	94.5	71	5.5	1,073	89.6	125	9.8
Control	1,234	96.7	42	3.3	1,095	91.2	98	8.2
Age Under 75 Years[2]								
Total	1,645	97.2	48	2.8	1,493	94.1	94	5.9
Treatment	826	96.7	28	3.3	744	93.7	50	6.3
Control	819	97.6	20	2.4	749	94.5	44	5.6
Age Over 75 Years[3]								
Total	800	92.5	65	7.5	675	84.0	129	16.0
Treatment	385	90.0	43	10.0	329	81.4	75	18.6
Control	415	95.0	22	5.0	346	86.5	52	13.5

[1] For 24-month followup, chi-square 7.645 (degrees of freedom = 1, p = 0.006); for 48-month followup, chi-square 3.482 (degrees of freedom = 1, p = 0.062).

[2] For 24-month followup, chi-square 1.230 (degrees of freedom = 1, p = 0.267); for 48-month followup, chi-square 0.399 (degrees of freedom = 1, p = 0.528).

[3] For 24-month followup, chi-square 7.817 (degrees of freedom = 1, p = 0.005); for 48-month followup, chi-square 3.827 (df = 1, p = .050)

SOURCE: Patrick et al., Seattle, Washington, 1999.

Table 4

Health-Status Changes for Participants Surviving to 24-Month Followup

Health Outcome	Treatment Group (n = 1,211)		Control Group (n = 1,234)		t-test
	Average	Number	Average	Number	p-value
Perceived Quality of Life[1]					
Baseline	7.92	1,114	7.98	1,152	0.303
Followup	7.97	1,114	7.93	1,152	0.402
Difference	0.06	1,114	-0.05	1,152	0.018 (t)
Self-Rated Health[1]					
Baseline	3.35	1,152	3.41	1,193	0.116
Followup	3.25	1,152	3.18	1,193	0.063
Difference	-0.11	1,152	-0.24	1,193	<0.000 (t)
Depression[2,3]					
Baseline	8.25	1,057	8.06	1,106	0.530
Followup	8.85	1,057	9.19	1,106	0.277
Difference	0.59	1,057	1.13	1,106	0.049 (t)
Health Worry[3]					
Baseline	3.09	1,089	2.94	1,144	0.215
Followup	3.51	1,089	3.63	1,144	0.340
Difference	0.42	1,089	0.69	1,144	0.047 (t)
Pain[3]					
Baseline	2.88	1,081	2.70	1,140	0.101
Followup	3.00	1,081	3.03	1,140	0.816
Difference	0.13	1,081	0.33	1,140	0.082 (t)
Quality of Well-Being Excluding Deaths					
Baseline	0.71	1,134	0.70	1,176	0.211
Followup	0.70	1,134	0.70	1,176	0.704
Difference	-0.01	1,134	0.00	1,176	0.412 (c)

[1] Large values indicate better health.

[2] Measured using the Centers for Epidemiological Studies Depression scale (Radloff, 1977).

[3] Large values indicate worse health.

(t) if favors treatment.

(c) if favors control.

NOTE: Difference is calculated as (followup score - baseline score) for each participant and averaged.

SOURCE: Patrick et al., Seattle, Washington, 1999.

Table 5

Average Total Cost per Treatment and Control Participant in Year Prior to Clinic Implementation and 24 to 48 Months After Clinic Implementation

Time Period	Treatment	Control	p-value[1]
Year Prior	$3,595	$3,414	0.392
24 Months	$3,564	$3,300	0.320
48 Months	$3,998	$4,010	0.800

[1] t-tests used to compare means.

