Follow-up Study of an Urban Family Medicine Home Visit Program

Donald J. Balaban, MD, MPH, Neil I. Goldfarb, Robert L. Perkel, MD, and Barbara Lepidus Carlson, MA
Philadelphia, Pennsylvania

A home visit program was established by a large urban family practice in an academic setting. At the program's inception, 198 patients were randomly assigned to either an experimental group, to be eligible for home visits, or a control group, to continue receiving only office-based care. Two years after this randomization, follow-up data were obtained on 194 of the 198 subjects to assess the program's effectiveness. Fifty-one of the subjects had died. There were an increased number of deaths in the experimental group (30 percent) compared with the control group (21 percent), although this difference was not statistically significant. No statistically significant differences were found between the remaining 143 experimental and control group patients in function or well-being. Patients in the experimental group had a significantly higher number of hospitalizations, although there was no difference in the number of days spent in the hospital. Although methodologic considerations limit the ability to draw policy conclusions from this follow-up study, this home visit program did not have a measurable sustained impact on health outcomes or utilization of health services.

A number of different models for house call or home visit programs have been described in the primary care literature.[1–6] The use of home visits has been proposed for various populations, including heart attack victims,[7–12] stroke patients,[13,14] psychiatric patients,[6] and different pediatric populations.[15–20] In particular, the literature contains numerous accounts of the potential positive impact such programs may have for the generalized elderly and debilitated population.[1–3,21–23] Although studies of home care have repeatedly attempted to uncover the benefits to patients, physicians, or the health care system,[1–3,5,24–26] none has determined conclusively whether making home visits has an impact on overall health outcomes.

In 1981 an urban family practice in an academic setting established a home visit program to provide medical care to patients who potentially could benefit from home visits.[13] Care provided by a program physician and nurse, often accompanied by family practice residents, medical students, nursing students, and other health care providers, included diagnostic and therapeutic medical care, posthospitalization follow-up, education and counseling, and social service referrals. The intention was to improve the function and well-being of the patient and the family.

Because of the large existing patient practice, it was decided to limit the home visit program to patients already enrolled in the practice. Eligibility criteria for the home visit program were (1) partial or total disability to the extent that mobility is seriously impaired, (2) living alone and aged over 65 years, (3) not likely to maintain contact with physician, (4) major expenditure of energy and resources required to get to physician, (5) chronic debilitating disease, (6) contact with social support network desirable but difficult to obtain through office visits, and (7) critical aspects of the patient database obtainable only through home visits. In June of 1981, before the home visit program formally began, all residents and faculty were asked to consider which patients would meet any one of these criteria.

At the inception of the home visit program, a study of program effectiveness and efficiency was planned. Eligible patients were then asked to participate in the study and, following consent, were randomly assigned to either the experimental group or the control group. Home visits were scheduled for experimental group patients based on an

Submitted, revised, October 7, 1987.

From the Greenfield Research Center, Department of Family Medicine, Jefferson Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania. Requests for reprints should be addressed to Dr. Donald J. Balaban, Greenfield Research Center, Department of Family Medicine, Jefferson Medical College, 1015 Walnut Street, Philadelphia, PA 19107.

assessment of medical and social needs; these patients were not precluded from office visits with their usual family physicians. Control patients, not eligible for home visits, received their office-based care with their usual family physicians.

Between July 1981 and June 1982, 198 patients were enrolled in the study and randomly assigned to experimental and control groups. Because of the lack of prior research experience of the home visit team and the limited availability of resources for research, neither baseline nor follow-up data were collected. In August 1983, following the formation of a research division in the department, funding was obtained to collect health outcome and utilization information on all randomized patients. The primary question for this follow-up study was, can differences between experimental and control group patients relating to functional status, psychosocial status and well-being, mortality, and utilization of health services be identified?

METHODS

The intention of this study was to collect follow-up information on each study subject. The following instruments were used in the assessment:

1. The Quality of Well-Being (QWB)[27] index of general health status
2. The Barthel Activities of Daily Living[28] measure of function
3. A questionnaire on utilization of health services (inpatient, outpatient, and long-term care) for the 365 days prior to the interview (developed by the investigators)
4. A questionnaire on patient characteristics and attitudes that may influence health outcomes, developed by the investigators based on methods developed at the National Center for Health Services Research
5. The Patient Satisfaction Questionnaire,[29] the humaneness of care, continuity of care, and general satisfaction with health care subscales
6. The 13-item version of the Beck Depression Inventory[30]
7. The Philadelphia Geriatric Center Morale Scale,[31] a measure of patient mood and motivation
8. A global health status visual analog scale developed by the investigators for this study.

