U.S. Department of Health and Human Services
Assistant Secretary for Planning and Evaluation
Office of Disability, Aging and Long-Term Care Policy

The Comparability of Treatment and Control Groups at Randomization

October 1983

Office of the Assistant Secretary for Planning and Evaluation

The Office of the Assistant Secretary for Planning and Evaluation (ASPE) is the principal advisor to the Secretary of the Department of Health and Human Services (HHS) on policy development issues, and is responsible for major activities in the areas of legislative and budget development, strategic planning, policy research and evaluation, and economic analysis.

ASPE develops or reviews issues from the viewpoint of the Secretary, providing a perspective that is broader in scope than the specific focus of the various operating agencies. ASPE also works closely with the HHS operating divisions. It assists these agencies in developing policies, and planning policy research, evaluation and data collection within broad HHS and administration initiatives. ASPE often serves a coordinating role for crosscutting policy and administrative activities.

ASPE plans and conducts evaluations and research--both in-house and through support of projects by external researchers--of current and proposed programs and topics of particular interest to the Secretary, the Administration and the Congress.

Office of Disability, Aging and Long-Term Care Policy

The Office of Disability, Aging and Long-Term Care Policy (DALTCP), within ASPE, is responsible for the development, coordination, analysis, research and evaluation of HHS policies and programs which support the independence, health and long-term care of persons with disabilities--children, working aging adults, and older persons. DALTCP is also responsible for policy coordination and research to promote the economic and social well-being of the elderly.

In particular, DALTCP addresses policies concerning: nursing home and community-based services, informal caregiving, the integration of acute and long-term care, Medicare post-acute services and home care, managed care for people with disabilities, long-term rehabilitation services, children's disability, and linkages between employment and health policies. These activities are carried out through policy planning, policy and program analysis, regulatory reviews, formulation of legislative proposals, policy research, evaluation and data planning.

This paper was prepared by HHS's Office of Social Services Policy (now DALTCP). The paper was written as part of contract #HHS-100-80-0157 between ASPE and Mathematica Policy Research, Inc. Additional work on this project was conducted by Temple University. For additional information about this subject, you can visit the DALTCP home page at http://aspe.hhs.gov/_/office_specific/daltcp.cfm or contact the office at HHS/ASPE/DALTCP, Room 424E, H.H. Humphrey Building, 200 Independence Avenue, S.W., Washington, D.C. 20201. The e-mail address is: webmaster.DALTCP@hhs.gov. The Project Officer was Robert Clark.

THE COMPARABILITY OF TREATMENT AND CONTROL GROUPS AT RANDOMIZATION

Randall S. Brown
Margaret Harrigan

Mathematica Policy Research, Inc.

October 27, 1983

Prepared for
Office of the Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services
Contract #HHS-100-80-0157

This report was prepared for the Department of Health and Human Services under contract no. HHS-100-80-0157, a competitively awarded contract in the amount of $12.4 million. The Department of Health and Human Services project officer is Mr. Robert Clark, Office of the Secretary, Department of Health and Human Services, Room 439F, Hubert Humphrey Building, Washington, D.C. 20201.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Throughout the design and implementation of the channeling demonstration, emphasis has been placed on the importance of random assignment of eligible applicants into treatment and control groups. Due to the random assignment, the resulting control and treatment groups should be composed of eligible individuals that on average are very similar at the time of application on any observed or unobserved characteristic. This lack of pre-existing differences between treatments and controls implies that the control group yields reliable estimates of what would have happened to clients in the absence of channeling, and when these estimates are compared to outcomes for clients, reliable estimates of channeling impacts are obtained.

Only two factors can lead to differences in the true mean values of the pre-application characteristics of the treatment and control groups: deviation from the randomization procedures and normal sampling variability. Deviations from the carefully developed randomization procedures could be either deliberate (e.g., site staff purposely misrecording as treatments some applicants who are randomly assigned to the control group, but who have especially pressing needs for assistance) or accidental (e.g., misrecording of a sample member's status). The dedication and professionalism of this site staff and the safeguards built into the assignment procedure make either occurrence very unlikely. Site staff were extremely cooperative in faithfully executing the procedures. Sampling variability, on the other hand, is the difference between the two groups that occurs simply by chance. For the sample sizes available at the model level, such differences between the two groups should be very small, and statistically insignificant.

Despite the expected small, chance differences between the two groups, the implications of any such differences for estimates of program impacts is so great that the issue of treatment/control group equivalence must be examined thoroughly. For example, if we find that the treatment group was more severely impaired at the time the screen was given, differences between the two groups in mean impairment level six months after randomization would reflect both initial differences and the effects of channeling. The relationship may also be more subtle. For example, if channeling were more effective for certain subgroups than for others (e.g., those living with relatives compared to those who are not), differences between the treatment and control groups in the proportion of cases that are living with relatives would then affect estimates of overall program impact. Regression procedures can help to control for initial differences such as these, but there is no guarantee that the variables available to include in the regression will control for all of the factors which are differentially represented in the two groups and which affect the post-randomization values of outcome variables. Furthermore, the appropriate relationship may not be linear, as would typically be assumed in regression. Thus, one of the primary virtues of experimental design, the ability to rely on simple, robust comparisons of treatment and control group means to obtain unbiased estimates of program impacts, is lost if treatment and control groups are not equivalent at the time of randomization. Unbiasedness is then dependent upon

assumptions about the correctness of the regression specification (i.e., explanatory variables used, independence of disturbance term and regressors, functional form, etc.).

The comparability of the treatment and control groups at randomization is also important because it is the first stage in our investigation of a set of methodological problems that could result in biased estimates of channeling's impact. Differences between treatment and control groups in the types of individuals who fail to respond to interviews could result in noncomparable groups in the sample being analyzed, even if the full samples were comparable. Differences in the way baseline data were collected for treatments and controls could lead to differential measurement error, which could cause regression estimates of program impacts to -be biased. In order to assess these other potential sources of bias, it is important to first determine whether the two groups were comparable before the baseline interview.

Because of these goals and concerns, in this paper we assess the equivalence of the treatment and control groups at the time of randomization by comparing the screen characteristics of the two groups. The data collected at the screen do not provide an ideal basis for the comparison in that differences between treatments and controls in the extent of item nonresponse and differences in the accuracy of the screen data ultimately recorded could lead to differences in the computed means at the screen, even if the two groups are comparable. However, these problems did not seem to occur, and in any case are much less significant than the problems of interpretation that would be caused by using baseline data to assess comparability of the two groups. The differences between the two groups in baseline data collection procedures and interview nonresponse are potential problems with that interview that are not problems with the screen. If our analysis of comparability of the two groups using screen data indicates no differences between the treatment and control groups, then comparisons of their data at baseline can be conducted to assess the issues of nonresponse and measurement bias described above. If preenrollment differences between the two groups are found, then these differences must be controlled for in the other methodological investigations.

Section I contains descriptions of the screen data used for this analysis, and the randomization process employed to assign treatment status to eligible elders expressing potential interest in channeling. Section II contains a description of the statistical tests performed and the results of these tests. The emphasis is on comparisons between treatment and control groups at the model level; however, site-specific comparisons are also examined. Section III concludes the paper, indicating the implications of the results for the analysis of channeling impacts.

I. SCREEN DATA AND RANDOMIZATION

The source and nature of the screen data on which this analysis is based are discussed below, and sample sizes are indicated. This is followed by a brief description of the randomization procedures.

A. THE SCREEN DATA

The screening instrument was developed to identify those elderly individuals who were at high risk of nursing home placement (those who in the absence of channeling would be in an institution). A set of objective criteria were established that were felt would distinguish such individuals. Data collected from the screen were used to establish whether a given applicant satisfied these criteria and should therefore be classified as eligible. The criteria incorporated the following dimensions: severe functional impairment; expected unmet need in two service categories (e.g., meal preparation, housework, administration of medication or medical treatment, etc.) for six months or more, or expected lack of sufficient help from family and friends in the coming months; residence in the community or, if institutionalized, certified as likely to be discharged into a noninstitutional setting within three months; residence within the project's geographical boundaries; age; and (for financial control sites only) Medicare Part A eligibility.[1]

The screening instrument was designed for a short telephone interview, to be administered in a uniform manner by each of the 10 demonstration projects. The telephone screening process was intended to reduce the cost of determining appropriateness for channeling compared to using a comprehensive in-person assessment for that purpose. Channeling project staff who conducted the screening interviews were in a separate administrative unit from assessment and case management staff. This was required chiefly to preserve the integrity of the experimental design--the potential for influencing the behavior of persons assigned to the control groups through contact with channeling staff was minimized by this administrative separation.