NOTES: Total costs include Group Health Cooperative and non-Group Health Cooperative costs based on mean cost per unit of service in the summary Group Health Cooperative utilization data calculated for all participants enrolled during the observation period. For detailed tables, please contact the primary author.
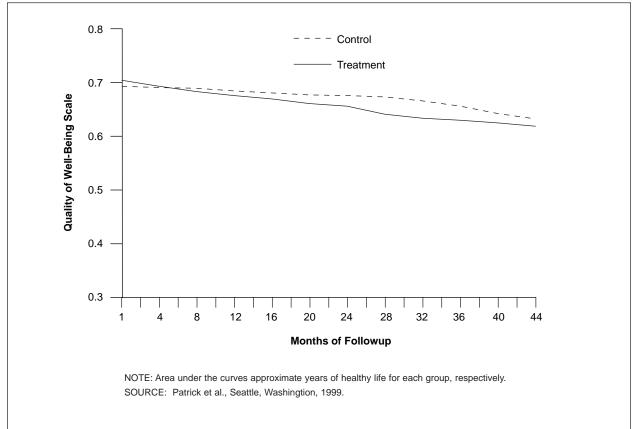
SOURCE: Patrick et al., Seattle, Washington, 1999.

of care from the societal perspective was higher for the treatment group than the control group at all points of observation.

DISCUSSION

We observed mixed results in behavior-change and negative-cost-utility results. Previously we reported our investigation of excess deaths among the treatment group (Patrick et al., 1995). The increased mortality risk existed primarily for treatment-group enrollees age 75 or over. Although some evidence supported the hypothesis that adverse selection occurred, i.e., persons with higher mortality risk were randomized to the treatment group, none of the measured variables could explain the differential. We also found no evidence of unintended adverse effects in the course of

Figure 2

Mean Quality-Adjusted Survival for Treatment and Control Groups in "A Healthy Future"

NOTE: Area under the curves approximate years of healthy life for each group, respectively.

SOURCE: Patrick et al., Seattle, Washingtion, 1999.

our investigation. Although some unmeasured effect remains a possibility, the excess mortality results could be explained in part by the increase in advance directives among treatment-group enrollees and lack of life-sustaining treatment in the event of a serious medical event. Finally, the excess mortality results in those 75 years of age or over may have occurred by chance alone.

Among survivors at the 24-month followup (n = 2,445; 95.6 percent of all participants), our evaluation results showed that short-term benefits in health-behavior change and in selected health-status outcomes (perceived quality of life, self-rating of health, health worry, and depression) were associated with the intervention. Few of these benefits remained at the 48-month evaluation. The fact that improvement was

maintained only for advance directives and flu shots is not surprising, because those risks require only a point-in-time change in behavior. In contrast, improvements in physical activity and dietary fat are more susceptible to decay because they require continuous maintenance, which may become increasingly difficult in an aging population.

Overall, no benefits in lower use of services or cost could be detected at either the short-term or long-term evaluation. These results are consistent with the overall cross-cutting evaluation results of different projects financed by HCFA to examine the cost-effectiveness of Medicare payment for preventive services (Abt Associates, 1995). A followup study in one site that observed modest gains in health and increased use of preventive services found that these

positive results could not be sustained (Burton, German, and Shapiro, 1997). In the Seattle demonstration setting, several reasons may explain the lack of any detected difference in long-term behavioral, health, or use outcomes between the treatment and control groups. GHC provided an existing set of preventive services, a senior health-promotion program, and other services to older adult enrollees. GHC also has a long-established tradition in preventive medicine with particular emphasis on older enrollees. In order to detect a treatment-control-group difference in this population, an intervention may have been needed that was substantially more intensive or included different components than what control participants could receive through usual care.

We offered a package of preventive services delivered over 24 months through annual health-promotion and disease-prevention visits with optional participation in classes. The 24-month followup data indicated that treatment participants indeed had more exposure to health-promotion services from providers than control participants. This difference in exposure, however, may not have been intense or sustained enough to improve health status or lower use of services among treatment participants in comparison to controls. A more intensive intervention, however, would have entailed higher costs with no guaranteed benefit.