Data collection began with an abstract of the Family Practice Center patient charts for all 198 study participants. The abstract included administrative data for contacting patients, data on utilization of family medicine physician services (office and home visits), hospitalizations and nursing home placements as documented in the chart, patient status (living, deceased), and major diagnoses as noted in the chart.

Two research assistants conducted in-person interviews for all patients who could be contacted. Proxy and telephone interviews, using a format consisting of the Barthel Activities of Daily Living, the Utilization Questionnaire, and an abbreviated version of the Patient Characteristics Questionnaire, were conducted when necessary. Data were collected on precoded forms and entered in a microcomputerized database and were analyzed using microcomputer statistical programs (descriptive statistics, chi-square, Pearson correlation, and *t* tests) developed in Metafile[32] by the investigators.

TABLE 1. SUMMARY OF CONTACTS FOR STUDY INTERVIEW (n = 198)

	Experimental (n = 103)	Control (n = 95)	Total (n = 198)
	No. (%)	No. (%)	No. (%)
Patient interviews	49 (48)	60 (63)	109 (55)
Family Practice Center	28 (27)	33 (34)	61 (31)
Patient's home	12 (12)	12 (13)	24 (12)
Over telephone	9 (9)	15 (16)	24 (12)
Proxy interviews	20 (19)	14 (15)	34 (18)
Deceased	31 (30)	20 (21)	51 (25)
Refused interview	2 (2)	0 (0)	2 (1)
Could not contact	1 (1)	1 (1)	2 (1)

RESULTS

Of the 198 study patients, 51 had died (Table 1). The interviewers were able to collect information on 143 of the remaining 147 patients: 109 interviews were conducted directly with the patient (61 in the Family Practice Center, 24 in patients' homes, and 24 over the telephone); 34 proxy interviews were conducted with a close relative, friend, or health care provider; two patients refused to be interviewed; and two could not be contacted. The reasons for proxy interviews included senility (six experimental and five control), retardation (two experimental and four control), severe psychiatric illness (two experimental and one control), severe hearing loss or deafness (two experimental and two control), aphasia and other communication problems (five experimental and one control), and could not be reached or was "too sick" (three experimental and one control).

Analyses and examination for differences between experimental and control patients were conducted on the entire population (n = 198) and on three study population subgroups defined by the amount and type of data available: the deceased patients (n = 51), the interview (in-person, telephone, and proxy) population (n = 143), and the in-person interview population (n = 86).

The entire study population included 198 subjects, 103 (52 percent) in the experimental group and 95 (48 percent) in the control group. Of these 198, 150 (76 percent) were female and 137 (70 percent) were black. The control group had slightly more female (78 vs 74 percent) and black patients (73 vs 66 percent) than the experimental group, although these differences were not statistically significant.

TABLE 2. SUMMARY OF MAJOR CHRONIC DIAGNOSES RESULTING FROM CHART AUDIT
(Many patients have more than one chronic condition)

Diagnosis	Experimental* No. (%)	Control* No. (%)	Total No. (%)	P (chi-square)
Hypertension	51 (50)	61 (64)	112 (57)	.052
Arthritis	30 (29)	31 (33)	61 (31)	>.7
Diabetes	29 (28)	23 (24)	62 (31)	>.6
Arteriosclerotic heart disease	28 (27)	10 (11)	38 (19)	.002
Depression	23 (22)	18 (19)	41 (21)	>.6
Congestive heart failure	24 (23)	18 (19)	42 (21)	>.5
Total	103 (52)	95 (48)	198 (100)	

** The average number of diagnoses per patient for the control group = 5.3, for the experimental group = 5.1.*

There was no statistically significant difference in the mean age of the experimental (69.2 years) and control group (67.7 years) patients. Patients ranged in age from 17 to 99 years at the time of study enrollment, with 143 (72 percent) of the patients aged 65 years or over. Younger patients tended to be less mobile, wheelchair bound, or terminally ill, whereas older patients were included based on broader social support and illness criteria.

