A Randomized Clinical Trial of Outpatient Comprehensive Geriatric Assessment Coupled with an Intervention to Increase Adherence to Recommendations

David B. Reuben, MD, Janet C. Frank, Dr PH,* Susan H. Hirsch, MPH,* Kimberly A. McGuigan, PhD,† and Rose C. Maly, MD, MSPH‡*

BACKGROUND: Although comprehensive geriatric assessment (CGA) has been demonstrated to confer health benefits in some settings, its value in outpatient or office settings is uncertain.

OBJECTIVE: To assess the effectiveness of outpatient CGA consultation coupled with an adherence intervention on 15-month health outcomes.

DESIGN: A randomized controlled trial.

SETTING: Community-based sites.

PATIENTS: 363 community-dwelling older persons who had failed a screen for at least one of four conditions (falls, urinary incontinence, depressive symptoms, or functional impairment)

INTERVENTION: A single outpatient CGA consultation coupled with an intervention to improve primary care physician and patient adherence with CGA recommendations.

MEASUREMENTS: Medical Outcomes Study Short Form-36 (MOS SF-36), restricted activity and bed days, Physical Perfomance Test, NIA lower-extremity battery.

RESULTS: In complete case analysis (excluding the five control group subjects who died during the follow-up period), the adjusted difference in change scores (4.69 points) for physical functioning between treatment and control groups indicated a significant benefit of treatment ($P = .021$). Similar benefits were demonstrated for number of restricted activity days and MOS SF-36 energy/fatigue, social functioning, and physical health summary scales. In analyses assigning scores of 0 to those who died, these benefits were greater, and significant benefits for the Physical Performance Test and MOS SF-36 emotional/well being, pain, and mental health summary scales were also demonstrated.

CONCLUSIONS: A single outpatient comprehensive geriatric assessment coupled with an adherence intervention can prevent functional and health-related quality-of-life decline among community-dwelling older persons who have specific geriatric conditions. J Am Geriatr Soc 47:269–276, 1999.

Key words: comprehensive geriatric assessment; consultation; functional status; geriatrics; prevention; adherence

Although previous studies have demonstrated the health benefits of comprehensive geriatric assessment (CGA) when provided in inpatient or rehabilitation units, or during home visits,[1,2] its value in outpatient or office settings is more controversial. Numerous randomized clinical trials of outpatient-based CGA[3–11] have provided inconsistent findings about its benefit in this setting. Possible explanations for its ineffectiveness have been that the population being assessed was too healthy, that the control group received better than usual care, or that the CGA intervention was not powerful enough (either because of poor implementation rates of CGA recommendations or lack of follow-up by the assessment team).[12]

The problem of low rates of implementation of CGA recommendations has been an obstacle to the effectiveness of inpatient consultative models of CGA,[13–16] and implementation rates in outpatient settings have generally been no better.[17,18] Even if physicians implement CGA recommendations, patient adherence may be worse in outpatient settings because older persons are less acutely ill and may perceive the value of such an evaluation and the resulting recommendations to be less beneficial.[19] In addition, refusing recommended evaluation and therapy may be less confrontational in outpatient settings because patients can easily cancel or not keep scheduled outpatient appointments rather than actively having to say "no" when a transporter or a therapist arrives.

For editorial comment, see p 371

In light of the potential benefits that CGA can provide when physician implementation and patient adherence are high, new strategies have focused on case management, with continuity by the geriatrics team[9,10,20] and CGA consultation coupled to an adherence intervention that emphasizes patient empowerment.[21] These new strategies have demonstrated high implementation and adherence rates in a case series,[21] reduced mortality and use of emergency room services in a

From the *Multicampus Program in Geriatric Medicine and Gerontology, and the ‡Division of Family Medicine, UCLA School of Medicine; and †RAND Corporation, Santa Monica, California.

Supported by the National Institute on Aging, Claude D. Pepper Older American Independence Center No. AG10415–01, and the Department of Veterans Affairs Medical Center-West Los Angeles and Sepulveda.

Address correspondence and reprint requests to David B. Reuben, MD, Multicampus Program in Geriatric Medicine and Gerontology, UCLA School of Medicine, 10945 Le Conte Ave., Suite 2339, Los Angeles, CA 90095-1687.

0002-8614/99/$3.50

nonrandomized study,[20] and better perceived health, life satisfaction, affective health status, and quality of health and social care in randomized clinical trials.[9–11] Although promising, the impact of these studies has been limited because of nonrandomized designs, small numbers of patients enrolled, or atypical settings of care (e.g., Department of Veterans Affairs medical centers).

In this randomized clinical trial, we evaluated the effectiveness of a single outpatient CGA consultation coupled with an adherence intervention for community-dwelling older persons who failed a screen for at least 1 of 4 conditions (falls, urinary incontinence, depressive symptoms, or functional impairment), which placed them at risk for functional or health-related quality-of-life (HRQOL) decline. Because the study focused on older persons who were living in the community, we hypothesized that CGA would preserve the more advanced levels of physical functioning and HRQOL that this population currently enjoys. We also hypothesized that patient satisfaction with medical care and patients' perceptions of the physician-patient relationship would be improved as a result of the adherence intervention.

