This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.'s in policy analysis. The dissertation has been supervised, reviewed, and approved by the graduate fellow's faculty committee.

DISSERTATION

Effects of Trial Design on Participation and Costs in Clinical Trials, with an Examination of Cost Analysis Methods and Data Sources

MEREDITH L. KILGORE

This document was submitted as a dissertation in March 2004 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Jacob Klerman (Chair), Dana Goldman, and Emmett Keeler.

PARDEE RAND GRADUATE SCHOOL

The Pardee RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

Published 2004 by the RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 202, Pittsburgh, PA 15213-1516
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

ABSTRACT

This dissertation comprises a series of studies conducted as part of the Cost of Cancer Treatment Study (CCTS). The specific aims include exploring theoretical issues concerning the problem of representativeness in trial design with an explicit investigation of the causes of the under-representation of older adults in clinical cancer trials; comparing sources of data and modeling approaches for estimating treatment costs in health services research; and estimating the impact of clinical trial participation on prescription drug costs.

An exploration of the sample size requirements for power and significance levels in clinical trials suggests that proportional representation of subpopulations in trials will often not allow valid inferences to be drawn about differential treatment effects. Where differential treatment effects in subpopulations are suspected, targeted trials should be undertaken. Under-representation of older cancer could be accounted for by exclusion criteria based on comorbid conditions that disproportionately afflict the elderly.

Data from patient interviews, medical records abstraction, provider billing records, and Medicare claims were compared as data sources for estimating health care utilization rates and costs; the data were compared in terms of completeness and accessibility. Medicare claims contain data on all covered services, including charges, and reimbursements. The costs of Medicare data compare favorably with other sources of comparable quality, but claims data are missing for individuals in managed care and do not include information on prescription drugs. Provider billing records, however, constituted a poor data source, primarily because providers were unwilling or unable to provide these records. Medical records provide accessible, detailed data on service utilization, but not costs. Self-reported health services utilization generally agreed with other sources on inpatient care but not with respect to outpatient services. Cost estimates for utilization measures were derived from administrative data using hedonic regression models.

Prescription drug costs and out-of-pocket drug expenditures were compared for patients enrolled in cancer trials and for similar cancer patients with who did not participate in trials. Trial participation was associated with higher prescription drug costs, but that did not result in any significant difference in out-of-pocket expenditures for participants. These results were robust to a variety of modeling approaches.

Table of Contents

Chapter I. Introduction

This dissertation comprises a series of studies conducted as part of the Cost of Cancer Treatment Study (CCTS). The CCTS sought to determine how and to what extent participation in clinical trials affects cancer treatment costs. The studies presented here use data gathered during the CCTS to investigate several topics related to the design of clinical trials, data collection, and economic analysis in the context of clinical trials.

Clinical trials represent the gold standard for translating biomedical theory into practical treatment for and prevention of disease. Clinical trials have led to curative treatments for a number of cancers (leukemias, lymphomas), prolonged life expectancy for others (breast, colo-rectal) and new treatments with fewer and less severe side effects (NIH 1990 & 1991; Fisher et al. 1989 & 1997; Perez et al. 1998). Carefully designed trials allow investigators to assess new treatments or treatment combinations. Such studies are required to obtain Food and Drug Administration (FDA) approval for new drugs and medical devices; without this approval products cannot be marketed. Trials are conducted in phases. Phase 1 and 2 trials are typically small and often do not have control arms. The purpose of these trials is to evaluate dosage schedules, measure pharmacokinetics, and provide preliminary information on adverse events. Phase 3 trials are larger, are almost always randomized-controlled trials (RCTs), and are designed to determine the safety and efficacy of the treatment under investigation.

Clinical trials are also expensive undertakings. The costs of trials can be divided into research costs and incremental treatment costs. The research costs include salary support for investigators and support personnel, the costs of data collection and management, and other costs related to the administration of the research project. Incremental treatment costs are associated with more intensive treatment that results from trial participation (i.e. more diagnostic tests, more

frequent physician visits). These incremental treatment costs have customarily been borne by third party payers—government or private sector insurers. The CCTS was conducted to definitively estimate the magnitude of those costs. If trial participation results in substantially increased costs, then decisions need to be made about who should bear those costs.

In this introduction, we first describe the CCTS, and then examine three questions relating to the validity of inference from both randomized controlled trials and from uncontrolled observational studies. At the end of the chapter, we describe the sequence and content of the remaining chapters.

THE COST OF CANCER TREATMENT STUDY

The design of the CCTS has been described elsewhere (Goldman et al, 2000 & 2001) and a report on the principal findings was published in *JAMA* (Goldman et al, 2003), but it is worthwhile to provide a brief description here. The CCTS sets the context for the studies reported below and supplies the core policy relevance motivating them. Preliminary studies found trial participation associated with only modestly higher treatment costs (Bennett et al. 2000; Fireman et al. 2000; Wagner et al. 1999). These studies, however, were small, localized, and most were conducted at academic medical centers. The CCTS sought to produce a generalizable estimate for incremental treatment costs by analyzing costs for a national probability sample of cancer patients, using a retrospective case-control design.

The sampling strategy employed a database containing enrollment for all NCI-sponsored trials at all participating institutions. The sampling frame was restricted to adult, phase 3, cancer treatment trials, and a two stage sampling method was use to select which trials and institutions would be included in the CCTS. The restriction to adult trials was made for practical reasons

related to the difficulty of including children and the differences in how pediatric trials are designed and run. The restriction to phase 3 trials resulted from the fact that institutions do not uniformly report accrual into phase 1 and 2 trials. Thirty-five trials were chosen with probability of selection proportional to enrollment. Fifty-five institutions were chosen from a second stage sampling frame made up of all institutions participating in the trials sampled in the first stage. Institutions were also selected with probabilities proportional to their enrollment. This sampling design allowed us to draw a national probability sample of cancer trial participants while limiting the number of trials and institutions to a reasonable number.

For the purposes of the CCTS, trial participants are referred to as "cases" regardless of the trial arm in which they were enrolled. CCTS "controls" are non-participants who met the protocol enrollment criteria for the sampled trials and were being treated at sampled institutions—thus controls matched cases on such variables as cancer type and stage, absence of comorbid conditions, and cancer care provider. Controls were identified using administrative datasets, tumor registries, or lists of patients who had previously been approached to participate in trials but never enrolled.

Personnel at the sites that agreed to participate identified cases and controls and asked if CCTS personnel could contact them about the study. Table 1.1 shows the distribution of institutions that agreed to participate and the number of cases accrued into sampled trials at both participating and non-participating institutions. Note that the number of institutions is greater than 55. This is because sampled institutions often included a network of affiliate providers where the actual care was delivered. Each of those affiliates was approached directly about participating in the CCTS. In all, 83 out of 149 providers, with 66% of the total accrued cases

agreed to participate. Participating providers were also asked to approach any patients they had participating in phase 1 or 2 trials along with appropriate control candidates.

Table 1.1 Site Enrollment Status

	Number	Phase 3 Accrual	% of Accrual
Participating Sites	83	1756	66%
Refusing Sites	65	921	34%
Total	148	2677	100%

Table 1.2 shows the numbers and percent of potential participants identified who agreed to be approached for the CCTS. These numbers include deceased patients for whom medical records were provided. Of participants in phase 3 trials, 849 (57%) agreed to be contacted, along with 712 (50%) of the potential controls. Rates of agreement were higher for phase 1 and 2 trial cases and controls. Of patients that agreed to be contacted. Individuals that gave consent were interviewed about their health care providers and their utilization of health care services. They were also asked for permission to obtain medical and billing records from all their inpatient and outpatient health service providers. Those who had Medicare coverage were asked permission to access their Medicare billing records as well.