Diagnostic information abstracted from the patient charts (Table 2) indicated that more control patients had hypertension (P = .052) and arthritis (P > .7) than experimental patients, whereas more experimental patients had arteriosclerotic heart disease (P = .002), congestive heart failure, diabetes, and depression (P > .5 for each) than control patients. That 28 percent of the experimental patients had heart disease, compared with only 9 percent of the control patients, indicates that the experimental patients may have had a greater degree of illness. There was no significant difference between the groups in the average number of diagnoses recorded per person.

Of the 103 experimental patients, 95 received one or more home visits over the study period, as determined by reviewing family medicine patient charts and home-visit program administrative records. The remaining eight patients either dropped out of the practice, moved, or died before any home visits could be scheduled. Of the 72 surviving experimental patients, 42 (58 percent) were still receiving home visits at the time of this follow-up study. A total of 35 home visits were made for 29 of the 95 control group patients (30 percent) some time after the initial randomization; most of these visits were made early in 1983, prior to the decision to conduct this follow-up study.

More deaths occurred in the experimental group (n = 31) than in the control group (n = 20). The difference in mortality rates, as reported in Table 1—30 percent for experimental vs 21 percent for control group—is not statistically significant (chi-square = 1.67, P = .20). Analysis of the dates of death, age at death, and time from study enrollment to death did not reveal any significant differences between experimental and control patients.

There were no significant differences in demographic characteristics between the experimental (n = 69) and control patients (n = 74) in the interview cohort (n = 143), nor were there differences between the interviewed and noninterviewed populations. The average age of the interviewed population at randomization was 66.6 years, slightly younger than the mean age for the entire study population (68.4 years).

The mean score on the Barthel Activities of Daily Living for the experimental group patients (77.9, SD = 26.8) was slightly worse than that for the control group (83.0, SD = 23.4), but the difference was not statistically significant (P > .2).

The number of Family Practice Center office visits and home visits, based on information recorded in patients' charts, is displayed in Table 3. The 1980 figures indicate that prior to the program's inception, the experimental patients had slightly fewer visits to the Family Practice Center per year than the control patients. After the home visit program began, overall utilization by experimental patients rose above that by control patients (6.9 vs 5.4 total home and office visits in 1982).

Information on number and length of hospitalizations, based on patient and proxy interviews, is displayed in Table 4. Experimental group patients had significantly more hospitalizations (mean of 1.17 per patient) than the control group patients (.64) over the year prior to the interview (two-sided *t* test, P < .003), although there was no significant difference between groups in the total number of days hospitalized.

The in-person interview population included all patients who completed the full battery of questionnaires (four of whom were interviewed over the telephone). Forty experimental group patients (39 percent of all patients randomly assigned to the experimental group) and 46 control group patients (48 percent of all patients randomly assigned to the control group) are represented in this cohort. There were no statistically significant differences in sex, race, or age between the experimental and control patients in the interview cohort. Comparing the interview population (n = 86) with the entire study population (n = 198), the interview population had slightly more female patients (83 vs 76 percent), more black patients (72 vs 70 percent), and was younger than the entire study population (mean age of 66.7 years vs 68.4 years).

TABLE 3. MEAN NUMBER OF HOME AND OFFICE VISITS, AS RECORDED IN MEDICAL CHART, OF INTERVIEW, AND PROXY POPULATIONS (n = 143)

	Experimental (n = 69)			Control (n = 74)		
Year	Home	Office	Total	Home	Office	Total
1980	0.0	4.2	4.2	0.1	4.9	4.9
1981	2.0	5.2	7.2	0.1	6.6	6.6
1982	3.8	3.1	6.9	0.2	5.2	5.4
1983	2.5	1.7	4.2	0.9	3.8	4.6

A comparison of the experimental and control groups on selected outcome measures is summarized in Table 5. There were no statistically significant differences between the groups. Comparison between the control group patients who did and did not receive home visits did not reveal any significant differences between these two groups. Excluding the control group subjects who received one or more home visits from the experimental-control comparison did not change any of the findings significantly.