Because treatment and control participants received services in the same clinics and because doctors and other primary care clinicians treated treatment as well as control participants, it is possible that contamination occurred between the two groups. Treatment participants may have shared their intervention materials with control participants, or clinicians may not have differentiated strongly enough between treatment and control patients.

Finally, it also is possible that secular trends in behavior change may have reduced the differential in behavior and outcomes between treatment and control participants. This intervention was conducted during a period of greater emphasis in the Seattle area on fitness among seniors, with newspaper articles, radio, and television programs directed to seniors. These changes may have been experienced equally between treatment and control participants, diluting any differences between the groups.

For participants in the intervention age 75 or over, years of healthy life were significantly less for the treatment group than for the control group at both evaluation time points. Because of this finding and our experience of adverse mortality, we conclude that a uniform preventive-services package may not be efficacious for all potential participants in a health-promotion/disease-prevention program for older adults. There may be adverse consequences associated with health-promotion and disease-prevention programs among the oldest old population and among those in less than good health. If this differential mortality reflects an increase in advance directives, however, the adverse consequences may be interpreted in a more positive light.

Although Medicare has expanded coverage for several specific preventive services for which effectiveness information is available, HCFA has not expanded coverage for a general preventive visit or package of services under the Medicare program. Preventive services as part of a covered benefit may need to be targeted to the specific needs of the young-old, older old, and oldest old participants. Younger seniors might benefit more from early identification of those physical and mental conditions for which there are efficacious interventions available to modify risk factors for

disease. Some interventions, such as advanced directives (planning ahead), treatment for depression, or medication monitoring may be more important for older and sicker seniors than the initiation of long-range exercise or nutrition programs that may be more appropriate for younger seniors. Modifying the social and physical environment to make it as supportive as possible in maintaining functional independence as long as possible is most likely relevant to all age groups.

This project demonstrated that integration of health promotion into primary care was possible to achieve for older enrollees in this HMO and that offering of the preventive-services package strengthened some components of a healthy lifestyle and improved subjective well-being. The material resources and communications networks that encouraged the intervention were possible to build. The project also demonstrated that higher users of services were not reluctant to participate in health-promotion programs for older adults. Even though the project had older and more frail participants, the most likely group to participate remains that segment of the senior population who are the healthiest; thus, this study represents a best case scenario for participation in a benefit package such as evaluated here. Although age accounts for lower participation in the very oldest age groups, older adults who go to their clinics regularly appear willing to make additional visits for health-promotion purposes. Higher participation, a sustained and reinforced intervention, and a benefit package more targeted to the individual needs of seniors may be necessary to provide health benefits and to reduce costs in a staff-model HMO.

TECHNICAL NOTE

Measures Glossary

This glossary documents the measures described in the Conceptual Model of "A Healthy Future" (Figure 1).

Health Risk

Behavioral health risk was defined for seven areas: advance directives, alcohol consumption, breast self-exam, dietary fat and fiber, physical activity, seat belt use, and smoking. Physical and mental risk was defined for nine areas: body mass index (obesity), depression, flu shot, hearing, incontinence, life events/stress, medications, sleep, and vision. The single environment risk consisted of home safety, which focused on the prevention of falls.

Definitions and prevalence rates are shown for the behavioral, physical and mental status, and environment areas in Table A. The remaining measures are discussed in the text of this note.

Self-Efficacy

Self-efficacy measures were developed based on the theory developed by Bandura (1977b). Efficacy expectations were defined by a participant's perceived ability