Table 1.2 Patient Enrollment						
	Phase 3		Phase 2 (At 20 Sites)		Phase 1 (At 5 Sites)	
	Cases	Controls	Cases	Controls	Cases	Controls
Total Identified	1482	1415	220	58	28	16
Refused	566	571	72	17	6	1
Agreed	849 (57%)	712 (50%)	148 (67%)	41 (71%)	22 (79%)	15 (94%)

The main outcome measure was the incremental direct treatment cost of care; research design, administration and analysis costs were excluded. Participation in clinical trials was found to be associated with a 6.5% increased in treatment costs over a 2.5 year period, but the costs difference was not statistically significant ($35,418 for trial participants versus $33,248 for non-participants, $P = 0.22$). The CCTS had been powered to detect a cost difference of 10% or more. Treatment cost differences were higher for subjects who died ($39,420 vs $33,432, respectively, $P = 0.20$). The cost differences found were consistent with other smaller studies and the magnitude of the difference suggests that financing routine care for trial participants does not impose an undue burden on third party payers. All of the work presented in this dissertation was conducted in the context of the CCTS. Our examination of a wide variety of clinical trials and review of the literature on trial design and conduct allows us to comment on some relevant issues in the remainder of this introduction.

Drawing Inferences about Particular Populations from Randomized Controlled Trials

Patients and their providers often would like to know whether the results of a trial apply to them. There are two concerns here, even within randomized controlled trials sampling bias may leave their type of patient under-represented, and differences in treatment effectiveness between subpopulations in the trial would mean that the benefits and side effects of a treatment differ for different types of patients.

Sampling Bias and Under-representation

Critics have lamented the lack of external validity in clinical trials. They contend that trials are conducted at elite institutions on selectively chosen participants and follow protocols of care more rigorous than are found in more typical care settings. This section addresses with the selection issue—the concern that trial participants should be representative of the general population for which a treatment or program is being evaluated. There is an extensive literature documenting the under-representation of subgroups in clinical trials, particularly women, minorities, and the elderly. Lately attention has been given to the inclusion of children as well. The Congress has passed legislation mandating the inclusion of women and minorities in trials (Public Law 103-43, §492B). There are essentially two rationales given for these concerns (Lumley and Bastian, 1996): subgroups not included in clinical trials are effectively denied access to some treatments; and, failure to make trials representative of the general population compromises the generalizability of results.

The first argument is certainly true—there are experimental treatments available only in the context of clinical trials. It is, however, hard to demonstrate that a lack of access to such treatments results in harm. If the trial is properly designed, and there is equipoise as to the effectiveness of the experimental treatment, then persons lacking access to trial participation are

precluded from receiving treatment of questionable efficacy. It might be argued that trials are conducted only for therapies expected to produce better outcomes than currently standard treatments, but in fact only about one in five drugs that enter clinical trial testing receives FDA approval (Tufts 2001). The second concern, related to external validity, is the primary issue here.

Why are people excluded from trials? Two disparate rationales are in play: beneficence and efficiency. In the first case, persons should be excluded from trials if they cannot be expected to receive no benefit or may be harmed by the treatment. Cancer treatments in particular often involve significant bodily insult from surgery, radiation, or toxic agents. Patients with pre-existing organ system failure or impaired functional status may not be able to tolerate such treatments (NCI 2003). Their exclusion from trials is appropriate if they would not be candidates for therapy in typical practice. Patients with impaired mental function, as from Alzheimer's disease or psychosis, may be unable to provide informed consent and thus be ineligible for randomization.

Efficiency is quite a different rationale for exclusions, and relates primarily to the interests of investigators and organizations funding research. Unrepresentative enrollment can arise from convenience sampling (e.g. trials conducted in single institutions or locales). Exclusion criteria can be incorporated into protocols for the explicit purpose of increasing the power of the trial to detect treatment effects for a given number of participants (Finn 1999). For industry sponsored trials the objective is to get drugs to market, establishing safety and efficacy is a means to that end. If individuals with poor prognoses and co-morbid conditions are excluded, fewer will be lost to follow-up from deaths due to unrelated causes. Furthermore, the more homogeneous the trial sample is, the less likely are unobserved confounding factors to influence the results.

Whether the trial is sponsored by industry, the government, or a non-profit entity, investigators need to be cognizant of scare resources and will want to maximize the value and minimize the acquisition cost of information produced. Suppose that costs of clinical trials correlate with the number of individuals enrolled. When individuals in the trial die from extraneous causes (unrelated to the condition or treatment under investigation), information is lost. Proper study design then will need to calculate the actual sample size, *s*, taking the baseline non-disease-specific mortality rate, *m*, into account: $s = (1 + m) * n$, where *n* is the hypothetical sample size for a given power and confidence level. Minimizing the extraneous mortality rate is obviously desirable, so investigators would be inclined to exclude subjects with comorbidities that carry risks of death or complications that might cause them to drop out of the study.

A cursory search of the National Library of Medicine's MEDLINE database shows that concerns about representativeness are quite current. Hutchins et al. (1999) reported that the elderly are enrolled in cancer clinical trials in numbers far below what would be expected based on cancer incidence rates. Fossa and Skovlund (2002) found differences in survival between cancer trial participants and eligible non-participants receiving similar therapies. They concluded the "Results and treatments recommendations from a trial can be transferred to daily practice only if eligibility criteria and selection of patients are taken into account."

Bandyopadhyay, Bayer, and O'Mahony (2001) found age and gender bias in patient recruitment for statin (treatment for hypercholesterolemia) trials and concluded that this bias cast doubt on extrapolating results to under-represented groups. Similarly, studies of cardiac trials have found lack of representation for women, minorities, and the elderly (Lee et al. 2001; Heiat, Gross, and Krumholz, 2002). Moore et al. (2000) found disparities in routes of HIV transmission for patients in antiretroviral therapy trials compared with the distributions of HIV/AIDS patients

in the general population. Each of these studies concluded under-representation posed a problem for generalizing the trial results.

Alongside studies of under-representation has arisen a literature concerned with barriers to trial enrollment. Putative barriers to entry include attitudes of patients (Madsen et al., 2002; Schain 1994) and providers (Mansour 1994); toxicity, protocol requirements, and health status in elderly patients (Kornblith et al., 2002); socioeconomic factors (Saterne et al., 2002); distrust of research on the part of African Americans (Shavers 2001); reimbursement problems (Fleming 1994); and the presentation of information to obtain informed consent (Cox 2002).

In each of the studies cited, the problem of external validity was asserted as a given or probable problem. This may not be quite so obvious. Other researchers have questioned the desirability of constructing trials to permit subgroup analysis (SCT 1993); some have gone so far as to dismiss such concerns as mere political correctness (Piantodosi and Wittes 1993). There is clearly a spectrum of views. In the medical literature on inference we can anchor the ends of the skeptical spectrum on one end with Sheldon et al. (1998) who conclude:

> "it is probably more appropriate to assume that research findings are generalizable across patients unless there is strong theoretical or empirical evidence to suggest that a particular group of patients will respond differently."

