Final Program Review

Gulf War Veterans' Illness Demonstration Projects

October, 2000

1. RESPONSIVENESS TO INITIAL PROPOSAL

a. What is the expected number of patients in the Demonstration Project and in the comparison group? Is the VA medical center Gulf War population large enough to give adequate sample size to meet study goals?

Table 1 summarizes recruitment and enrollment. Recruitment began in February 1999 by mailing out several waves of letters to veterans who had a Gulf War registry exam in a VISN 8 facility and to non-registry veterans who were receiving care at the Tampa VAMC. **Table 2** summarizes the final sample size

Table 1. Recruitment and enrollment into Gulf War Demonstration Project

	GROUP 1 At risk Rehab + case management	GROUP 2 At risk Rehab, + routine care	GROUP 3 Low risk Case manage-ment only	GROUP 4 Low risk Routine care	GROUP 5 At risk Routine care	GROUP 6 Low risk Routine care
2/99: 1st wave mail out	**343 (24% response rate)**				**548 (56% response rate)**	
Number responses	**39**		**43**		**309**	
Number enrolled	**7**	**6**	**17**	**15**	**104**	**78**
4/99: 2nd wave mail out	**500 (45% response rate)**					
6/99: 3rd wave mail out	**500 (25% response rate)**					
Cumulative # enrolled	**10**	**9**	**32**	**33**		
7/99: 4th wave mail out	**500 (anticipated 60% response rate)**					
Enrollment goal	**64**	**64**	**64**	**64**	**64**	**64**

Early on we initiated a marketing plan including development and distribution of both veterans and provider brochures and posters, local e-mail announcements, announcement at the local VSO meeting, personal visits to veterans service organizations and to provider groups, and advertisements in a local veterans newspaper. We also established an ambassador's program whereby prior program participants assist in recruiting new participants.

In June 1999 we evaluated non-responders to the screening questionnaire by randomly calling 25 veterans. Of these 25 non-responders, we were unable to contact 13, 5 reported that they did not receive a letter, 2 said they had returned a screening questionnaire, and 5 were not interested in the program. Based on feedback from program participants we modified our mail out system.

In August 1999, a review of reasons for refusal indicated that a large number were because veterans could not take two weeks off from family or work responsibilities to attend the residential rehabilitation program, yet they indicated that they would be more willing to attend a one-week program. In response, the clinical staff developed a one-week version of the program. The first one-week program was held September 20, 1999. This programmatic change to one week had little effect on subject recruitment in the experimental group.

b. Is a method for valid measurement of the study subject's health care utilization outlined?

Yes, see Results Section of Report.

c. Are assessment measures utilized that can quantitatively or qualitatively demonstrate study outcomes?

Yes, all quantitative measures are reliable and valid instruments that will demonstrate study outcomes.

1) The Self-Management Behaviors Scale (Lorig et al., 1997) measured **health maintenance behaviors** . 2) The Self-Efficacy to Perform Self-Management Behaviors Scale measured **confidence in performing health**

maintenance activities including: exercising regularly, getting information about the illness, obtaining help, communication with physicians, and managing illness (Lorig et al., 1997). 3) **Fatigue** , defined as both a physiological response to activity and a subjective experience of tiredness was measured using numeric rating scales. Veterans were asked to rate on a scale from 0 to 10 (0=no fatigue, 10=the worst fatigue imaginable) current fatigue, fatigue at its worst during the past two months, and the most severe fatigue experienced. 4) **Sleep Quality** was measured using the Verran/Snyder-Halpern Sleep Scale (VSH). 5) **Pain,** defined as an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage, was measured using numeric rating scales. Veterans were asked to rate on a scale form 0 – 10 (0=no pain, 10=worst pain possible) current pain, pain at its worst during the past two months, and the most severe pain experienced. 6) **Anxiety,** defined as feelings of dread, apprehension, worry, or uneasiness, was measured by the State-Trait Anxiety Inventory (STAI). The STAI consists of two 20-item, self-report scales designed to measure anxiety proneness (trait) and current level of tension and apprehension (state). 7) **Depression,** defined as depressive symptoms, was measured using the Center of Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977).

d. Are patient satisfaction and functional status determined using VA's national customer satisfaction survey form "1998 About Your VA Clinic Visits" and the SF-36 (modified for veterans) respectively? (Note: Satisfaction and functional status measurements were to be made for each participant before the onset of the Demonstration Project, after one year, and at the end of the two year project.)

Yes, these two instruments were used to collect baseline data, 4 weeks post intervention and follow up at 6 months. We were unable to measure outcomes the onset of the Demonstration Project, after one year, and at the end of the two year project, as specified in the request for applications because (1) there was 7 months start up time to launch this complex clinical program, and (2) recruitment was more difficult that anticipated, necessitating a change in our data collection time points. Two year's worth of data was unrealistic for a two year demonstration project.

e. Have compliance requirements for staff experience, and expertise in clinical research, been reviewed?

Yes, our project has been reviewed by our facility's Research & Development committee upon initiation of the project, and by the University of South Florida Institutional Review Board upon initiation and yearly, as required.

2. SCIENTIFIC MERIT

Sample

This analysis is based on a total 711 completed survey questionnaires available as of August 31, 2000. **Table 2** provides a description of the number of Gulf War (GW) veterans responding by risk/intervention group and time of questionnaire administration. Despite post card and telephone reminders, the percent of respondents lost to follow-up between the baseline survey and six-month follow-up ranged from 22% to 55% with the highest rates being in the at-risk groups. Participants in at-risk groups were asked to complete the questionnaire three times rather than just twice for the low-risk groups. All respondents included in the analysis provided responses to the vast majority of items on the questionnaire. However, given the complexity of the instrument, individual items or small groups items were occasionally left blank by a small number of respondents. This results in slight variation in the number of responses available for analysis of specific variables. When responses to individual items were summed to form scale scores, the rules established by the developer of the scale were followed in determining whether the score should be included in the analysis.

Table 2. Number of Survey Questionnaires Available for Analysis

Group	Baseline (1 week prior)	Follow-up (4 weeks post)	Follow-up (6 month post)	Lost to Follow-up At 6 months[1]
At-Risk				
Experimental (Original 1&2)	44	28	20	55%
Usual Care (Original 5)	116	92	75	35%
Low Risk				
Experimental (Original 3&4)	105	NA	82	22%
Usual Care (Original Group 6)	86	NA	64	26%

1. Rounded to the nearest whole percent

Risk Status: Risk status was determined using a 2-item risk status questionnaire specially developed for this project. Satisfaction with VHA care and presence of unexplained illness was assessed using 7-point numeric rating scales. Those who rated health in the lowest three response categories and who rated their satisfaction in the lowest three response categories were assigned to the at-risk group. The remainder was assigned to the low risk group.

Statistical Analysis

The analysis is organized to provide answers to a series of questions regarding responses to the baseline survey, a four-week follow-up and a six-month follow-up. More than fifty dependent variables and a number of potential covariates were measured in the survey questionnaire for this study. While this strategy was necessary in light of the complex nature of the issues surrounding at-risk GW veterans, it does necessitate a strategy to control for violations of family-wide error rates that may occur when simultaneous multiple statistical inferences are conducted. To control for this potential problem, two steps were taken. First, families of variables were defined as per Miller (1981) to include like variables. We defined five families: (1) Physical and Social Health (SF-36V), (2) Patient Satisfaction (VHA Patient Satisfaction Survey), (3) Clinical Outcomes (depression, anxiety, sleep, fatigue, pain), (4) Self-efficacy, and (5) Self-Management. Secondly, within each family of variables a modified Bonferoni approach was taken to adjust the nominal alpha level ($p \leq .05$) for effect of multiple statistical inferences (Holland & Copenhaver, 1998).

BASELINE SURVEY

Comparison of At-risk and Low-risk Respondents

Demographics

Table 3 provides a comparison of at-risk and low-risk GW respondents base on age, education level, gender, race, employment status, income level, and whether the respondent had any health care insurance or not. Chi-square tests for differences in proportions were conducted for each of the variables separately. No statistically significant differences were found between at-risk and low-risk respondents for these variables.

Table 3. Demographic Background of Low-Risk and At-Risk Gulf War Veteran Respondents at Baseline[1]

Demographic Variables	Low-Risk n (%)	At-risk n (%)	Total n (%)
Age			
25-34	32 (20.9)	36 (19.2)	68 (19.9)
35-44	46 (30.1)	61 (32.5)	107 (31.4)
45-54	54 (35.3)	67 (35.6)	121 (35.5)
55-64	17 (11.1)	21 (11.2)	38 (11.2)
> 65	4 (2.6)	3 (1.6)	7 (2.1)
Education Level			
< HS Graduate	3 (1.9)	1 (0.5)	4 (1.14)
HS Graduate	49 (30.6)	43 (22.5)	92 (26.2)
Some College	64 (40.0)	85 (44.5)	149 (42.5)
College Graduate	23 (14.4)	32 (16.8)	55 (15.7)
Graduate School	21 (13.1)	30 (15.71)	51 (14.5)

Demographic Variables	Low-Risk n (%)	At-risk n (%)	Total n (%)
Gender			
Male	164 (86.3)	137 (85.6)	301 (86)
Female	26 (13.7)	23 (14.4)	49 (14)
Race			
Black (not Hispanic)	38 (23.9)	31 (16.5)	69 (19.9)
White (not Hispanic)	91 (57.2)	132 (70.2)	223 (64.3)
Hispanic	23 (14.5)	20 (10.6)	43 (12.4)
Other	7 (2.0)	5 (1.4)	12 (3.5)
Employed			
Yes, full-time	152 (79.6)	112 (70.0)	264 (75.2)
Yes, part-time	30 (15.7)	43 (26.8)	73 (20.8)
No, not currently	9 (4.7)	5 (3.13)	14 (4.0)
Income Level			
$4,999 or less	8 (4.3)	7 (4.4)	15 (4.3)
$50,000-19,999	29 (15.4)	34 (21.5)	63 (18.2)
$20,000-49,999	117 (62.2)	96 (60.8)	213 (61.6)
$50,000-74,999	23 (12.2)	19 (12.0)	42 (12.1)
$75,000 or more	11 (5.9)	2 (1.3)	13 (3.76)
Health Insurance			
Yes	171 (90.0)	142 (89.9)	313 (89.9)
No	19 (10.0)	16 (10.1)	35 (10.1)

1. Rounded to the nearest 0.0%

Dependents Variables

Tables 4 – 8 provide comparisons of the low risk versus the high risk GW veterans. *As expected compared to low risk veterans, high risk veterans report more health problems, are less satisfied with VA care, have poorer health, and lower levels of self-efficacy in managing their health. In three areas, at-risk veterans engaged in more self-management behaviors.*

Table 4 provides a comparison of at-risk and low-risk GW respondents on the eight subscales of the SF-36V health related quality of life instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. As expected, the at-risk respondents mean scores were statistically significantly lower than the low-risk respondents on all eight of the subscales. Results of the Physical subscales include: Physical Functioning t (345) = 4.6, $p < .0001$; Role Physical t (345) = 6.1, $p < .0001$; Bodily Pain t (343) = 7.4, $p = .0001$; General Health t (346) = 9.4, $p < .0001$. Results of the Social subscales include: Vitality t (345) = 6.9, $p < .0001$; Social Functioning t (312) = 6.7, $p = .0001$; Role Emotional t (342) = 4.8, $p = .0001$; Mental Health t (345) = 5.6, $p < .0001$

Table 4. Mean Scores Reported by Low-Risk and At-Risk Gulf War Veteran Respondents on the SF-36V Sub-scales at Baseline[2]

SF-36V Outcome Variables	Low-Risk Respondents Mean (SD)	Low Risk N	At-risk Respondents Mean (SD)	At-risk n	p Value
Physical					
Physical Functioning	65.9 (23.8)	189	54.4 (22.3)	158	< .0001[2]
Role Physical	67.3 (29.5)	189	47.7 (30.0)	158	< .0001[2]
Bodily Pain	55.7 (25.7)	187	36.6 (21.9)	158	. 0001[2]
General Health	54.6 (23.5)	190	31.5 (22.3)	158	< .0001[2]
Social					
Vitality	46.6 (23.0)	189	29.4 (23.2)	158	< .0001[2]
Social Functioning	69.1 (27.0)	189	47.9 (31.3)	158	.0001[2]
Role Emotional	72.8 (27.3)	185	57.9 (30.4)	159	<. 0001[2]
Mental Health	65.5 (22.2)	189	51.2 (25.3)	158	< .0001[2]

1. Values rounded to nearest 0.0 %; Higher values represent better health status.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 8)

Table 5 provides a comparison of at-risk and low-risk GW respondents on nine subscales of the Veterans Health Administration Patient Satisfaction instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. Higher scores indicate more dissatisfaction. As expected, the at-risk respondents mean scores were statistically significantly higher (more dissatisfied) than the low-risk respondents on seven of the nine subscales. Subscales upon which the at-risk respondents scored higher include: Courtesy $t(266) = -2.9$, $p = .003$; Preferences $t(324) = -4.3$, $p < .0001$; Emotional Support $t(325) = -5.4$, $p < .0001$; Patient Education/Information y $t(301) = -4.6$, $p = .0001$; Overall Coordination of Care $t(293) = -6.6$, $p < .0001$; Coordination of Most Rest Visit $t(285) = -4.7$, $p = .0001$; and Specialist Provider Access $t(275) = -4.0$, $p = .0001$. Low-risk respondents scored significantly higher (more satisfied) than at-risk respond on the Overall Satisfaction subscale $t(328) = 3.8$, $p < .0002$.

