

The Utilization of a Chronic Disease Self-Management Program To Improve Health Outcomes and Reduce Health Services Utilization for African-Americans with Diabetes: A Muskegon Community-Based Project

Final Report to the Blue Cross Blue Shield Foundation of Michigan
September 30, 2004

Cynthia Coviak, Ph.D., R.N.

Project history

The Muskegon Community Study of Diabetes Self-Management in African Americans was begun in the summer of 2002, under the direction of Carmen Eribes, Ph.D., R.N. This project, which utilized the Chronic Disease Self-Management Program (CDSMP) of researchers at Stanford University (Lorig, et al., 1999) was designed to achieve

1. Improved health behaviors, measured by stretching and strengthening exercises, cognitive symptom management, and communication with the health care provider.
2. Improved health status, measured by self-rated health, disability, social/role activity limitations, psychological well-being, energy/fatigue, and health distress.
3. Decreased health services utilization, measured by healthcare provider and emergency room visits, number of hospital stays, and nights in a hospital.
4. Improved glycosylated hemoglobin (HgbA1c) levels.

To achieve these objectives, a randomized controlled intervention study was designed. The intervention group participants (a) underwent pre-testing, which consisted of questionnaire completion and measurement of glycosylated hemoglobin A1C levels, (b) attended the 6-week course of standardized class sessions, led by trained facilitators, and (c) completed post-testing of the above measures 6 months after they were recruited. The control group was pre-tested in the same manner as the intervention group, waited 6 months, were post-tested, and then invited to participate in the classes. Both groups received \$15.00 for completing the pre-testing and post-testing, and those who attended classes received the book, *Living a Healthy Life with Chronic Conditions* (Lorig, Holman, Sobel, Laurent, Gonzalez, and Minor, 2000) and the audio relaxation tape demonstrated in class sessions.

Recruitment

The first participants in the study were recruited in the summer of 2002. Participants were recruited from churches, worksite paycheck memos, and flyers posted at various community locations. It was determined that these were ineffective recruitment efforts, and therefore, they were discontinued in October 2002. New recruitment strategies included placement of flyers in nine MacDonal'd's Corporation restaurants and in the MacDonal'd's employees' time sheets, placement of 1300 flyers in medical facilities, and posting flyers at the prescription pickup windows of several local pharmacies. Area beauticians attending the National Kidney Foundation diabetes

education class were taught to make referrals to the program, and *Access Health*, a medical coverage program enrolling low income workers, allowed the Muskegon Community Health Project (MCHP) staff to use their database to target African-Americans with diabetes, and to contact them by phone.

Review of recruitment efforts as of March, 2003 revealed that of all the varied attempts to connect with the African-American community, “word of mouth” recommendations from former participants and from the beauty salons involved in the National Kidney Foundation program were most effective in enrolling study participants. Other recruitment activities included presentations at various African-American meetings and events, a radio and television guest appearance, attendance at Muskegon Community College’s Martin Luther King community breakfast with circulation of flyers, and attendance and personal recruitment efforts at three health fairs. Volunteers organized by *Volunteer Muskegon* were also enlisted to help with recruitment.

The outreach workers made good use of the various festivals and events that occur in Muskegon during the summer, to reach additional African-American community members. The events attended by the recruiters included a concert by L.L. Cool J. at the Muskegon Summer Celebration, a Healthy Hair conference at the Muskegon Holiday Inn, a food commodities give-away, a boxing match, and a gospel fest.

These recruitment activities met with a moderate level of success. However, the long time the control group participants were required to wait before participation in the classes contributed to a great deal of participant attrition. Those who were recruited were excited about the opportunity, and didn’t wish to wait before attending classes. Therefore, there were a number of individuals who originally were recruited but were later lost to follow-up.

Early in the project, individuals who were recruited did not necessarily have their blood drawn immediately and did not engage in the processes of pretest data collection until a few weeks after expressing interest. Often, before having their blood drawn or completing questionnaires, they withdrew from the project when finding out they were randomized to the control group and would have to wait to attend the classes. Some attended disease-specific classes in local hospitals, churches and health clinics, some moved from the area, some lost interest in participation, and others died. Later recruitment efforts (from the summer of 2003, onward) took this attrition into account. The HgbA1c and initial questionnaire data were collected immediately when individuals expressed interest in the study. This led to the situation of having many more potential participants recruited than the classes accommodated, but most who were recruited in this manner did not wish to attend. Overall, usable pretest data were obtained from 291 participants. This represented 71 of the intervention group subjects and 220 control group subjects. Forty-six of the control group subjects with usable data later completed the classes. A large proportion of the control group participants (58.4% of the total sample) never completed classes, even though they were offered the opportunity. The percentage of participants who completed all components of the study was ___% in the intervention group and ___% in the control group.

In the process of collecting data, the research staff contributed a great deal of case-finding to the community and even to those who did not participate in the project. Blood pressure measurements were taken at the time of study entry (although not recorded) as well as the HgbA1c levels. High blood pressure readings were found in several individuals who had no knowledge of their illnesses. In addition, at least one person who had been told by her physician that she was a “borderline” diabetic several

months before interacting with the research team was found to have a high HgbA1c when she was tested. This allowed her to seek care very early in the course of her full-fledged disease, rather than several months later, which is when she was ordinarily scheduled for a return health visit.