At the other extreme, Julian and Pocock (1997) list that criteria trials must meet to be deemed externally valid, the primary criterion be representativeness of the clinically relevant population in the trial. The general terms of these arguments can be formalized in such a way as to render the issues relating to representativeness subject to hypothesis testing.

Chapter 2 addresses causes for the observed under-representation of elderly subjects in cancer clinical trials. The specific issue is how much this fact can be explained by the presence of exclusion criteria based on comorbid disease states, and life expectancy and functional status

requirements. A separate issue is whether low rates of elderly participation has implications for making treatment decisions for older cancer patients based on clinical trial results.

Subpopulation Differences in Treatment Effectiveness

Even if there is a lack of representation in trials, if we want evaluate whether and how much this is a problem we need to consider two things. First, lack of representation is an issue only if the treatment effects are different among different subpopulations. As discussed below, evidence for such variation is weak. Second, even if there is heterogeneity among subpopulations the cost of accurately ascertaining the magnitude of the differences in the context of a clinical trial may be prohibitive. We have to decide what we need to know. Is the treatment effective on average? Is it effective for large subpopulations (e.g. women)? Do we need to estimate the treatment effectiveness separately for specific subgroups (e.g. minorities, children)? The desirability of having answers to these questions then needs to be weighed against the cost of obtaining them.

Although the great majority of studies examining external validity present only descriptions of how certain groups are under-represented, there have been some that actually explored the hypothesis that heterogeneity in patient characteristics produces differences in treatment effects. Zimmermann, Mattia, and Posternak (2002) found that an anti-depressant efficacy study had exclusion criteria that would have screened out 88% of persons suffering from clinical depression. The exclusion criteria included a prior history of substance abuse and current suicidal ideation. They suggest reasons why patients typically included in trials might respond very differently from the majority of patients presenting with clinical depression, though they lack the data to make the needed parameter estimates.

Rocha Lima et al. (2002) conducted a subgroup analysis by age of two chemotherapy trials for lung cancer treatment. Patients were grouped into four age cohorts: <50, 50-59, 60-69, and 70-79. There was no difference in toleration of treatment, response, or survival among the different age groups. It should be noted that one of the trials had exclusion criteria for patients with impaired functional status, and hematological, hepatic, renal, or pulmonary co-morbid conditions (CLB-9130 protocol abstract, NCI cancer trial search website). A study of acute myeloid leukemia (AML) compared outcomes for elderly and younger AML patients and explored the reasons for those differences (Leith et al. 1997). The authors found that differences in disease characteristics (unfavorable cytogenetics, MDR1 protein expression, and functional drug efflux) between older and younger patients accounted for differences in outcomes. When disease characteristics were controlled for, elderly AML patients were as likely as younger patients to experience remission and enjoyed similar periods of disease free survival. Muss (2001) compared outcomes for breast cancer treatment by age, race, and socioeconomic status. One key finding was that, matched for disease stage, histological and cytological characteristics had equivalent outcomes given comparable treatments.

In these studies differential treatment effects by age group and ethnicity arose because the subgroups were proxies for disease characteristics. Older AML patients had more resistant leukemias, African American breast cancer patients presented with later stage and/or more aggressive carcinomas. In a study of treatment for heart failure (Carson, Ziesche, Johnson, and Cohn, 1999) whites responded to treatment with enalapril, but blacks did not. On the other hand, blacks responded to treatment with hydralazine plus isosorbide dinitrate, and whites received no benefit. So it is possible for treatment effects to differ based on patient characteristics where no difference in disease could be established.

We now turn to the question of how an attempt to capture differences in treatment effects could affect the design of clinical trials and the expense of conducting them. Let us pose the simplest possible example for evaluating the effect of a hypothetical treatment. Assume that a RCT is conducted to evaluate some treatment, *t*, in terms of *X*, a beneficial outcome either in relation to placebo or to some alternative treatment. Assume further that a t-test of the difference in the means between the treatment and control groups is appropriate. Following the method of power calculation in Lipsey (1990, p. 34), the magnitude of the standardized effect size is $ES = (\bar{X}_t - \bar{X}_c) / S$, the difference in the means divided by the standard deviation. A properly designed study will have a sample size large enough to detect an anticipated treatment effect size with an appropriate degree of power. The test statistic for the significance of a difference will then be $t = \frac{ES}{\sqrt{1/n_t + 1/n_c}}$. Here n_t and n_c represent the sample sizes for the treatment and control groups, respectively. For convenience we can assume that these numbers are equal, so the equation becomes $t = \frac{ES}{\sqrt{2/n}}$.

Now let us consider the possibility that there exists a subpopulation that benefits from the treatment, but only by half as much, so $ES_s = \frac{1}{2} ES$ expresses the effect size for the subpopulation in relation to the general population. How large would the trial need to be to detect the treatment effect for this subpopulation with the same power? The subpopulation sample size would need to be four times that of the general population, implying that a trial generalizable for the subpopulation would need to be nearly five times as large (depending on the proportion of the general population contained in the subgroup).

$$\frac{ES}{\sqrt{2/n}} = \frac{1/2\,ES}{\sqrt{2/n_s}} \rightarrow \sqrt{2/n_s} = 1/2\sqrt{2/n}$$

$$2/n_s = 1/2n$$

$$n_s = 4n$$

The principle lesson to be drawn from this exercise is this: if there is reason to be skeptical about the applicability of trial results for some sub-population, then simple representativeness is likely to be insufficient to allay that skepticism. In most trials, even representation of numerous subgroups in numbers proportionate to their presence in the general population would not provide sufficient power to make a valid test of interaction effects as even important differences could lack statistical significance. This would have significant implications for the costs of designing and conducting presumptively valid clinical trials. Test the hypothesis that treatment effects differed among groups (rather than differing from zero) would require even larger sample sizes.

The formulation set out remains over-simplified. The level of abstraction is useful for framing the problem, but some crucial information is elided. Treatments being evaluated in clinical trials are not, one hopes, arbitrary interventions. Drug compounds are evaluated because theory independent of clinical trials suggests that they should produce beneficial effects. Such compounds go through considerable preliminary testing before the involvement of human subjects. So a simple frequentist statistical approach is inadequate; there is prior information that suggests a Bayesian framework would be more appropriate. That level of modeling is beyond the scope of this project, but suggests a potential fruitful direction for future research.

We have seen examples of studies that found differences in subpopulations, but where those differences could be explained as differences in underlying disease states. Other studies

have found differences between groups that could not be clinically explained, and still others have found a complete absence of differential effects. This diversity of findings suggests that it is important to conduct research involving subpopulations. This does not suggest that clinical trials need to be designed to mirror the diversity of the general population. To derive significant findings on subpopulations, studies must be focused on those groups specifically. When randomized designs are not practical for identifying differential outcomes, it may be necessary to turn to observational studies.

Making Non-Randomized Observational Studies More Rigorous

Although the randomized controlled design is the gold standard for clinical research, there are often very good reasons for not conducting an RCT. Randomization may be impractical and/or unethical. The CCTS is itself an example of this problem. A theoretically stronger study would randomize people to participate in clinical trials or not. However, such a study design would be unethical—one cannot force some people into trials without their consent and deny participation to others who might wish to participate.