METHODS

Protocol

We conducted a controlled clinical trial in which subjects randomized to the treatment group received a CGA consultation *and* an intervention to achieve adherence to recommendations from the CGA.[21] Subjects randomized to the control arm received usual care from their primary care physician plus nonmedical recruitment incentives (a 1-year membership to the American Association of Retired Persons, newsletters, and refrigerator magnets) to enhance the completion of study outcome measures. The study was approved by the UCLA Human Subjects Protection Committee.

Our recruitment strategy has been described previously.[21,22] In brief, we recruited subjects 65 years of age or older at community-based sites where older persons congregate and used an interviewer-assisted, self-administered, medical and functional screen to identify subjects who were appropriate for CGA. All subjects included in this trial were recruited expressly for the purposes of this study and are not the same subjects that have been reported on in preliminary studies.[19,21,22] The screen took an average of 10.8 minutes to complete and contained three subscales of the Functional Status Questionnaire,[23] a one-item screen for depression taken from the Yale battery,[24,25] validated specific questions about urinary incontinence,[26] and a probe question about falls[27] followed by three supplemental questions that identify risk of injurious falls.[28] Subjects who scored within the warning zones on one or more of the functional status subscales,[23] answered affirmatively to both incontinence questions, or answered affirmatively to the falls screening question and at least one supplemental falls question were considered to have failed the screen. Those who answered the depression question affirmatively were given a second screen using the 30-item Geriatric Depression Scale (GDS) at the time of the baseline measures interview.[29] A score of 11 or greater on the GDS was considered indicative of depressive symptoms substantial enough to be eligible for the study. Subjects were excluded if they did not speak English, did not have a telephone, did not have a primary care physician, were demented (or had Mini-Mental State scores[30] less than 24), or had other mental, emotional, or physical disorders to the extent that they could not be expected to complete the questionnaires and protocol required for the study.

Potential subjects who were eligible by screening evaluation and met no exclusionary criterion were administered informed consent. Those who agreed to participate were then interviewed in their homes or, occasionally, at a meal site or senior center by a research assistant who collected the baseline self-report and observed (performance-based) measures. Measures included in this report were the Medical Outcomes Study Short Form-36 (MOS SF-36),[31,32] Mini-Mental State Examination,[30] items from the Patient Satisfaction Questionnaire,[33] the NIA battery of lower extremity function,[34] the Physical Performance Test,[35] the Perceived Efficacy in the Patient-Physician Interaction (PEPPI) scale,[36] and questions on restricted activity days and "bed days" during the past 4 weeks based on the National Health Interview Survey.[37] After baseline measures were obtained, the research assistant opened an envelope that indicated group assignment (see "Assignment" below).

Treatment group subjects received (usually within 2 weeks) an in-depth, standardized, comprehensive geriatric assessment (instruments available upon request from the first author) from a social worker, a gerontologic nurse practitioner/geriatrician team, and a physical therapist (when indicated by falls or impaired mobility) at a community-based clinic. A short interdisciplinary case conference followed the evaluations. Six different board-certified geriatricians served on the team on a rotating basis; the same nurse practitioner, social worker, and physical therapist participated in the team throughout the duration of the study.

Intervention group subjects then received a previously described adherence intervention aimed both at patients and their physicians,[21] which consisted of the following elements. The geriatrician leading the assessment telephoned the subject's primary care physician to convey the CGA recommendations. This highly personal approach to CGA consultation allowed the primary care physician to provide input regarding the appropriateness of CGA recommendations. This telephone call was followed by a letter describing the recommendations, a copy of the dictated consultation, and copies of full-text references specific to the patient's conditions. The patient component included receiving a written list of recommendations at the time of the CGA and subsequently receiving a mailed copy of the dictated CGA consultation accompanied by a duplicate copy of the list of recommendations and a "How to Talk to Your Doctor" booklet. The patient was also contacted by telephone by a health educator approximately 2 weeks after receiving the CGA to review the team's recommendations and to help prepare the patient for discussion of the proposed recommendations with her or his physician. This preparation included ensuring that the patient understood the recommendations, assessing the level of patient agreement with recommendations, and empowering the patient to interact proactively with their physicians to implement and adhere to the recommendations.

Three months after the CGA, subjects were contacted by telephone, and information about physician implementation of (as reported by the subject) and patient adherence with recommendations was collected. At 15 months after randomization, all baseline measures were again administered to all participants, and patient adherence to recommendations was again determined. Mortality during the 15-month follow-up

period was determined and confirmed with official state records.

All analyses were completed on an intention-to-treat basis. The primary outcome of the study was change between baseline and end-of-study physical function (as measured by the MOS SF-36, 10-item physical function scale). Physical function change scores[38] were compared between treatment and control groups. Regressions predicting change scores using an indicator variable for treatment group status were performed adjusting for age, gender, ethnicity, and martial status. A robust form of least-squares regression was employed to account for intrahousehold correlation (i.e., nonindependence of responses between spouses within households).[39] Similar regression analyses were conducted for change in other functional status measures and health-related quality-of-life scales contained in the MOS SF-36. We also used two SF-36 summary scales, representing physical health and mental health,[40] which have been standardized to have a mean of 50 and a standard deviation of 10 in the general US population.