Correlations between the outcome measures and age were investigated. The direction of the correlation for each pair of variables was as expected. The strongest correlation was between the Philadelphia Geriatric Center Morale Scale and the Beck Depression Inventory ($r = -.69$: depression is inversely related to morale). The morale scale also correlated highly with the Quality of Well-Being ($r = .41$) and Global Health Status Visual Analog ($r = .41$).

DISCUSSION

This follow-up study was undertaken because a patient population had been randomly assigned to experimental and control groups, and to learn as much as possible about this intervention seemed appropriate, given the increasing attention to development of programs that provide in-home care to the elderly and chronically ill and the limited rigorous information on the benefits and costs of such programs.

This follow-up study did not demonstrate a beneficial effect for home visit patients in comparison with patients receiving traditional office-based care. The intervention did not seem to have a measurable impact on mortality or use of health services, and although measurements of health status and function are less precise than measures of mortality or utilization of health services, the small differences noted between the surviving experimental and control group patients suggest that the intervention had little impact on these outcomes.

The increased number of deaths in the experimental group, although not statistically significant, is difficult to explain. The control group was slightly younger (67.7 vs 69.2 years) and had more female patients (78 vs 74 percent), both of which would favor a lower mortality rate,

TABLE 4. MEAN NUMBER AND LENGTH OF HOSPITALIZATIONS FOR 365 DAYS PRIOR TO INTERVIEW, AS RECALLED BY RESPONDENTS, INTERVIEW, AND PROXY POPULATIONS (n = 143)

	Experimental (n = 69)	Control (n = 74)	
Number and Duration	Mean (SD)	Mean (SD)	P*
Number of hospitalizations	1.2 (1.2)	0.6 (0.8)	<.003
Total days in hospital, all subjects	6.2 (11.1)	7.7 (21.7)	>.60
Total days in hospital, subjects with one or more hospitalizations	7.5 (12.2)	11.5 (27.4)	>.30

SD—Standard deviation
** Two-sided t test*

although the control group also had a higher percentage of blacks (73 vs 66 percent), which would reduce any expected difference in mortality between groups. In any case, these demographic differences are small (none are statistically significant) and cannot in themselves explain the difference in mortality. Comparing the mean age of death between the experimental group (83) and the control group (77) suggests that the older members of the experimental group were responsible for the increased mortality, but it does not help explain the excess number of deaths in the experimental group. It is noteworthy that significantly more experimental group patients had arteriosclerotic heart disease than did control group patients, and that such patients may have been more seriously ill. It is also possible that the home visit program had deleterious effects, such as patients not seeking office-based care when such care would have been appropriate.

Differences in utilization of outpatient services were as expected: the experimental group used more home care and the control group had more office visits. The experimental group had significantly more hospitalizations than did the control group, although there was no difference between the groups in the total number of days spent in the hospital. Again, the difference may be because of initial differences between the study groups (eg, the experimental patients had a greater incidence of heart disease) or because of deleterious effects of the home visit program. It is also possible that the home visit team appropriately identified patients requiring hospitalization; however, the long-term benefits of earlier or more appropriate hospitalization resulting from the program could not be identified from this study.

The results indicating that the home visit program had no effect or a deleterious effect need to be balanced against methodologic limitations of this follow-up study. Many of the study's potential methodologic problems, such as the limited availability of baseline information, the limited value of existing records for research needs, and the difficulty in locating and collecting information on all subjects, were anticipated in designing and implementing this

TABLE 5. COMPARISON BETWEEN GROUPS ON OUTCOME MEASURES OF INTERVIEW POPULATION (n = 86)

Instrument	Experimental Mean (SD)	Control Mean (SD)	P*
Quality of Well-Being (0 = death, 1 = complete well-being)	.608 (0.14)	.632 (0.10)	>.30
Barthel Activities of Daily Living (0 = dependence, 100 = independence)	86.8 (17.6)	89.7 (15.9)	>.40
Philadelphia Geriatrics Center Morale Scale (0 = low, 18 = high)	9.8 (4.8)	9.8 (4.7)	>.90
Global Health Status Visual Analog (0 = death, 1 = completely well)	5.7 (2.4)	6.0 (2.3)	>.50
Beck Depression Inventory (0 = no symptoms, 39 = severe)	6.6 (7.2)	7.2 (7.6)	>.70
Patient satisfaction questionnaire (0 = none, 75 = complete)	54.0 (6.8)	53.0 (7.7)	>.50

SD—Standard deviation
** Two-sided t test*

follow-up study. Furthermore, because formal power calculations were not possible, it was understood that any findings other than a statistically significant difference between groups would be inconclusive; the absence of statistically significant positive results could be attributed to the study not having sufficient statistical power to detect small yet meaningful differences.