Horton (2000) provides an excellent example of practical constraints on conducting RCTs to answer important questions concerning the effectiveness of treatments for coronary artery disease. While trials of coronary stents yielded favorable results, the eligibility criteria limited participation to subjects with very specific coronary lesions and used only one type of stent. Subsequently more than 30 thirty types of stents have come into use, and are used a wide variety of lesions not represented in the trials. It would require thousands of RCTs to evaluate every type of stent in every type of lesion to which they have been applied. Subsequent research

has used registry data to estimate the effectiveness of stents in lesions and vessels not studied in RCTs (Saha et al. 2001).

There are a variety of observational study designs and data sources. In some cases simple observation of program outcomes sheds light on an issue, particularly when prior research provides a basis for hypothesis testing. Several studies of automated external defibrillators (AEDs) have taken this form (Groh et al. 2001; Cobb et al. 1999; MacDonald, Mottley and Weinstein, 2002; Calle et al. 1997). These studies tested the hypothesis that the use of AEDs would lead to better outcomes for heart attack victims. The studies have found strong evidence to support this hypothesis, leading to the deployment of AEDs in high-traffic public areas such as airports and shopping malls. Patient registries have provided data for a variety of studies. These include studies of coronary stenting outcomes (Kimura et al. 1996, Laham et al 1996, and Moussa et al. 1997, quoted in Horton 2000). The Surveillance, Epidemiology, and End Results (SEER, 2003) tumor registry and the SEER-Medicare linked database have provided similar resources for studying outcomes in cancer patients (Warren et al. 2002).

Simple observational studies suffer from the disadvantage that there is no true comparison group; observed outcomes are compared with expected outcomes. A variety of strategies have arisen to attempt to minimize potential biases arising from differences between groups who receive treatments and those who do not. One example is the natural experiment. Lu-Yao et al. (2002) used the geographic variation in the deployment of prostate cancer screening to estimate the effect of screening on treatment decisions and outcomes for prostate cancer (finding greater rates of diagnosis and treatment did not affect disease specific mortality).

A more common approach is the case-control study. Here a group of individuals treated (or exposed) in some way are compared to another group without such treatment or exposure,

with or without adjustment for observed covariates (Schlesselman 1982). While there are problems with this sort of design, considerable knowledge has been gained from such studies. Most of the studies linking smoking and lung cancer have been and continue to be case-control studies (Yan et al. 2002). This design continues to be widely used in epidemiological studies (Caballero-Granado et al. 2001).

The problem with case control studies is the lack of randomization. In the example of the CCTS, patients chose beforehand whether or not to participate in clinical trials, so presumably trial participants differed in important ways from non-participants, and those differences could have effects on the costs of the care they received independent of trial participation. We used two methods to minimize the selection bias arising from these differences. First, controls for the CCTS received care from the same providers for the same conditions as trial participants. So differences in health status and provider practice patterns were minimized between the two groups. Second, the weight given to each observation was adjusted by a propensity score, described below.

Propensity Scores in Cohort Studies

Propensity scores were first described by Rosenbaum and Rubin (1984). Consider two non-random cohorts of individuals, one treated in some way and the other not, and a set of observed variables, **x**, presumed to have some affect on an outcome of interest. In the absence of randomization there is no presupposition that the expected value of **x** given treatment should equal the expected value given no treatment. It is usual to present tables of such variables indicating the ways that the treatment group differs from the control group. Propensity scores are obtained by regressing treatment status on all observable covariates to obtain the conditional

probability that an individual would be expected to receive the treatment. The form of the regression is usually a logit or probit model.

Propensity scores have been deployed in a variety of ways to reduce selection bias: matching cases to controls, stratification, and regression adjustment (D'Agostino 1998). We consider each of these applications in turn.

When treatment and control groups are known to differ, a stratified analysis can be used to compensate. When there are differences along several dimensions, however, strata proliferate exponentially. Here propensity scores can provide a univariate means for stratifying units of observation (Coyte, Young, and Croxford, 1998). One application of propensity scores is to improve matching in case-control studies (D'Agostino 1998). It is often the case with registry or administrative data to have a relatively small number of treatment cases and a very large number of controls who were not treated or exposed. In this case, propensity scores can be used to match controls to cases in such a way as to insure similarity between the two groups along a wide range of observable characteristics.

A second use of propensity scores is in sub-classification of subjects in case control cohort studies. Analysis is conducted among cases and controls within propensity score quantiles (Rosenbaum and Rubin, 1983 & 1984, Rose et al. 2000). Successful stratification is often evaluated by the degree to which differences between treatment groups is reduced after propensity score adjustment (D'Agostino 1998). Finally, propensity scores can be incorporated into a regression model either directly (D'Agostino 1998) or through a weighting scheme (Hirano, Imbens, and Ridder, 2000).

D'Agostino (1998) provides examples of propensity score matching in a March of Dimes study of the effects of post-term delivery on perinatal outcomes and of stratification in the

context of the Active Management of Labor Trial (ACT, Frigoletto et al. 1995). Numerous examples of propensity score use can be found in the recent literature. Mehta et al. (2002) used propensity score adjustment to analyze the effects of diuretic therapy on outcomes in acute renal failure. Other examples include studies of coronary artery bypass surgery (Stamou et al. 2002; Magee et al. 2002), methods of repairing aortic aneurysms (Teufelsbauer et al. 2002), arthritis treatments (Rhame, Pettitt and LeLorier, 2002), and cancer screening (Iwashyna and Lamont, 2002).

Finally, it has been suggested that inverse propensity weights can provide a useful means of reducing selection bias (Hirano, Imbens, and Ridder 2000). The CCTS (Goldman et al, 2003) used propensity score weights to adjust for differences in a variety of factors between cases and controls. The precise method for calculating these weights is discussed in Chapter 5. Table 1.3 gives weighted and unweighted mean values for several factors that differed for the two groups. When propensity score weights were used the differences were narrowed or eliminated.

Table 1.3 Weighted & Unweighted Means

	Unweighted		Weighted	
	Cases	Controls	Cases	Controls
Age	57.9	60.5	58.9	58.8
Male	24%	23%	23%	23%
Wealth	$330,633	$404,997	$352,648	$375,338
Medicare	32%	38%	34%	35%
Private Ins	67%	64%	66%	67%
Diabetes	13%	9%	11%	11%
Arthritis	37%	40%	38%	38%
Oth_Cancer	9%	14%	10%	13%
HTN	33%	34%	33%	33%

Propensity scores are frequently compared, often unfavorably, with instrumental variables (IV). A paper by Posner et al. (2002) compares OLS, IV, and propensity scores for estimating the effect of mammography screening on breast cancer stage at diagnosis. Unfortunately, this turned out to be a poor example as the three methods produced very similar

estimates, indicating that selection and endogeneity were not significant problems. Propensity scores cannot remove omitted variable biases except to the extent to which unobserved factors are correlated with measured covariates. This makes propensity scores look like estimation with weak instruments (Staiger and Stock 1997).

An alternative perspective to seeing propensity score adjustment as a poor substitute for IV is to view it as a way to improve the efficiency and reduce the bias in case-control studies. As noted above, it is possible to gain knowledge and even test hypotheses using observational studies, even studies of very crude design. Propensity scores provide a means of reducing observable biases. The last chapter provides an example of the use of propensity scores as does the CCTS.