In the initial analyses, only participants who completed baseline and end-of-study measures were included (i.e., those who declined follow-up interviews or died during the study were censored from these analyses). We also used an alternative strategy for including deaths that assigned scores of 0 to those who died during the follow-up period.[41,42] Differences between treatment and control groups in percentages experiencing dichotomous outcomes (e.g., any restricted activity days) were tested using Wald's chi-square test within logistic regression analysis, and mortality differences were assessed using Fisher's exact test.

Post-hoc subgroup analyses were conducted to assess the possibility that certain subgroups (e.g., those meeting specific entry criteria or more than one entry criteria) received differential benefit from the treatment. Because they were based on smaller sample sizes and were not a priori hypotheses, these analyses were considered exploratory. Because of imbalances in baseline functional status between the two groups despite randomization, we conducted analyses that included baseline physical functioning as an interaction term (baseline functional status by treatment group). In this manner, we could examine differential benefit among those with varying functional status. Because of the possibility that change may have been an artifact of "regression toward the mean," in this analysis, we estimated the magnitude of treatment effect that would exceed naturally occurring regression toward the mean by control group subjects. This was accomplished by entering main effects terms for treatment group and baseline measures as well as an interaction term, treatment × baseline measure. The interaction term is the estimate of the magnitude of treatment effect that would exceed naturally occurring regression toward the mean by control group subjects.

Adherence to recommendations was tabulated and classified using a previously published methodology[21] that assessed physician implementation and patient adherence to physician-initiated and self-care recommendations. Recommendations that were implemented in a modified form but addressed the same problem were coded as having been implemented.

The study sample size was calculated to have a statistical power of .83 (beta) to detect a 15% difference (8.9 points) between groups on the MOS-SF 36 physical functioning scale. For all analyses, $P < .05$ (two-tailed) was regarded as statistically significant. All statistical analyses were performed using the Statistical Analysis System (Carey, NC) or Stata 5.0 (College Station, TX).

Assignment

The unit of randomization was the individual subject, except married couples, who were randomized in pairs if both were eligible and chose to participate. Randomization occurred in blocks of 8 and was stratified based on type of health care insurance coverage (fee-for-service versus capitated managed care) and on "couple" status. A computer program generated random group assignment using a set seed. The assignment was printed and placed in sealed unmarked opaque envelopes numbered in the order they were generated. The sets of envelopes were then given to research assistants who conducted baseline interviews and, once completed, opened the envelopes.

Masking

All participants were aware of their assignment group. The clinical team conducting the CGA saw only treatment group subjects and was unaware of baseline and outcome measures of subjects in either group. Research assistants were unaware of the group assignment at the time baseline measures were obtained and were unaware of study hypothesis and scoring of outcome measures at follow-up. All interviewer-administered instruments were read verbatim, and answers were recorded without interpretation. All performance-based instruments were administered according to strict protocol.

RESULTS

A diagram of participant flow and follow-up is provided in Figure 1. During the 18-month recruitment period, 1304 potential subjects were screened. Of the 601 who met eligibility criteria, 363 (60%) were eligible and agreed to participate in the study; 180 were assigned randomly to the experimental group, and 183 were assigned to the control group. Of those assigned to the experimental group, 173 (96%) actually received the CGA and adherence intervention; the remainder refused the CGA. Among treatment group sub-

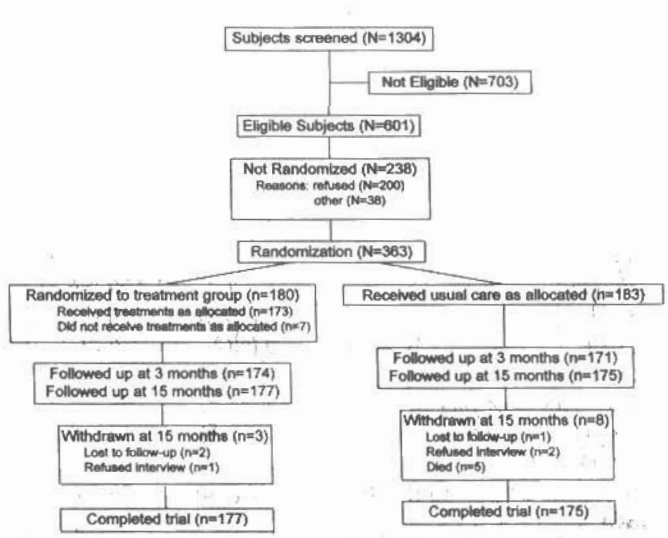

Figure 1. Progress through various stages of the trial and timing of primary and secondary outcome measures.

jects, primary care physicians implemented 59% of physician-initiated CGA recommendations within 3 months. During the 15 months after the CGA, patients adhered to 67% of all physician-initiated recommendations and 61% of all self-care recommendations.