Applicants for channeling services came to the attention of the screening unit primarily in two ways: elderly individuals (or family, friends, clergy, neighbors, or other persons acting on their behalf) contacted the screening unit directly, or formal provider organizations contacted channeling to make a referral. Hospitals, home health agencies, senior centers, and nursing homes were among the formal referral sources. Screeners were instructed to conduct the interview directly with potential clients where possible, but could also accept reports from formal referral sources, families, friends, and other proxies.

[1] For more details on the eligibility criteria, see The Planning and Implementation of Channeling: Early Experiences of the National Long Term Care Demonstration (April, 1983). [http://aspe.hhs.gov/daltcp/reports/implees.htm]

Projects imposed guidelines of generally no more than 72 hours from referral to screen completion (although it was not always possible to meet these guidelines). Most screens were conducted by telephone, but in a very small proportion of the cases in-person screens were performed instead. Major reasons reported for the use of in-person screens included applicants who had hearing impairments, difficulty understanding the project, or no access to telephones.

The analysis presented here is based on the screen data for 6,327 research sample observations--3702 treatments and 2625 controls (see Table 1). Those who enrolled either before or after the March 1982 to June 1983 period[2] during which randomization of eligible applicants occurred are not included in this analysis, nor are individuals residing in the same household as a previously assigned sample member. A small number of eligible applicants (15 control group members) are excluded because their screening instruments were lost in the mail.

TABLE 1. Number of Research Sample Members with Completed Screen Interviews

Sites	Treatments	Controls	Total
BASIC CASE MANAGEMENT			
Baltimore	417	271	688
Eastern Kentucky	246	242	488
Houston	401	273	674
Middlesex County	451	299	750
Southern Maine	264	260	524
Total	1,779	1,345	3,124
FINANCIAL CONTROL			
Cleveland	388	191	579
Greater Lynn	309	308	617
Miami	450	297	747
Philadelphia	581	288	869
Rensselaer County	195	196	391
Total	1,923	1,280	3,203
ALL SITES	3,702	2,625	6,327

NOTE: An additional 15 control group members were randomized and completed the screen, but the instruments were lost in the mail and are therefore unavailable for analysis.

B. THE RANDOMIZATION PROCESS

After the screen was completed, eligible applicants were randomly assigned to either the treatment or the control group. The randomization process was designed to be as error-free and easy to implement as possible. A random number generator was used to create a string of ones and zeroes for each site, designating treatment and control status, respectively.[3] Sequential research identification numbers and the

[2] The start and end dates of the randomization period varied by site. March 1982 was the earliest start date; June 1983 was the latest end date.

[3] The ratio of ones to zeroes generated varied by site, ranging from 2:1 to 1:1. See the Research Design of the National Long Term Care Demonstration (November, 1982). [http://aspe.hhs.gov/daltcp/reports/designes.htm]

corresponding randomly assigned treatment/control status were then preprinted on labels, which were attached to applicants' randomization recording forms by MPR's survey staff in Princeton. The process is summarized below:

1. Sites called in daily with a list of eligible applicants. The clerk recorded the following information on a recording form for each applicant:
   - time of day applicant's name was transmitted (no two times could be the same)
   - name of applicant
   - address
   - phone number
   - social security number, Medicare number, or equivalent identification number
   - names of individuals over age 65 in applicants' household
   - birth date

   This information was then read back or spelled, as necessary, by the MPR clerk to ensure that it was recorded accurately.

2. The clerk arranged these forms by time of day, then placed the preprinted labels on the recording forms in sequential order. The labels contained a research ID number and randomly assigned treatment/control status.

3. A search of two card files (one arranged by social security number, one alphabetically by name) was conducted to determine whether the individual had ever previously applied. The search of the name file, which also contained the names of other members of households containing an applicant, was also used to determine whether the current applicant was living in or had previously lived in a household containing a person who had already been randomized.[4]

4. If a match was found in either of these files, the new status assignment, if different from that obtained previously or from that of the household member who had previously applied, was changed to eliminate this difference, and information about any match found was entered on the recording form.

5. The site was called back to inform them of the research ID number and treatment/ control status of each eligible applicant from the previous day. The site staff were required to read back this information to ensure that it had been recorded accurately. Checks of the applicant's name and social security number were also done at this time to make certain the information was attached to the correct person's record. Sites sent a copy of their record to MPR, which was later checked against MPR's records to ensure that no errors occurred in recording status over the phone.

6. Cards were prepared and inserted in the card files for the applicant, and a card was also prepared for every other household member over age 65.

[4] This search process has continued even after the end of the randomization period and will continue as long as sites are accepting new clients. Thus, persons previously assigned to experimental or control status and those living in the households of previously assigned persons will maintain or be assigned to the appropriate status.

7. At the end of each month, a list of new clients was sent to each site to confirm in writing the status of each client.

This structured process leaves little room for error or ambiguity and has worked well throughout the course of the project. Although three instances of misrecorded treatment status were discovered,[5] no general problems with these procedures were encountered that could compromise the integrity of the experimental design. Thus, the procedures used are not likely to result in differences between the treatment and control groups.[6] However, significant differences between the groups could result by chance. Only empirical analysis of the data, as discussed below, can reveal whether either sampling error or procedural mistakes have produced non-equivalent treatment and control groups.

[5] The three cases were all cases in which a control group member was erroneously listed as a treatment and allowed to participate. These cases are treated as controls in the analysis.

[6] It is possible for the procedures indirectly to result in unequal groups. For example, referral agencies that do not fully understand randomization could become disturbed about the proportion of their clients that, by chance, are assigned to the control group, and stop referring clients to channeling. This type of behavior could result in statistically significant differences between the distributions of the two groups by referral source. If there are substantive differences between clients from different referral sources, this could result in significant differences between treatments and controls on other observed or unobserved traits. We know of no such behavior by referral agencies, however, and the data do not seem to support such a conclusion.

II. ASSESSMENT OF EQUIVALENCE OF TREATMENT AND CONTROL GROUPS

To assess whether the treatment and control groups created by the randomization procedures were equivalent at the time of randomization, variables describing the characteristics of the sample members were constructed from the screen data. Mean values of these variables were obtained for treatment and control groups at each site and a standard statistical test of the difference between these means[7] was conducted. This statistical test provides us with an indication of whether any observed differences between the two groups on average should be considered "large" relative to what would be expected as the result of chance sampling variability. If the difference between the means is so great that randomly drawn samples would produce a difference that large fewer than one time in 10, we may not be very confident that the two groups being compared are alike enough that mean outcomes for the control group can be assumed to provide a good indication of what would have happened to treatments in the absence of channeling. For screen values of outcome variables, such as ability to perform activities of daily living (ADL), statistical significance of differences is also important because they imply that even if channeling had no impact at all, a comparison of treatment and control group means on ADL at six months after randomization would appear to indicate that channeling had had a statistically significant impact, because the difference in initial values would also be reflected in the values at six months. Statistically significant differences will occur by chance, especially when many different variables are being examined. However, the differences are not expected to be pervasive or large.

The statistical tests of the treatment/control difference in mean values of a set of variables will indicate whether such problems exist for any given site. However, because of the relatively small number of observations at each site, most of the analysis of channeling will be based on treatment/control differences at the model level, to ensure a high level of precision (i.e., the ability to distinguish between fairly small impacts of channeling and differences between treatment and control groups arising simply by chance).