The CCTS frames the overall context in which the following studies were conducted. Two are tangential, the examination in Chapter 2 of participation rates for the elderly in clinical trials, and the comparison of data sources for cost estimation in Chapter 3. The study in Chapter 4 on developing prices for health care utilization measures provided a direct input for the main results. Finally, Chapter 5 examines the effect of trial participation on the use and costs of prescription drugs, a subject that has not been previously addressed and one that should be of particular interest to individual trial participants as much as to third party payers.

References

Bandopadhyay S, Bayer AJ, O'Mahony MS. "Age and Gender Bias in Statin Trials." *QJ Med* 2001;94:127-32.

Bennet CL, Stinson TJ, Vogel V, Robertson L, Leedy D, O'Brien P, Hobbs J, Sutton T, Ruckdeschel JC, Chirikos TN, Weiner RS, Ramsey MM, Wicha MS 'Evaluating the Financial Impact of Clinical Trials in Oncology: Results from a Pilot Study from the Association of American Cancer Institutes/Northwestern University Clinical Trials Costs and Charges Project.' J Clin Oncol 2000;18:2805-10.

Berlin I. *The Magus of the North* Fontana: London, 1993.

Caballero-Granado FJ, Becerril B, Cisneros JM, Cuberos L, Moreno I, Pachon J. "Case-Control Stucy of Risk Factors for the Development of Enterococcal Bacteremia." *Eur J Clin Microbiol Infect Dis* 2001;20:82-90.

Calle PA, Verbeke A, Vanhouate O, Van Acker P, Martens, P, Buylaert W. "The Effect of Semi-Automatic External Defibrillation by Emergency Medical Technicians on Survival after Out-of-Hospital Cardiac Arrest: an Observational Study in Urban and Rural Areas in Belgium." *Acta Clin Belg* 1997;52:72-83.

Carson P, Ziesche S, Johnson G, Cohn JN. "Racial Differences in Response to Therapy for Heart Failure: Analysis of the Vasodilator-Heart Failure Trials." *J Card Fail* 1999;3:178-87.

Cawley J. "An Instrumental Variables Approach to Measuring the Effect of Body Weight on Employment Disability." *Health Serv Res* 2000;35:1159-79.

Cobb LA, Fahrenbruch CE, Walsh TR, Copass MK, Olsufka M, Breskin M, Hallstrom AP. "Influence of Cardiopulmonary Resuscitation Prior to Defibrillation in Patients with Out-of-Hospital Ventricular Fibrillation." *JAMA* 1999;281:1220-2.

Cox K. "Informed Consent and Decision-Making: Patients' Experiences of the Process of Recruitment to Phases I and II Anti-Cancer Drug Trials." *Patient Education and Counseling* 2002;46:31-8.

Coyte PC, Young W, Croxford R. "Cost and Outcomes Associated with Alternative Discharge Strategies Following Joint Replacement Surgery: Analysis of an Observational Study Using a Propensity Score."

D'Agostino RB. "Tutorial in Biostatistics: Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group." *Stat Med* 1998;17:2265-81.

Ettner SL, Hermann RC. "The Role of Profit Status under Imperfect Information: Evidence from the Treatment Patterns of Elderly Medicare Beneficiaries Hospitalized for Psychiatric Diagnoses." *J Health Econ* 2001;20:23-49.

Farley DE, Ozminkowski RJ. "Volume-Outcome Relationships and In-Hospital Mortality: the Effect of Changes in Volume over Time." *Med Care* 1992;30:77-94.

Finn R. *Cancer Clinical Trials: Experimental Treatments & How They Can Help You.* O'Reilly & Assoc, Sabastopol, CA, 1999.

Fireman BH, Fehrenbacher L, Gruskin EP, Ray GT. "Cost of care for patients in cancer clinical trials," *J Natl Cancer Inst* 2000:92:136-42.

Fisher B, Redmond C, Poisson R, Margolese R, Wolmark N, Wickerman L, et al. "Eight-year Results of a Randomized Clinical Trial Comparing total Mastectomy and Lumpectomy With or Without Irradiation in the Treatment of Breast Cancer. *N Engl J Med* 1989;320:822-8.

Fisher B, Brown A, Mamounas E, Wieand S, Robidoux A, Margolese RG, et al. "Effect of Preoperative Chemotherapy on Local-Regional Disease in Women with Operable Breast Cancer: Findings from National Surgical Adjuvant Breast and Bowel Project B-18." *J Clin Oncol* 1997;15:2483-93.

Fleming ID. "Barriers to Clinical Trials Part I: Reimbursement Issues." *Cancer* 1994;74:2662-5.

Fossa SD, Skovlund E. "Selection of Patients may Limit Generalizability of Results from Cancer Trials." *Acta Oncol* 2002;41:131-7.

Frances CD, Shlipak MG, Noguchi H, Heidenreich PA, McClellan M. "Does Physician Specialty Affect the Survival of elderly Patients with Myocardial Infarction?" *Health Serv Res* 2000;35:1093-116.

Frigoletto FD, Lieberman E, Lang JM, Cohen A, Barss V, Ringer S, Datta S. "A Clinical Trial of Active Management of Labor." *NEJM* 1995;333:745-50.

Goldman DP, Schoenbaum ML, Potosky AL, Weeks JC, Berry SH, Escarce JJ, Weidmer B, Kilgore ML, Wagle N, Adams JL, Figlin RA, Lewis JH, Rosen MR, Cohen J, Kaplan R, McCabe M. "Measuring the Incremental Cost of Clinical Cancer Research," *Journal of Clinical Oncology*, 2001;19:105-110.

Goldman DP, Adams JL, Berry SH, Escarce JJ, Kilgore ML, Lewis JH, Rosen MR, Schoenbaum ML, Schonlau M, Wagle N, Weidmer B. "The Cost of Cancer Treatment Study's Design and Methods," RAND Health Science and Technology Policy Institute, 2000.

Goldman DP, Berry SH, McCabe M, Kilgore ML, Potosky AL, Schoenbaum ML, Schonlau M, Weeks JC, Kaplan R, Escarce JJ. "Incremental Treatment Costs in NCI-Sponsored Trials." *JAMA*, 2003;289:2970-2977..

Groh WJ, Newman MN, Beal PE, Fineberg NS, Zipes DP. "Limited Response to Cardiac Arrest by Police Equipped with Automated External Defibrillators: Lack of Survival Benefit in Suburban and Rural Indiana—the Police as Responder Automated Defibrillation Evaluation (PARADE)." *Acad Emerg Med* 2001;8:324-30.

Heiat A, Gross CP, Krumholz HM. "Representation of the Elderly, Women, and Minorities in Heart Failure Clinical Trials." *Arch Intern Med* 2002;162:1682-88.

Hauck WW, Anderson S, Marcus SM. "Should We Adjust for Covariates in Nonlinear Regression Analyses of Randomized Trials?" *Controlled Clin Trials* 1998;19:249-56.

Hellmann D. "Evidence, Belief, and Action: the Failure of Equipoise to Resolve the Ethical Tension in the Randomized Clinical Trial." *J Law Med Ethics*. 2002;30(3):375-80.

Hirano K, Imbens GW, Ridder G. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." NBER Technical Working Paper 251, March 2000. http://www.nber.org/papers/T0251.