Descriptions of the participants

Total sample

The participants in the study represented a wide range of ages, income levels, and years of experience in dealing with the disease. More than a third of the sample (n = 110, 37.8%) was married, but there were many single participants (n = 92, 31.6%). The remaining subjects reported being separated (n = 14, 4.8%), divorced (n = 34, 11.7%) or widowed (n = 41, 14.1%). In addition, they varied in their years of education, numbers of chronic illnesses they had to manage, and insurance coverage available to facilitate disease management. Tables 1 through 4 illustrate the sample's responses to questions exploring these variables.

Table 1

Personal characteristics and health visits of total sample.

Sample characteristic	Minimum	Maximum	Mean (SD)	n ^a
Age (in years)	21	89	50.56 (15.02)	291
Years diagnosed with diabetes	0	61	6.82 (8.70)	271
Age at diagnosis with diabetes	0	79	43.01 (15.06)	267
Visits to health care provider In last year	0	36	4.06 (4.05)	208
Number of chronic illnesses	1	6	2.05 (1.05)	291
Height (in inches)	55	80	66.64 (4.16)	281
Weight (in pounds)	110	373	202.48 ^b (47.56)	272

^a Sample size varies, due to missing and unreported data.

^b Calculated after outliers (2 people with weight > 400 lbs.) removed from analysis.

Table 2

Educational level of total sample

Educational level completed	Count	Percentage of sample
No school education	3	1.0
1-4 years	0	0.0
5-8 years	22	7.6
9-12 years	148	50.9
Graduate Equivalency Diploma	16	5.5
Trade School	27	9.3
Community college	34	11.7
University/ 4-year college	40	13.7
Missing data	1	0.3

Table 3

Income levels of total sample

Annual income Category	Count	Percentage of sample
Less than \$4,999	40	13.7
Between \$5,000 and \$9,999	49	16.8
Between \$10,000 and \$14,999	48	16.5
Between \$15,000 and \$19,999	37	12.7
Between \$20,000 and \$29,999	43	14.8
Between \$30,000 and \$39,999	31	10.7
Between \$40,000 and \$49,999	21	7.2
\$50,000 or more	10	3.4
No response (missing data)	12	4.1

Table 4

Medical insurance coverage of total sample

Insurance type/carrier	Count	Percentage of sample
No insurance	30	10.3
Medicare	97	33.3
Health Maintenance organization	35	12.0
Medicaid	77	26.5
Private insurance	81	27.8
Other types of insurance (<i>Access Muskegon</i> , veteran's benefits, and others)	34	11.7
No response (missing data)	1	.3

Note: Percentages total more than 100% because some participants had more than one insurance type (e.g., Medicare and a private insurance).

Finally, the respondents were asked to indicate their adherence to or failure to adhere to instructions from health care providers regarding blood glucose testing. In relation to this, the amounts of insurance coverage for diabetic medications and supplies, physician and emergency room visits, and hospitalization charges were assessed. Overall, 76.3 % indicated that they had been told to check their blood sugar at home. Of these, all but 12 respondents indicated they did carry out this disease management task. Although respondents were not asked why they did not check blood sugar at home if they had not been advised to do so by their health care provider, 37 people replied to the options provided. The largest group of respondents to this item indicated they did not want to check their blood sugar ($n = 25$), while 4 had no glucometer, 7 had no glucometer strips, and 1 saw "no use" in doing it.

Insurance coverage for health care visits, medications, diabetic supplies and laboratory tests were covered fully for only 42-52.5% of the sample (depending on the health care need). Only 122 (46.9%) respondents indicated their physician visits were fully covered, while 133 (45.7%) enjoyed coverage for emergency room visits and only 141 (48.5%) had full benefits for hospitalization.

Diabetic medications were covered fully for 152 (52.2%) of the participants. Another 78 (26.8%) reported having partial coverage, creating a total of 79% of the sample having coverage of these expenses. Diabetic supplies were covered fully for 148 (50.9%), and partially for 71 (24.4%) of the group. Lab tests were fully covered for 142 (48.8%) of the participants, and partial coverage was enjoyed by 105 (36.1%) members of the study.

The number of visits to health care providers made by the study participants was assessed with two different questions, one of them asking the number of visits in the last 12 months and the other asking the same, but for the last 6 months (this shorter time frame was used for the value evaluated at pretest and posttest). Of interest is that 271 people responded to the first item, but only 208 to the later item. In response to the 12-month time frame the minimum number of visits was 0 and the highest, 52. The mean number of visits, which was 5.45 (SD = 6.84) suggested a skewed response pattern, and indeed, 93% of the sample indicated making 12 visits or less, which is an amount roughly equivalent to once each month. In the later question the minimum number was again 0, with the maximum decreasing to 36 (M = 4.06, SD = 4.05). In comparing the responses to those from the earlier item, a similar visit frequency of one per month (6 visits or fewer in the 6 month period) was identified. The percentage of respondents reporting 6 visits or fewer was 87.5%, which was within 5% of the 12-month response. This presented evidence of a good degree of reliability in the responses offered, even though the numbers of respondents completing the second question were fewer.

Group differences in demographic variables

The intervention group pretest data were usable for 22 men and 49 women. Similarly, there were 107 packets from control group men and 109 packets from control group women. The differences in group characteristics can be examined in Tables 5, 6, 7 and 8.

Table 5

Group differences in personal characteristics and health visits.