Mean values of variables for the treatment and control, groups could be computed and dated for statistically significant differences at the model level; however,

---

[7] The test was a two-tailed student t-test. The test statistic, based on the assumption that the variable (X, say) is normally distributed, has the following form:

$$t = (X_T - X_C) / [S^2(X_T) + S^2(X_C)]^{1/2}$$

where X_T, X_C refer to the sample means for the variable X for the treatment and control groups, respectively, and $S^2(X_T) + S^2(X_C)$ refer to their estimated sample variances. Critical values for assessing statistical significance at any desired confidence level are readily available. The test statistics are reliable for variables not normally distributed as well, given the large sample sizes available here.

the results could be very misleading. This is because the ratio of treatment to control group members is different in different sites,[8] ranging from about 2:1 in the larger sites to 1:1 in the smaller sites. Simple means at the model level for each group are equal to a weighted average of the five site means for the group, with the weight for each site being the proportion of observations for the group which come from that site. Thus, in estimating the model level mean for the treatment group, the treatment group mean at a 2:1 site will have a larger weight than will the corresponding site control group mean in estimating the model level mean for controls.

This different weight applied to treatment and control groups from a given site can lead to anomalous results and can eliminate the very advantages that a randomized design offers. For example, suppose that randomization "worked" perfectly in that at every site the treatment/control group differences were zero, but that Site A had applicants with much lower functional ability (ADL, say) than other sites. Suppose further that 25 percent of the treatment group came from Site A but only 15 percent of the control group did. This would result in a treatment group mean at the model level that was lower than the control group mean, simply because the site with low ADL comprised a greater proportion of the treatment group, and in spite of the fact that the randomization process produced equivalent treatment and control groups in every site.

What is required is a procedure that preserves the equivalence of the two groups in comparisons at the model level. That is, the estimated model-level difference between treatments and controls should be a weighted average of the site-level difference. An attractive choice for a set of weights would be one in which the site differences that were measured most precisely received the largest weights. That is the procedure implemented in this report. In practice, this weighted average is obtained by regressing each variable being examined (e.g., age, ADL, etc.) on a treatment/control binary variable and five site binaries. It can be shown (see Appendix) that the estimated regression coefficient on the treatment status variable will be a weighted average of the treatment/control differences at the five sites, with the weights being largest for the sites with the largest total sample size (N_i) and the most even proportional split between treatment and control groups. The weight for the i^{th} site is:[9]

$$W_i = N_i r_i (1 - r_i) / \Sigma N_j r_j (1 - r_j),$$

where r_i is the proportion of observations from the i^{th} site that belong to the treatment group. Standard errors and t-statistics of these estimates are readily obtained from the computer printout.

[8] These differences arose from the design change enacted in May, 1983 to account for the initial underestimate of the number of eligible applicants, thereby boosting the overall sample size back to the level necessary to obtain the desired precision. See the Research Design of the National Long Term Care Demonstration (November, 1982). [http://aspe.hhs.gov/daltcp/reports/designes.htm]

[9] The weights for the screen sample are presented in the Appendix and in the tables containing the site-specific comparisons in Section B.

These estimates and test statistics are presented below for screen data on a variety of variables. Treatment group means are also presented for reference.[10] For continuous variables (e.g., age, income) discrete categorical variables have been defined to help identify any differences in distribution between the two groups that might be obscured by simply comparing the variable means for the two groups. Results are presented for the model level differences followed by a brief discussion of site-specific differences in means.

A. TREATMENT/CONTROL DIFFERENCES AT THE MODEL LEVEL

The screen contains data on respondents' demographic characteristics, financial resources, living arrangement, health and functioning, help received, and referral source. The variables in each of these categories that were used in the comparisons of treatment and control groups were:

- Demographic: age, sex, ethnic background.
- Financial Resources: monthly income, types of insurance coverage.
- Living Arrangement: proportion in long-term care institution; proportion living alone, with spouse, with others, or with spouse and others.
- Health and Functioning (see below): activities of daily living (ADL) index, cognitive impairments affecting functioning, unmet needs for service.
- Help Received: whether help is received in the areas of meal preparation, household or shopping, taking medicine, medical treatments at home, and personal care; expected lack of sufficient support from family and friends in coming months (fragile informal supports).
- Referral Source: whether referred to channeling by family, by a hospital, by a home health agency, etc.
- Nursing Home Application: whether have applied for admission, to nursing home or currently on nursing home waiting list.

The health and functioning variables require some further explanation. A modified version of the activities of daily living scale (Katz et al., 1970) which consists of questions in six areas--bathing, dressing, toileting, transfer, continence, and eating--was used as the primary determinant of functional disability. This version of the scale relies on client self-reports and uses a three-level classification for each area--independent, moderate, and severe--with total scores ranging from 0 to 12 (a low score indicating

[10] The treatment group means for the two models are also weighted averages of the site means for the treatment group. The same weight is used in these constructions. See the Appendix for details.

severe disability). Instrumental activities of daily living, also included on the screening instrument, are not considered in this analysis, because these questions were required to be asked only of those applicants who had fewer than two moderate or severe impairments on the ADL scale (about 13 percent of the sample). "Cognitive impairments" include disorientation, confusion, impairment of judgment, memory loss, or inappropriate behavior, and are reported if they affect the daily functioning of the applicant. Service areas of meal preparation, housework or shopping, taking medicine, medical treatments at home, and personal care are assessed for unmet needs. If these are expected to be unfulfilled for six months or more, they are included in the count of unmet needs.

TABLE 2. Percent of Cases with Missing Data on Screen Characteristics

Screen Characteristics	Basic Case Management Model		Financial Control Model	
	Percent Missing of Total	Treatment/Control Difference	Percent Missing of Total	Treatment/Control Difference
DEMOGRAPHICS				
Age	[a]	0.1	0.0	0.0
Sex	0.0	0.0	0.0	0.0
Ethnic Background	0.2	-0.3*	0.2	0.2
FINANCIAL RESOURCES				
Income (categories)	3.8	0.6	8.0	-0.9
Income	22.8	0.0	21.7	0.4
Insurance Coverage	2.6	0.1	0.0	0.0
LIVING ARRANGEMENT				
LTC Institution	0.0	-0.1	0.0	0.0
Community Living Arrangement	0.1	0.1	0.1	0.0
HEALTH AND FUNCTIONING				
Activities of Daily Living	2.4	-0.5	2.7	-0.2
Cognitive Impairments Affecting Functioning	8.4	-0.7	3.0	0.0
Number of Unmet Needs	5.0	0.3	5.0	0.3
EXISTING CARE AND CONTACTS				
Current Health With Services Received:				
Meal Preparation	20.8	-1.5	28.4	-1.4
Housework/Shopping	20.9	-1.6	28.5	-1.3
Taking Medicine	21.4	-1.4	29.3	-1.3
Medical Treatments	21.7	-1.7	29.3	-1.5
Personal Care	21.0	-1.4	28.9	-1.5
Fragile Informal Supports	10.9	-0.3	5.0	0.5
Nursing Home Waiting List	20.7	-1.8	28.4	-2.3
Referral Source	0.9	0.4	0.7	0.9***

NOTE: Variable definitions are contained in text and in footnotes to Table 3.

a. Less than .05 percent (one case) had missing data.

* Significantly different from zero statistically at the 90 percent confidence level (using a two-tailed test).
** Significantly different from zero statistically at the 95 percent confidence level (using a two-tailed test).
*** Significantly different from zero statistically at the 99 percent confidence level (using a two-tailed test).

Before turning to the results, it is also instructive to examine the extent of missing data at the screen. Differences between the two groups in the amount of missing data could result by chance, or by sites amending the screen data on clients subsequent to randomization, in the interest of having data on clients that are as accurate and complete as possible. Table 2 contains the percent of cases with missing data for the full sample, and the difference between treatments and controls on this dimension. In general, there are very little missing data, and virtually no differences between

treatment and control groups. Exact income data are missing for nearly a quarter of the cases for both models, but information on the range in which income fell is available for all but 3.8 percent of the sample in the basic sites and 8 percent in the financial control sites. Data on current help with services and on nursing home application are missing for about 21 percent of sample members in basic sites and about 29 percent of those in financial control sites. These high rates are because these questions were not asked of sample members who were in nursing homes or in the hospital at the time the screen was given; less than 3 percent of sample members who were asked these questions failed to respond. The only significant treatment-control differences in the percent with missing data are for ethnic background in basic sites and referral source in the financial control sites. For both of these variables, the overall percent missing is less than one percent, and the differences are small in absolute size. There is no evidence of systematic augmenting of screen data for clients.