Ho V, Hamilton BH, Roos LL. "Multiple Approaches to Assessing the Effects of Delays for Hip Fracture Patients in the United States and Canada." *Health Serv Res* 2000;34:1499-518.

Holmberg L, Baum M, Adami AO. "On the Scientific Inference from Clinical Trials." *J Eval Clin Practice* 1999;5:157-62.

Horton R. "common Sense and Figures: the Rhetoric of Validity in Medicine." *Statist Med* 2000;19:3149-64.

Hume D. *An Enquiry Concerning Human Understanding*. Originally published 1748, Oxford University Press 1999.

Hutchins LF, Unger JM, Crowley JJ, Coltman CA Jr, Albain KS. Under-representation of Patients 65 Years of Age or Older in Cancer-Treatment Trials. *N Engl J Med.* 1999;341(27):2061-2067.

Iwashyna TJ, Lamont EB. "Effectiveness of Adjuvant Fluorouracil in Clinical Practice: A Population-Based Cohort Study of Elderly Patients with Stage III Colon Cancer." *J Clin Oncol* 2002;20:3992-98.

Julian DG, Pocock SJ. "Interpreting a Trial Report." In *Clinical Trials in Cardiology*, Pitt B, Julian D, Pocock S (Eds.). WB Saunders: London, 1997; 33-42.

Kornblith AB, Kemeny M, Peterson BL, Wheeler J, Crawford J, et al. "Survey of Oncologists' Perceptions of Barriers to Accrual of Older Patients with Breast Carcinoma to Clinical Trials." *Cancer* 2002;95:989-96.

Lee PY, Alezander KP Hammill, Pasquali SK, Peterson ED. "Representation of elderly Persons and Women in Published Randomized Trials of Acute Coronary Syndromes." *JAMA* 2001;286:708-13.

Leith CP, Kopecky KJ, Godwin J, McConnell T, Slovak ML, Chen IM, Head DR, Appelbaum FR, Willman CL. "Acute Myeloid Leukemia in the Elderly: Assessment of Multi-Drug Resistance (MDRI) and Cytogenetics Distinguishes Biologic Subgroups with Remarkably Distinct Responses to Standard Chemotherapy." *Blood* 1997;9:3323-29.

Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, Housman MG, Escarce JJ. "The Participation of Patients 65 Years of Age and Older in Cancer Clinical Trials," *J Clin Oncol*, Vol 21, Issue 7 (April), 2003: 1383-1389.

Lipsey MW. *Design Sensitivity: Statistical Power for Experimental Research*. Newbury Park: Sage Publications, 1990.

Lumley J, Bastian H. "Competing or Complementary? Ethical Considerations and the Quality of Randomized Trials." *Int J Technol Assess Health Care* 1996;12:247-63.

Lu-Yao G, Albertsen PC, Stanford JL, Stukel TA, Walker-Corkery ES, Barry MJ. "Natural Experiment Examining Impact of Aggressive Screening and Treatment on Prostate Cancer Mortality in Two Fixed Cohorts from Seattle and Connecticut." *BMJ* 2002;325:740-45.

MacDonald RD, Mottley JL, Weinstein C. "Impact of Prompt Defibrillation on Cardiac Arrest at a Major International Airport." *Prehosp Emerg Care* 2002;6:1-5.

Madsen SM, Mirza MR, Holm S, Hilsted KL, Kampmann K, Riis P. "Attitudes towards Clinical Research Amongst Participants and Non-Participants." *J Intern Med* 2002;251:156-68.

Magee MJ, Jablonski KA, Stamou SC, Pfister AJ, Dewey TM, Dullum MK, Edgerton JR, Prince SL, Acuff TE, Corso PJ, Mack MJ.

Mansour EG. "Barriers to Clinical Trials Part III: Knowledge and Attitudes of Health Care Providers." *Cancer* 1994;74:2672-75.

McClellan M, McNiel BJ, Newhouse JP. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *JAMA* 1994;272:859-66.

McMahon AD. "Study Control, Violators, Inclusion Criteria and Defining Explanatory and Pragmatic Trials." *Statist Med* 2002;21:1365-76.

Mehta RL, Pascual MT, Soroko S, Chertow GM. "Diuretics, Mortality, and Non-Recovery of Renal Function in Acute Renal Failure." *JAMA* 2002;288:2547-53.

Moore DAJ, Goodall RL, Hooker M, Gazzard BG, Easterbrook PJ. "How Generalizable are the Results of Large Randomized Controlled Trials of Antiretroviral Therapy?" *HIV Med* 2000;1:149-554.

Muss HB. "Factors used to Select Adjuvant Therapy of Breast Cancer in the United States: and Overview of Age, Race, and Socioeconomic Status." *Monogr Natl Cancer Inst* 2001;30:52-5.

National Cancer Institute (NCI), *Investigator's Handbook: a Manual for Participation in Clinical Trials of Investigational Agents*. http://ctep.cancer.gov/handbook/index.html, accessed Jan 2003.

National Institutes of Health Consensus Conference. "Adjuvant Therapy for Patients with Colon and Rectal Cancer," *JAMA* 1990;264:1444-50.

NIH Consensus Conference. "Treatment of Early-Stage Breast Cancer" *JAMA* 1991;265:391-5.

Perez EA, Hesketh P, Sandbach J, Reeves J, Chawla S, Markman M, et al. "Comparison of Single-Dose Oral Granisetron versus Intravenous Ondansetron in the Prevention of Nausea and Vomiting Induced by Moderately Emetogenic Chemotherapy: A Multi-Center, Double-Blind, Randomized Parallel Study." *J Clin Oncol* 1998;16:754-60.

Piantadosi S, Wittes J. "Politically Correct Trials." *Controlled Clin Trials* 1993;14:562-7.

Pierce CS. "Abduction and Induction." In *Philosophical Writing of Peirce*, Buchler J (ed.) Dover: NY, 1955.

Rahme E, Pettitt D, LeLorier J. "Determinants and Sequelae Associated with Utilization of Acetaminophen versus Traditional Non-Steroidal Anti-Inflammatory Drugs in an Elderly Population." *Arthritis Rheum* 2002;46:3046-54.

Rocha Lima CM, Herndon JE, Kosty M, Clamon G, Green MR. "Therapy Choices among Older Patients with Lung Carcinoma: an Evaluation of Two Trials of the Cancer and Leukemia Group B." *Cancer* 2002;94:181-7.

Rose JH, O'Toole EE, Dawson NV, Thomas C, Connors AF, Wenger N, Phillips RS, Hamel MB, Reding DT, Cohen HJ, Lynn J. "Generalists and Oncologists Show Similar Care Practices and Outcomes for Hospitalized Late-Stage Cancer Patients." *Med Care* 2000;38:1103-18.

Rosenbaum PR, Rubin DB. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 1983;70:41-55.

Rosenbaum PR, Rubin DB. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *JASA* 1984;79:516-24.

Saha S., Gibson M, Torrie EPH, Magee TR, Galland RB. "Stenting for Localized Arterial Stenoses in the Aorto-Illiac Segment." *Eur J Endovasc Surg* 2001;22:37-40.

Sateren WB, Trimble EL, Abrams J, Brawley O, Breen N, Ford L, McCabe M, Kaplan R, Smith M, Ungerleider R, Christian MC. "How Sociodemographics, Presence of Oncology Specialists, and Hospital Cancer Programs Affect Accrual to Cancer Treatment Trials." *J Clin Oncol* 2002;20:2109-17.