Data on the study population at 15-months were available for 356 of the 363 (98%) recruited subjects. Of these, end-of-study interviews were conducted on 177 of 180 treatment group subjects (98%) and 175 of 183 control group subjects (96%). One treatment group subject refused the 15-month end-of-study interview, and two subjects were lost to follow-up. Two control group subjects refused the 15-month interview, and one was lost to follow-up. During the follow-up period, five control group subjects died; none of the treatment group subjects died ($P = .061$).

The baseline sociodemographic characteristics of the study groups are shown in Table 1 and did not differ between experimental and control group subjects on any characteristic. Baseline and 15-month functional status measures and change scores are presented in Table 2. At baseline, the control group had higher scores on the 10-item physical function scale (MOS-PF 10) and role functioning as a result of physical health scales; a higher percentage of the treatment group had any restricted activity days (38% vs 25%, $P = .015$). At 15 months, physical functional status scores in the control group had dropped significantly, whereas the treatment group had maintained its functional status. The adjusted (for age, gender, ethnicity, and marital status) difference (4.69 points on the MOS-PF 10) in change scores for physical functioning between treatment and control groups was significant at $P = .021$. In the analysis that assigned 0 as a functional status score for those who died, the adjusted difference was 5.73 points ($P = .007$). Compared with baseline, the percentage of control group subjects with any restricted activity days and any bed days increased, whereas the percentage in the treatment group did not change. The difference in change in average number of restricted activity days between treatment and controls (−2.84 days, $P = .006$) also indicated benefit from the treatment. Scores on performance-based measures (the Physical Performance Test and the NIA Battery) did not differ between groups at baseline and declined in both groups. In the analysis that assigned 0 as a score for those who died, the adjusted differences on the Physical Performance Test and NIA Battery were 1.90 and .324 ($P = .019$ and $P = .293$, respectively).

Subgroup analyses based on specific entry criteria or number of entry criteria that were met did not indicate a clear pattern of patients whose functional status demonstrated greater or lesser benefit from the treatment. However, based on the inclusion of the interaction term of baseline functional status by treatment group, treatment group subjects with poorer physical functioning, as demonstrated by the MOS PF-10 at baseline, demonstrated greater improvement ($P = .046$) at follow-up compared with other treatment group patients, and this difference exceeded the natural regression toward the mean demonstrated by control group subjects.

The effects of the treatment on other health-related quality-of-life measures captured by the SF-36 are presented in Table 3. Change scores on the physical health summary scale demonstrated a significant treatment effect (adjusted difference 1.99, $P = .043$). If those who died were assigned values of 0, change scores on both the physical and mental health summary scales indicated a treatment effect (adjusted differences 2.98 and 3.55, $P = .005$ and $P = .006$, respectively). At baseline, energy/fatigue scale scores, social functioning, and pain (indicating less pain) scores were higher in the control than in the treatment group. At 15 months, the adjusted differences in change scores for the energy/fatigue scale (6.59 points, $P = .001$) and social functioning scale (7.34 points, $P = .010$) scores between treatment and control groups indicated a significant benefit of the treatment. Assigning those who died scores of 0, change scores for emotional well-being (adjusted difference 4.75, $P = .016$) and pain (adjusted difference, $P = .043$) also indicated significant benefits of the treatment.

Changes in measures of patient satisfaction and perceived patient efficacy in patient-physician interaction are shown in Table 4. In adjusted analyses, the differences in change scores for any of the three scales were not significant.

Table 1. Baseline Characteristics of the Study Subjects*

Sociodemographic Characteristics	Treatment n = 180 (%)	Control n = 183 (%)
Mean age in years (SD)	75.8 (6.1)	75.9 (5.7)
Female	150 (83.3)	147 (80.3)
Living alone	114 (63.3)	110 (60.1)
Widowed	86 (47.8)	79 (43.2)
Ethnicity		
White	146 (81.1)	159 (86.9)
Black	25 (13.9)	19 (10.4)
Other	9 (5.0)	5 (2.7)
Education		
High school or higher	159 (88.3)	168 (91.8)
College or higher	50 (27.8)	59 (32.2)
HMO members	106 (58.9)	112 (61.2)
Mini-Mental State Examination (SD)	28.2 (1.4)	28.2 (1.5)

*There were no significant differences between treatment and control group subjects on any variable.