Intervention group				
Sample characteristic	Minimum	Maximum	Mean (SD)	n ^a
Age (in years)	23	83	51.2 (12.12)	71
Visits to health care provider In last year	0	40	6.68 (6.66)	68
Years diagnosed with diabetes	.083	40	7.03 (7.36)	67
Age at diagnosis with diabetes	8	71	44.03 (12.80)	67
Number of chronic illnesses	1	6	2.44 (1.10)	71
Height (in inches)	55	80	66.38 (4.79)	69
Weight (in pounds)	116	450	210.62 (59.55)	68

Control group				
Sample characteristic	Minimum	Maximum	Mean (SD)	n ^a
Age (in years)	21	89	50.34 (15.97)	216
Visits to health care provider In last year	0	52	5.06 (6.92)	199
Years diagnosed with diabetes	0	61	6.8 (9.18)	200
Age at diagnosis with diabetes	0	79	42.7 (15.91)	196
Number of chronic illnesses	1	6	1.92 (1.01)	216
Height (in inches)	57	76	63.95 (66.71)	208
Weight (in pounds)	110	470	201.63 (4.49)	202

^a Sample size varies, due to missing and unreported data.

Table 6

Group differences in educational level.

Educational level completed	Intervention group n (%)	Control group n (%)
No school education	1 (1.4)	2 (0.9)
1-4 years	0 (0)	0 (0)
5-8 years	3 (4.2)	19 (8.8)
9-12 years	39 (54.9)	107 (49.5)
Graduate Equivalency Diploma	5 (7.0)	10 (4.6)
Trade School	5 (7.0)	22 (10.2)
Community college	3 (4.2)	30 (13.9)
University/ 4-year college	15 (21.1)	25 (11.6)
Missing data	0 (0)	1 (0.5)

Note: Percentages are calculated from the sample size of the group being reported.

Table 7

Income levels of total sample

Annual income Category	Intervention group n (%)	Control group n (%)
No income	1 (1.4)	0 (0)
Less than \$4,999	11 (15.5)	28 (13.0)
Between \$5,000 and \$9,999	12 (17.6)	37 (17.1)
Between \$10,000 and \$14,999	14 (19.7)	34 (15.7)
Between \$15,000 and \$19,999	8 (11.3)	28 (13.0)
Between \$20,000 and \$29,999	8 (11.3)	33 (15.3)
Between \$30,000 and \$39,999	7 (9.9)	23 (10.6)
Between \$40,000 and \$49,999	2 (2.8)	19 (8.8)
\$50,000 or more	5 (7.0)	5 (2.3)
No response (missing data)	3 (4.2)	9 (4.2)

Note: Percentages are calculated from the sample size of the group being reported.

Table 8

Group differences in medical insurance coverage.

Insurance type/carrier	Intervention group n (%)	Control group n (%)
No insurance	14 (19.7)	20 (9.3)
Medicare	22 (31.0)	74 (34.3)
Health Maintenance organization	8 (11.3)	26 (12.0)
Medicaid	16 (22.5)	60 (27.8)
Private insurance	16 (22.5)	64 (29.6)
Other types of insurance (Access Muskegon, veteran's benefits, and others)	17 (23.9)	17(7.9)
No response (missing data)	1 (1.4)	0 (0)

Note: Percentages total more than 100% because some participants had more than one insurance type (e.g., Medicare and a private insurance).

Examination of study variables

Measurement of the variables

Health behaviors. Health behaviors measured in this study were the estimated minutes of strengthening and stretching exercises and aerobic physical activity, cognitive symptom management, and communication with the health care provider. As self-efficacy is reputed to be a belief system that precedes the performance of health behaviors, and the CDSMP uses theory associated with the concept to effect changes in the management of the participants' diseases, scores on this measure were assessed along with the health behavior scores. The questionnaire items associated with each of these variables are described in Appendix A.

To account for a few limited number of omitted responses in a scale, Lorig and colleagues developed scoring instructions for self-reported health behaviors, health status, and health system utilization that employ the means of the completed items to represent the scale score. In examining the tables where the participants' responses are summarized, the maximum possible score for any scale would usually be less than 5, but in some cases as high as 7 or 10. However, answers related to duration of exercise (stretching or aerobic) were converted to an estimation of minutes, which can be as high as several hundreds. Table 9 summarizes the pretest scores of the total sample on the health behavior measures.

Table 9

Health behavior scores of total sample at pretest

	Minimum	Maximum	Mean (SD)	n ^a
Stretching exercises	0	180	27.68 (46.38)	291
Aerobic physical activity	0	900	92.32 (2.41)	291
Cognitive symptom management	0	4.5	1.53 (.95)	289
Communication with health care provider	0	5.0	2.43 (1.41)	290
Self-efficacy	1	10	6.61 (2.70)	287

In examination of the means obtained for the minutes of stretching and aerobic exercise, their low magnitude are indicative of a sample that is quite sedentary. Although the number of minutes engaged in these activities was quite high for a few individuals, most reported no to little physical activity on a regular basis.

Health status. Health status was measured by self-rated health, self-reported disability, social/role activity limitations, psychological well-being, energy/fatigue, and levels of health distress. As discussed previously, means were calculated for participants' scores on the scales when a specified number of completed answers were

available in the respondents' data. For the disability, social limitations, fatigue and health distress scales, a higher score indicates a higher degree of difficulty in dealing with one's chronic disease. For the self-rating of health, higher scores indicate a rating of health as poorer, while a score of 1 indicates a rating of excellent health. Additionally, psychological well-being is probably better conceptualized using Lorig et al.'s terminology of "impact" to view the scores in the "direction" higher scores would indicate. Someone with "well-being" would have a lower score, and a higher score would indicate greater "impact". Finally, higher energy scores can be interpreted as being more favorable. The questionnaire items associated with each of these variables are described in Appendix B.