Table 5. Mean Scores Reported by Low-Risk and At-Risk Gulf War Respondents on the VHA Patient Satisfaction Survey Sub-scales at Baseline[1]

VHA Patient Satisfaction Outcome Variables	Low-Risk Respondents Mean (SD)	Low-Risk n	At-risk Respondents Mean (SD)	At Risk n	p Value
Overall Satisfaction	0.36 (0.29)	185	0.25 (0.21)	154	.0002[2]
Courtesy	0.09 (0.23)	173	0.19 (0.33)	153	.003[2]
Preferences	0.25 (0.29)	173	0.40 (0.33)	153	< .0001[2]
Emotional Support	0.25 (0.29)	174	0.43 (0.33)	153	< .0001[2]
Patient Education/ Information	0.30 (0.31)	175	0.48 (0.36)	153	.0001[2]
Overall Coordination of Care	0.30 (0.28)	157	0.53 (0.31)	138	< .0001[2]
Visit (most recent) Coordination of Care	0.18 (0.22)	157	0.32 (0.29)	140	.0001[2]
Specialist Provider Access	0.14 (0.28)	164	0.28 (0.33)	142	.0001[2]
Pharmacy Access	0.10 (0.22)	173	0.15 (0.26)	153	.1112

1. Values rounded to nearest 0.00 %; Higher scores represent more problems in each area with the exception of Overall Satisfaction. Higher score for Overall satisfaction means higher levels of satisfaction.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 9)

Table 6 provides a comparison of at-risk and low-risk GW respondents on the five measurement scales related to clinical outcomes included in the questionnaire. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the clinical scales separately. The at-risk respondents mean scores were statistically significantly different than the low-risk respondents on five of the six scales. Results of the Clinical scales include: the CEDS Depression scale $t(308) = -6.4$, $p = .0001$; State Anxiety Index **t** $(347) = -5.8$, $p < .0001$; the Trait Anxiety Index $t(347) = -5.5$, $p < .0001$; the Fatigue Index $t(344) = -6.1$, $p < .0001$; and Pain Scale $t(345) = -7.1$, $p < .0001$.

Table 6. Mean Scores Reported by Low-Risk and At-Risk Gulf War Veteran Respondents on the Short Term Clinical Measures at Baseline[1]

Short Term Clinical Outcome Variables	Low-Risk Respondents Mean (SD)	Low-Risk n	At-risk Respondents Mean (SD)	At Risk n	p Value
Center for Epidemiology Studies Depression Scale	15.7 (11.8)	190	24.8 (14.1)	159	< .0001[2]
State Anxiety Inventory	39.6 (14.6)	190	48.4 (15.4)	159	< .0001[2]
Trait Anxiety Inventory	39.8 (13.4)	190	48.5 (14.7)	159	< .0001[2]
Verran/Snyder-Halpern Sleep Scale	95.9 (31.5)	157	94.4 (27.6)	157	.6569
Fatigue Scale	4.8 (2.6)	188	6.6 (2.6)	158	< .0001[2]
Pain Scale	3.9 (2.6)	189	6.0 (2.7)	158	< .0001[2]

1. Values rounded to nearest 0.0 %. Higher values represent higher levels of depression, anxiety, sleep disturbance, fatigue and pain.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 5)

Table 7 provides a comparison of at-risk and low-risk GW respondents on the 10 subscales of the Lorig Self-efficacy instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. The at-risk respondents mean scores were statistically significantly lower than the low-risk respondents on all10 subscales including: Exercise Regularly $t(308) = 6.7$, $p = .0001$; Information About Disease $t(343)$

= 6.0, p < .0001; Help from Community, Family, Friends t (338) = 5.1, p < .0001; Communicate with Physicians t (346) = 3.6, p = .0004; Manage Disease t (347) = 5.7, p < .0001; Do Chores t (287) = 6.6, p = .0001; Social/Recreational Activities t (308) = 6.7, p = .0001; Manage Symptoms (344) = 6.6, p < .0001; Manage shortness of Breath (341) = 4.9, p < .0001; and Control/Manage Depression (341) = 4.9, p < .0001.

Table 7. Mean Scores Reported by Low-Risk and At-Risk Gulf War Veteran Respondents on the Self-Efficacy Survey Sub-scales at Baseline[1]

Self-Efficacy Outcome Variables	Low-Risk Respondents Mean (SD)	Low-Risk n	At-risk Respondents Mean (SD)	At Risk n	p Value
Self-efficacy related to…					
Exercising Regularly	7.3 (2.7)	188	5.6 (2.8)	157	.0001[2]
Obtaining information About Disease	6.1 (3.4)	184	4.3 (3.2)	156	< .0001[2]
Obtaining Help for Community, Family, Friends	6.9(2.3)	158	5.8 (2.5)	158	< .0001[2]
Communicating with Physicians	7.9 (2.4)	190	6.9 (2.7)	158	.0004[2]
Managing disease	7.7 (2.0)	190	6.3 (2.3)	159	.0001[2]
Doing Chores	8.4 (2.3)	188	6.4 (3.0)	158	.0001[2]
Engaging in Social/Recreational Activities	7.7 (2.6)	190	5.7 (3.0)	159	.0001[2]
Managing Symptoms	6.6 (2.6)	189	4.7 (2.8)	157	< .0001[2]
Managing Shortness of Breath	7.3 (3.0)	186	5.7 (3.2)	157	< .0001[2]
Controlling and Managing Depression	7.0 (2.5)	190	5.7 (2.8)	158	< .0001[2]

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-efficacy.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 10)

Table 8 provides a comparison of at-risk and low-risk GW respondents on the nine subscales of the Lorig Self-management instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. The at-risk respondents mean scores were statistically significantly higher than the low-risk respondents on three of the nine subscales. Results of the Self-management subscales include: Cognitive Symptom Management t (294) = -6.4, p = .0001; Use of Community Services for Tangible Services t (214) = -3.3, p = .001; Use of Community Services For Emotional Support t (226) = -2.9, p = .003.

Table 8. Mean Scores Reported by Low-Risk and At-risk Gulf War Veteran Respondents on the Self-Management Survey Sub-scales at Baseline[1]

Self-Management Outcome Variables	Low-Risk Respondents Mean (SD)	Low-Risk n	At-risk Respondents Mean (SD)	At Risk n	p Value
Exercise: Stretching/ Strengthening	47.6 (58.8)	187	42.6 (57.1)	158	.42
Exercise: Aerobic	92.4 (89.5)	187	91.2 (106.9)	158	.91
Cognitive Symptom Management	1.1 (0.97)	186	1.9 (1.24)	158	.0001[2]
Mental Stress: Management/Relaxation	1.6 (0.63)	185	1.6 (0.65)	154	.31
Use of Community Services For Tangible Help	0.9 (1.2)	141	1.4 (1.5)	114	.001[2]
Use of Community Services For Emotional Support	0.3 (0.5)	127	0.5 (0.5)	101	.003[2]
Use of Community Services For Educational Services/Support Groups for Health Problems	1.3 (0.8)	69	1.4 (0.9)	78	.25
Use of Organized Exercise Programs	1.5 (0.9)	65	1.3 (0.7)	80	.06
Communication With Physicians	2.2 (1.1)	189	2.5 (1.2)	157	.01

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-management.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 9)

Comparison of At-Risk Intervention and At-Risk Usual Care Respondents

Demographics

Table 9 provides a comparison of at-risk intervention group and at-risk usual care group respondents base on age, education level, gender, race, employment status, income level, and whether the respondent had any health care insurance or not. Chi-square tests for differences in proportions were conducted for each of the variables separately. Only one statistically significant difference was found between at-risk and low-risk respondents on these variables. Respondents in the intervention group were more like to be employed part time (rather than full time) than did those in the usual care group χ^2 (2, n = 160) = 16.5, p = .001.

Table 9. Demographic Background of At-Risk Experimental and Usual Care Gulf War Veterans[1]

Demographic Variables	Intervention Respondents n = (%)	Usual Care Respondents n = (%)	Total n = (%)
Age			
25-34	1 (2.7)	2 (1.7)	3 (1.9)
35-44	11 (25.0)	38 (32.8)	49 (30.1)
45-54	17 (38.6)	47 (40.5.6)	64 (40.0)
55-64	11 (25.0)	12 (10.3)	23 (14.4)
> 65	4 (9.1)	17 (14.7)	21 (13.1)
Education Level			
< HS Graduate	1 (2.3)	2 (1.7)	3 (1.9)
HS Graduate	11 (6.9)	38 (32.8)	49 (30.6)
Some College	17 (38.6)	47 (40.5)	64 (40.0)
College Graduate	11 (25.0)	12 (10.3)	23 (14.4)
Graduate School	4 (9.1)	17 (14.6)	21 (13.2)
Gender			
Male	34 (77.3)	103 (88.8)	137 (85.6)
Female	10 (22.7)	13 (11.2)	23 (14.4)
Race			
Black (not Hispanic)	10 (23.3)	28 (24.2)	38 (23.9)
White (not Hispanic)	22 (13.8)	69 (59.5)	91 (57.2)
Hispanic	7 (16.3)	16 (13.8)	23 (14.5)
Other	4 (2.5)	3 (1.8)	7 (4.4)
Employed[2]			
Yes, full-time	21 (47.7)	91 (78.5)	112 (70.0)
Yes, part-time	22 (50.0)	21 (18.1)	43 (26.9)
No, not currently	1 (2.3)	4 (3.5)	5 (3.1)
Income Level			
$4,999 or less	3 (6.8)	4 (3.5)	7 (4.4)
$50,000-19,999	12 (27.3)	22 (19.3)	34 (21.5)
$20,000-49,999	23 (52.3)	73 (64.1)	96 (60.8)
$50,000-74,999	5 (11.4)	14 (12.3)	19 (12.0)
$75,000 or more	1 (2.7)	1 (0.9)	2 (1.3)
Health Insurance			
Yes	5 (11.4)	11 (9.6)	16 (10.1)
No	39 (88.6)	103 (90.4)	142 (89.9)

1. Rounded to the nearest 0.0%
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 8)

Dependent Variables

Tables 10 – 14 provide comparisons of the at-risk intervention versus the at-risk usual care groups at baseline. *Compared to usual care respondents, intervention respondents had less pain, better role/emotional functioning, lower depression, lower levels of anxiety, less fatigue, lower levels of self-efficacy in two areas, and different use of self-management strategies in two areas.*

Table 10 provides a comparison of at-risk intervention and at-risk usual care GW respondents on the eight subscales of the SF-36V health related quality of life instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. The at-risk respondents mean scores were statistically significantly lower than the low-risk respondents on two of the eight subscales. The at-risk respondents scored lower at a statistically significant level on the Bodily Pain $t\,(156) = 3.1$, $p = .002$ and the Role Emotional $t\,(157) = 2.9$, $p = .004$ subscales.