Schain WS. "Barriers to Clinical Trials Part II: Knowledge and Attitudes of Potential Participants." *Cancer* 1994;74:2666-71.

Schlesselman JJ. *Case-control Studies. Design, Conduct, Analysis.* Oxford University Press, NY, 1982.

Shavers VL, Lynch CF, Burmeister LF. "Racial Differences in Factors that Influence the Willingness to Participate in Medical Research Studies." *Ann Epidemiol* 2002;12:248-56.

Society for Clinical Trials [SCT]. "Recommendations to the Director of NIH from the Membership of the society for Clinical Trials." *Controlled Clin Trials* 1993;14:559.

Staiger D, Stock JH. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 1997;65:557-86.

Stamou SC, Jablonski KA, Pfister AJ Hill PC, Dullum MK, Bafi AS, Boyce SW, Petro KR, Corso PJ. "Stroke after Conventional versus Minimally Invasive Coronary Artery Bypass." *Ann Thorac Surg* 2002;74:394-9.

Teufelsbauer H, Prusa AM, Wolff K, Poterauer P, Nanobashvili J, Prage M, Holzenbein T, Thurnher S, Lammer J, Schemper M, Kretschnmer G, Huk I. "Endovascular Stent Grafting versus Open Surgical Operation in Patients with Infrarenal Aortic Aneurysms: A Propensity Score-Adjusted Analysis." *Circulation* 2002;106:782-7.

Trimble EL, Carter CL, Cain D, Freidlin B, Ungerleider RS, Friedman MA. Representation of older patients in cancer treatment trials. *Cancer.* 1994;74:Suppl 7:2208-2214.

Tufts Center for the Study of Drug Development. "BACKGROUNDER: How New Drugs Move through the Development and Approval Process." http://csdd.tufts.edu/NewsEvents/RecentNews.asp?newsid=4, accessed 1/7/2003.

Wagner JL, Alberts SR, Sloan JA, Cha S, Killian J, O'Connell MJ, Grevenhof PV, Lindman J, Chute CG. "Incremental costs of enrolling cancer patients in clinical trials: a population-based study." *J Natl Cancer Inst* 1999;91:847-53.

Warren JL, Klabunde, CN, Schrag D, Bach P, Riley GF. "Overview of the SEER-Medicare Data: Content, Research Applications, and Generalizability to the United Sates Elderly Population." *Med Care* 2002;40:3-18.

Weaver DA. "Labor force Participation, Income, and the Use of Short-Term Hospitals by the Elderly." *Med Care* 1996;34:989-1002.

Yang P, Cerhan JR, Vierkant RA, Olson JE, Vachon CM, Limburg PJ, Parker AS, Anderson KE, Sellers TA. "Adenocarcinoma of the Lung is Strongly Associated with Cigarette Smoking: Further Evidence from a Prospective Study of Women." Am J Epidemiol 2002 Dec 15;156(12):1114-22.

Zimmerman M, Mattia JI, Posternak MA. "Are Subjects in Pharmacological Treatment Trials of Depression Representative of Patients in Routine Clinical Practice?" *Am J Psychiatry* 2002;159:469-7

Chapter 2. Factors Affecting the Participation of Patients 65 Years of Age or Older in Cancer Clinical Trials [*]

[*] An earlier analysis of these data has been accepted for publication in the *Journal of Clinical Oncology* as Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, Housman MG, Escarce JJ, "The Participation of Patients 65 Years of Age and Older in Cancer Clinical Trials." Original work in this dissertation includes an exploration of the theoretical basis for testing the hypothesis that exclusion criteria would be expected disproportionately to affect elderly cancer patients, and an *a priori* estimate of the expected effect size. Alternative models are examined to determine whether having a proportion as the dependent variable constitutes a problem in the use of ordinary least squares regression.

Background and Theory

This study tests the hypothesis that lower cancer clinical trial participation rates for elderly (people aged 65 or older) patients can be explained by the presence of protocol exclusion criteria based on comorbid conditions, functional status, and life expectancy. In 1999, cancer was second only to heart disease as a leading cause of death (NCHS, 1999). The elderly account for approximately 61% of all incident cases of cancer and 70% of all cancer deaths (Yancik & Ries, 2000), and it is estimated that they have 11 times the cancer risk of people under age 65. By 2030 approximately 20% of the U.S. population will be aged 65 or older (Muss, 2001). Consequently, cancer care will become increasingly important, particularly for the elderly.

As a result of continuing advances in cancer care, cancer patients are living longer and experiencing better quality of life. Clinical studies have resulted in curative treatments for leukemias, lymphomas, and germ cell tumors and decreased morbidity and mortality from colorectal and breast cancer (NIH, 1991). Other clinical trials have helped establish better ways of caring for cancer patients, minimizing the side-effects of therapies, and reducing invasive procedures (Fisher et al., 1989; Perez et al., 1998).

For these reasons, concerns have been raised that clinical trials should include representative samples of patients to ensure that results are generalizable to the afflicted population. Considerable effort has gone into studying participation rates for elderly (herein defined as individuals 65 years or older) cancer patients (Goodwin et al., 1988; Hutchins et al., 1999; Trimble et al., 1994; Wardle et al., 2000). Furthermore, studies have been conducted to examine barriers to trial participation for elderly patients and for others, an excellent review of which may be found in Ross et al. (2001).

As noted in the introduction, federal law requires that NIH supported enroll representative samples of women and members of minority groups. These mandates have had some success: research suggests that racial and ethnic minorities and women are proportionately enrolled in National Cancer Institute (NCI)-sponsored cooperative group treatment trials (Tejeda et al., 1996; Chamberlain et al., 1998; Klabunde et al., 1999).

In contrast, studies suggest that the elderly are under-represented in cancer clinical trials (Goodwin et al., 1988; Trimble et al., 1994; Hutchins et al., 1999). A recent study of Southwest Oncology Group (SWOG) clinical trials active between 1993 and 1996 found that while approximately 63% of U.S. cancer patients were over age 65, the elderly comprised only 25% of trial participants (Hutchins et al., 1999). This study evaluated the elderly's participation using data from only one cooperative group. Moreover, the investigators did not evaluate whether the elderly's participation differed by phase of the trial or stage of disease, or what the reasons were for under-representation among the elderly. Recent federal efforts have focused on expanded Medicare coverage for clinical trials. However, to assess the likely impact of improved insurance coverage, it is important to understand the numerous factors that may affect the representation of elderly persons in cancer clinical trials.

One reason that the elderly may be under-represented in cancer trials is that protocol entry criteria may disproportionately impact these patients. Trials are designed to maximize confidence in the results found within constraints imposed by sample size and budget. Enrolling healthier patients decreases probability that subjects die or fall out of the study for causes unrelated to the disease or therapy being evaluated.

Older patients are more likely to have medical histories and conditions that make them ineligible for cancer treatment trials that include protocol exclusions. Table 2.1 details the

relative prevalence of potentially excludable conditions among adults 18-64 years old and those 65 or older (CDC, 2002). The table also shows a simple simulation of a hypothetical cancer trial with increasing numbers of exclusion criteria. Assume a hypothetical population of 2000 adult cancer patients, with equal numbers of whom are older and younger than 65. A trial with each of the listed organ system exclusions that screened equal numbers of elderly and non-elderly subjects would be expected to enroll only 27% of elderly participants. The proportion would fall to 15% if participants were required to have no activity impairments. This is a crude simulation as there is no data included on the likely joint distributions of comorbid conditions and impaired activity.