DISCUSSION

A major goal in the care of older persons has been to apply the principles of effective inpatient models of comprehensive geriatric assessment to outpatient settings, using fewer resources while achieving similar benefits. In this randomized clinical trial, a single outpatient CGA consultation, coupled with a modest intervention to ensure adherence with the resulting recommendations, was able to preserve physical and social functioning in community-dwelling older persons who were at risk for functional decline. Differences between treatment and control groups consistently favored the treatment group across a variety of measures of physical function. Moreover, our primary analyses excluded those who died during the study and, therefore, were conservative in estimating the treatment's effectiveness. In analyses that assigned subjects who died a score of 0,[41,42] the benefits of the treatment were even more profound. Although results of subgroup analyses should be interpreted cautiously, the benefit of treatment was greatest among those who were the most functionally impaired at baseline. The clinical significance of change scores of summary measures may be somewhat difficult to interpret. The magnitude of difference in our primary

Table 2. Baseline and 15-Month Functional Status Measures

Measure	Baseline Unadjusted Means		15-month Unadjusted Means		Complete Case Analysis		Analyses Assigning 0 to Dead	
	Treatment (n = 176)	Control (n = 175)	Treatment (n = 176)	Control (n = 175)	Difference in Change Score‡	*P* Value	Difference in Change Score‡	*P* Value
MOS SF-36 PF10*	52.4	62.5†	52.7	58.5	4.69 (.63, 8.75)	.021	5.73 (1.59, 9.87)	.007
MOS SF-36 role functioning/physical*	48.2	62.7†	49.7	55.4	8.87 (−1.03, 18.78)	.074	10.77 (.85, 20.69)	.034
Any restricted activity days (%)	38	25†	38	37				
No. of restricted activity days	4.04	2.22	3.89	4.89	−2.84 (−.75, −4.93)	.006		
Any bed days (%)	20	16	26	25				
No. of bed days	1.22	.73	1.79	1.52	−0.35 (.77, −1.47)	.533		
PPT score (0–28)	22.0	22.6	20.4	19.7	1.58 (−.12, 2.98)	.066	1.90 (.31, 3.48)	.019
NIA Battery score (0–12)	8.0	8.0	7.2	7.1	.14 (−.45, .72)	.634	.32 (−.28, .93)	.293

*Scored 0–100, with 100 indicating best functioning.
†$P < .05$ between experimental and control groups at baseline.
‡Differences in change scores (15 month − baseline) between treatment and control groups adjusted for age, gender, ethnicity, and marital status. Positive scores indicate a favorable effect of the treatment, except for restricted activity days and bed days where negative scores indicate a favorable effect. Subjects who died during the follow-up period are excluded from these analyses. 95% confidence intervals are included in parentheses.

Table 3. Baseline and 15-Month Health Status Measures

Measure	Baseline Unadjusted Means		15-month Unadjusted Means		Complete Case Analysis		Analyses Assigning 0 to Dead	
	Treatment (n = 176)	Control (n = 175)	Treatment (n = 176)	Control (n = 175)	Difference in Change Score‡	*P* Value	Difference in Change Score‡	*P* Value
MOS Summary Scales*								
Physical health	37.8	41.8	37.4	39.7	1.99 (.07, 3.91)	.043	2.98 (.88, 5.10)	.005
Mental health	50.1	51.4	52.4	51.5	2.07 (−.10, 4.24)	.062	3.55 (1.05, 6.06)	.006
MOS SF-36 subscales*								
Emotional/well-being	68.7	71.2	72.3	72.2	2.62 (−.82, 6.05)	.127	4.75 (.88, 8.61)	.016
Role functioning/emotional	73.6	78.5	81.3	80.4	5.39 (−3.14, 13.94)	.192	7.57 (−1.08, 16.22)	.086
Energy/fatigue	48.7	56.7†	50.5	52.1	6.59 (2.66, 10.53)	.001	7.92 (3.81, 12.04)	<.001
Social functioning	75.2	82.2†	77.2	76.1	7.34 (1.66, 13.03)	.010	9.40 (3.50, 15.29)	.002
Pain	59.6	64.9†	60.4	62.4	3.84 (−1.59, 9.26)	.160	5.80 (.17, 11.4)	.043
General health	58.1	61.8	59.5	62.2	1.98 (−12.8, 5.25)	.228	3.19 (−.26, 6.63)	.070

*All scales are scored 0–100, with 100 indicating best functioning. Summary scales have been standardized to have a mean of 50 and a standard deviation of 10.
†$P < .05$ between experimental and control groups.
‡Differences in change scores (15-month − baseline) between treatment and control groups adjusted for age, gender, ethnicity, and marital status. Positive scores indicate a favorable effect of treatment. Subjects who died during the follow-up period are excluded from these analyses. 95% confidence intervals are included in parentheses.

Table 4. Baseline and 15-Month Patient Satisfaction/Efficacy Measures

Measure	Baseline Unadjusted Means		15-Month Unadjusted Means		Difference in Change Score*
	Treatment	Control	Treatment	Control	
Patient satisfaction/general	56.9	57.1	58.4	58.0	.38
Patient satisfaction/doctor	65.9	67.6	69.2	69.6	1.85
PEPPI†	74.2	75.5	78.0	79.1	−.19

All scales scored from 0–100, with 100 indicating best function.
*Differences in change scores (15-month − baseline) between treatment and control groups adjusted for age, gender, ethnicity, and marital status. Positive scores indicate a favorable effect of the treatment. No differences were statistically significant.
†PEPPI = Perceived Efficacy in the Physician Patient Interaction scale.

outcome measure, self-reported physical functioning, was 8 to 10%, depending on how those who died were handled in the analysis. Intervention group subjects would also have experienced approximately three fewer restricted activity days in the past 4 weeks, a difference that is likely to be meaningful to patients as well as to healthcare providers.