Table A

Definitions and Prevalence Rates for Risk Areas

Risk Area	Risk Criteria	Baseline Prevalence	Number at Risk
Behavioral			
Advance Directives	No living will, or has a living will but has not informed anyone, or low expected ability to care for self in 2 years (Danis et al., 1994)	67	1,700
Alcohol Consumption	More than 4 drinks in one sitting, or more than 2 drinks at least 3 times a week, or at least 1 drink a day with possible medication interaction (Ewing, 1984; Ewing and Rouse, 1970; Mayfield, McLeod, and Hall, 1974)	16	418
Breast Self-Exam	Female participants who do not perform monthly breast self-exams, or have out-of-date mammograms (Kane, Kane, and Arnold, 1985)	73	1,134
Dietary Fat and Fiber	Daily calories from fat are 40 percent or higher, or daily insoluble fiber is less than 19 grams (Kristal et al., 1990)	59	1,504
Physical Activity	None or insufficient aerobic exercise, defined as physical (aerobic) activity at least 3 times per week, at least 20 minutes per session at sufficient intensity (Buchner et al., 1991)	61	1,557
Seat Belts	Does not always wear seat belt (Goldbaum et al., 1986; Mawson, 1985)	26	676
Smoking	Currently smokes cigarettes or recently quit (Adams and Benson, 1990)	9	236
Physical and Mental Status			
Medications	Use of drugs that interact with alcohol or report of any side effects (Shimp, 1985)	65	1,650
Sleep	Waking up too early, waking up at night frequently, difficulty falling asleep, sleeping more often during the day (Kane, Kane, and Arnold, 1985)	30	774
Vision	Reports problems with vision (Sekuler et al., 1982)	22	568
Environment			
Home Safety	Age 75 or over and fell at least once in year, or female age 75 or over and lives alone, or has vision problems (Cooper, 1981)	22	564

SOURCE: Patrick et al., Seattle, Washingtion, 1999.

and likelihood to control a specific health behavior. Using exercise as an example, the general form of the question was as follows: "How sure are you that you will exercise regularly (at least three times a week) during the next year?" Subjects responded by circling a number on a 0-to-10 scale with 0 defined as "not at all sure" and 10 as "very sure." Outcome expectations were defined as the perceived harmfulness of a risky behavior. This definition was based on evidence indicating that beliefs about the consequences of not performing a health behavior may reflect outcome expectations (Strecher et al., 1986). Using exercise as an example, the general form of the question was as follows: "Do you believe that not exercising regularly (at least three times a week) is harmful to your health?" Participants responded on a 0-to-10 scale anchored by "not at all harmful" at 0 and "very harmful" at 10. Efficacy and outcome expectations were developed for alcohol intake, exercise, dietary fat, smoking and weight control (Grembowski et al., 1993).

Intervention

The group variable defines whether a consenting participant was assigned randomly to either the treatment group or the control group. For treatment participants, binary (0,1) variables indicated whether a participant completed each of the four health-promotion and disease-prevention visits. Other binary variables indicated whether a treatment participant attended each health-promotion class. If a health risk was discussed in the health-promotion or disease-prevention visits, a participant in the treatment group was defined as exposed to the intervention for that risk area. Exposure was measured for alcohol consumption, breast self-exam, depression, dietary fat and fiber, exercise, hearing, hypertension, incontinence, medication awareness, planning ahead (advanced directives, housing plans, and LTC insurance), seat belt use, sleep disturbance, smoking, stress, and vision.

To verify that differential exposure to preventive services occurred between the control and treatment groups, at the 2-year followup, all participants indicated whether they had received specific preventive services provided in the intervention. For each risk area, a binary measure indicated whether the participant developed a behavior-change plan to reduce one or more health risks at the health-promotion visits. At the second health-promotion visit, participants reported their attempts and successes in changing health behaviors, and responses were reviewed by the health-promotion nurse during the visit. For each health risk, separate binary measures were created, indicating whether the participant reported attempts or successes in changing her or his behavior.