Table 10

Health status scores of total sample at pretest

Scale (possible range of scores)	Minimum	Maximum	Mean (SD)	n ^a
Self-rated health (1 to 5)	1	5	2.89 (1.24)	290
Disability (1 to 4)	1	2.88	1.21 (.37)	290
Social limitations (1 to 7)	1	7.0	2.57 (1.63)	290
Psychological well-being (1 to 7)	1	6.7	2.58 (1.44)	290
Energy (1 to 5)	1	4	2.58 (.45)	291
Fatigue (0 to 10)	0	10	4.01 (2.88)	284
Health distress (1 to 5)	0	5	1.66 (1.42)	290

Examination of the values in Table 10 reveal that as a whole, the sample had little difficulty dealing with their disease. Mean scores for each of the scales tended to be in the median of scores indicating a more favorable view of the effects of the illness on their health.

Health services utilization. Health services utilization was measured by the reported number of healthcare provider and emergency room visits, number of hospital stays, and nights in a hospital. The questionnaire items associated with each of these variables are described in Appendix C.

The reported number of visits with health care providers were discussed with the descriptions of the participants. Table 11, below, provides the descriptive statistics obtained for the items concerned with utilization of emergency care, hospital visits, and the length of hospital stays.

Table 11

Health services utilization scores of total sample at pretest

	Minimum	Maximum	Mean (SD)	n ^a
Number of visits to health care provider	0	36	3.01 (3.92)	281
Number of emergency room visits	0	50	.92 (3.20)	287
Number of urgent care visits	0	28	.26 (1.90)	269
Number of hospitalizations	0	5	.28 (.74)	287
Number of hospital night stays	0	90	1.20 (7.43)	284

In the variables above, it should be noted that several (urgent care visits, number of hospitalizations, number of nights in the hospital) were applicable to only a few individuals, creating very skewed distributions. Most of the sample had values of zero (0) entered for these scores. Later analyses involving these variables should be viewed with this in mind.

Glycosylated hemoglobin A1c (HgbA1c). Finally, glycosylated hemoglobin A1c was measured by fingerstick samples analyzed with Bayer Corporation DCA 2000+ Hemoglobin A1 C Machine Model 5031 C, which utilizes a DCA 2000 Reagent Cartridge. For the total sample, the minimum pretest HgbA1c value was 4.3 and the maximum was 14.0. The mean HgbA1c was 7.19 (SD = 2.01, n = 289). HgbA1c levels are categorized with respect to the degree of “control” of blood sugar the values represent. Values of 4 to 6 are considered normal, 6 to 9 is good control for diabetics, and values over 9 are considered to represent poor control. This sample, therefore, had participants with values at both extremes of the scale, but the measures for most were within the range constituting good control for diabetics.

Descriptive comparisons of group differences in variables at time 1

At intake, the responses of participants to individual items of the various scales were examined to determine whether a pattern of bias was evident in any of the response distributions. All of the items measuring the variables of interest were found to elicit the full range of responses allowed for that question. However, many responses regarding methods of dealing with the disease and its symptoms, as well as the items addressing exercise, were skewed to the lower end of the scales. Therefore, it was not surprising that the composite scores for the measures yielded skewed distributions. Handling of these variables for accuracy in data analysis is described under *Hypothesis testing*. The descriptive analyses of health behaviors, health status, and health services utilization variables are summarized in Tables 12, 13, and 14.

Table 12

Health behavior scores of study groups at pretest

Intervention group				
	Minimum	Maximum	Mean (SD)	n ^a
Stretching exercises	0	180	25.99 (49.23)	71
Aerobic physical activity	0	900	91.06 (144.57)	71
Cognitive symptom management	0	4.5	1.50 (0.98)	70
Communication with health care provider	0	5	2.48 (1.46)	71
Self-efficacy *	1	10	6.02 (2.65)	71
Control group				
Stretching exercises	0	180	27.64 (44.69)	216
Aerobic physical activity	0	600	91.60 (115.41)	216
Cognitive symptom management	0	4.5	1.54 (0.95)	215
Communication with health care provider	0	5	2.43 (1.40)	215
Self-efficacy *	1	10	6.76 (2.70)	212

* Indicates a variable found to have significantly different group means (using t-test for independent samples) between the intervention and control groups at pretest.

Table 13

Health status scores of study groups at pretest

Intervention group				
	Minimum	Maximum	Mean (SD)	n ^a
Self-rated health *	1	5	3.49 (1.02)	70
Disability	1	2.88	1.28 (0.44)	71
Social limitations *	1	7	3.08 (1.70)	70
Psychological well-being (impact) *	1	6.4	3.17 (1.53)	71
Energy *	1.6	4.0	2.44 (0.48)	71
Fatigue *	0	10	4.68 (2.51)	65
Health distress *	0	5.0	1.97 (1.44)	71
Control group				
Self-rated health *	1	5	2.70 (1.25)	216
Disability	1	2.88	1.19 (0.34)	215
Social limitations *	1	7	2.41 (1.58)	216
Psychological well-being (Impact) *	1	6.7	2.39 (1.36)	215
Energy *	1	4	2.63 (0.43)	216
Fatigue *	0	10	3.87 (2.92)	215
Health distress *	0	5	1.57 (1.40)	215

* Indicates a variable found to have significantly different group means (using t-test for independent samples) between the intervention and control groups at pretest.