1. Basic Case Management Model

There is very little difference between treatments and controls in the basic case management model. Of the 53 variables examined in Table 3, the only statistically significant difference between treatments and controls was in the proportion of referrals from case management agencies. Treatment/control differences tended to be small in relation to the mean for the treatment group, with very low test statistics. Furthermore, a joint test that the multiple correlation' between treatment/control status and all of the variables (controlling for site) is zero could not be rejected.[11]

The average age of both groups at the basic case management sites is 79 years, and treatments and controls are equally likely to be male (28.7 percent). The average monthly income of treatments is only 1 percent higher than that of controls (532 dollars versus 526 dollars). Treatments report Medicaid coverage slightly more often than do controls (20.4 percent versus 19.7 percent, respectively) and declare Medicare insurance-slightly less often (96.3 percent of treatments compared to 97.0 percent of controls).

Similarly, no significant differences between treatments and controls exist in the area of health and functioning. An average of three unmet needs are reported for both treatments and controls at the basic case management sites. Control and treatment group members are about equally impaired on the ADL scale.[12] There are also no

[11] This test is a summary test of whether there are any differences between treatment and control groups. Treatment/control status is regressed on a set of binary site variables and all other characteristics in Table 3. An F-test is then conducted of whether all coefficients other than those on the site variables are equal to zero. If this hypothesis cannot be rejected, the probability that a sample member belongs to the treatment group is not significantly higher for some types of individuals than for others. One implication of this for the analysis is that regression estimates of treatment/control differences that control only for differences in distribution across sites will not differ substantially from estimates obtained from regressions which control for many other factors.

[12] Individual components of the ADL index were also examined for treatment/control equivalence. No significant differences were found. The instrumental activities of daily living (IADL) scale, which was present for only 13 percent of the sample, was also examined; the treatment/control differences were small and statistically insignificant.

substantive differences between the two groups in the proportion receiving help with most services at the time of the screen. These proportions range from about .30 percent to over 70 percent, depending upon the service. The proportion with fragile informal supports was about 85 percent for both groups.

TABLE 3. Screen Characteristics of Treatment Group and Treatment/Control Differences

Screen Characteristics	Basic Case Management Model		Financial Control Model	
	Treatment Group	T/C Difference	Treatment Group	T/C Difference
DEMOGRAPHICS				
Age (percent):				
Less than 75	29.7	0.2	25.4	-1.3
75-84	44.3	-1.2	44.8	-0.4
85 and over	26.0	1.0	29.8	1.7
Mean age	79.1	0.1	80.2	0.3
Sex (percent):				
Male	28.7	0.0	29.0	1.6
Female	71.3	0.0	71.0	-1.6
Ethnic Background (percent):				
Black (not of Hispanic origin)	20.5	-1.8	20.3	-1.1
Hispanic	1.8	0.1	5.3	0.0
White (not of Hispanic origin)	77.6	1.8	74.2	1.1
Other (American Indian, Asian, other)	0.1	-0.1	0.2	0.1
FINANCIAL RESOURCES				
Income (percent):				
Less than $500	57.5	-1.1	58.9	-0.3
$500 to $999	33.9	-0.2	35.5	1.9
$1,000 or more	8.5	1.3	5.7	-1.6*
Mean monthly income (dollars)	532	5.6	513	-13.0
Insurance coverage (percent):				
Medicare only	37.3	0.0	27.6	2.7*
Medicare and private insurance	40.5	-0.9	49.2	-2.6
Medicare and Medicaid	18.5	0.2	23.2	0.0
Medicaid only	1.9	0.5	0.0	-0.1
Private insurance only	1.5	0.5	0.0	0.0
No insurance	0.3	-0.3	0.0	0.0
LIVING ARRANGEMENT				
Type of Living Arrangement (percent):				
Nursing home or LTC facility	4.1	0.2	2.1	-0.1
Community	95.9	-0.2	97.9	0.1
Community Living Arrangement (percent):[a]				
Alone	35.7	-0.4	39.8	-0.7
With spouse only	28.0	2.6	27.9	0.6
With spouse and others	3.2	-0.7	3.1	0.3
With others	33.1	-1.5	29.2	-0.2
HEALTH AND FUNCTIONING				
Activities of Daily Living[b] (percent):				
Mild	5.9	0.0	3.4	-0.9
Moderate	23.1	0.6	18.8	-2.2
Severe	40.6	1.4	43.6	2.0
Very severe	30.4	-2.0	34.3	1.1
Mean ADL score	6.2	0.1	5.9	-0.1
Cognitive Impairments Affecting Functioning (percent)	58.5	0.7	60.0	0.3
Number of Unmet Needs[c] (percent):				
0-1	8.0	-0.3	3.9	-0.2
2-3	58.1	-0.8	65.3	-0.6
4-5	33.9	1.0	30.8	0.8
Mean number of unmet needs	3.0	0.0	3.0	0.0

TABLE 3 (*continued*)				
Screen Characteristics	**Basic Case Management Model**		**Financial Control Model**	
	Treatment Group	**T/C Difference**	**Treatment Group**	**T/C Difference**
EXISTING CARE AND CONTACTS				
Current Help With Services Received (percent):				
Meal Preparation	68.5	-1.8	73.6	-2.2
Housework/Shopping	73.3	-1.1	77.1	-1.5
Taking Medicine	45.9	-1.1	52.0	-1.2
Medical Treatments	29.3	1.5	36.9	-0.4
Personal Care	61.5	-1.0	69.3	-2.7
Fragile Informal Supports	84.8	0.9	89.2	0.5
Nursing Home Waiting List (percent)	6.9	1.4	5.5	0.6
Referral Source (percent):				
Family/friend/self-referral	33.4	-1.7	20.7	-0.7
Hospital	19.9	-0.1	27.1	-0.9
Home health agency	11.9	0.2	22.9	-0.9
Senior center/nutrition	3.2	-0.4	9.0	0.4
Case management agency	6.7	2.0**	5.1	1.5*
Welfare/Medicaid	4.9	10.0	2.5	0.4
Information and referral agency	4.2	-0.6	0.9	0.3
Nursing home	2.7	0.6	1.6	-0.1
Channeling outreach	0.9	0.2	1.4	-0.8*
Other	12.3	-0.1	8.7	0.8
MAXIMUM SAMPLE SIZE[d]	3,124		3,203	

NOTE: Estimated treatment group means and treatment/control differences are weighted averages of site level treatment group means and treatment/control differences, respectively. See Appendix for further explanation.

a. "Community living arrangement" includes those living in the community or hospitalized at the time of the screen. For those hospitalized, living arrangements prior to hospitalization are reported.
b. Activities of daily living include bathing, dressing, toileting, transfer, continence, and eating. In each of the six areas, sample members are classified by degree of impairment in performing these functions (severe, moderate, none) and assigned a "score" (0, 1, 2, respectively). The six scores are summed, and individuals are classified into one of four categories based on their total score: mild (11-12), moderate (9-10), severe (5-8), very severe (0-4).
c. "Unmet needs" include meal preparation, housework or shopping, taking medicine, medical treatments at home, and personal care.
d. The maximum sample size is the number of cases with screens. Means for any particular variable may be based on smaller samples due to item nonresponse. See Table 2 for the proportion of cases with missing data for each of the variables in this table.

* Significantly different from zero statistically at the 90 percent confidence level using a two-tailed test.
** Significantly different from zero statistically at the 95 percent confidence level using a two-tailed test.
*** Significantly different from zero statistically at the 99 percent confidence level using a two-tailed test.

As already noted, treatments are significantly more likely than controls to be referred by a case management agency--6.7 percent, compared to 4.7 percent, respectively. There are no other noteworthy differences in referral source. The most common referral source for both treatments and controls is family, friend, or self-referral. About one-fifth of both treatments and controls are referred by hospitals.