Another consideration in interpreting the results is that the program being evaluated continued to develop throughout the evaluation period, and program activities at the time of this follow-up study may have been significantly different from those at the time of randomization two years earlier. This problem is not unique to this study and is often not addressed in evaluations of health services. Evaluation of a specific program is only an evaluation at one point in time for a particular implementation. The rationale for the program may be sound, and other implementations (eg, other sites, personnel, or subjects) may have measurable benefit.

Also to be considered is the concern that 30 percent of the control group patients received one or more visits. The majority of these visits occurred shortly before this follow-up study was conducted and would not be expected to have much impact on the findings regarding mortality and utilization of health services. Although these visits could have a greater or more immediate impact on functional status, emotional status, and well-being, comparisons among the control group patients who received visits, those who did not, and the experimental group patients did not reveal any statistically significant differences. Nevertheless, this problem reduces the ability of this study to detect subtle differences resulting from the experimental treatment.

Another methodologic factor to consider in evaluating this (and any other) health care delivery program is identifying the appropriate target population. Many of the patients who were initially judged eligible for the program and randomized for the study could not have been expected to show improvement in the outcomes of interest identified for this study (eg, mortality, utilization of health services). For example, terminally ill patients or otherwise healthy patients with limited social supports who were included in this study population would dilute any beneficial effects of the home visit program on these outcomes.

A last methodologic concern that was not anticipated was the possibility that the initial random assignment was not maintained for all subjects. Although randomization tables were used to assign the patients to experimental and control groups, the director of the home visit program recalled that home visit staff had allowed "one or two" patients to be transferred from the experimental group to the control group shortly after the initial randomization because of special requests from their primary physicians. This information was not known until the follow-up study had been completed. It was not possible to identify those patients who had been shifted. While the disturbance of randomization cannot be ignored, assuming that the disturbance was of a minimal order of magnitude, as suggested by the home visit program staff, the results would not be expected to be affected in a statistically significant way, but the positive effects of the program, if any, would be diluted. The equality of the experimental and control groups at baseline in terms of sex, race, and age supports but does not confirm the belief that the disturbance to randomization was minimal. The lack of comparable baseline information on diagnosis or functional status limits the ability to state unequivocally that any disturbance to random assignment was minor.

CONCLUSIONS

Despite methodologic limitations apparent at the outset, it seemed appropriate to learn as much as possible about home visit program effects from the previously random-

ized study population. This study was relatively inexpensive to complete and would have detected relatively large benefits attributable to the home visit program. Additionally, the description of this study population should prove valuable in restructuring this or other home visit programs and in designing more rigorous outcome investigations.

Because health care delivery programs similar to this one have been developing throughout the country, obtaining more evidence on the effectiveness of these programs is essential. Providing home care for individuals who cannot come to physicians' offices or to the hospital seems humane and represents an alternative for decreasing hospitalizations, nursing home placements, and utilization of other services, and it neither should be accepted nor rejected without thorough investigation. Studies of home care conducted to date have not demonstrated unequivocally health outcome or economic benefit. The health care community needs to recognize the methodologic difficulties inherent in these studies and continue to define criteria for translating individual research findings into generalizable health care policies. Although the present study did not demonstrate benefits attributable to this home visit program, it does not preclude the possibility that this type of program can be effective when made available to an appropriately selected population.

Acknowledgment

Funding for this project was provided by the Robert Wood Johnson Foundation. Ms. Merle Happe and Ms. Jean Higley interviewed patients.