Why did this study show benefits resulting from outpatient CGA when many others have failed to do so? The answer is likely to be multifactorial, relating to the population studied, the design of the study, the combination of CGA plus an adherence intervention, and the outcomes that were chosen. In this study, we focused on older persons who had geriatric conditions (functional impairment, depressive symptoms, urinary incontinence, and falls) that were likely to benefit from treatment[43] and were markers of other treatable health problems.[44,45] Such targeting to identify the older persons who are most likely to be helped by CGA has been associated with greater benefits in other controlled trials of CGA.[1]

The design of the study also differed from several previous trials[3,5] of outpatient CGA in that usual care, rather than another "second opinion" consultation, was provided to the control group. As such, it assesses the additional contribution that CGA consultation plus the adherence intervention can make beyond the care that older persons currently receive.

The treatment approach used in the present study differs conceptually in several respects from those used in previous trials of one-time outpatient CGA consultation.[3,8] We used standardized assessments formulated by the team's health professionals[21] and applied consistently to all treatment group subjects. We also sought to reduce the problem of poor implementation of CGA recommendations by adding an adherence intervention that relied on principles of academic detailing,[46] patient empowerment, and improved patient-physician communication. In this clinical trial, physician implementation and patient adherence rates were lower than in a previous study.[21] Nevertheless, they were higher than previously reported for outpatient CGA studies.[17,18] Moreover, patient adherence to both physician-initiated and self-care recommendations was relatively high, both exceeding 60%.

We also selected outcome measures that were appropriate to capture the changes that were anticipated from this specific treatment in this specific population. Because these patients were less frail than those in other studies, we expected that the chief benefit would be in preventing decline rather than in restoring function. The functional status measures used in the study better reflect function at the instrumental or intermediate activities of daily living level than at the more basic activities of daily living level.[47] We also noticed benefits on restricted activity days, which are highly correlated with illnesses and hospitalizations in the past 12 months,[48] suggesting health benefits beyond functional status. The relevance of these measures is supported further by noting that the MOS-SF 36 is currently being administered as the functional status measure in the Health Plan Employer Data and Information Set (HEDIS) 3.0, which is being used by the Health Care Financing Administration to assess at-risk Medicare contract health plans.

The study's results also support the following sequence that may help understand the mechanism by which the program conferred its benefits. The program began by administering a screening instrument that detected inadequately evaluated or treated geriatric conditions and identified those who might benefit from a more comprehensive assessment. The CGA generated further diagnostic and therapeutic recommendations for these and other medical and social problems. The adherence intervention helped prepare physicians to accept these recommendations and empowered patients to discuss these issues and recommendations. Physicians implemented a large percentage of these recommendations, and patients adhered to most of these physician-initiated, as well as the self-care, recommendations and achieved health benefits as a result. We had hoped that the intervention would also lead to increased patient satisfaction and improve perceived self-efficacy in relating to their physicians, but we were unable to demonstrate such benefits.

The study also yielded some unanticipated findings. Perhaps the most important of these was the reduction in mortality, which achieved borderline significance in two-tailed testing. The trial was inadequately powered to detect differences in mortality, but such findings suggest the treatment may have greater benefit than anticipated. Other limitations of the study must also be recognized. Perhaps the most important of these is the imbalance in some baseline measures between groups. Although the randomization resulted in virtually identical sociodemographic characteristics in each group, such imbalances raise questions about the equivalence between groups. Imbalances can occur by chance despite randomization. In our analyses, we have attempted to correct for such imbalances by utilizing change scores, which include baseline status for each individual. We also conducted interaction analyses of baseline functional status by treatment group to address the competing hypothesis of regression toward the mean. The consistency of our findings across many outcomes, including those that did not differ between

groups at baseline, supports our conclusions. Moreover, the higher baseline functional status in the control group would have led to an expected fewer deaths in this group; in fact, the opposite occurred, which supports further the effectiveness of the intervention. Finally, as a single site trial, the reproducibility of the treatment may be questioned. This concern is mitigated by the detailed structured assessments and protocols for the adherence intervention. Nevertheless, the academic base of the clinical team may have contributed to its effectiveness, and similar efforts by community-based CGA teams may be less effective.

These findings have several implications for incorporating CGA into clinical settings and for future directions of CGA programs. First, our approach was designed to collaborate with primary care physicians to provide better geriatric care for their older patients. Accordingly, it is less resource intensive than programs that assume primary care for the patient[9,10,20] and does not set up a second parallel system of healthcare delivery. Such integration within existing healthcare delivery systems makes this strategy particularly suitable for at-risk contract Medicare health maintenance organizations. Second, identification of appropriate subjects relied on community-based screening rather than referral or case-finding. Since all of the items on the screening instrument were self-administered, such an instrument could be administered by mail or via telephone to a defined patient population, such as members of a health plan or a physician's practice. They could also be used for case-finding when older patients have medical care visits for other reasons. Third, the targeting approach focused on patients who are healthier than those studied in inpatient CGA and rehabilitation units. Because this population is considerably larger than the frail population of older persons,[49] from a public health perspective the overall yield from this less intensive approach may be greater than for more intensive programs that are focused on fewer persons.