Table 10. Mean Scores Reported by At-Risk Intervention and Usual Care Gulf War Veteran Respondents on the SF-36V Sub-scales at Baseline[1]

SF-36V Outcome Variables	Intervention Respondents Mean (SD) (n=44)	Usual Care Respondents Mean (SD) (n=114)	p Value
Physical			
Physical Functioning	50.2 (18.8)	56.0 (23.4)	.15
Role Physical	37.9 (24.9)	51.5 (31.0)	.01
Bodily Pain	28.1 (18.8)	39.8 (22.2)	.002[2]
General Health	24.4 (19.1)	34.2 (22.9)	.01
Social			
Vitality	23.3 (16.4)	31.7 (25.0)	.01
Social Functioning	38.4 (27.5)	51.5 (32.0)	.02
Role Emotional	47.0 (30.0)	62.1 (29.9)	.004[2]
Mental Health	44.1 (25.2)	54.0 (24.9)	.03

1. Values rounded to nearest 0.0 %; Higher values represent better health status.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 8)

Table 11 provides a comparison of at-risk intervention and at-risk usual care GW respondents on nine subscales of the Veterans Administration Patient Satisfaction instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. Higher scores indicate more dissatisfaction, except for overall satisfaction. There were no statistically significant differences between mean scores of the at-risk intervention and at-risk usual care GW respondents on any the nine subscales.

Table 11. Mean Scores Reported by the At-Risk Intervention and Usual Care Gulf War Veteran Respondents on the VHA Patient Satisfaction Survey Sub-scales at Baseline[1]

VHA Patient Satisfaction Outcome Variable	Intervention Respondents Mean (SD)	Intervention Respondents n	Usual Care Respondents Mean (SD)	Usual Care Respondents n	p Value
Overall Satisfaction	0.33 (0.25)	43	0.23 (0.18)	43	.0184
Courtesy	0.14 (0.31)	43	0.20 (0.34)	110	.2804
Preferences	0.36 (0.30)	43	0.41 (0.34)	110	.3549
Emotional Support	0.40 (0.34)	43	0.45 (0.33)	110	.3711
Patient Education/Information	0.50 (0.40)	43	0.47 (0.35)	110	.5950
Overall Coordination of Care	0.56 (0.35)	41	0.51 (0.30)	41	.4646
Visit Coordination of Care	0.31 (0.24)	43	0.32 (0.31)	110	.9047
Specialist Provider Access	0.30 (0.35)	41	0.27 (0.32)	101	.6726
Pharmacy Access	0.16 (0.24)	41	0.14 (0.27)	41	.7222

1. Values rounded to nearest 0.00 %; Higher scores represent more problems in each area with the exception of Overall Satisfaction. Higher score for Overall satisfaction means higher levels of satisfaction.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 10)

Table 12 provides a comparison of the at-risk and low-risk GW respondents on the five measurement scales related to clinical outcomes included in the survey. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the clinical scales separately. The at-risk intervention respondents mean scores were statistically significantly different than the at-risk usual care respondents on five of the six scales. Results of the Clinical scales include: the CEDS Depression scale t (308) = -6.4, p = .0001; State Anxiety Index **t** (157) = -2.5, $p < .01$; the Trait Anxiety Index t (157) = -2.9, $p < .003$; the Fatigue Index t (156) = -2.2, $p < .03$; and the Pain Scale t (156) = -2.4, $p < .02$.

Table 12. Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Short Term Clinical Measures at Baseline[1]

Short Term Clinical Outcome Variables	Intervention Respondents Mean (SD)	Intervention Respondents n	Usual Care Respondents Mean (SD)	Usual Care Respondents n	p Value
Center for Epidemiology Studies Depression Scale	15.7 (11.8)	159	24.8 (14.1)	190	.005[2]
State Anxiety Inventory	39.6 (14.6)	159	48.4 (15.4)	190	.01[2]
Trait Anxiety Inventory	39.8 (13.4)	159	48.5 (14.7)	190	.003[2]
Verran/Snyder-Halpern Sleep Scale	95.9 (31.5)	157	94.4 (27.6)	190	.63
Fatigue Scale	4.8 (2.6)	158	6.6 (2.6)	188	.03[2]
Pain Scale	3.9 (2.6)	158	6.0 (2.7)	189	.02[2]

1. Values rounded to nearest 0.0 %; Higher values represent higher levels of depression, anxiety, sleep disturbance, fatigue and pain.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 5)

Table 13 provides a comparison of the at-risk and low-risk GW respondents on the nine subscales of the Lorig Self-management instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk intervention and at-risk usual care groups on each of the subscales separately. The at-risk intervention respondents mean scores were statistically significantly different than the at-risk usual care respondents on two of the nine subscales. At–risk intervention respondents reported significantly higher mean scores on the Use of Community Services for Emotional Support subscale t (99) = -3.0, p = .003. At–risk intervention respondents reported significantly lower mean scores on the use of Exercise for Strengthening/ Stretching subscale t (135) = 4.6, p = .0001.

Table 13. Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Self-Efficacy Sub-scales at Baseline[1]

Self-Efficacy Outcome Variables	At-Risk Intervention Respondents Mean (SD)	Intervention Respondents n	At-Risk Usual Care Respondents Mean (SD)	Usual Care Respondents n	p Value
Self-efficacy with respect to…					
Exercising Regularly	4.8 (2.4)	44	5.9 (2.9)	44	.03
Obtaining Information About Disease	4.3 (3.1)	44	4.2 (3.3)	112	.86
Obtaining Help for Community, Family, Friends	5.5(2.5)	44	5.9 (2.6)	44	.34
Communicating with Physicians	6.4 (3.0)	44	7.1 (2.6)	44	.18
Managing disease	5.5 (2.2)	44	6.7 (2.3)	115	.003[2]
Doing Chores	5.2 (3.1)	44	6.9 (2.8)	44	.001[2]
Social/Recreational Activities	4.8 (3.0)	44	6.0 (3.1)	115	.03
Managing Symptoms	3.8 (2.3)	44	5.0 (2.9)	44	.02
Managing Shortness of Breath	4.4 (2.9)	44	6.1 (3.1)	44	.02
Controlling/Managing Depression	5.0 (2.7)	44	6.0 (2.7)	44	.05

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-efficacy.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 10)

Table 14 provides a comparison of at-risk and low-risk GW respondents on the 10 subscales of the Lorig Self-efficacy instrument. Independent mean t-tests were conducted comparing mean scores of the at-risk and low-risk groups on each of the subscales separately. The at-risk intervention respondents mean scores were statistically significantly lower than the at-risk usual car respondents on two of the 10 subscales. At-risk respondents scored lower on the Manage Disease subscale t (157) = 3.0, p = .003 and the Do Chores subscale t (156) = 3.3, p = .001.

Table 14. Mean Scores Reported by High Risk Intervention and Usual Care Respondents on the Self-Management Survey Sub-scales at Baseline[1]

Lorig Self-management Outcome Variables	Intervention Respondents Mean (SD)	Intervention Respondents n	Usual Care Respondents Mean (SD)	Usual Care Respondents n	p Value
Exercise: Stretching/ Strengthening	17.0 (34.2)	44	52.5 (61.1)	114	.0001[2]
Exercise: Aerobic	85.9 (91.9)	44	93.3 (112.5)	114	.70
Cognitive Symptom Management	1.9 (1.1)	44	2.0 (1.3)	114	.87
Mental Stress Management/Relaxation	1.5 (0.6)	44	1.7 (0.7)	110	.08
Use of Community Services: Tangible Help	1.5 (1.4)	34	1.4 (1.5)	80	.64
Use of Community Services: Emotional Support	0.7 (0.5)	29	0.4 (0.5)	72	.003[2]
Use of Community Services: Educational Services/Support Groups for Health Problems	1.5 (0.9)	27	1.4 (0.9)	51	.50
Use of Organized Exercise Programs	1.3 (0.8)	29	1.2 (0.6)	51	.37
Communication With Physicians	2.4 (1.3)	44	2.6 (1.2)	113	.36

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-management.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 9)

COMPARISON BASELINE SURVEY WITH FOUR-WEEK FOLLOW-UP

Dependent Variables

Tables 15 and 16 show comparisons of the at-risk intervention group versus the at-risk usual care group. Means scores on the dependent variables of the at-risk intervention and usual care groups were compared at baseline and four week follow-up using one way repeated measures ANOVA. The time by group hypothesis was interpreted. *Compared to the usual group, the intervention group improved in four areas of self-efficacy (obtaining information about their disease, managing their disease, managing symptoms, and managing shortness of breath), and in one area of self management (mental stress management and relaxation. No statistically significant differences were found for the SF-36V measures, the short-term clinical outcome measures, or the Veteran Administration Patient Satisfaction instrument.*

Table 15 provides a comparison of the at-risk intervention and usual care GW respondents on the nine subscales of the Lorig Self-efficacy instrument. One way repeated measures ANOVA were conducted on each of the subscales separately with the test for the group by time interaction reported. Statistically significant differences were found on four of the nine subscales. Respondents in the intervention group reported mean greater scores at follow-up compared with respondents in the usual care group for the: Information About Disease subscale $f(1, 117) = 7.53$ $p = 0.007$; Manage Care Subscale $f(1, 120) = 10.2$ $p = 0.002$; Manage Symptoms subscale $f(1, 118) = 7.7$ $p = 0.006$; and the Manage Shortness of Breath subscale $f(1, 118) = 9.7$ $p = 0.002$.

Table 15. Comparison of Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Self - Efficacy Subscales at Baseline and Four Week Follow-up[1]

Self-Efficacy Outcome Variable	Baseline Mean (SD)	Follow-up (4 Week) Mean (SD)	p Value
Self-efficacy with respect to...			
Exercising Regularly			
Intervention (n= 28)	5.7 (2.9)	5.5 (2.7)	
Usual Care (n = 92)	5.1 (2.6)	5.4 (2.3)	.17
Obtaining Information About Disease			
Intervention (n= 28)	3.6 (2.9)	5.1(3.4)	
Usual Care (n = 91)	4.4 (3.3)	4.1(3.1)	.007[2]
Obtaining Help for Community, Family, Friends			
Intervention (n= 27)	5.8 (2.5)	5.3(2.6)	
Usual Care (n = 92)	5.8 (2.5)	6.3 (2.1)	.05

Self-Efficacy Outcome Variable	Baseline Mean (SD)	Follow-up (4 Week) Mean (SD)	p Value
Communicating with Physicians			
Intervention (n= 28)	6.3 (3.1)	6.8 (2.4)	
Usual Care (n = 93)	7.1 (2.7)	6.8 (2.6)	.16
Managing disease			
Intervention (n= 28)	5.3 (2.2)	6.3 (2.3)	
Usual Care (n = 94)	6.6 (2.3)	6.1 (2.4)	.001[2]
Doing Chores			
Intervention (n= 28)	5.8 (3.3)	6.5 (2.9)	
Usual Care (n = 92)	6.8 (2.9)	6.7 (2.6)	.10
Social/Recreational Activities			
Intervention (n= 28)	5.3 (3.2)	5.5 (2.7)	
Usual Care (n = 94)	6.0 (3.1)	6.1 (2.7)	.68
Managing Symptoms			
Intervention (n= 28)	4.0 (2.7)	5.0 (2.2)	
Usual Care (n = 92)	5.0 (2.9)	4.6 (2.6)	.006[2]
Managing Shortness of Breath			
Intervention (n= 28)	4.5 (3.0)	5.9 (2.7)	
Usual Care (n = 92)	6.1 (3.1)	5.7 (2.8)	.002[2]
Controlling/Managing Depression			
Intervention (n= 28)	5.3 (2.9)	5.6 (2.4)	
Usual Care (n = 93)	6.0 (2.7)	6.1 (2.5)	.49

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-efficacy.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 10)

Table 16 provides a comparison of the at-risk intervention and usual care GW respondents on the nine subscales of the Lorig Self-management instrument. One way repeated measures ANOVA were conducted on each of the subscales separately with the test for the group by time interaction reported. Statistically significant differences were found on one of the nine subscales. Respondents in the intervention group reported a mean increase in the likelihood to utilize relaxation techniques for stress reduction $\underline{f}$ (1, 112) = 10.3 $\underline{p}$ = 0.002 when compared with the usual care group.