2. Financial Control Model

As for the basic sites, very few differences between treatments and controls were found. As shown in Table 3, statistically significant differences (at the 90 percent level) were identified for only four of the 53 variables: incomes over 1,000 dollars per month (but not average income), Medicare-only insurance, coverage, referral by case management agencies, and referral by channeling outreach. These differences tended to be small in absolute terms, and none were significant at the 95 percent level. As for

the basic sites, the joint test that the multiple correlation between treatment/control status and all of the variables in Table 3 is zero could not be rejected.

Demographics and living arrangements show no significant differences between treatments and controls for the financial control model. Slightly more treatments than controls are male; slightly more controls than treatments are black. The proportion of treatments with income in excess of 1,000 dollars per month was significantly lower for treatments than controls (5.7 versus 7.3 percent, respectively); however, the difference is not large in absolute terms and the average incomes of the two groups do not differ significantly. Just over 2 percent of both treatments and controls lived in long term care institutions at the time the screen.

Although equal numbers of treatments and controls report Medicare insurance coverage (99.9 percent), treatments are significantly more likely than controls to report only Medicare insurance (27.6 percent versus 24.9 percent, respectively). Medicare combined with private insurance covers more controls (51.8 percent) than treatments (49.2 percent), although this difference is not significant.

In the area of health and functioning, only small and insignificant differences between treatments and controls are observed. Controls are slightly less disabled than treatments on the ADL scale (the average score for controls is 6.0; for treatments it is 5.9).[13] Both treatments and controls average three unmet service needs. No significant differences between treatments and controls are observed for current help with services. For four of the services reported, over half of the treatments and controls received help at the time of the screen. Differences in the proportion with fragile informal supports was negligible.

As already noted, two referral sources show significant differences between treatments and controls. More treatments (5.1 percent) than controls (3.6 percent) were referred by a case management, agency. Controls were significantly more likely to be referred by channeling outreach--2.2 percent of controls versus 1.4 percent of treatments. These differences represent such a small proportion of the sample that they are not considered especially important. The differences by referral source do not seem to have resulted in differences between the individuals in the two groups.

B. TREATMENT/CONTROL DIFFERENCES AT THE SITE LEVEL

The bulk of the analysis of the effects of channeling will be conducted at the model level; hence, the discussion above has focused on differences between the groups at this level. However, since some of the analysis will be conducted at the site level, we have also examined differences between treatments and controls at the screen for each site. In addition, if there were systematic problems with the procedures

[13] As for the basic sites, the components of ADL and the IADL scale were examined for treatment/control differences. These differences were all small and all but one (the proportion with no bathing impairment) were statistically insignificant.

they are more likely to exist at specific sites. Because of the smaller sample sizes, large differences between the groups have a much higher probability of occurring simply by chance at the site level than at the model level. Again, statistical tests guide us in determining what should be considered a large difference. It should be kept in mind, however, that since there are five times as many comparisons being made in each model, we expect to find many more statistically significant differences occurring simply by chance in the site level comparisons.

1. Basic Case Management Sites

Out of over 250 comparisons at the five basic sites, we find 15 statistically significant differences between treatments and controls.(at the 90 percent or greater confidence level). This is substantially less than the 25 that might be expected to occur simply by chance. As shown in Table 4, the significant differences were more prevalent in Kentucky than in other sites, but tended to be scattered rather than concentrated in specific variables. Thus, there is no indication of systematic tampering with the random assignment process.

Kentucky shows significant differences in the proportion of channeling appropriates aged 75-84 (but not in average age), the proportion living alone, the fraction with mild or severe ADL impairments (but not in average ADL score), and the percent of persons referred by a case management agency. The proportion of individuals for whom private insurance is their only source of coverage is significantly higher for treatments than controls in Kentucky, Southern Maine, and Middlesex County, but is very small (less than 3 percent) for both groups in each site. In Baltimore, controls are more likely than treatments to be covered by both Medicare and Medicaid. Treatments are significantly less likely than controls in Houston to be referred by an information and referral agency. (It is this difference that accounts for the statistically significant difference in referral source at the model level.) In Middlesex County, controls are significantly more likely than treatments to live with others and to receive help with meal preparation, housework or shopping, and taking medicine. Treatments at the Middlesex County site are significantly more likely to be on a nursing home waiting list.

The scattered differences found at the site level are likely to be due to chance sampling variability. However, the differences indicate that site-specific impact estimates will have to be interpreted with greater care than the model results.

2. Financial Control Sites

Significant differences were somewhat more frequent for financial control sites, with 28 of the 255 comparisons being statistically significant at the 90 percent confidence level. This is about the number that would be expected to occur by chance. In Table 5 it can be seen that Greater Lynn and Cleveland had more such differences than other sites, but nothing which indicates that clients were systematically more or less disabled than controls.

The significant differences are scattered across the variables examined. None of the financial control sites shows, significant differences in demographics. Mean income is significantly higher for controls in Greater Lynn. In Cleveland and Greater Lynn, some differences in insurance coverage occur. Significant differences in living arrangements are confined to Rensselaer County and Greater Lynn.

Controls are significantly more likely to be rated as only mildly or moderately impaired on the ADL scale in Rensselaer County, which results in a mean ADL score for controls' that is significantly higher than that of their treatment counterparts. In Miami, treatments are judged very severely impaired significantly more often than controls. In Greater Lynn the opposite is true--controls are significantly more likely to be rated very severely impaired. Nonetheless, neither of these differences leads to significant differences in mean ADL. Unmet needs are significantly higher for treatments than controls in Miami.

In both Greater Lynn and Philadelphia, treatments are significantly less likely than controls to receive help with various services. Scattered statistically significant differences between treatments and controls in referral sources are found in Cleveland, Philadelphia, and Rensselaer County.

As for the basic sites, the differences found at the site level are probably due to chance sampling variability. However, regression procedures should be used to control for the effect of pre-existing differences between the groups at the site level.