HEALTH AND QUALITY-OF-LIFE OUTCOMES

Functional health status was measured by the Quality of Well-Being scale (Kaplan et al., 1976; Kaplan and Bush, 1982), which assigns each participant a numeric, point-in-time value between 1.0 (asymptomatic optimum functioning) and 0.0 (death). The scale provides an integrative assessment of physical activity, mobility, social activity, self-care, and symptoms. Perceived general health status (Adams and Benson, 1990) was measured on a 1-to-5 scale, ranging from excellent (1) to poor (5). Depression was measured by the 20-item Center for Epidemiological Studies Depression scale (Radloff, 1977). The Perceived Quality of Life Scale is a 20-item instrument designed to measure participants' satisfaction on a 0-to-10 scale, with the major categories of fundamental life needs (Patrick et al., 1988). The extent of pain and health worry were measured on 0-to-10 scales, where large values indicate greater levels of pain or worry (Ware and Sherbourne, 1992).

Service Use and Cost Outcomes

Service-use outcomes were calculated from utilization data files at GHC for inpatient services (percent admitted, number of admissions, hospital days), number of surgical procedures, outpatient visits (number of visits, number of acute emergency visits), pharmacy (prescriptions filled), home health services, and ancillary services (laboratory tests, radiology exams, optometry visits). Service use was calculated separately for those provided by GHC and those provided by other health care professionals or institutions. Cost outcomes were calculated by applying a standard fee schedule against units of service.

REFERENCES

Abt Associates: *COBRA Medicare Prevention Demonstration Summary Evaluation Report.* Cambridge, MA. Abt Associates, 1995.

Adams, P.F., and Benson, V.: Current Estimates from the National Health Interview Survey, 1989. National Center for Health Statistics. *Vital and Health Statistics* 10(176), 1990.

Bandura, A.: *Social Learning Theory.* Englewood Cliffs, NJ. Prentice Hall, 1977a.

Bandura, A.: Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84:191-215, 1977b.

Buchner, D., Casperson, C., Carter, W., et al.: *A New Physical Activity Questionnaire for Older Adults.* Manuscript. Seattle: University of Washington, 1991.

Burton, L.C., Steinwachs, D.M., German, P.S., et al.: Preventive services for the elderly: Would coverage affect utilization and costs under Medicare? *American Journal of Public Health* 85(3):387-391, 1995.

Burton, L.C., German, P.S., and Shapiro, S.: A preventive services demonstration: Health status, health behaviors, and cost outcomes 2 years after intervention. *Medical Care* 35(11):1149-1157, November 1997.

Cooper, S.: Common Concern: Accidents and Older adults. *Geriatric Nursing* 2(4):287-290, July-August 1981.

Danis, M., Garrett, J., Harris, R., and Patrick, D.L.: Stability of choices about life-sustaining treatments. *Annals Internal Medicine* 120(7):567-573, April 1, 1994.

Day, J.C.: Population projections of the United States by Age, Sex, Race, and Hispanic Origin: 1995-2050. *Current Population Reports*, P25-1130. Washington, DC. U.S. Bureau of the Census, 1996.

Dillman, D.: *Mail and Telephone Surveys: The Total Design Method.* John Wiley & Sons, 1978.

Durham, M., Beresford, S., Diehr, P., et al.: Participation of higher users in a randomized trial of Medicare reimbursement for preventive services. *The Gerontologist* 31(5):603-606, October 1991.

Ewing, J.A.: Detecting alcoholism: the CAGE questionnaire. *Journal of American Medical Association* 252(14):1905-1907, October 12, 1984.

Ewing, J.A., and Rouse, B.A.: Identifying the "Hidden Alcoholic." Paper presented at the 29th International Congress on Alcohol and Drug Dependency. Sidney, Australia. 1970.

Gold, M., Siegel, J., Russell, L., and Weinstein, M., eds.: *Cost Effectiveness in Health and Medicine.* New York. Oxford, 1996.

Goldbaum, G.M., Remington, P.L., Powell, K.E., et al.: Failure to use seatbelts in the United States. *Journal of the American Medical Association* 255(18):2459-2462, May 9,1986.

Grembowski, D., Patrick, D., Diehr, P., et al.: Self-efficacy and health behavior among older adults. *Journal of Health and Social Behavior* 34(2):89-104, June 1993.