Table 16. Comparison of Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Self-Management Subscales at Baseline and Four-Week Follow-up[1]

Self-Management Outcome Variable	Baseline Mean (SD)	Follow-up (4 Week) Mean (SD)	p Value
Exercise Stretching/ Strengthening			
Intervention (n= 28)	17.1 (36.2)	57.3 (67.3)	
Usual Care (n = 93)	50.6 (62.4)	59.4 (66.9)	.01
Exercise Aerobic			
Intervention (n= 28)	76.6 (78.5)	100.7 (97.4)	
Usual Care (n = 91)	98.4 (120.3)	107.4 (112.1)	.59
Cognitive Symptom Management			
Intervention (n= 27)	1.8 (1.3)	2.1 (1.1)	
Usual Care (n = 87)	1.9 (1.3)	1.8 (1.2)	.08
Mental Stress Management/Relaxation			
Intervention (n= 27)	1.6 (0.64)	2.1 0.64)	
Usual Care (n = 87)	1.7 (0.67)	1.7 (0.7)	.002[2]
Use of Community Services For Tangible Help			
Intervention (n= 21)	1.5 (1.4)	1.1 (1.2)	
Usual Care (n = 63)	1.4 (1.5)	1.3 (1.5)	.44
Use of Community Services For Emotional Support			
Intervention (n= 3)	1.6 (0.6)	2.3 (1.5)	
Usual Care (n = 3)	3.3 (0.6)	2.6 (0.6)	.23

Self-Management Outcome Variable	Baseline Mean (SD)	Follow-up (4 Week) Mean (SD)	p Value
Use of Community Services For Educational Services/Support Groups for Health Problems			
Intervention (n= 28)	2.3 (1.2)	2.5 (1.1)	
Usual Care (n = 92)	2.6 (1.2)	2.6 (1.1)	.65
Use of Organized Exercise Programs			
Intervention (n= 2)	1.0 (0.0)	1.5 (0.7)	
Usual Care (n = 5)	1.6 (0.5)	2.2 (0.4)	84.
Communication With Physicians			
Intervention (n= 28)	2.3 (1.2)	2.5 (1.1)	
Usual Care (n = 92)	2.6 (1.2)	2.6 (1.1)	.65

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-management.
2. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 10)

COMPARISON BASELINE SURVEY WITH SIX-MONTH FOLLOW-UP

Dependent Variables

Means scores on the dependent variables of the low-risk intervention and usual care groups were compared between baseline and six-month follow-up using one way repeated measures ANOVA. The time by group hypothesis was interpreted. *No statistically significant differences were found for the SF-36V measures, the short-term clinical outcome measures, the Lorig Self-management or the Lorig Self-efficacy instruments.*

One way repeated measures ANOVA were conducted on each of the Veterans Administration Patient Satisfaction instrument subscales separately with the test for the group by time interaction reported. Statistically significant differences were found on one of the nine subscales. *Respondents in the intervention group reported mean lower problem scores at follow-up compared with respondents in the usual care group for the Specialist Access subscale* $f(1, 70) = 7.2$ $p = 0.009$.

COMPARISON OF BASELINE SURVEY WITH FOUR-WEEK AND SIX-MONTH FOLLOW-UP

Dependent Variables

Tables 17 and 18 show comparison of baseline data with four-week and six-month follow-up. Means scores on the dependent variables of the at-risk intervention and usual care respondents were compared at baseline, four week and six-month follow-up using a mixed model repeated measures ANOVA. The SAS Proc Mixed procedure was employed to estimate values for missing data using the expectation-maximization algorithm under the missing at random (MAR) assumption. The time by group hypothesis was interpreted. *No statistically significant differences were found for the SF-36V measures, the VHA Patient Satisfaction instrument or the short-term clinical outcome measures. Differences in the usual care group versus the intervention group were found for self-efficacy to manage their disease, and in the use of mental stress management and relaxation.*

Table 17 provides a comparison of the at-risk intervention and usual care GW respondents on the Lorig Self-Efficacy instrument. Mixed model repeated measures ANOVAs were conducted on each of the subscales separately with the test for the group by time interaction reported. Statistically significant differences were found on one of the nine subscales. Respondents in the intervention group were shown to have statistically significantly increase in their Manage Your Disease subscale scores compared with respondents in the intervention group reported $f(1, 213) = 4.8$ $p = 0.009$. A post hoc contrast analysis (ANOVA) was conducted to compare intervention and usual care respondents least squares means at baseline with their scores at four week and six-month follow-up. At four weeks respondents form the intervention group showed a highly significant increase in lease squares mean scores the Manage Your Disease scale on the intervention $f(1, 213) = 9.6$ $p = 0.002$. However this relationship was not statistically significant when comparing baseline to six-month follow-up $f(1, 213) = 2.1$ $p = 0.14$.

Table 17. Comparison of Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Lorig Self-Efficacy Subscales at Baseline, Four Week and Six Month Follow-up[1,2]

Self-Efficacy Outcome Variable	Baseline LS Mean (SE)	Follow-up (4 Week) LS Mean (SE)	Follow-up (6 Months) LS Mean (SE)	P Value
Exercise Regularly				
Intervention	4.8 (0.4)	5.4 (0.5)	4.6 (0.6)	
Usual Care	5.9 (0.3)	5.6 (0.3)	5.9 (0.3)	.47
Information About Disease				
Intervention	4.3 (0.5)	5.5 (0.6)	4.6 (0.7)	
Usual Care	4.3 (0.3)	4.0 (0.3)	4.7 (0.4)	.04
Help from Community, Family, Friends				
Intervention	5.5 (0.4)	6.1 (0.4)	5.5 (0.5)	
Usual Care	5.9 (0.2)	5.6 (0.3)	6.1 (0.3)	.02
Communicate with Physicians				
Intervention	6.4 (0.4)	6.9 (0.5)	7.2 (0.6)	
Usual Care	7.1 (0.2)	6.9 (0.3)	7.5 (0.3)	.44
Manage Disease				
Intervention	5.5 (0.3)	6.4 (0.4)	6.0 (0.5)	
Usual Care	6.7 (0.2)	6.2 (0.2)	6.4 (0.3)	.009[3]
Do Chores				
Intervention	5.2 (0.4)	6.0 (0.5)	6.0 (0.5)	
Usual Care	7.0 (0.3)	6.9 (0.3)	6.8 (0.3)	.14
Social/Recreational Activities				
Intervention	4.8 (0.4)	5.2 (0.5)	4.7 (0.6)	
Usual Care	6.0 (0.3)	6.2 (0.3)	6.0 (0.3)	.73
Manage Symptoms				
Intervention	3.8 (0.4)	4.9 (0.5)	4.5 (0.5)	
Usual Care	5.0 (0.3)	4.8 (0.3)	4.9 (0.3)	.02
Manage Shortness of Breath				
Intervention	4.4 (0.4)	5.9 (0.6)	5.0 (0.7)	
Usual Care	6.1 (0.3)	5.8 (0.3)	6.0 (0.4)	.02
Control/Manage Depression				
Intervention	5.0 (0.4)	5.6 (0.5)	5.4 (0.6)	
Usual Care	5.6 (0.3)	6.1 (0.3)	6.1 (0.3)	.83

1. Values rounded to nearest 0.00 %; Higher scores represent higher levels of self-efficacy.
2. Least sqaures means generated based on responses from intervention (n=44 at baseline) and usual care (n=114 at baseline)
3. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple test = 10)

Table 18 provides a comparison of the at-risk intervention and usual care GW respondents on the Lorig Self-Management instrument. Mixed model repeated measures ANOVAs were conducted on each of the subscales separately with the test for the group by time interaction reported. Statistically significant differences were found on one of the nine subscales. Respondents in the intervention group were shown to have statistically significantly increase in their Mental Stress Reduction Through Relaxation subscale scores compared with respondents in the intervention group reported $f(1, 204) = 4.8$ $p = 0.002$. A post hoc contrast analysis (ANOVA) was conducted to compare intervention and usual care respondents least squares means at baseline with their scores at four week and six-month follow-up. At four weeks respondents form the intervention group showed a highly significant increase in lease squares mean scores the scale on the Mental Stress Reduction Through Relaxation intervention $f(1, 204) = 12.6$ $p = 0.0005$ and this relationship continued to be statistically significant at six-month follow-up $f(1, 204) = 4.6$ $p = 0.03$.

Table 18. Comparison of Least Squares Mean Scores Reported by At-Risk Intervention and Usual Care Respondents on the Self-Management Subscales at Baseline, Four-Week and Six Month Follow-up[1,2]

Self-Management Outcome Variable	Baseline LS Means (SE)	Follow-up (4 Week) LS Means (SD)	Follow-up (6 Months) LS Means (SD)	P Value
Exercise Stretching/Strengthening				
Intervention	17.0 (9.0)	57.3 (11.2)	16.5 (13.3)	
Usual Care	52.5 (5.6)	60.0 (6.2)	49.1 (7.1)	.11
Exercise Aerobic				
Intervention	85.9 (15.5)	103.7 (19.1)	92.8 (22.7)	
Usual Care	92.6 (9.7)	105.7 (10.6)	89.5 (11.8)	.94
Cognitive Symptom Management				
Intervention	1.9 (0.1)	1.8 (0.1)	1.6 (0.2)	
Usual Care	1.9 (0.18)	2.2 (0.2)	1.8 (0.3)	.15
Mental Stress Management/Relaxation				
Intervention	1.5 (0.1)	2.1 (0.1)	1.8 (0.1)	
Usual Care	1.7 (0.1)	1.7 (0.1)	1.6 (0.1)	.002[3]
Use of Community Services: Tangible Help				
Intervention	1.5 (0.2)	0.9 (0.3)	1.3 (0.3)	
Usual Care	1.3 (0.2)	1.0 (0.1)	1.2 (0.2)	.52
Use of Community Services: Emotional Support				
Intervention	0.7 (0.1)	0.4 (0.1)	0.5 (0.1)	
Usual Care	0.4 (0.1)	0.2 (0.0)	0.4 (0.1)	.21
Use of Community Services (Ed. Serv./Supp. Groups)				
Intervention	1.5 (0.2)	2.0 (0.4)	2.1 (0.5)	
Usual Care	1.3 (0.1)	1.7 (0.4)	2.0 (0.2)	.96
Use of Organized Exercise Programs				
Intervention	1.4 (0.1)	1.8 (0.3)	***	
Usual Care	1.2 (0.1)	2.0 (0.2)	1.4 (0.2)	.44
Communication With Physicians				
Intervention	2.4 (0.2)	2.5 (0.2)	2.5 (0.3)	
Usual Care	2.6 (0.1)	2.6 (0.1)	2.6 (0.1)	.90

1. Values rounded to nearest 0.0 %; Higher scores represent higher levels of self-management.
2. Least sqaures means generated based on responses from intervention (n=44 at baseline) and usual care (n=114 at baseline)
3. Statistically significant based on the modified Bonferoni adjustment to experiment wise error (nominal alpha = .05, multiple tests = 9)

*** Missing data does not allow for estimation of least squares mean

Sample Size and Power

Recruitment of subjects for this study did not meet goal of the study resulting in fewer responses than planned for the analysis of the survey questionnaire. This situation was exacerbated by the high lost to follow-up rate particular in those patients measured at three time points. Attempts were made to compensate for this situation by merging study groups and employing statistical models that attempt to adjust for missing data. *Even so, it is important to note that the sample sizes available for the analysis might have a negative impact on the statistical power many of the tests made, particularly the repeated measures ANOVAs.*

To estimate to potential impact of small samples, a series of post hoc power analyses were conducted on the repeated measures ANOVA tests conducted on the high-risk group between baseline and four-week follow-up (those with nominal alpha of .05 or less but not considered significant in the study). These analyses resulted in power values ranging between .50 and .65, well below levels usually sought by investigators. Had more participants completed the intervention, we may have found significant effects on clinical indicators.