Table 14

Health services utilization scores of study groups at pretest

Intervention group				
	Minimum	Maximum	Mean (SD)	n ^a
Number of visits to health care provider	0	20	3.69 (3.27)	70
Number of emergency room visits	0	8	.83 (1.54)	69
Number of urgent care visits	0	28	.76 (3.71)	63
Number of hospitalizations	0	5	.41 (.99)	69
Number of hospital night stays	0	90	2.78 (11.96)	68
Control group				
Number of visits to health care provider	0	36	2.80 (4.12)	207
Number of emergency room visits	0	50	.97 (3.60)	214
Number of urgent care visits	0	7	.10 (.69)	202
Number of hospitalizations	0	4	.24 (.64)	214
Number of hospital night stays	0	70	.72 (5.24)	212

Note: No health utilization variables differed between groups at pretest.

Hemoglobin A1c

In comparing the intervention and control groups, it was found that the control group's A1c (M = 6.93, SD = 1.74, n = 215, range of 4.3 to 14.0) was significantly lower ($t = 2.94$, $df = 91.49$, $p = .004$) than the A1c demonstrated by the intervention group (M = 7.89, SD = 2.52, n = 70, range of 4.8 to 13.8). However, as the pretest values were to be used as a covariate in the final analysis, these group differences were not of concern, as they would be controlled for in assessing change as a result of the intervention.

Hypothesis testing

As the analysis plan for this project was a factorial (intervention group and control group) repeated measures design (pretest variables vs. posttest variables), there were

a number of statistical assumptions that should be met by the distribution of the scores obtained to enable an accurate analysis. Therefore, the distributions for the health behavior, health status, health care utilization, and glycosylated hemoglobin variables were examined to verify these assumptions were met and, if not, to determine if transformations of the data through mathematical operations was possible. Specifically, it was important to determine whether skew or kurtosis was excessive for any of the variables.

Distributions obtained with group data were found to include some variables that were, indeed, skewed. To reduce these problems, variables were transformed through square root and logarithmic functions, depending on the severity of the deviation from normal, in accord with the recommendations of Tabachnick and Fidell (2001). Then a factorial (intervention group vs. control group) repeated measures (pretest vs. posttest) analysis of variance (ANOVA) was performed with either the raw score or transformed score. An alpha (p) of .05 was established as the level of significance, but values between .01 and .05 were interpreted with caution due to the number of tests run, and the possibility of an “inflated” degree of error. However, when significance was found, the alpha was usually less than .01. Thus, most of the analyses finding significant results can be viewed with confidence.

Health behavior scores

The first group of variables tested were the health behavior variables—stretching exercise frequency, aerobic exercise frequency, cognitive symptom management, communication with health care provider, and self-efficacy. Of these variables, the two physical activity variables required square root transformations to correct the skew of the distributions. In the analyses, 71 intervention group participants and 160 (aerobic activity) or 159 (stretching exercises) control group participants had complete data sets allowing hypothesis testing. Neither variable showed significant change from pretest to posttest, nor were there differences between the groups. Power was extremely low, (less than .063 in all analyses) so it is likely that the sample size was too small or group sizes too disparate for testing these variables. However, as exercise frequency is only one of the goals a CDSMP participant can decide to concentrate on, and it was not a specific target of the classes, these results are consistent with the low degree of emphasis on these behaviors in the classes. Further, a large body of literature as well as the U.S. Public Health Service’s *Healthy People 2010* identify exercise as a particularly problematic health behavior to address.

The other three health behavior variables did not require transformations prior to analysis of results. To evaluate *communication with health care providers* 71 intervention and 158 control group participants’ data were analyzed. Although the change in the scale scores from pretest to posttest was significant (Wilk’s Lambda = .950, $p < .001$, power = .928 with alpha = .05), both groups realized an increase in this score. To control for the variation accounted for by pretest values, an analysis of covariance (ANCOVA) was then performed with the pretest communication score as the covariate, the posttest score as the dependent variable and the project group as the factor. No group differences were found in the posttest score ($F = 2.77$, $df = 1$, $p = .097$).

The analysis of *Cognitive symptom management* included 69 intervention group and 159 control group participants. Repeated measures analyses did not reveal any changes, either from pretest to posttest, or between groups. Power again was found to

be low (.192, calculated for $p = .05$), indicating a sample size too small to allow definitive conclusions about the hypothesis testing.

An additional consideration in the analysis of the cognitive management variable was the overall low prevalence of this behavior in the sample as a whole, and in the two groups. As the score for this behavior was low, it can be assumed that there was little opportunity or need for the strategies the scale represents (“not applicable” cases would have been converted to 0, thereby lowering the sample and group means). If the need is low, the need to change the behavior was low, making the finding understandable.

Among the health behavior variables, the final measure tested was self-efficacy. This cognition regarding the self is felt to be foundational to the performance of a number of health behaviors. It is a belief system that is explicitly targeted during the Chronic Disease Self-Management classes, thus, it is a variable that is expected to be one of the first to show change following class attendance.

The self-efficacy variable in this study did not require a transformation to normalize the distribution, so the raw scores were analyzed. The intervention group (71 participants) and control group ($n = 156$) both showed an increase in this score. As was found for the pretest scores, the control group’s posttest score was higher than the intervention group’s score. To statistically eliminate any effects of pretest differences in the final analysis, an ANCOVA was performed, with pretest score as covariate, and group membership the factor. With this it was found that the control group’s score ($M = 7.15$) was still significantly higher than the intervention group’s score ($M = 6.25$; $F = 5.6$, $df = 1$, $p < .019$). This remains a contradictory finding.

Health status scores

The measures of health status for the study were the next group of variables analyzed. Self-rated health, disability, social limitations, disease impact, and levels of energy, fatigue, and health distress were examined for change at posttest. All of these variables except disability had been found at pretest to be significantly different between groups, with greater social limitations, fatigue and distress, and lower psychological well-being in the intervention group, while the control group had increased energy scores and higher ratings of health than the intervention group.