TABLE 4. Screen Characteristics of Treatment Group and Treatment/Control Differences: Basic Case Management Sites												
Screen Characteristics	**Baltimore**		**Eastern Kentucky**		**Houston**		**Middlesex County**		**Southern Maine**		**Total**	
	Treatment Group	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**
DEMOGRAPHICS												
Age (percent):												
65-74	30.5	-0.2	31.3	4.0	30.9	-2.8	27.5	1.7	28.9	-1.1	29.7	0.2
75-84	44.8	-2.8	43.1	-9.0**	42.9	-2.5	45.2	2.8	45.2	4.1	44.3	-1.2
85 and over	24.7	2.9	25.6	4.9	26.2	5.3	27.3	-4.5	25.9	-3.0	26.0	1.0
Mean age	79.1	0.7	79.2	0.1	78.9	0.7	79.5	-0.3	78.9	-0.7	79.1	0.1
Sex (percent):												
Male	28.1	0.4	25.6	-0.4	28.7	-1.7	26.2	0.7	36.0	1.0	28.7	0.0
Female	71.9	-0.4	74.4	0.4	71.3	1.7	73.8	-0.7	64.0	-1.0	71.3	0.0
Ethnic Background (percent):												
Black (not of Hispanic origin)	50.6	-2.8	2.4	0.0	39.5	-3.1	3.1	-1.9	0.0	-0.4	20.5	-1.8
Hispanic	0.2	0.2	0.0	0.0	5.3	0.8	2.7	-0.7	0.0	0.0	1.8	0.1
White (not of Hispanic origin)	48.9	2.7	97.6	0.0	55.3	2.7	94.0	2.4	100.0	0.4	77.6	1.8
Other (American Indian, Asian, other)	0.2	-0.1	0.0	0.0	0.0	-0.4	0.2	0.2	0.0	0.0	0.1	-0.1
FINANCIAL RESOURCES												
Monthly income (percent)												
Less than $500	65.9	-2.3	60.2	0.8	62.2	-3.2	54.4	1.3	42.9	-2.3	57.5	-1.1
$500 to $999	28.3	2.9	34.9	-0.3	29.1	0.3	37.7	-3.5	40.9	0.1	33.9	-0.2
$1,000 or more	5.8	-0.7	5.0	-0.5	8.7	2.9	7.9	2.3	16.3	2.3	8.5	1.3
Mean monthly income (dollars)	489	11.7	506	-22.5	514	43	550	8.9	611	-47.8	532	5.6
Insurance coverage (percent):												
Medicare only	43.2	4.1	49.4	0.4	45.5	-2.1	27.6	0.1	22.5	-2.5	37.3	0.0
Medicare and private insurance	32.5	1.9	17.6	-5.5	29.6	-0.7	55.6	-3.3	63.4	3.0	40.5	-0.9
Medicare and Medicaid	18.4	-6.2*	31.8	4.6	20.9	2.5	13.6	2.5	10.3	-2.8	18.5	0.2
Medicaid only	3.5	1.0	0.0	-0.8	2.0	1.3	2.2	0.2	1.1	0.8	1.9	0.5
Private insurance only	1.9	-0.6	1.2	1.2*	1.3	-0.6	0.9	0.9**	2.7	1.9*	1.5	0.5
No insurance	0.5	-0.3	0.0	0.0	0.7	-0.4	0.2	-0.4	0.0	-0.4	0.3	-0.3
LIVING ARRANGEMENT												
Type of Living Arrangement (percent):												
Nursing home or LTC facility	5.0	1.7	0.4	-1.2	4.2	1.7	1.3	-0.3	9.8	-1.7	4.1	0.2
Community	95.0	-1.7	99.6	1.2	95.8	-1.7	98.7	0.3	91.2	1.7	95.9	0.2
Community Living Arrangement (percent):[a]												
Alone	39.9	2.5	36.3	-7.8*	36.7	-3.7	31.7	3.8	34.0	1.4	35.7	-0.4
With spouse only	22.0	2.9	33.5	4.1	31.3	4.5	23.1	2.4	33.6	-1.2	28.0	2.6
With spouse and others	3.0	-2.7	0.0	-0.4	2.1	-1.3	6.5	0.4	3.4	0.8	3.2	-0.7
With others	35.1	-2.7	30.2	4.2	29.9	0.5	38.7	-6.6*	29.0	-1.0	33.1	-1.5
HEALTH AND FUNCTIONING												
Activities of Daily Living (percent):[b]												
Mild	7.9	2.1	9.5	-5.1*	2.5	1.0	3.4	0.0	7.8	1.1	5.9	0.0
Moderate	20.2	2.9	29.3	-2.9	23.5	0.0	19.4	-1.1	25.2	3.9	23.1	0.6
Severe	37.9	-2.5	37.6	8.3*	43.0	-2.0	40.2	4.9	44.2	-0.7	40.6	1.4
Very severe	34.0	-2.5	23.6	-0.3	30.9	1.0	37.0	-3.8	22.9	-4.3	30.4	-2.0
Mean ADL score	6.0	0.2	7.0	-0.4	6.1	-0.1	5.6	0.2	6.8	0.3	6.2	0.1
Cognitive Impairments Affecting Functioning (percent)	55.1	0.7	57.7	2.8	68.4	1.4	59.3	-5.1	54.2	6.1	58.5	0.7

TABLE 4 (*continued*)

Screen Characteristics	Baltimore		Eastern Kentucky		Houston		Middlesex County		Southern Maine		Total	
	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference
Number of Unmet Needs (percent):[c]												
0-1	11.7	-4.7	4.1	0.8	0.3	-0.1	0.9	-0.4	27.1	4.1	8.0	-0.3
2-3	56.3	4.8	54.5	-4.6	44.6	-2.6	73.1	0.3	59.3	-2.8	58.1	-0.8
4-5	32.0	-0.1	41.5	3.9	55.2	2.7	26.0	0.2	13.6	-1.2	33.9	1.0
Mean number of unmet needs	2.9	0.1	3.3	0.0	3.6	0.0	3.0	0.1	2.2	-0.1	3.0	0.0
EXISITING CARE AND CONTACTS												
Current Help With Services Received (percent):												
Meal preparation	79.9	-0.1	67.5	1.0	21.9	-3.5	85.3	-5.0*	91.5	0.3	68.5	-1.8
Housework/shopping	83.4	-0.6	74.9	-0.2	24.1	-0.4	92.0	-3.3*	95.2	-0.1	73.3	-1.1
Taking medicine	54.4	-4.2	44.2	4.8	13.0	4.0	59.2	-6.5*	60.6	-3.3	45.9	-1.1
Medical treatments	36.4	2.7	24.7	1.1	24.7	2.1	35.5	1.1	22.2	0.1	29.3	1.5
Personal care	67.8	0.8	53.3	0.8	24.4	-3.4	81.7	-1.9	80.9	-0.3	61.5	1.0
Fragile Informal supports	89.6	-1.6	94.9	1.0	54.4	-0.6	86.9	2.1	91.1	3.4	84.8	0.9
Nursing Home Waiting List (percent)	4.5	-0.7	5.2	1.5	5.8	2.1	10.4	4.5**	8.0	-2.2	6.9	1.4
Referral Source (percent):												
Family/friend/self-referral	28.7	-0.7	24.7	-3.9	37.2	0.8	41.6	-5.6	31.4	1.2	33.4	-1.7
Hospital	24.5	-1.1	4.9	-0.5	21.1	0.5	26.2	0.5	17.8	0.0	19.9	-0.1
Home health agency	18.8	0.2	8.2	-1.9	13.9	3.2	5.8	0.8	12.5	-2.6	11.9	0.2
Senior center/nutrition	0.7	-0.4	16.5	-2.9	1.5	0.0	0.4	0.4	0.0	0.0	3.2	-0.4
Case management agency	1.0	-0.1	14.8	8.1***	0.3	0.3	16.9	2.8	0.0	0.0	6.7	2.0**
Welfare/Medicaid	4.0	0.6	2.1	-0.9	13.3	0.8	1.8	-0.2	2.3	-0.4	4.9	0.0
Information and referral agency	0.7	0.0	0.0	0.0	1.8	-2.7*	1.6	-0.8	18.9	0.7	4.2	-0.6*
Nursing home	3.7	1.5	1.6	-0.9	1.5	0.4	0.2	0.2	7.6	1.8	2.7	0.6
Channeling outreach	0.2	-0.5	4.5	1.6	0.0	0.0	0.4	0.4	0.0	-0.4	0.9	0.2
Other	17.6	0.5	22.6	1.2	9.5	-3.3	1.4	1.4	9.5	-0.2	12.3	-0.1
MAXIMUM SAMPLE SIZE[d]	688		488		674		750		524		3,124	
SITE WEIGHT[e]	.2163		.1606		.2139		.2367		.1725			

NOTE: Estimated "total" treatment group means and treatment control differences are weighted averages of the site level treatment group means and differences, respectively. See text and Appendix for further explanation.

a. "Community living arrangement" includes those living in the community or hospitalized at the time of the screen. For those hospitalized, living arrangements prior to hospitalization are reported.
b. Activities of daily living include bathing, dressing, toileting, transfer, continence, and eating. In each of the six areas, sample members are classified by degree of impairment in performing these functions (severe, moderate, none) and assigned a "score" (0, 1, 2, respectively). The six scores are summed, and individuals are classified into one of four categories based on their total score: mild (11-12), moderate (9-10), severe (5-8), very severe (0-4).
c. "Unmet needs" include meal preparation, housework or shopping, taking medicine, medical treatments at home, and personal care.
d. The maximum sample size is the number of cases with screens. Means for any particular variable may be based on smaller samples due to item nonresponse. See Table 2 for the proportion of cases with missing data for each of the variables in this table.
e. The site weight is the weight applied to the treatment/control difference at a given site in obtaining the estimated overall difference for the model, for variables with no item nonresponse. Actual weights will differ slightly from these values due to different rates of item nonresponse across sites and between treatment and control groups. These weights are also used in constructing the overall treatment group means as weighted averages of the treatment group means from the five sites.

* Significantly different from zero statistically at the 90 percent confidence level using a two-tailed test.
** Significantly different from zero statistically at the 95 percent confidence level using a two-tailed test.
*** Significantly different from zero statistically at the 99 percent confidence level using a two-tailed test.