Healthcare Utilization and Expense Analysis

OBJECTIVE

The objective of the cost-impact sub-study is to determine whether the intervention(s) made an impact on the Gulf War veterans' consumption of healthcare resources, and the associated expense to the VHA. Thus, the perspective of the analysis is that of the VHA, and direct expenses for care are the focus.

BACKGROUND

To address the hypothesis that the intervention will be less costly to the VHA than usual care we examine resource use and then health care expenses, in total, and by categories and components of costs. The resource use and costs hypothesis seek to confirm a lower health care use rate and / or lower expenditure values, (net of intervention costs), in the intervened subjects in the months following intervention, as compared to enrolled "routine care controls".

Logically, an intervention designed to improve a chronic conditions could result in several changes for participants receiving this intervention. A patient, in this case a veteran, could: (a) increase or reduce use of all VHA service types; or (b) reduce use of some VHA service types while increasing use of others; or (c) decrease VHA use and increase non-VHA use overall or by component, or (d) all of the above. Such changes would not be expected immediately, so a comparison of 18-month utilization profiles and valuation totals is important. Comparisons of the early cohorts over the entire study period would provide additional insight about the decay of any effect. These long-term comparisons are longitudinal analyses.

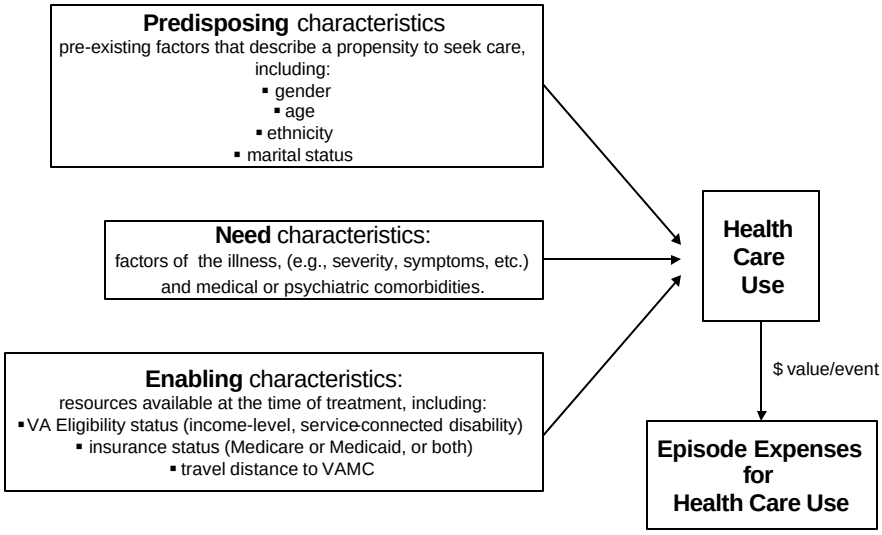
Longitudinal episodes of health care use (Hornbrook et al., 1985; Bradham et al, 2000) and their expenses will be the analytic unit for these analyses. Episodes of care for one-year and four-year periods will be assembled. Because we are concerned with sustainability of the effect into the "post-intervention" period, we must examine utilization beyond the four-years, if possible. Furthermore, because the intervention may lower selected types of health care use, we must examine the profile of use over these periods. All health care use and estimated expense will be summarized by health care settings (e.g., ER, ambulatory visit, hospital day, etc.), by service components (e.g., Inpatient, Outpatient, Medical, Pharmacy, etc.) and by overall total. Clearly, the comparison of total estimated expenses is a summary measure.

In such an analysis, the primary analytic tasks include: (a) tracking the use of health services by individuals in the five study groups, (b) placing an economic value on each event, (c) adjusting for inflation and time-value of money (i.e., discounting), if necessary, and (d) comparing group means by t-statistic, while adjusting for exogenous and patient characteristics, or covariates, if sample size permits.

The **theoretical underpinnings** for the estimation of health care use in the cost analysis is guided by the framework describing health services use by Aday and Andersen (1981; 1974). This foundation assists in selection of patient-level variables that may influence variance in the dependent variable, based on published empirical efforts. Within this framework, health services use is an outcome that is a function of three groups of characteristics -- predisposing, enabling, and need. **Predisposing** characteristics are pre-existing variables before the illness and relate to a propensity to seek care. We will examine: gender, age, ethnicity, and marital status. **Enabling** characteristics are available resources at the time of treatment. In the ACCORD study, these variables will be: patient's VA Eligibility status (based on income or the presence of a service-connected disability), alternative insurance status (Medicare, Medicaid, or both), and travel distance to the "home" VAMC. **Need** characteristics are factors directly related to the illness, including severity and comorbidities. If there are sufficient observations, this framework lends itself to a multiple-regression form of ANCOVA or hierarchical regression following the conceptual specification indicated below.

Figure 1: Theoretical Underpinning for Cost-Impact Analysis

Aday and Anderson's Model of Health Services Use



STUDY APPROACH

The intervention in this case was composed of two treatment experiences (i.e., Primary Care with Case Coordination and a more aggressive Case Management). Two types of dependent variables are used. Medical event utilization (e.g., number of events, days of inpatient care, etc.) and total resource consumption, (as measured by estimates of total costs of care for a period), are summarized for each veteran over a 24-month period pre- and post-intervention. Then group means are calculated and compared for similar quarterly periods. The comparisons of group means are statistically examined using a matched t-statistic with both equal and unequal variance tests applied to adjust for the non-normality of these measures. The research design is shown in Figure 2.

Final comparisons are stratified by risk so that the means are compared between: each intervention group, which is defined as "high-risk" (i.e., Group 1 and 2) or "low risk" (i.e., Group 3 and 4) and a group receiving "routine care" (i.e., Group 5 or 6). Original intervention groups of similar risk stratification have been collapsed for this analysis due to small samples and the lack of a distinction between the interventions when implemented.

Utilization Events and Expenditures - Dependent Variables

VHA health care use events and Direct Expense Estimates: A Veteran's utilization episode from an 18-month pre-enrollment date to the end of the study period would represent the analytic window. Categorized VHA utilization events during this period (e.g., physician visits, hospitalizations, ER visits, pharmacy units, etc.) are captured from the VHA Austin Databases. These VHA care components are expected to account for nearly all of the health care use by these veterans with chronic diseases, since data from two recent VHA studies, (Wasson and Reda, 1995; and Weinberger, et al, 1996) indicate that only 5 to 7% of veterans in the studies used non-VHA care. Non- VHA utilization is tracked to determine whether it should be added to the direct expense analysis before valuation efforts.

Scope of the Cost Analysis: The scope of the cost analysis should be comprehensive to cover total health care use and expenses. There are two reasons. First, the possibility of substitution with non-VHA service stems from the unique opportunity available to many elderly veterans who qualify for Medicare and VA coverage. Typically, chronic diseases imply a loyalty to a single provider system; however, there may be some "out-of-plan" consumption. Second, an intervention of this type may create additional use of services beyond those directly related to the disease of focus, requiring that our scope include total care, not just disease-specific care. These health care use data constitute direct medical expenses (Bradham and O'Shea, 1997; Eisenberg, 1989).

Indirect Expenses: The estimated health care expenditures based on visits, procedures and hospital stays describe the **direct expenses for medical care** . Patient out-of-pocket expenses to acquire medical care compose the traditional **indirect medical and non-medical expenses** that should be captured to describe societal costs (Bradham and O'Shea, 1998; Gold et al, 1996). In some cases (e.g., service-connected disability), the VHA pays a portion or all of these indirect expenses, so from the VHA perspective these typical indirect expenses actual migrate into the direct costs of care category in the analysis. When these are patient or family expenses they are treated as indirect costs of care.

Copayments are indirect medical expenditures that must be captured to broaden our perspective to a societal one (Donaldson et al., 1990; Eisenberg, 1989). Because the DVA has instituted means testing and is about to achieve subvention status for Medicare-eligibles, we must ask the patient and collect copayment amounts during the episode of treatment under study. ***Travel time and mileage expenses*** can be calculated for each subject by using estimated highway distances and the income values recorded in VISTA. ***Income*** for those missing data or retired will be estimated from occupation-specific estimates for the Bureau of Labor Standards, or zip-code-specific census data. We will acquire additional (e.g., meals, over night motels, etc.) self-reported expenses from the subjects and families, if the proportion of these expenses warrants that collection. Because of the nature of these patients' condition, there may be ***caregiver expenses*** that should also be captured. Using detailed interviews from prior work (Sevick and Bradham, 1997), we could interview the informal caregiver for those patients who indicate that consistent caregiver time is allocated to facilitate their home-based care, and their accessing the respite care intervention. There are additional ***indirect expenses that are borne by the VHA***. The DVA also reimburses some indirect expenses, (e.g. travel costs to the medical center if over 50 miles for those with service-connected status) so we must include these as the payer's direct non-medical expenses.

Valuations in for VHA events–Medicare Fee Structures

Utilization events in the VHA or outside the VHA require valuation, which is the task of assigning a reasonable market-level dollar for the healthcare expense event. For each patient, outpatient events (visits, procedures, labs, medications, etc.) and inpatient DRGs and events (procedures, labs, medications, etc.) are captured patient-specific clinical data sets of VISTA, the local electronic medical record, or the Austin national databases, which are derived from VISTA. Key elements of these local files are transferred to the national Patient Treatment File (PTF, or inpatient file) and the Outpatient Care Files (OPC), which allows tracking of health events within the VHA. Thus, the VHA databases provide sufficient outpatient and inpatient procedure and associated treatment classifications (CPT-4[1] and ICD-9 codes) to allow valuation at Medicare fee rates. We have constructed national average Medicare fee rate translation tables and have used them in other studies (Bradham, et al., 2000 and forthcoming). Use of charge ("asking price") information to value services is acknowledged to be an inferior evaluation method; however, use of Medicare rates, which are calculated from standardized cost reports are less likely to distort costs than are market prices and are more acceptable valuations (Finkler, 1982; Chapko, 1991). Any non-VA hospitalizations will also use the appropriate Medicare event code to estimate its

[1] Clinical Procedures Terminology, 4th revision is a standard outpatient classification used in abstracting medical records for reimbursement.

value. Since measurement of VHA health care use and estimated expenditures will be more complete than the assessment of non-VHA costs it is reassuring that, as noted above, the majority of the health care costs in a VHA HSR&D study can be expected to be VHA utilization-based data.

However, these Medicare-derived, event-driven values do not include the professional fees, which would normally be present in society's expenditure, if care is provided in the private sector. In an effort to estimate the broader societal expense for health care that is expected of the PHS Task Force on Cost-Effectiveness (Gold, et al., 1996), and to address the details of estimating the intervention's cost of production, we examine the internal cost accounting systems of the local VAMC.

The cost of non-VHA health care cannot be estimated with as much uniformity or precision as VHA costs for two reasons. First, because each non-VHA provider maintains its own, often distinct, accounting system; and second, because non-VHA cost data cannot realistically be obtained with the same level of completeness as that of VHA data, since the only readily available source is the patient's recall.

DATA and ANALYSIS

Data describing each veteran's individual health care use before and following the intervention were extracted from the VHA national databases stored at Austin Automation Center (AAC). Numbers of physician visits, inpatient admissions and inpatient days of care were summarized by quarter and annually. Total costs of care are estimated using a national average Medicare DRG- (for inpatient care) and procedure-based fee structure for each year. This obtains a summarization of the overall consumption weighted by "market costs" that is region-and local-wage neutral. Such a value under-estimates the VHA's true costs of care since physician and nursing personnel costs are not included. However, the figures are a reasonable estimate of the VHA's "facility cost" of delivering care.