In view of a shrinking Medicare budget, CGA programs will have to be more streamlined than the early programs[50,51] that first demonstrated the effectiveness of this method of healthcare delivery for older persons. This study demonstrates that a relatively inexpensive CGA program that incorporates features of successful but more intensive programs can confer many of the same benefits. Moreover, the functional status outcomes achieved by the program are among the most relevant in maintaining the independence of older persons.

ACKNOWLEDGMENTS

The authors acknowledge the additional clinical team members: Nancy Weintraub, MD; Susan Melchiore, MD; P. Nina Shah, MD; Scott Sherman, MD; Marcia Gold, GNP; Barbara Engleman, MA; and Karen Jessum, RPT. The research associates included Janice Chernoff, Kathleen Adams, Kathleen Walsh, Daniel Millner, and Jennifer Levin. We also thank Gail A. Greendale, MD; Emmett B. Keeler, PhD; Laurence Z. Rubenstein, MD; and John C. Beck, MD, for reviewing earlier drafts of the manuscript.

REFERENCES

1. Stuck AE, Siu AL, Wieland GD et al. Comprehensive geriatric assessment: A meta-analysis of controlled trials. Lancet 1993;342:1032–1036.
2. Stuck AE, Aronow HU, Steiner A et al. A trial of annual in-home comprehensive geriatric assessments for elderly people living in the community. N Engl J Med 1995;333:1184–1189.
3. Epstein AM, Hall JA, Fretwell M et al. Consultative geriatric assessment for ambulatory patients: A randomized trial in a health maintenance organization. JAMA 1990;263:538–544.
4. Rubin CG, Sizemore MT, Loftis PA et al. A randomized, controlled trial of outpatient geriatric evaluation and management in a large public hospital. J Am Geriatr Soc 1993;41:1023–1028.
5. Williams ME, Williams TF, Zimmer JG et al. How does the team approach to outpatient geriatric evaluation compare with traditional care: A report of a randomized controlled trial. J Am Geriatr Soc 1987;35:1071–1078.
6. Tulloch AJ, Moore V. A randomized controlled trial of geriatric screening and surveillance in general practice. J R Coll Gen Pract 1979;29:733–742.
7. Yeo G, Ingram L, Skurnick J, Crapo L. Effects of a geriatric clinic on functional health and well-being of elders. J Gerontol 1987;42:252–258.
8. Silverman M, Musa D, Martin DC et al. Evaluation of outpatient geriatric assessment: A randomized multi-site trial. J Am Geriatr Soc 1995;43:733–740.
9. Burns R, Nichols LO, Graney MJ, Cloar T. Impact of continued geriatric outpatient management on health outcomes of older veterans. Arch Intern Med 1995;155:1313–1318.
10. Toseland RW, O'Donnell JC, Engelhardt JB et al. Outpatient geriatric evaluation and management: Results of a randomized trial. Med Care 1996;34: 624–640.
11. Engelhardt JB, Toseland RW, O'Donnell JC et al. The effectiveness and efficiency of outpatient geriatric evaluation and management. J Am Geriatr Soc 1996;44:847–856.
12. Moore AA, Reuben DB. Comprehensive Geriatric Assessment. In: Reuben DB, Yoshikawa TT, Besdine RW, eds. Geriatrics Review Syllabus: A Core Curriculum in Geriatric Medicine, 3rd Ed. Dubuque, IA: Kendall/Hunt Publishing Company, for the American Geriatrics Society, 1996, pp 78–80.
13. Allen CM, Becker PM, McVey LJ et al. A randomized, controlled clinical trial of a geriatric consultation team: Compliance with recommendations. JAMA 1986;255:2617–2621.
14. McVey LJ, Becker PM, Saltz CC et al. Effect of a geriatric consultation team on functional status of elderly hospitalized patients. Ann Intern Med 1989; 110:79–84.
15. Winograd, CH, Gerety MB, Lai NA. A negative trial of impatient geriatric consultation: Lessons learned and recommendations for future research. Arch Intern Med 1993;153:2017–2023.
16. Reuben DB, Borok GM, Wolde-Tsadik G et al. A randomized trial of comprehensive geriatric assessment in the care of hospitalized patients. N Engl J Med 1995;332:1345–1350.
17. Reed RL, Kligman EW, Weiss BD. Comprehensive geriatric assessment recommendations: Adherence of family practice residents. J Fam Pract 1990;31: 398–392.
18. Devor M, Wang A, Renvall M et al. Compliance with societal safety recommendations in a comprehensive geriatric assessment program. J Gerontol Med Sci 1994;49:M168–173.
19. Frank JC, Hirsch SH, Chernoff J, et al. Determinants of patient adherence to consultative comprehensive geriatric assessment recommendations. J Gerontol: Med Sci 1997;52A:M44–51.
20. Boult C, Boult L, Murphy C et al. A controlled trial of outpatient geriatric evaluation and management. J Am Geriatr Soc 1994;42:465–470.
21. Reuben DB, Maly RC, Hirsch SH et al. Physician implementation of and patient adherence to recommendations from comprehensive geriatric assessment. Am J Med 1996;100:444–451.
22. Reuben DB, Hirsch SH, Chernoff JC et al. Project Safety Net: A health screening outreach and assessment program. Gerontologist 1993;33:557–560.
23. Jette AM, Davies AR, Cleary PD et al. The functional status questionnaire: Reliability and validity when used in primary care. J Gen Intern Med 1986; 143–149.
24. Lachs MS, Feinstein AR, Cooney LM et al. A simple procedure for general screening for functional disability in elderly patients. Ann Intern Med 1990; 112:699–706.
25. Mahoney J, Drinka TJK, Abler R et al. Screening for depression: Single question versus GDS. J Am Geriatr Soc 1994;42:1006–1008.
26. Diokno AC, Brown MB, Brock BM et al. Clinical and cystometric characteristics of continent and incontinent noninstitutionalized elderly. J Urol 1988; 140:567–571.
27. Cummings, SR, Nevitt MC, Kidd S. Forgetting falls: The limited accuracy of recall of falls in elderly. J Am Geriatr Soc 1988;36:613–616.
28. Nevitt MC, Cummings SR, Hudes ES. Risk factors for injurious falls: A prospective study. J Gerontol Med Sci 1991;46:M164–170.
29. Yesavage JA, Brink TL, Rose TL et al. Development and validation of a geriatric depression screening scale: A preliminary report. J Psychiatr Res 1983; 17:37–49.

30. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
31. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): 1. Conceptual framework and item selection. Med Care 1992;30:473–483.
32. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-item health survey 1.0. Health Econ 1993;2:217–227.
33. Ware JE, Snyder MK, Wright WR, Davies AR. Defining and measuring patient satisfaction with medical care. Eva Program Planning 1983;6:247–263.
34. Guralnik JM, Simonsick EM, Ferrucci L et al. A short performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol Med Sci 1994;49:M85–94.
35. Reuben DB, Siu AL. An objective measure of physical function of elderly patients: The physical performance test. J Am Geriatr Soc 1990;38:1105–1112.
36. Maly RC, Frank JC, Marshall GN. Perceived efficacy in patient-physician interactions (PEPPI): Validation of an instrument in older persons. J Am Geriatr Soc 1998;46:889–899.
37. National Center for Health Statistics. Current estimates from the National Health Interview Survey: United States 1986. Vital Health Stat 1986;10:160.
38. Maxwell SE, Howard GS. Change scores: Necessarily anathema? Educ Psychol Meas 1981;41:747–756.
39. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. Proceedings of the Fifth (1965) Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, CA: University of California Press, 1967, pp 221–233.
40. Ware JE, Kosinski M, Keller SK. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: The Health Institute, New England Medical Center, 1994.
41. Diehr P, Patrick D, Hedrick S et al. Including deaths when measuring health status over time. Med Care 1995;33:AS164–172.
42. Ware JE, Bayliss MS, Rogers WH et al. Difference in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems: Results from the medical outcomes study. JAMA 1996;276:1039–1047.
43. Reuben DB. Defining and refining targeting criteria. In: Rubenstein LZ, Wieland D, Bernabei R, eds. Geriatric Assessment Technology: The State of the Art. Milan, Italy: Editrice Kurtis, pp 265–269.
44. Maly RC, Hirsch SH, Reuben DB. The performance of simple instruments in detecting geriatric conditions and selecting community-dwelling older persons for geriatric assessment. Age Ageing 1997;26:223–231.
45. Rubenstein LZ, Josephson KR, Wieland GD et al. Effectiveness of a geriatric evaluation unit: A randomized clinical trial. N Engl J Med 1984;311:1664–1670.
46. Soumerai SB, Avorn J. Principles of educational outreach (academic detailing) to improve clinical decision making. JAMA 1990:263:549–556.
47. Reuben DB, Valle LA, Hays RD Siu AL. Measuring physical function in community-dwelling older persons: A comparison of self-administered, interviewer-administered, and performance-based measures. J Am Geriatr Soc 1995;43:17–23.
48. Scholes D, LaCroix AZ, Wagner EH et al. Tracking progress toward national health objectives in the elderly: What do restricted activity days signify? Am J Public Health 1991;81:485–488.
49. American Association of Retired Persons. A profile of older Americans: 1996. PF3049 (1296) D996. Administration on Aging, US Department of Health and Human Services. 1996.
50. Rubenstein LZ, Josephsen KR, Wieland GD et al. Effectiveness of a geriatric evaluation unit: A randomized clinical trial. N Engl J Med 1984;311:1664–1670.
51. Applegate WB, Miller ST, Graney MJ et al. A randomized, controlled trial of a geriatric assessment unit in a community rehabilitation hospital. N Engl J Med 1990;322:1572–1578.