TABLE 5. Screen Characteristics of Treatment Group and Treatment/Control Differences: Financial Control Sites												
Screen Characteristics	**Cleveland**		**Greater Lynn**		**Miami**		**Philadelphia**		**Rensselaer County**		**Total**	
	Treatment Group	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**	**Treatment Group**	**T/C Difference**
DEMOGRAPHICS												
Age (percent):												
65-74	32.2	0.8	20.7	-3.6	20.7	2.1	29.1	-4.9	25.1	0.6	25.4	-1.3
75-84	42.5	0.6	46.0	-1.4	43.3	-4.8	46.0	4.6	46.7	-1.8	44.8	-0.4
85 and over	25.3	-1.4	33.3	5.1	36.0	2.7	25.0	0.3	28.2	1.2	29.8	1.7
Mean age	79.0	-0.4	80.8	0.9	81.3	-0.2	79.4	0.7	80.1	0.4	80.2	0.3
Sex (percent):												
Male	30.2	5.0	23.6	-3.6	29.8	4.5	29.8	-0.1	33.3	3.7	29.0	1.6
Female	69.8	-5.0	76.4	3.6	70.2	-4.5	70.2	0.1	66.7	-3.7	71.0	-1.6
Ethnic Background (percent):												
Black (not of Hispanic origin)	34.0	-1.1	1.0	-0.3	12.9	0.1	43.1	-3.8	1.5	0.5	20.3	-1.1
Hispanic	0.0	-0.5	0.3	0.3	21.4	0.5	0.5	-0.5	0.0	0.0	5.3	0.0
White (not of Hispanic origin)	65.5	1.1	98.7	0.0	65.5	-0.8	56.2	4.5	98.5	-0.5	74.2	1.1
Other (American Indian, Asian, other)	0.5	0.5	0.0	0.0	0.2	0.2	0.2	-0.2	0.0	0.0	0.2	0.1
FINANCIAL RESOURCES												
Monthly income (percent)												
Less than $500	53.9	-6.4	53.8	7.3	71.7	-1.1	60.5	1.3	42.8	-4.1	58.9	-0.3
$500 to $999	38.3	1.8	38.9	0.5	24.9	2.5	37.3	1.6	44.6	3.8	35.5	1.9
$1,000 or more	7.8	4.6**	7.3	-7.8***	3.4	-1.4	2.2	-2.8**	12.7	0.3	5.7	-1.6*
Mean monthly income (dollars)	546	38.6	574	-68.6**	427	22.2	481	-51.7	623	27.1	513	-12.6
Insurance coverage (percent):												
Medicare only	27.6	7.2*	21.7	2.2	30.4	-2.2	35.6	5.0	15.9	1.6	27.6	2.7*
Medicare and private insurance	57.5	-4.3	52.1	-7.0*	38.0	4.7	40.8	-5.0	70.8	-1.7	49.2	-2.6
Medicare and Medicaid	14.7	-2.6	26.2	4.8	31.6	-2.5	23.6	0.0	13.3	0.1	23.2	0.0
Medicaid only	0.0	-0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1
Private insurance only	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No insurance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LIVING ARRANGEMENT												
Type of Living Arrangement (percent):												
Nursing home or LTC facility	3.9	-0.3	3.9	-1.3	2.9	1.5	0.0	0.0	0.0	-1.0	2.1	-0.1
Community	96.1	0.3	96.1	1.3	97.1	-1.5	100.0	0.0	100.0	1.0	97.9	0.1
Community Living Arrangement (percent):[a]												
Alone	33.5	-5.0	46.8	8.1**	47.1	-2.7	29.9	1.1	43.1	-8.5*	39.8	-0.7
With spouse only	28.4	3.7	22.6	-8.6**	30.2	2.2	27.7	1.3	31.8	6.5	27.9	0.6
With spouse and others	4.8	2.1	2.0	0.3	0.9	-0.1	4.5	-1.8	4.1	2.6	3.1	0.3
With others	33.2	-0.9	28.6	0.2	21.7	0.6	37.9	-0.7	21.0	-0.6	29.2	-0.2
HEALTH AND FUNCTIONING												
Activities of Daily Living (percent):[b]												
Mild	2.9	-0.4	3.0	-0.3	6.2	-0.1	1.4	-1.1	3.2	-3.7*	3.4	-0.9
Moderate	21.6	2.6	20.0	2.0	23.0	-4.5	13.0	-3.0	16.6	-9.6**	18.8	-2.2
Severe	42.7	-2.4	55.7	4.3	37.1	-1.6	40.0	4.0	44.4	6.4	43.6	2.0
Very severe	32.7	0.1	21.3	-5.9*	33.6	6.1*	45.6	0.0	35.8	7.0	34.3	1.1
Mean ADL score	5.9	0.2	6.6	0.3	6.2	-0.4	4.9	0.0	5.9	-0.7**	5.9	-0.1
Cognitive Impairments Affecting Functioning (percent)	65.7	3.4	65.4	1.4	58.7	-0.4	56.7	-1.1	53.4	-1.2	60.0	0.3

TABLE 5 (*continued*)

Screen Characteristics	Cleveland		Greater Lynn		Miami		Philadelphia		Rensselaer County		Total	
	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference	Treatment Group	T/C Difference
Number of Unmet Needs (percent):[c]												
0-1	0.9	-0.3	1.3	0.3	3.0	-0.5	5.9	-2.2	10.3	3.6	3.9	-0.2
2-3	60.5	2.1	58.4	2.1	61.0	-6.5*	79.0	1.4	64.6	-0.9	65.3	0.6
4-5	38.6	-1.7	40.3	-2.4	36.0	7.0*	15.0	0.8	25.1	-2.7	30.8	-0.8
Mean number of unmet needs	3.2	0.0	3.3	0.0	3.2	0.1*	2.6	0.0	2.8	-0.2	3.0	0.0
EXISITING CARE AND CONTACTS												
Current Help With Services Received (percent):												
Meal preparation	76.9	-4.1	79.9	-6.6*	65.9	6.2	71.7	-8.9**	80.3	4.3	73.6	-2.2
Housework/shopping	80.6	-6.1	93.3	-2.7	64.1	2.7	72.9	-6.2*	82.9	7.7	77.1	-1.5
Taking medicine	58.0	-9.8*	54.9	0.4	44.9	5.6	50.1	-8.6**	59.1	8.7	52.0	-1.2
Medical treatments	26.7	-4.7	25.9	-5.5	35.1	-1.3	54.9	4.1	35.7	7.8	36.9	-0.4
Personal care	67.5	-5.8	74.1	-10.2***	56.1	2.3	79.0	-5.0	72.2	8.9	69.3	-2.7
Fragile Informal supports	90.0	4.3	76.5	2.4	94.5	-2.0	93.5	-1.1	90.4	0.4	89.2	0.5
Nursing Home Waiting List (percent)	5.0	0.9	6.6	1.6	5.6	0.9	4.8	-0.6	5.1	0.0	5.5	0.6
Referral Source (percent):												
Family/friend/self-referral	35.2	-1.9	6.2	2.6	41.0	-0.1	0.0	0.0	28.2	-6.5	20.7	-0.7
Hospital	29.3	-7.4*	26.7	3.7	16.8	-2.7	29.2	-1.8	39.5	5.3	27.1	-0.9
Home health agency	10.4	-1.2	41.4	-3.8	12.3	-1.2	29.9	0.6	15.9	1.1	22.9	-0.9
Senior center/nutrition	4.9	3.4**	0.0	0.0	2.7	-1.3	28.5	1.3	2.1	-1.0	9.0	0.4
Case management agency	5.7	1.5	10.7	2.0	1.3	0.0	6.0	2.8**	0.5	0.5	5.1	1.5*
Welfare/Medicaid	0.5	0.5	0.0	-0.3	7.4	2.0	0.0	0.0	5.1	-0.5	2.5	0.4
Information and referral agency	3.1	1.0	1.3	0.3	0.2	0.2	0.0	0.0	0.5	0.0	0.9	0.3
Nursing home	0.5	0.0	2.0	-0.6	4.3	0.6	0.2	-0.5	0.5	0.5	1.6	-0.1
Channeling outreach	0.0	0.0	6.5	-2.9	0.2	-0.1	0.0	0.0	0.0	-1.5*	1.4	-0.8*
Other	10.4	4.1*	5.2	-1.0	13.7	2.6	6.3	-2.4	7.7	2.1	8.7	0.8
MAXIMUM SAMPLE SIZE[d]	579		617		747		869		391		3,203	
SITE WEIGHT[e]	.1703		.2053		.2381		.2562		.1301			