When the quarterly amounts are examined, the dependent variables are transformed into log-normal values, which removes any effect of skewedness, and brings the values into conformance with normality assumptions of the parametric statistics used for the comparisons. Comparisons between group means are made using matched t-test on the logged values.

LIMITATIONS

The current report has been compiled before sufficient time has transpired to capture a complete 18-month period for all enrolled subjects, which reduces the comparative samples in the post-intervention periods. Additionally, the non-VHA utilization has not been validated in any manner.

FINDINGS

Tables 19 through 24 present the statistically significant findings for various dependent variables measuring healthcare consumption through events and total cost of care. The quarterly pre- and post-intervention costs of care are compared in **Tables 20 and 21** for the high-risk subjects, and then in **Tables 23 and 24** for the low-risk patients. **Figures 2 and 3** display quarterly data from **Tables 20 and 23.** In Tables 1 through 6, no utilization is treated as a missing value; therefore, the means in these figures and tables represents the average number of VHA events per quarter for an individual, if VHA services were used.

NOTE: Additional tables for the case where no utilization is treated as a zero value are available. These cases would respond to the question of what level of use would be expected in a population represented by the study groups.

HIGH-RISK GROUPS The findings for the high-risk groups are shown in **Tables 19 through 22**. Both the intervened and comparison groups exhibit considerable **outpatient healthcare events**, as measured by average clinic stops, visits and procedures over the entire cost analysis study period, 18 months pre- and 9-months post intervention, and shown in **Table 19**. In each category of outpatient use, the intervened sample averages more visits, with more clinic consultations (various clinic stops) and experiencing more procedures during these encounters. These outpatient differences are significant and represent the total outpatient use. No differences were noted among various inpatient events. The number of procedure and diagnostic codes measure the **intensity of outpatient services** delivered to these veterans, since each code represent the provider's intervention and diagnostic work-up, and each is associated with additional expenses of care. Again the intervened high-risk group exhibits receipt of more intense outpatient services. In

summary fashion, **overall healthcare expenses** for these high-risk patients are found to be only slightly more costly among the intervened patients than for the usual care group. Because of small samples, a more liberal criterion value for significance ($p \leq 0.10$) is used, and confirmed by the comparison of the logged values.

We conclude that throughout the cost-study period, a 27-month interval of 18-months pre-intervention and 9-months post-intervention, the intervened high-risk group demanded more outpatient care and received more services, resulting in higher VHA costs of care. The anticipated positive impact of the intervention (VHA cost-savings) is not manifested in these data from a shortened follow-up period. Surveillance beyond the 9-month periods may reveal more benefits. This implies that the intervened high-risk group received more care, which would be expected. The fact that the overall average costs of care for this high-risk group are only marginally different from that of the similar usual care group is promising.

This finding could be swamped by the intervention period of services, or per-intervention increased use associated with the characteristics that define high-risk. To isolate the impact from these confounding artifacts, we examined the quarterly sequence of these total costs of care in **Table 20,** and **Figure 1.** Careful examination of the quarterly means suggests that the intervened group may have been characterized by higher-rates of utilization even as much as 18-months before the intervention. Thus, another traditional confounder in these analyses is presented – the "hypochondria effect" – some individuals are simply heavy users of healthcare. Because this difference may be the result of skewedness in the data for a small group, we examine the means of logged values of these total VHA healthcare costs in **Table 21**. The significances that were found in the untransformed data vanish, indicating that the differences are likely due to the small samples in the quarterly comparisons. Thus, again a larger sample and more complete follow-up could reveal a different finding.

LOW-RISK GROUPS The findings for the low-risk groups are shown in **Tables 4 through 6**. Few differences are noted between the intervened and comparison groups for **outpatient healthcare events**, (e.g., average clinic stops, visits and procedures) over the entire cost analysis study period shown in **Table 22**. Only one category of outpatient use – CAT scans is different, and in this case the intervened sample averages fewer events during outpatient encounters. No differences were noted among inpatient events or measures of **intensity of outpatient services** delivered. **Overall healthcare expenses** for the entire study period were similar between the study groups among the low-risk subjects. Quarterly total healthcare expense means also show no significant differences (**Table 23** and **Table 24**). We conclude that there was no measurable cost of care difference for the low-risk group.

CONCLUSIONS:

The currently available healthcare utilization data from VHA for the veterans in the intervened and comparison groups show that the intervention does not achieve the expected cost-savings benefit by the 9-month post-intervention quarterly period. Additional data from more subjects and a longer follow-up period may reveal the anticipated positive VHA cost-sharing effect.

Figure 2: High-Risk Gulf War Average Healthcare Expense per Patient per Quarter

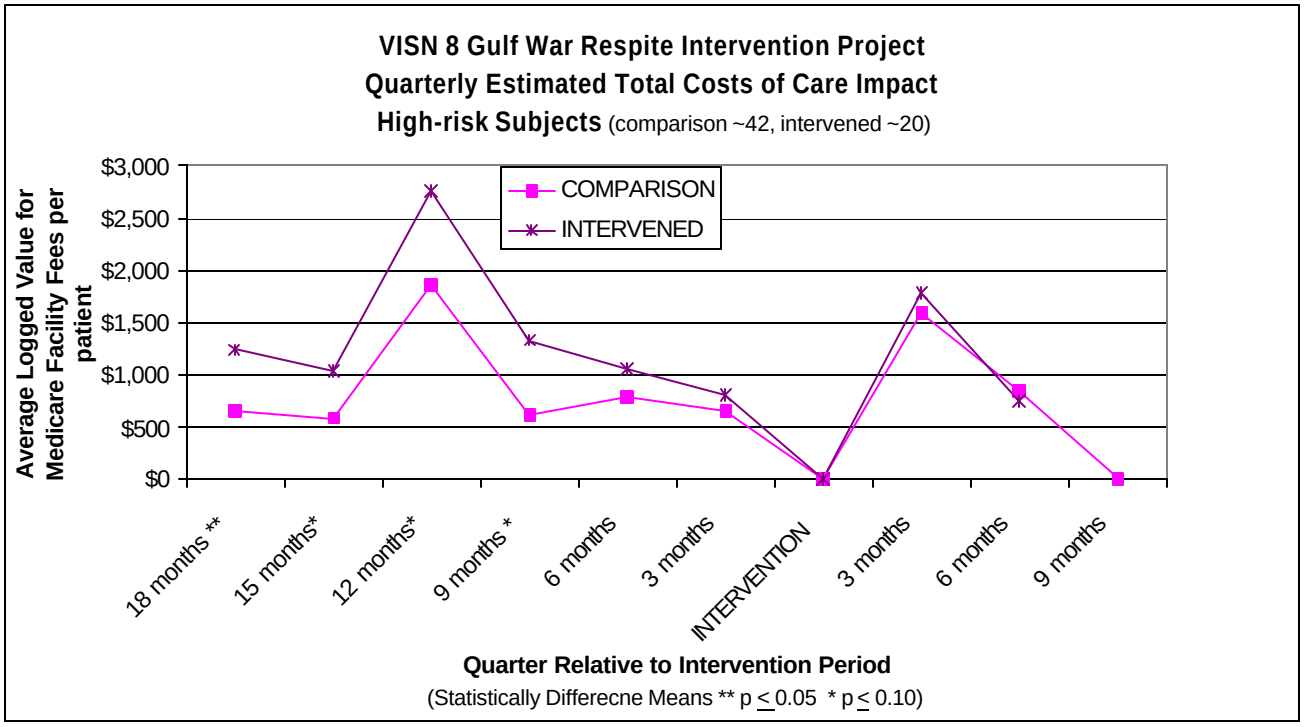


Figure 3. Low-Risk Gulf War Average Healthcare Expense per Patient per Quarter

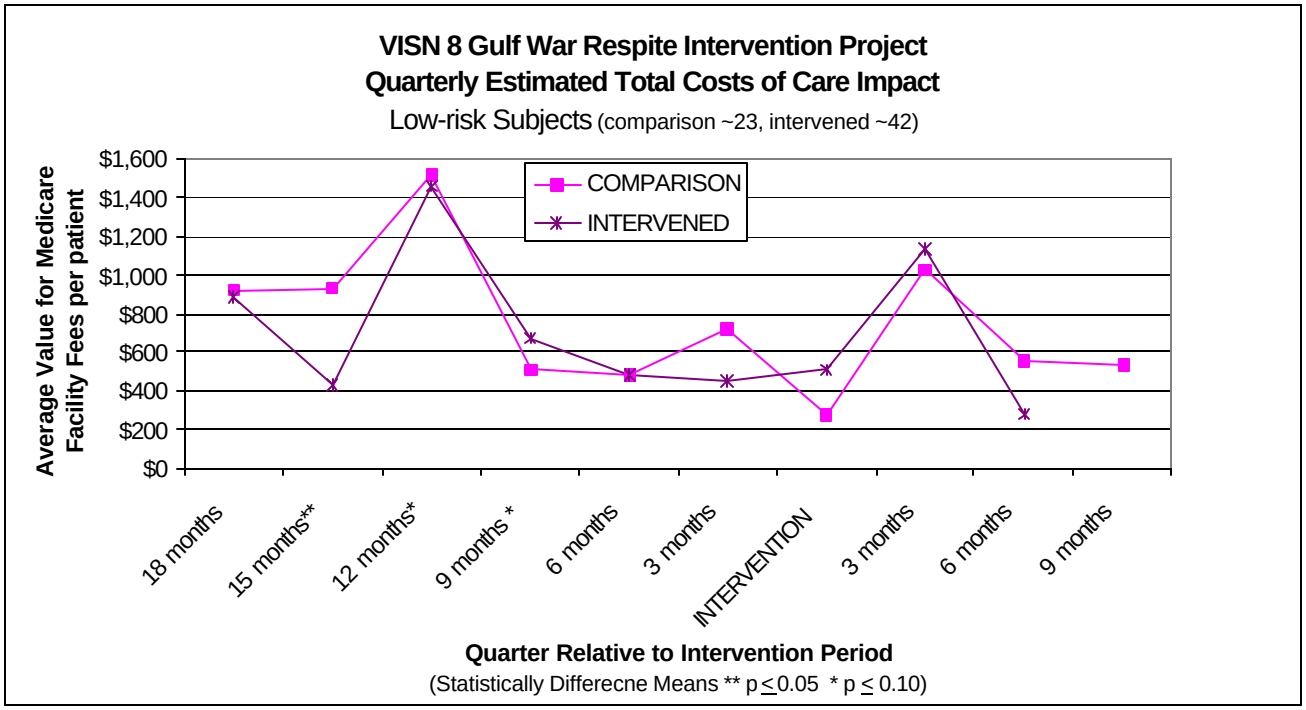


TABLE 19 – HIGH-RISK COMPARISON

General Healthcare Resource Consumption, if Used

Conceptual area of healthcare resources consumption

VARIABLE Variable definition

	N (usual care) / N (intervened)	MEAN (STD)		Intervened = Usual Care ? t-statistic PROB VALUE Assuming variance is:	
				Equal	***Unequal****
Healthcare use events measured over a 36-month pre- and post-intervention period per patient					
TOT_NSTP Total # Opt Clinic Stops	667	13.49	17.92	0.0104	**0.0112**
	314	16.67	18.39		
TOT_NVIS Total # Opt Visits	667	7.31	10.01	0.0143	**0.0119**
	314	8.96	9.26		
TOT_PRC Total # Opt Procedures Performed	667	14.15	18.43	**0.0001**	0.0002
	314	19.50	21.41		
Healthcare resource intensity measured over a 36-month pre- and post-intervention period per patient					
TOT_CPT Total # Opt CPT codes	653	10.96	14.59	0.0041	**0.0044**
	310	13.87	14.81		
TOT_DIAG Total # Outpatient Diagnoses	661	9.17	11.96	0.0001	**0.0001**
	312	12.55	12.39		
Healthcare overall expense measures over 36-month period					
TOT_FS99 Total Cost-All CPT-level Procedures	631	1122.29	1490.28	0.0993	**0.1105**
	310	1298.44	1635.05		
HCCOST Total Health Care Cost (excludes RX)	631	1172.05	1640.93	**0.0824**	0.0931
	310	1376.75	1807.65		
LGHCCST Log base10 for Total Health Care Cost	631	2.72	0.56	**0.0003**	0.0001
	310	2.86	0.49		