Self-rated health did not require a transformation operation for appropriate analysis. The scores for the intervention group ($n = 70$) were significantly higher ($M = 3.49$, $SD = 1.02$) than the scores for the control group ($n = 160$, $M = 2.96$, $SD = 1.18$), indicating a rating of personal health poorer than in the control group. Both the within subjects’ effects (Wilk’s $\lambda = 18.54$, $p < .000$, power = .99) and between subjects’ effects were highly significant ($F = 30.03$, $df = 1$, $p < .000$, power = 1.0). Further analysis using ANCOVA was performed to control for the effects of pretest differences in the testing, and the intervention group’s scores remained significantly higher than those of the control group ($F = 18.98$, $df = 1$, $p < .000$).

A log transformation was performed on disability scores before groups were compared. An intervention group size of 71 and control group of 159 participants were available to analyze. To assess within-subjects changes, repeated measures ANOVA was performed, then the effects of significant differences in scores at pretest were controlled for with ANCOVA using the pretest score as the covariate. No within-subjects or between subjects differences were found.

The next cluster of health status measures were analyzed using raw scores, as their distributions were not significantly skewed. They were analyzed in the manner described previously, and it was found that the groups differed in most of the measures. Table 15 summarizes the results of the tests of differences among these variables.

Table 15

Group differences in health status at posttest

Within-subjects analyses						
Variable	Statistic	Value	F	Hypothesis df	Error <i>d.f.</i>	<i>p</i>
Social Limitations	Pillai's trace*	.060	14.64	1.00	228.0	.000
Energy	Wilks' lambda	.974	6.05	1.00	229.0	.015
Fatigue	Wilks' lambda	.934	15.01	1.00	214.0	.000
Health distress	Wilks' lambda	.970	6.95	1.00	228.0	.009
Psychological Well-being	Pillai's trace*	.068	16.51	1.00	227.0	.000
Between subjects analyses (ANCOVA)						
Variable	Intervention M (SD)	Control M (SD)	F	<i>d.f.</i>	<i>p</i>	
Social Limitations	2.61 (1.52)	1.91 (1.25)	13.27	1	.000	
Energy	2.58 (.44)	2.66 (.36)	3.01	1	.084 (n.s.)	
Fatigue	3.67 (2.43)	2.99 (2.55)	3.47	1	.064 (n.s.)	
Health distress	1.89 (1.37)	1.20 (1.09)	16.18	1	.000	
Psychological well-being	2.61 (1.30)	2.05 (1.10)	10.38	1	.001	

* Pillai's Trace was used when the Box's M test of equality of covariance matrices was significant (assumption violated) as it is more robust to violations of assumptions.

The within-subjects analyses consistently revealed that there were changes in scale scores from pretest to posttest for each group. However, it was unexpected that the group differences at posttest would continue to favor the control group in the social limitation, health distress and psychological well-being measures.

Health services utilization.

Health services utilization measures were another group of variables found to have very skewed distributions. Log transformations were performed, but it was found that in the mathematical calculation a number of cases were dropped by SPSS because the operations required were not always valid. Square root transformations were then tried, and successful, although they did not decrease the skew to the extent the log transformations should have. The transformed variables were then analyzed with repeated measures and ANCOVA techniques, similar to those done with health status variables. The procedures were repeated with untransformed variables, and compared to the data obtained with transformed. All of the procedures produced similar results, so untransformed data are reported here for ease of interpretation of the actual values.

Table 16

Group differences in health status at posttest

Within-subjects analyses						
Variable	Statistic	Value	F	Hypothesis df	Error d.f.	p
Health care visits	Pillai's trace*	.002	.540	1.00	219.0	.463 (n.s.)
Emergency room visits	Pillai's trace*	.004	.925	1.00	224.0	.337 (n.s.)
Urgent care visits	Pillai's trace*	.034	6.41	1.00	183.0	.012
Number of hospitalizations	Pillai's trace*	.002	.401	1.00	225.0	.527 (n.s.)
Number of nights in hospital	Pillai's trace*	.004	.993	1.00	220.0	.320 (n.s.)
Between subjects analyses (ANCOVA)						
Variable	Intervention M (SD)	Control M (SD)	F	d.f.	p	
Health care visits	4.59 (5.92)	1.83 (2.09)	25.83	1	.000	
Emergency room visits	.826 (1.54)	1.07 (4.14)	.746	1	.389 (n.s.)	
Urgent care visits	.122 (.526)	.037 (.308)	.667	1	.415 (n.s.)	
Number of hospitalizations	.430 (.790)	.133 (.393)	15.50	1	.000	
Number of nights in hospital	2.06 (8.42)	.333 (2.06)	5.88	1	.016	

* Pillai's Trace was used when the Box's M test of equality of covariance matrices was significant (assumption violated) as it is more robust to violations of assumptions.

As Table 16 reveals, only the urgent care visits variable exhibited within-subjects changes from pretest to posttest, while 3 of 5 variables showed group differences in the types of utilization the variables represented. Despite controlling for pretest differences, the number of visits to health care providers, the number of hospitalizations, and the number of nights spent in the hospital were higher in the intervention group than in the control group. This was another unexpected finding. Trips to urgent care settings did decrease for both the intervention and control group, but in numbers that kept both groups' utilization similar for this variable at posttest.

Power estimates for the variables not found to achieve significance were, as would be expected, poor. Most had values less than .40. In the variables attaining significance, the power for the health care visit between subjects result was .96 and for the number of hospitalizations between subjects result it was .93. The power for the significant within-subjects result for urgent care visits and for the between subjects result for number of nights of hospitalization were .71 and .73, respectively.