NOTE: Estimated "total" treatment group means and treatment control differences are weighted averages of the site level treatment group means and differences, respectively. See text and Appendix for further explanation.

a. "Community living arrangement" includes those living in the community or hospitalized at the time of the screen. For those hospitalized, living arrangements prior to hospitalization are reported.
b. Activities of daily living include bathing, dressing, toileting, transfer, continence, and eating. In each of the six areas, sample members are classified by degree of impairment in performing these functions (severe, moderate, none) and assigned a "score" (0, 1, 2, respectively). The six scores are summed, and individuals are classified into one of four categories based on their total score: mild (11-12), moderate (9-10), severe (5-8), very severe (0-4).
c. "Unmet needs" include meal preparation, housework or shopping, taking medicine, medical treatments at home, and personal care.
d. The maximum sample size is the number of cases with screens. Means for any particular variable may be based on smaller samples due to item nonresponse. See Table 2 for the proportion of cases with missing data for each of the variables in this table.
e. The site weight is the weight applied to the treatment/control difference at a given site in obtaining the estimated overall difference for the model, for variables with no item nonresponse. Actual weights will differ slightly from these values due to different rates of item nonresponse across sites and between treatment and control groups. These weights are also used in constructing the overall treatment group means as weighted averages of the treatment group means from the five sites.

* Significantly different from zero statistically at the 90 percent confidence level using a two-tailed test.
** Significantly different from zero statistically at the 95 percent confidence level using a two-tailed test.
*** Significantly different from zero statistically at the 99 percent confidence level using a two-tailed test.

III. SUMMARY AND IMPLICATIONS FOR FUTURE ANALYSES

The overriding conclusion from all of the comparisons made between treatment and control groups is that the randomization procedure has resulted in groups that are very similar on observable characteristics. Very few significant differences are found at the model level, and these were judged to be inconsequential; nearly all differences were quantitatively small and had very small test statistics. Even for the site level comparisons, where larger differences were expected because of the smaller sample sizes, the number of statistically significant differences was no larger than would be expected by chance and no patterns of differences were found to indicate that noncomparable groups were obtained in any site. Thus, although there may be unobserved differences between the two groups, the comparisons on observed characteristics made here provide no evidence of either systematic deviations from the random assignment procedures or important treatment/control differences arising by chance.

This conclusion bas a number of important implications for the analysis of channeling's impacts. First and foremost, it implies that the control group provides a reliable measure of what would have happened to the treatment group its the absence of channeling, and therefore, that simple comparisons of outcomes for treatment and control groups (controlling for differences in distribution across sites) will yield unbiased estimates of channeling impacts. Second, the site level differences are probably small enough that controlling for them in a regression model should be sufficient to yield unbiased estimates of channeling's impact at specific sites. However, these estimates will have rather wide confidence intervals because of the small sample size at each site, and should be interpreted with considerable caution. Third, it implies that our investigations which rely on screen data to assess other possible sources of noncomparability of data for treatment and control groups will not be confounded by differences between the groups at randomization. These investigations include analysis of whether there is differential attrition at the baseline, and whether the baseline data, that are collected on the two groups are comparable. Finally, the treatment/control equivalence at the screen implies that if the baseline data do differ for treatments and controls, comparable variables from the screen can be substituted.

The tests of differential attrition at baseline and the comparability of baseline data will be conducted over the next several months and the results will be presented in the data comparability report in March. In assessing baseline attrition we will examine the differences in screen characteristics between baseline responders and nouresponders, separately for treatment and control groups. This examination will enable us to determine whether baseline attrition was random or related to certain characteristics, and whether any treatment/control screen differences among baseline respondents are due to systematic attrition for one group but not the other or for both groups. After baseline attrition differences are identified, baseline data on treatments and controls will

be examined for comparability. Differences arising from differential attrition (if any) will be controlled for in order to identify those treatment/control differences that are due to the differences in collection of the baseline data.

APPENDIX A. ESTIMATION METHODOLOGY

While simple differences in grand means for the treatment and control groups could be used to estimate treatment/control differences on any variable, the potential differences across sites in these variables and in the ratio of treatments to controls could lead to distorted estimates. For example, sites with 2:1 ratios of treatments to controls would have a heavier weight in the estimate of the overall treatment group mean than in the estimate of the overall control group mean. Thus, it is necessary to use an estimation procedure that avoids this problem and takes full advantage of the random assignment of experimental status. This attachment describes the methodology used to estimate the average treatment/control differences and the treatment group means for each model.

For each model, the following regression equation was estimated:

$$Y = aT + b_1S_1 + b_2S_2 + b_3S_3 + b_4S_4 + b_5S_5,$$

where

Y = the variable being examined, e.g., the ADL index;

T = 1 if the sample member is in the treatment group and zero if the sample member is in the control group;

S_i = 1 if the sample member is at site i and zero otherwise;

a = the estimated coefficient on T, i.e., the estimate of the average treatment/control difference;

b_i = the estimated coefficient on S_i, i.e., the site-specific intercept.

The treatment/control difference is given by the estimate of the coefficient "a," and its standard error was used to calculate significance levels. The mean value for the treatment group was calculated as a weighted average of the individual site means for the treatment group.

This approach for calculating the treatment/control difference has an intuitively appealing interpretation: the estimated overall difference between treatments and controls (the coefficient a) is a weighted average of the five site-specific treatment/control differences in means, with the weights being inversely proportional to the variance of the estimated treatment/control difference at each site. Thus, the more precise is the estimated difference in any site, the greater the weight this site difference receives in the estimate of the overall (model level) treatment-control difference. Formally, the result can be written as:

$$a = \sum_{i=1}^{5} (Y_{Ti} - Y_{Ci})\, W_i,$$

Y_{Ti} and Y_{Ci} = the mean values of the dependent variable for treatments (T) and controls (C), respectively, at the i^{th} site, and

$$W_i = \frac{P_{Ti}\,(1 - P_{Ti})\, N_i}{\sum_{j=1}^{5} P_{Tj}\,(1 - P_{Tj})\, N_j}$$

where

W_i = the weight applied to each site's treatment/control difference,

P_{Ti} = proportion of site i observations belonging to the treatment group,

N_i = total sample size in site i.

The weight W_i for any site increases (1) as the total sample size in the i^{th} site (N_i) increases, and (2) as the ratio of treatments to controls in any site approaches 1:1--both factors which reduce the variance of an individual site difference. Because P_{Ti} ($1-P_{Ti}$) varies relatively little across sites--from .22 at sites with 2:1 treatment/control ratios to .25 at sites with 1:1 ratios--the sample size is the more important determinant of the site weights. The number of completed screens, contained in Table 1 of the text, result in the following set of weights for comparison of treatments and controls on variables with no missing data:

Basic Sites	Weight	Financial Control Sites	Weight
Baltimore	.2163	Cleveland	.1703
Eastern Kentucky	.1606	Greater Lynn	.2053
Houston	.2139	Miami	.2381
Middlesex County	.2367	Philadelphia	.2562
Southern Maine	.1725	Rensselaer County	.1301

For any particular comparison of treatments and controls, the actual weights will depend on the number and mix of observations with valid data on the variable being examined.

Since the estimated difference is a weighted average of site specific treatment/ control differences, a logical choice for the estimate of the treatment group mean is to use a similar weighted average of the site treatment group means. This has the advantage of treating treatment and control groups symmetrically in computing group means and yielding a set of estimates that are internally consistent. Thus we have,

$$a = \sum_{i=1}^{5} (Y_{Ti} - Y_{Ci})\, W_i$$

$$= \sum_{i=1}^{5} Y_{Ti} W_i - \sum_{i=1}^{5} Y_{Ci}\, W_i$$

$$= E(Y_T) - E(Y_C),$$

where $E(Y_T)$ and $E(Y_C)$ denote the "expected values" for treatments and controls, i.e, the weighted average of the site means for each group. These are the estimates reported for the treatment group means.