Notes: a. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"

b. Unit of analysis is per patient

c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

TABLE 20 – HIGH-RISK COMPARISON Quarterly Healthcare Resource Consumption, if used Conceptual area of healthcare resources consumption VARIABLE Meaning of variable	**N (usual care) MEAN (STD)** **N (intervened) MEAN (STD)**			**Intervened = Usual Care ? t-statistic PROB VALUE** **Assuming variance is:**	
				Equal	***Unequal****
Healthcare overall expense measures for specific 3-month period relative to intervention					
PRE18MO HC$s/Quarter @18 months pre-intervention	38	660.78	683.17	**0.0244**	0.0535
	23	1238.12	1270.70		
PRE15MO HC$s/Quarter @15 months pre-intervention	45	583.60	601.79	**0.0924**	0.1829
	25	1039.24	1606.34		
PRE1YR HC$s/Quarter @12 months pre-intervention	61	1863.06	2016.36	**0.0871**	0.1326
	27	2769.68	2757.56		
PRE9MO HC$s/Quarter @9 months pre-intervention	48	617.75	828.04	**0.0506**	0.1429
	23	1325.29	2168.87		
PRE6MO HC$s/Quarter @6 months pre-intervention	44	784.65	982.12	**0.4162**	0.4851
	24	1047.75	1674.50		
PRE3MO HC$s/Quarter @3 months pre-intervention	46	647.72	638.86	0.4124	**0.4573**
	26	798.15	901.76		
INTRV HC$s/Intervention Period	28	2.20	0.47	**0.0000**	0.0001
	29	2.69	0.23		
PST3MO HC$s/Quarter @3 months post-intervention	55	1578.44	1565.44	0.6006	**0.6162**
	26	1781.87	1750.25		
PST6MO HC$s/Quarter @6 months post-intervention	38	834.39	814.97	**0.8198**	0.8445
	6	750.59	950.12		
PST9MO HC$s/Quarter @9 months post-intervention	15	1.69	1.24	**No variance**	
	0	0	0		
PST1YR HC$s/Quarter @12 months post-intervention				**No obs**	
PST15MO HC$s/Quarter @15 months post-intervention				**No obs**	
PST18MO HC$s/Quarter @18 months post-intervention				**No obs**	

Notes: **a. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"**
b. Unit of analysis is per patient
c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

TABLE 21 – HIGH-RISK COMPARISON

Logged Values of Quarterly Healthcare Resource Consumption, if used

Conceptual area of healthcare resources consumption
VARIABLE Meaning of variable

	N (usual care) MEAN (STD) N (intervened) MEAN (STD)			Intervened = Usual Care ? t-statistic PROB VALUE Assuming variance is:	
				Equal	**Unequal**
healthcare overall expense measures for specific 3-month period relative to intervention					
PRE18MO HC$s/Quarter @18 months pre-intervention	38	2.55	0.56	**0.0622**	0.0602
	23	2.83	0.54		
PRE15MO HC$s/Quarter @15 months pre-intervention	45	2.58	0.40	**0.3415**	0.3773
	25	2.69	0.52		
PRE1YR HC$s/Quarter @12 months pre-intervention	61	3.01	0.51	**0.0547**	0.0505
	27	3.23	0.47		
PRE9MO HC$s/Quarter @9 months pre-intervention	48	2.54	0.45	**0.0542**	0.0803
	23	2.78	0.57		
PRE6MO HC$s/Quarter @6 months pre-intervention	44	2.59	0.54	**0.2411**	0.2241
	24	2.74	0.47		
PRE3MO HC$s/Quarter @3 months pre-intervention	46	2.59	0.45	**0.5657**	0.5736
	26	2.66	0.48		
INTRV HC$s/Intervention Period	28	2.20	0.47	**0.0000**	0.0001
	29	2.69	0.23		
PST3MO HC$s/Quarter @3 months post-intervention	55	2.93	0.54	**0.3222**	0.2827
	26	3.06	0.43		
PST6MO HC$s/Quarter @6 months post-intervention	38	2.69	0.49	**0.8589**	0.8553
	6	2.65	0.45		
PST9MO HC$s/Quarter @9 months post-intervention	15	2.11	3.30	**No variance**	
	0	0.0	0.0		
PST1YR HC$s/Quarter @12 months post-intervention				**No obs**	
PST15MO HC$s/Quarter @15 months post-intervention				**No obs**	
PST18MO HC$s/Quarter @18 months post-intervention				**No obs**	

Notes:
a. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"
b. Unit of analysis is per patient
c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

TABLE 22 – LOW-RISK COMPARISON ***General Healthcare Resource Consumption, if Used*** Conceptual area of healthcare resources consumption **VARIABLE Variable definition**	**N (usual care) MEAN (STD)** **N (intervened) MEAN (STD)**			**Intervened = Usual Care ? t-statistic PROB VALUE** **Assuming variance is:**	
				Equal	***Unequal****
healthcare use events measured over a 36-month pre- and post-intervention period per patient					
TOT_CTSN Total # Cat Scans	10	1.90	8.75	0.0019	0.0100
	12	1.00	0.00		
healthcare resource intensity measured over a 36-month pre- and post-intervention period per patient					
None significant					
healthcare overall expense measures over 36-month period					
None significant					

Notes: a. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"

b. Unit of analysis is per patient

c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

TABLE 23 – LOW-RISK COMPARISON Quarterly Healthcare Resource Consumption, if used Conceptual area of healthcare resources consumption VARIABLE Meaning of variable	**N (usual care) MEAN (STD)** **N (intervened) MEAN (STD)**			**Intervened = Usual Care ? t-statistic PROB VALUE** **Assuming variance is:**	
				Equal	***Unequal****
healthcare overall expense measures for specific 3-month period relative to intervention					
PRE18MO HC$s/Quarter @18 months pre-intervention	25	917.28	2406.19	**0.9319**	0.9422
	43	879.55	1224.58		
PRE15MO HC$s/Quarter @15 months pre-intervention	27	932.28	1351.99	**0.0338**	0.0744
	41	430.40	501.49		
PRE1YR HC$s/Quarter @12 months pre-intervention	36	1511.47	2138.06	0.8823	**0.8894**
	67	1453.48	1745.18		
PRE9MO HC$s/Quarter @9 months pre-intervention	25	509.03	651.05	**0.5134**	0.4447
	46	674.29	1160.05		
PRE6MO HC$s/Quarter @6 months pre-intervention	26	475.37	442.16	0.9776	**0.9750**
	52	479.21	617.83		
PRE3MO HC$s/Quarter @3 months pre-intervention	23	716.78	1230.46	**0.2062**	0.3279
	55	447.31	634.57		
INTRV HC$s/Intervention Period	9	277.43	269.86	0.3359	0. 1657
	26	503.39	670.23		
PST3MO HC$s/Quarter @3 months post-intervention	34	1021.75	1611.85	**0.7444**	0.7526
	64	1126.50	1452.93		
PST6MO HC$s/Quarter @6 months post-intervention	22	554.48	636.80	0.0678	**0.0712**
	22	277.82	271.33		
PST9MO HC$s/Quarter @9 months post-intervention	3	539.32	606.37	0.1788	**0.3208**
	4	81.14	17.55		
PST1YR HC$s/Quarter @12 months post-intervention				**No obs**	
PST15MO HC$s/Quarter @15 months post-intervention	1	7.42	0.00	**No variance**	
	0	0.00	0.0		
PST18MO HC$s/Quarter @18 months post-intervention	1	2.51	0.00	**No variance**	
	0	0.00	0.0		

Notes: a**. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"**
b. Unit of analysis is per patient
c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

TABLE 24 – LOW-RISK COMPARISON

Logged Values of Quarterly Healthcare Resource Consumption, if used

Conceptual area of healthcare resources consumption
VARIABLE Meaning of variable

	N (usual care) MEAN (STD) N (intervened) MEAN (STD)			Intervened = Usual Care ? t-statistic PROB VALUE Assuming variance is:	
				Equal	**Unequal**
healthcare overall expense measures for specific 3-month period relative to intervention					
PRE18MO HC$s/Quarter @18 months pre-intervention	25	2.43	0.64	**0.2576**	0.2776
	43	2.60	0.55		
PRE15MO HC$s/Quarter @15 months pre-intervention	27	2.68	0.50	**0.0165**	0.0188
	41	2.39	0.46		
PRE1YR HC$s/Quarter @12 months pre-intervention	36	2.90	0.50	**0.8016**	0.8080
	67	2.93	0.45		
PRE9MO HC$s/Quarter @9 months pre-intervention	25	2.40	0.56	**0.3441**	0.3702
	46	2.52	0.47		
PRE6MO HC$s/Quarter @6 months pre-intervention	26	2.49	0.44	**0.6381**	0.6318
	52	2.43	0.46		
PRE3MO HC$s/Quarter @3 months pre-intervention	23	2.50	0.54	**0.3999**	0.4395
	55	2.40	0.45		
INTRV HC$s/Intervention Period	9	2.30	0.35	0.4801	0.4225
	25	2.42	0.46		
PST3MO HC$s/Quarter @3 months post-intervention	34	2.54	0.69	**0.1200**	0.1543
	64	2.74	0.53		
PST6MO HC$s/Quarter @6 months post-intervention	22	2.54	0.41	**0.0312**	0.0312
	22	2.27	0.38		
PST9MO HC$s/Quarter @9 months post-intervention	3	2.48	0.62	0.1155	**0.2463**
	4	1.90	0.09		
PST1YR HC$s/Quarter @12 months post-intervention				**No obs**	
PST15MO HC$s/Quarter @15 months post-intervention				**No obs**	
PST18MO HC$s/Quarter @18 months post-intervention				**No obs**	

Notes: **a. Zero utilization is a missing value, which corresponds to "What is the average use, when healthcare is used?"**
b. Unit of analysis is per patient
c. Data is from National VHA Utilization Data (Inpatient and Outpatient files) with "no use" as missing value.
d. Unequal variance t-test are the *Satterthwaite & Cochran tests*

RELEVANCE TO GULF WAR VETERANS' HEALTH

a. Do results contribute to the scientific body of knowledge in the areas of: 1) testing new approaches to health care delivery; and 2) improving the treatment satisfaction of Gulf War veterans suffering from undiagnosed and ill-defined illnesses, or disability?

Yes, our program took an existing rehabilitation approach designed for veterans with chronic pain and adapted it for use with Gulf War veterans. We shortened the program, and redesigned it to address multiple symptoms including pain, fatigue, sleep disturbance, and impaired social and physical functioning.

Forty-four veterans completed the residential rehabilitation portion of the program. Formative feedback told us that veterans like this rehabilitation approach and they report very positively about: 1) the interdisciplinary aspect of the program, 2) learning to be in more control of their care, 3) the input from physical therapy and kinesiotherapy, 4) one-to-one consultation with the Clinical Coordinator (e.g. case manager), and 5) the camaraderie that develops within groups. Moreover, they have told us that they appreciate being listened to, and as a result of this program, have more trust in the VA system.

We have run our program using high customer service standards and have developed several mechanisms for obtaining veteran input.

3. INNOVATION

a. Have any innovative or unique approaches to treatment been developed, or changes been undertaken, since the Demonstration Project was initiated?