Kane, R.L., Kane, A.R., and Arnold, S.B.: Prevention and the elderly: risk factors. *Health Services Research* 19(6):945-1006, February 1985.

Kaplan, R.M., Bush, J.W., and Berry, C.C.: Health Status: Types of validity and the Index of Well-Being. *Health Services Research* 11(4):478-507, Winter 1976.

Kaplan, R.M., and Bush, J.W.: Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychology* 1(1):61-80, 1982.

Kristal, A.R., Shattuck, A.L., Henry, H.J., and Fowler, A.S.: Rapid assessment of dietary intake of fat, fiber, and saturated fat: Validity of an instrument suitable for community intervention research and nutritional surveillance. *American Journal of Health Promotion* 4(4):288-295, March-April 1990.

Lave, J.R., Ives, D.G., Traven, N.D., and Kuller, L.H.: Evaluation of a health promotion demonstration program for the rural elderly. *Health Services Research* 31(3):261-281, August 1996.

Mawson, A.R.: Contrasting beliefs and actions of drivers regarding seatbelts: A study in New Orleans. *Journal of Trauma* 25(5):433-437, May 1985.

Mayfield, D., McLeod, G., and Hall, P.: The CAGE Questionnaire: Validation of a new alcoholism screening instrument. *American Journal of Psychiatry* 131(10):1121-1123, October 1974.

McKusick, D.R.: Demographic issues in Medicare reform. *Health Affairs* 18(1):194-207, January-February 1999.

McNeil, J.: Disabilities affect one-fifth of all Americans. *Census Brief* 97-5. Washington, DC. U.S. Bureau of the Census, 1997.

Morrissey, J.P., Harris, R.P., Kincade-Norburn, J., et al.: Medicare Reimbursement for preventive care: changes in performance of services, quality of life, and health care costs. *Medical Care* 33(4):315-331, April 1995.

Omenn, G.S., Anderson, K.W., Kronmal, R.A., and Vlietstra, R.E.: The temporal pattern of reduction of mortality risk after smoking cessation. *American Journal of Preventive Medicine* 6(5):251-257, September-October 1990.

Patrick, D.L., Beresford, S.A., Ehreth, J., et al.: Interpreting excess mortality in a prevention trial for older adults. *International Journal of Epidemiology* 24(Suppl 1):S27-S33, 1995.

Patrick, D.L., Danis, M.L., Southerland, L.I., and Hong, G.: Quality of Life Following Intensive Care. *Journal of General Internal Medicine* 3(3):218-223, May-June 1988.

Radloff, L.S.: The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement* 1(3):385-401, Summer 1977.

Roccella, E.J., and Lenfant, C.: Considerations regarding the cost and effectiveness of public and patient education programs. *Journal of Human Hypertension* 6(6):463-467, December 1992.

Sekuler, R., Kline, D., Dismukes, K., et al., eds.: Aging and Human Visual Function. *Modern Aging Research.* Vol. 2. New York. Alan R. Liss, Inc., 1982.

Shimp, L.A., Ascione, F.J., Glazer, H.M., and Atwood, B.F.: Potential medication-related problems in noninstitutionalized elderly. *Drug Intelligence and Clinical Pharmacology* 19(10):772-776, October 1985.

Somers, A.: Why not try preventing illness as a way of controlling Medicare costs? *New England Journal of Medicine* 311(13):853-856, September 27, 1984.

Strecher, V.J., DeVellis, B.E., Becker, M.H., and Rosenstock, I.M.: The role of self-efficacy in achieving health behavior change. *Health Education Quarterly* 13(1):73-91, Spring 1986.

Vita, A.J., Terry, R.B., Hubert, H.B., and Fries, J.F.: Aging, health risks, and cumulative disability. *New England Journal of Medicine* 338(15):1035-1041, April 9, 1998.

Ware, J.E., and Sherbourne, C.D.: The MOS-36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Medical Care* 30(6):473-483, June 1992.