References

1. Master RJ, Feltin M, Jainchill J, et al: A continuum of care for the inner city—Assessment of its benefits for Boston's elderly and high-risk populations. N Engl J Med 1980; 302:1434–1440
2. Brickner PW, Duque T, Kaufman A, et al: The homebound aged—A medically unreached group. Ann Intern Med 1975; 82:1–6
3. Perkel RL, Plumb J: The urban community family medicine home visit program at Thomas Jefferson University Hospital. Pride Institute J 1983; 2:1–8
4. Mims RB, Thomas LL, Conroy MV: Physician house calls: A complement to hospital-based medical care. J Am Geriatr Soc 1977; 25:28–34
5. Cauthen DB: The house call in current medical practice. J Fam Pract 1981; 13:209–213
6. Currie CT, Moore JT, Friedman SW, et al: Assessment of elderly patients at home: A report of fifty cases. J Am Geriatr Soc 1981; 29:398–401
7. Killip T, Kimball JT: Treatment of myocardial infarction in a coronary care unit—A 2-year experience with 250 patients. Am J Cardiol 1967; 20:457–464
8. Armstrong A, Duncan B, Oliver MF, et al: Natural history of acute coronary heart attacks—A community study. Br Heart J 1972; 34:67–80
9. Curtis P: Myocardial infarction—Home or hospital care. J Fam Pract 1978; 6:643–648
10. Mather HG, Pearson NG, Read KLQ, et al: Acute myocardial infarction: Home and hospital treatment. Br Med J 1971; 3:334–338
11. Hill JD, Hampton JR, Mitchell JRA: Home or hospital for myocardial infarction—Who cares? Am Heart J 1979; 98:545–547
12. Home or hospice care for acute myocardial infarction. NZ Med J 1978; 88:15–16
13. Mulley G, Arie T: Treating stroke: Home or hospital? Br Med J 1978; 2:1321–1322
14. Bryant NH, Candland L, Loewenstein R: Comparison of care and cost outcomes for stroke patients with and without home care. Stroke 1974; 5:54–59
15. Larson CP: Efficacy of prenatal and postpartum home visits on child health and development. Pediatrics 1980; 66:191–197
16. Currie AL, Gehlbach SH, Massion C, Thompson S: Newborn home visits. J Fam Pract 1983; 17:635–638
17. Berger LR, Samet KP: Home visits—Extending the boundaries of comprehensive pediatric care. Am J Dis Child 1981; 135:812–814
18. Kempe CH: Approaches to preventing child abuse—The health visitors concept. Am J Dis Child 1976; 130:941–947
19. Eddington C, Lee T: Sensory-motor stimulation for slow-to-develop children—A home-centered program for parents. Am J Nurs 1975; 75:59–62
20. Green M: The team approach in home care of mentally retarded children. Child Welfare 1972; 51:178–181
21. Lang SL, Ritchie MT: "Home" for aged. NY State J Med 1973; 73:1698–1699
22. Bell WG: Community care for the elderly: An alternative to institutionalization. Gerontologist 1973; 13:349–354
23. Dunlop BD: Expanded home-based care for the impaired elderly: Solution or pipe dream? Am J Public Health 1980; 70:514–519
24. Taylor RB: Home care. In Taylor RB (ed): Family Medicine: Principles and Practice. New York, Springer-Verlag, 1978, pp 112–116
25. Fry J: Common Dilemmas in Family Medicine. London, MTP Press, 1983, pp 185–201
26. Stone JR, Patterson E, Felson L: The effectiveness of home care for general hospital patients. JAMA 1968; 205:95–98
27. Kaplan RM, Bush JW, Berry CC: Health status: Types of validity and the index of well-being. Health Serv Res 1976; 11:478–507
28. Mahoney FI, Barthel DW: Functional evaluation: The Barthel Index. Md State Med J 1965; 14:61–65
29. Ware JE Jr, Snyder MK, Wright WR: Development and Validation of Scales to Measure Patient Satisfaction with Health Care Services. Volume 1: Final Report. Carbondale, Ill. Southern Illinois University Press, 1976
30. Beck AT, Rial WY, Rickels K: Short form of depression inventory: Cross-validation. Psychol Rep 1974; 34:1184–1186
31. Lawton MP: The Philadelphia Geriatric Center morale scale: A revision. J Gerontol 1975; 30:85–89
32. Metafile, Version 9.1, 1986. Chatfield, Minn, Metafile Information Systems, 1986