Hemoglobin A1c levels

The glycosylated hemoglobin A1c levels at posttest were available for 70 intervention group and 159 control group participants. At that point, the mean value for those in the intervention group was 7.3 (SD = 1.91) and 6.89 (SD = 2.09) for the control group. As previously noted, the control group had a lower mean HgA1c at pretest (6.93), which was to be controlled for in the final analysis. When the ANCOVA was performed with the pretest value as covariate, the difference between the two groups was no longer significant. The control group's value remained lower than that of the intervention group, but the control group's level at posttest was not significantly lower than the intervention group's level at posttest when corrected by the covariate. In effect, the control group had a significantly lower A1c at pretest, but the intervention group "made up for" that difference at posttest. The groups were no longer different in respect to this important index of diabetic control.

In addition to the findings from hypothesis testing, there were other changes detected at posttest. As indicated, the intervention group's mean A1c was decreasing. In addition, the variability of the scores in that group was decreased at posttest with standard deviations changing from 2.52 to 1.91, while in the control group the change in mean A1c was from 6.93 to 6.89 and variability was increasing from 1.74 to 2.09. Further, the maximum A1c value at posttest for the intervention group was 13.7, while in the control group the maximum increased to the extremely high level of 19.7. (This individual was referred to physician for immediate care.) Therefore, although groups did not differ, the trends seen in the intervention group were promising.

Discussion of findings

The most striking finding for this study was the number of times the control group was found to have significantly better scores on the study variables than the intervention group. These differences often originated at pretest and prevailed throughout the 6 months the participant was followed. Particularly in the case of the health status measures, the differences did not change to favor the intervention group participants, and therefore the effect of the intervention is best evaluated through attention to the

within-subjects analyses. For the most part, they indicated that the health status of the class participants improved during the study. What is less clear is why this also occurred with the control group.

It should be noted that the differences in the groups may have originated somewhat from the participants' patterns of enrollment in the study. Nearly all measures of health status indicated the intervention group had perceptions of greater difficulty with the disease effects. Early on, participants who were quick to join the study may have been doing so because of motivation to try to make their burden of disease management easier. When randomized to the class or wait-group, many participants who had been eager to take the classes were disappointed to find out they had to wait, and many in the meantime went to classes elsewhere. Some of these may have remained in the control group, but others were never to return for classes in this project, and some never to complete data collection. Therefore, groups of "sicker" control group participants may have been eliminated from the sample, through attrition. It was found that participants recruited later in the study, who required more efforts of the staff to become involved, were not easily persuaded to attend classes, even if they had been randomized to the intervention group. It is likely that these individuals may not have had the motivation to attend, because of their perceptions of few limitations or difficulties because of their disease. In effect they became control group members because they never attended classes. However, because they were repeatedly contacted to have arrangements made for their classes, the interaction of the staff with these participants may have been enough positive experiences to foster continued perceptions of minimal disease difficulty.

Similar dynamics may have been operative in the case of the self-efficacy and communication with health care providers variables. In the first testing period the participants may have scored higher on self-efficacy because of their more favorable health status at that time. Later improvements could have occurred because of testing effects (essentially, recall of the earlier questions), but also because of interaction with the recruitment staff. They were enthusiastic about the program, and their enthusiasm may have been communicated to the control group participants as they waited for classes to begin. Possibly, personal pride in participating in an effort that could be helpful for improving the health of other African Americans may have also increased their perceptions of self efficacy.

Circumstances contributing to the scores obtained for communication with health professionals are more difficult to discern. To a certain extent, contact with the recruitment staff may have influenced control group members. Many recruitment events occurred in public venues, and at times blood pressure or HgbA1c values were discovered that prompted the staff to be firm in their advisement to the participant to seek medical care as soon as possible. These interactions may likely have been witnessed, and the confident urgings of the staff to the individual may have influenced the people present to interact with their health care professionals more confidently. However, these ideas are primarily conjecture.

Fewer differences were found between the intervention and control groups in the health services utilization variables, however some were seen in the realms of health care visits, numbers of hospitalizations, and nights of hospitalization. In retrospect, for this population, an increase in the number of health visits would be a positive change, for many of its members. Those who usually spent little time carrying out preventive health procedures may have increased their efforts as a result of attending the classes, and these efforts may be what is reflected in the posttest results. This is not, however,

likely to be true of the hospital utilization variables. The number of hospitalizations represented by the indices obtained constitute less than ½ of one hospitalization episode, in each case. There were also few people who were hospitalized, overall. The number of hospital nights in the posttest measure were greatly increased by the influence of the long hospital stays of just a handful of people. There were outliers for this measure, in the intervention group. Overall, given that the intervention group had reported more difficulty from their disease throughout the study, it is not surprising that some individuals experienced hospitalizations for care.

As was discussed to a certain degree in the report of findings, the promising result obtained in the study was the downward trend for the HgbA1c levels of the intervention group. Considering that the intervention group began the study with a higher mean A1c, to move the group average to equivalency with the lower values found in the control group was very positive. In contrast to many of the other variables measured for the study, a great deal of emphasis on the HgbA1c would likely be perceived by the participants. Unlike a number of the other health indices of interest in the study (e.g. height and weight), the A1c was measured directly, rather than being collected through self-report. It represented a variable of common emphasis, rather than the “menu” of other areas that participants could choose from to make changes in their habits. In addition, this is a variable that would be of great importance to health care providers, and for the individual’s long term prognosis. Therefore, the changes seen in the study are sensible.