Three changes from original protocol have been made. First, seven months after funding we determined that we were about 4-5 months behind schedule and we submitted a formal request to the funding agency to revise our time line. We determined that a one-year follow-up was not be possible, therefore, we changed in the data collection time points from baseline-four weeks-one year to baseline-four weeks-six months (Low risk groups do not receive a four-week follow-up). Based on a concurrent survey in VISN 8 of Gulf War veterans, the shortened length of follow-up is methodologically more sound because of high mobility of this population and loss to follow-up. Second, because of the in-depth screening process the clinical team questioned whether or not the screening itself produced a change in outcomes. Therefore, patients enrolled into the residential rehabilitation program are administered a second screening questionnaire at the completion of screening/just prior to their coming into the program to detect screening effects. Third, the residential rehabilitation program was changed from two weeks to one week based on veteran feedback.

b. CLINICAL LESSONS LEARNED

1. **FIBROMYALGIA:** Many patients within our population could be formally diagnosed as having fibromyalgia based upon the American Rheumatology Association guidelines. Realization and acknowledgment/ awareness of this has helped us in the development of a more focused treatment protocol.
2. **FEMALE VETERANS:** Request by our first set of female veterans for counseling in specific GYN issues led us up to establish a lecture/session with ARNP/Coordinator of Women's Health Program where different questions could be addressed. This was well received.

3. **MALE VETERANS:** We recognize that there are concerns, not only from the patient(s), but also from their spouses, regarding sexual issues/dysfunction.

4. **AGING PATIENT POPULATION:** Because many of our program participants were "older" at their time of service compared to other veteran groups, we developed a lecture on the topics of dealing with "getting older." What we need to look at, and have them recognize, is that some of their symptoms may be only secondary to the fact that they are getting older, *or* they may have gotten the symptoms anyway but have gotten them earlier secondary to the stress experienced in the Gulf War, *or* the symptoms are only secondary to the Gulf War. We may or may not be able to really tell, but this allows us to educate them on different aspects of medical care (screenings, etc.) as one ages and to help them to put some of their symptoms in context.

5. **SCHEDULE CHANGE—10 Days to 5 Days:** Changes focus from "treatment" to "evaluation and recommendation." We had difficulty with the recruitment section in one aspect secondary to the length of stay of the program. Many of the veterans were interested but could not come for two weeks, and many that participated felt that one week would be better. We changed our focus to a more condensed version of the therapies/lectures, etc. In the shortened program a large amount of the actual treatment (P.T., pool, etc.) was accomplished on the outside, so case management follow-up became more important to assure ourselves that that appointments were scheduled and that patients kept appointments. A residential program is probably not the most effective method of delivering services to Gulf War veterans. Outpatient models should be assessed.

6. **CASE MANAGEMENT:** Low Risk groups of veterans were pleased with case management but did not view themselves as needing it. High Risk veterans were pleased with case management. They used case management for health information and assistance in obtaining services, thus insuring continuity of care.

7. **PROVIDER/VETERAN RELATIONS:** Veterans were generally satisfied with their experiences with VHA clinicians, however, they continued to be frustrated by some providers who did not acknowledge their symptoms, and with some clerical staff who were perceived as being more concerned with rules and procedures that with customer service. Veterans in our program had commented on the positive and empathic attitudes of program staff.

8. **HEALTH EDUCATION PROGRAM COMPONENT:** Veterans were very satisfied with the health education component of our program because it helped them to manage their own symptoms more effectively.

9. **INTERDISCIPLINARY REHABILITATION APPROACH**: We now have a strong interdisciplinary rehabilitation team with expertise working with Gulf War veterans who have chronic unexplained symptoms. Summative evaluation indicates that the veterans valued the interdisciplinary perspective of the program.

10. **USE OF NEEDS ASSESSMENT AND ONGOING FORMATIVE EVALUATION TECHNIQUES IN PROGRAM PLANNING:** Needs assessment techniques and continual input from Gulf War veterans were valuable in program planning. Major findings from veteran input was that they experienced low levels of support from employers and families, and they had prevailing negative feelings and distrust of the VA, Department of Defense, and of the government. In our program they responded positively to health education, support strategies, and attention to pain management.

REFERENCES

1. Aday LA, and Anderson R. 1974. A framework for the study of access to medical care. Health Ser, Res 9.2011-220.
2. Aday LA, and Anderson R. I981, Equity of access to medical care: A conceptual and empirical overview. Med Care 19 (Suppl. No. 12):4-27.
3. Bradham DD, O'Shea M. Resource consumption and cost outcomes in health services research. In Jacobs MD, Nelson A., Berrio MW (Eds.) Outcomes Assessment Tools. Veterans Health Administration Nursing Research Constituency Center. 1998.
4. Bradham DD, South BR, Saunders HJ, Heuser MD, Pane KW, Dennis KE. Obesity-related hospitalization costs to the U.S. Navy -- 1993 to 1998. Military Medicine, modification under review, January 2000.
5. Bradham DD, South BR, Saunders HJ, Heuser MD, Pane KW, Dennis KE. The cost-effectiveness of a shipboard weight-control program. Submitted, 2000.
6. Chapko, MK, Ehreth, JL, & Hedrick, S, Methods of determining the cost of health care in the Department of Veterans Affairs Medial Centers and other non-priced settings. Evaluation of the Health Professions, 1991, 14 (3), 282-303.
7. Donaldson C. The State of the Art of Costing Health Care for Economic Evaluation. Community Health Studies 1990; 24(4): 341-356.
8. Eisenberg JM. Clinical economics - A guide to the economic analysis of clinical practices. Journal of the American Medical Association. 1990. 262(20), 2879-2886.
9. Finkler, SA, The Distinction Between Cost and Charges. Annals of Internal Medicine. 1982, 96,102-109.
10. Gold M.R, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in Health and Medicine. New York: Oxford University Press. 1996.
11. Hornbrook MC, Hurtado AV, Johnson RE. Health care episodes: Definition, measurement and use. Medical Care Review. 1985; 42 (2): 163-218.
12. Lorig, K., Stewart, A., Ritter, P., Gonzalez, V., Laurent, D. & Lynch, J., (1997). *Outcomes measures for health education and other Health Care Interventions.* Thousand Oaks, CA: Sage Publications.
13. Miller, RG. Simultaneous statistical inference. Springer-Verlag. New York. 1981.
14. Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measures, 1*, 385-401.
15. Sevick MA, Bradham DD. Economic Value of Caregiver effort in Maintaining Long-term Ventilator-assisted Individuals at Home. Heart and Lung. 1997; 26 (2): 148-157.
16. Wasson JH, Reda DJ, Bruskewitz RC, Elinson J, Keller AM, Henderson WG. A comparison of transurethral surgery with watchful waiting for moderate symptoms of benign prostatic hyperplasia. The Veterans Affairs Cooperative Study Group on Transurethral Resection of the Prostrate. N Engl J Med. 1995 Jan 12; 332(2):75-9.
17. Weinberger M, Oddone EZ, Henderson WG. Does increased access to primary care reduce hospital readmissions? Veterans Affairs Cooperative Study Group on Primary Care and Hospital Readmission. N Engl J Med. 1996 May 30; 334(22):1441-7.

Report Date October 2000

DETAIL SUMMARY SHEET

TITLE: Case Management and Residential Rehabilitation for Gulf War Veterans

KEYWORDS: Case Management, Rehabilitation, Gulf War

PRINCIPLE INVESTIGATOR: Gail M. Powell-Cope, PhD, ARNP
CO-PRINCIPAL INVESTIGATOR Robert Roswell, MD

VA SITE: James A. Haley Veterans Hospital

STATUS:

Ongoing

X

Complete

APPROVAL DATE: 6/30/98
REVIEW DATE: 10/00

FUNDING: Current FY: $0 Total: $500,000

STUDY OBJECTIVE:

The objective of this 2-year demonstration project was to test an innovative treatment program for Gulf War (GW) veterans. Key components of the program included: (1) unique residential rehabilitation program targeted for at risk GW veterans and (2) random assignment of both at-risk and low-risk GW veterans to either a GS-7 clerk case coordinator or professional case manager. *At-risk GW veterans* are defined as the 15 – 25% of GW veterans who are either dissatisfied with VHA health care or have undiagnosed illnesses. *Low-risk GW veterans* are defined as those who are satisfied with VHA care and who do not have unexplained illnesses. Residential rehabilitation is a one-week program that focuses on health promotion and symptom management through multiple modalities including physical therapy, kinesiotherapy, recreational therapy, stress management and relaxation, and vocational rehabilitation, if needed. Case management is the coordination of care by and advanced registered nurse practitioner focusing on health promotion, patient education, appropriate referrals, and timely follow-up.

A quasi-experimental design was used to assess the impact of this innovative two-component intervention program by comparing seven groups of GW veterans, including comparison groups of at-risk and low risk GW veterans from VISN 8. The groups were compared on critical clinical (self management, self efficacy, fatigue, sleep, pain, anxiety, depression) and health services outcome measures (patient satisfaction, health status, health care utilization, cost). A notable feature of the project design was that it would have provided evaluative data on the residential rehabilitation intervention, case management or case coordination interventions, and a combination of these two interventions. However we were unable to disentangle effects due to either intervention alone or a combination of the two through a factorial design due to small sample sizes.

TECHNICAL APPROACH:
Four changes were made from the original proposal. (1) The follow-up was changed from one year to six months. (2) Screening questions were repeated upon entry into the program. (3) The residential rehabilitation program was reduced from two weeks to one week. (4) Due to inadequate sample sizes the effects of case management cannot be adequately evaluated.

CONCLUSIONS:

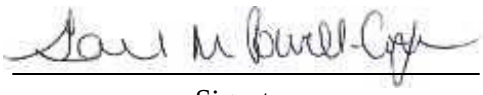
1. As expected, compared to low risk veterans, high risk veterans report more health problems, are less satisfied with VA care, have poorer health, and lower levels of self-efficacy in managing their health. In three areas, at-risk veterans engaged in more self-management behaviors. These data showed that we were able to successfully screen high risk patients using a simple two item screen.
2. For high risk GW veterans, compared to usual care group, comparing baseline to 4-week data the residential rehabilitation group improved in:
   - four areas of self-efficacy (obtaining information about their disease, managing their disease, managing symptoms, and managing shortness of breath), and one area of self management (mental stress management and relaxation)
   - No statistically significant differences were found for the SF-36V measures, the short-term clinical outcome measures, or the Veteran Administration Patient Satisfaction instrument.
3. For high risk GW veterans, compared to usual care group, at six months follow-up:
   - No statistically significant differences were found for the SF-36V measures, the VHA Patient Satisfaction instrument or the short-term clinical outcome measures.
   - Differences in the usual care group versus the intervention group were found for self-efficacy to manage their disease, and use of mental stress management and relaxation.

   Additional data from more subjects (being collected now) may reveal additional program effects.
4. With respect to the cost analysis, the currently available healthcare utilization data from VHA for the veterans in the intervened and comparison groups show that the intervention does not achieve the expected cost-savings benefit by the 9-month post-intervention quarterly period. Additional data from more subjects and a longer follow-up period may reveal the anticipated positive VHA cost-sharing effect.

FY00 FINAL REVIEW OF RESEARCH

Instructions: Please answer the following questions and sign at the bottom of the page. Give an explanation for all negative responses.

YES	NO		
_____	_____	1.	Research files are being maintained by the principal investigator
_____	_____	2.	These files are ready to be inspected as part of the continuing/periodic review process as required by VHA and other federal regulations.
_____	_____	3.	If human use, subject participation or risk has not been influenced by new developments or literature.
_____	_____	4.	If human use, the current risk/benefit ration is about the same (or lower) as when the study was first approved.
_____	_____	5.	If human use, I have reviewed the consent form during this report period to ensure its appropriateness. The consent form has been reviewed and updated, if required, to meet HUC/IRB guidelines.



Signature

10/13/00
Date

PROVIDE A COPY OF THE CURRENT CONSENT FORM AND, IF REQUIRED, A COPY OF THE REVISED/UPDATED VERSION.

Report Date 10/00

FY00 LIST OF PUBLICATIONS

Publications:

Crisfield, J., Callahan, P. The Gulf War Veteran: Biopsychosocial Management. Collaborative Practice, 2000. (incompete citation)

Presentations:

Crisfield, J., Spencer, J., & Powell-Cope, G. Residential Rehabilitation and Case Management for Gulf War Veterans. Gulf War Illnesses Conference, Washington DC, 1999.