The identified variables of interest for the project met with mixed levels of success in attaining statistical significance. However, in any health care research with human beings, there must be a consideration of not only the statistical significance found, but also the “clinical significance” of the study. This project does not differ from others in this respect.

The contributions of the recruiting staff to case-finding and outreach to the community for health concerns has already been mentioned. This fact requires further emphasis. In several cases, while the staff were recruiting participants, people were found who were experiencing dangerous blood pressure levels, out-of-control blood sugar levels, and other health concerns. Many times the individuals had no idea of the existence of the problem, much less the importance of immediate treatment. By providing the guidance needed to obtain appropriate care, the MCHP did much to enhance the health of the community.

In addition, the project served to stimulate greater interest and emphasis on health matters among the class participants. Following their formal 6-week sessions, at least two different class groups (possibly more) continued to meet on a weekly or monthly basis at the homes of the various class members. The meetings included continued support of each other’s health goals, sharing of health-conscious foods, and encouragement while members approached health issues of concern. These are just a few examples of “clinical significance” contributed by the project, which cannot be ignored. Perceptions of the helpfulness of the classes and the relationships with other participants led to long-term commitments to continue to work at disease management.

Limitations

Most longitudinal clinical studies suffer from the common problems of inadequate sample sizes, attrition, and inequality in the group sizes that result from the attrition and

other factors. This study shared these problems, despite intensive efforts to avoid them. Fortunately, for many analyses, acceptable levels of power were attained. For at least some of the cases where this statistic was not acceptable (e.g. in the analyses including exercises), the failure to achieve statistical significance during hypothesis testing makes sense when the broader literature concerning the variable is considered, therefore lending credibility to the acceptance of the null hypothesis.

There were a variety of additional limitations of the study. A number of them are summarized below.

- Most variables measured by self-report, rather than by direct measurement. This could have been altered for height and weight.
- Some measures of study effects could have been re-evaluated at posttest, but were not (weight, height, reported use of glucometer at home, reasons for use/non-use of glucometer). A number of these other measures could have picked up incremental changes in health behaviors, before the more robust changes were evident.
- Did not measure the knowledge of diabetic management techniques held by the participants. Although the classes are not designed to provide content, this variable would be expected to affect performance of activities necessary to achieving control, thus a factor affecting the HgbA1c levels.
- Responses to health utilization questions were confusing to many participants and data quality for these may have been compromised.
- It would have been helpful to add the number of classes attended to the database.
- Overall, the time frame for the project as originally planned was too short. Recruitment of African Americans for research is a task that requires time to develop a trusting relationship, allow for “word of mouth” spread of information, and interact with the community and families, as they are integral to the values of African Americans.
- In addition to increasing time for recruitment, training of a greater number of class facilitators should be planned. This would decrease the wait time before participants would enroll, and allow those who fulfill roles of both recruiter and class leader to rotate “off” from class schedule to devote time to the recruitment phase.
- Greater planning for class leader loss due to attrition should be considered.
- Consideration of alternate designs for research which don’t require randomization to groups. Wait time was discouraging to control group members and contributed to attrition. Essentially, many who didn’t get to classes as they desired, didn’t participate at all.

Implications

The CDSMP, as implemented for this study, holds promise for disease management of African Americans with diabetes. Although many health status variables were not found at posttest to be more favorable in the intervention group than in the control group, many did show improvement within the group. Trends seen in the changes in HgbA1c from pretest to posttest in the intervention group suggest the health management behaviors necessary for good diabetic control were developing, but may not have been fully operational when the follow-up evaluations were being done.

From a community standpoint, the project left some positive effects. Class participants have communicated their successes to the community. Informal follow-up support groups were established by some class groups. Greater awareness of the effects of the disease has resulted.

A final implication from the study results from the process of engaging African Americans in research. Researchers must plan careful methods to engage this minority group in their studies, and be mindful of the values of this community when choosing designs. Randomization methods, in particular, should be given careful consideration to avoid loss of participants who had specific reasons for engaging in the research.

Recommendations

A number of recommendations can be assumed from a review of the limitations of the study. But more substantive comments are put forward concerning study design.

First, the timeline for the study was unrealistic. From the standpoint of recruiting this minority group, time is necessary for establishing trust, engaging the community, and learning specific community priorities and values. There are subtleties of the culture that must be learned to successfully involve them in the research endeavor. For example, in this study, it was discovered that the waitlist control group contributed to attrition. In an early study by Baranowski and associates in the 1980s, it was found that randomization disrupted the ties that this culture values with family, friends, church, and thus, attrition was seen in that study even with a greater number of incentives for participation (such as child care, other benefits). Allowing self-selection to groups may not be the “scientific” standard, but may in the long run reduce attrition, which is as great or more of a threat to research as the attrition. The variables expected to confound the findings can be used for a regression model analysis plan; this uses the same general linear model as ANOVA, ANCOVA, and MANOVA, MANCOVA.

An additional recommendation for future consideration is the use of more specific activities and measures for diabetes, including new questionnaires from Stanford for the CDSMP specific to diabetes. At the start of this study, work on the disease-specific courses was still in progress. Future work can build on the findings Stanford has now obtained.

References

Lorig KR, Sobel DS, Stewart AL, Brown Jr BW, Ritter PL, González VM, Laurent DD, Holman HR. Evidence suggesting that a chronic disease self-management program can improve health status while reducing utilization and costs: A randomized trial. *Medical Care*, 37(1):5-14, 1999.

Tabachnick, B.G. & Fidell, L. S. (2001). Using multivariate statistics (4th ed.). Boston: Allyn and Bacon.