Table 2.1 Prevalence of Comorbid Conditions and Hypothetical Effects of Exclusions

		Prevalence							
Adult Population		**CAD**	**HTN**	**Pulmonary**	**Cancer**	**Diabetes**	**Renal**	**Hepatic**	**Impaired Activity**
18-64	163,269	4,939	22,377	43,777	6,059	5,831	1,762	1,396	47,439
65+	32,007	6,641	14,879	8,777	6,193	4,200	1,046	399	20,934
Percent									
18-64	83.6%	3.03%	13.71%	26.81%	3.71%	3.57%	1.08%	0.86%	29.06%
65+	16.4%	20.75%	46.49%	27.42%	19.35%	13.12%	3.27%	1.25%	65.40%
	Trial Simulation (Numbers Remaining after Exclusion for Comorbidity)								
18-64	1,000	970	837	612	590	569	563	558	396
65+	1,000	793	424	308	248	216	209	206	71
Total	2,000	1,762	1,261	920	838	784	771	764	467
%65+	50%	45%	34%	33%	30%	27%	27%	27%	15%

Frequencies for prevalence are in thousands (CDC, 2002). The simulation assumes that screening would reduce the numbers enrolled for each cohort proportionate to the prevalence of disease.

This study extends previous analyses by evaluating the participation of the elderly in a large sample of cancer clinical trials that were active from 1997 through 2000 and by using data from multiple cooperative groups. We examine the participation of elderly patients in clinical trials stratified by trial phase (II vs. III) and by stage of disease (early vs. late). Most important, we explore the impact of clinical trial protocol exclusions on elderly participation in trials.

Methods

Data Sources

We used three NCI databases: the Cancer Therapy Evaluation Program (CTEP, 2001), the Physician Data Query (PDQ, 2001) and the Surveillance Epidemiology and End Results Program (SEER, 2000). We used the CTEP and the PDQ data to detail the characteristics of NCI-sponsored clinical trials, including the age distribution of trial participants. We used the SEER data to compute national cancer incidence rates for the elderly, so that we could compare the proportion of patients enrolled in clinical trials who were elderly with the corresponding proportion of the population with cancer.

The CTEP Database

The Cancer Therapy Evaluation Program is operated within the Division of Cancer Treatment and Diagnosis of the NCI. Investigators report their progress with each protocol to the CTEP, which is responsible for planning, assessing and coordinating all aspects of clinical trials.

The CTEP data that we used included protocol identification numbers, trial phase, planned and actual trial accrual, date when the trial began to enroll patients, end date, and participation by age. Our study focused on 495 adult, phase II and III cooperative group cancer treatment trials that enrolled patients between 1997 and 2000. We chose to evaluate only cooperative group trials because of their strict reporting requirements: the CTEP database is considered complete for cooperative group trials active in 1997 and beyond. We assessed the participation of the elderly in these 495 clinical trials from 1997 through 2000. Table 2.2 describes the trials in the study by phase, cooperative group, and cancer type.

Table 2.2. Distribution of Sampled Trials by Phase, Cooperative Group and Cancer Type.

Category	Classification	Number	Accrual	Percent
Phase	Phase II	334	13,175	22%
	Phase III	161	46,125	78%
	Total	495	59,300	
Cooperative Group	CALGB	52	7,449	13%
	ECOG	112	13,311	22%
	GOG	72	6,766	11%
	INT	19	5,524	9%
	NCCTG	57	2,875	5%
	NSABP	11	7,435	13%
	RTOG	48	7,022	12%
	SWOG	96	5,859	10%
	Other Other: NABTC, NABTT, ACOSOG, EORTC, NCIC	28	3,059	5%
Cancer Type	Bladder	10	285	0.5%
	Breast	46	19,746	33.3%
	CNS	41	2,492	4.2%
	Cervical	26	1,335	2.3%
	Colorectal	26	6,431	10.8%
	Gastro-Esophageal	16	731	1.2%
	Head and Neck	23	2,006	3.4%
	Leukemia	38	1,989	3.4%
	Lung	62	6,873	11.6%
	Lymphoma	32	2,012	3.4%
	Melanoma	17	1,598	2.7%
	Myeloma	11	1,051	1.8%
	Ovarian	36	2,649	4.5%
	Pancreatic	12	1,121	1.9%
	Prostate	22	3,980	6.7%
	Renal	7	162	0.3%
	Soft Tissue Sarcoma	8	246	0.4%
	Uterine	30	3,466	5.8%
	Other	32	1,127	1.9%

*Clinical trials classified as "other" treated the following disorders: adrenocortical tumors, AIDS-related sarcomas and lymphomas, amyloidosis, carcinoid tumors, germ cell tumors, granulothrombocytopenia, hepatomas, mesotheliomas, mycosis fungoides, ostegenic sarcomas, penile tumors, testicular tumors, trophoblastic neoplasia, thymomas, urothelial tumors, vulvar tumors and Waldenstrom's macroglobulinemia.

The PDQ Database

The PDQ database contains detailed protocol exclusion criteria for NCI-sponsored clinical trials. For each of the 495 trials in the study, we determined the cancer type and stage, planned trial duration, and protocol exclusion criteria. Appendix 2.1 details the specific exclusions that we defined for each category of protocol exclusion criteria. Strict exclusions were those protocol exclusion criteria that required normal or nearly normal laboratory values or organ system function, whereas moderate exclusions allowed for mildly abnormal values, while still imposing restrictions.

To define functional status exclusions, we created a new performance score by matching the Karnofsky scores with the ECOG/Zubrod scores (Oken et al., 1982; Oncolink, 2000). The majority of the protocols used the ECOG/Zubrod score, where patients are assigned a score from zero to five based on their ability to carry on activities of daily living. However, a number of trials used the Karnofsky score, in which a person's functional status is rated from 0% to 100% of normal health. Appendix I provides a table relating our exclusion definitions to the ECOG/Zubrod and Karnofsky scales. We defined three levels of functional status restrictions. Each trial was coded to reflect the protocol requirement that participants be able to function at the specified level or better. Thus a trial with the most restrictive functional status requirements would require participants to be ambulatory and able to perform light work, whereas a trial with the most lenient functional status requirement would allow patients to enroll who were nonambulatory and had limited self-care capabilities.

Life expectancy requirements, where present, ranged from one month to ten years. We defined two categories of life expectancy criteria: less than or equal to 6 months, and greater than six months. Other exclusions included the requirements that patients have no history of

psychiatric problems specific to the elderly such as organic brain syndrome, Alzheimer's disease or senility; have no history of other neurologic or psychiatric disorders, other cancers, HIV/AIDS, other severe disease, or active infections; and not be pregnant.

We stratified cancer trials according to the stage of the cancer being treated in order to determine if the elderly were more or less likely to be represented in trials for treatment of early stage or late stage cancers. Appendix II details the stage categories used for each cancer type. In general, stage I and II cancers were considered early stage and stage III and IV cancers were considered late stage. Some protocols treated patients with varying stages of cancer that crossed over this division and, therefore, could not be classified.

The SEER Database

The Surveillance, Epidemiology, and End Results Program of the NCI is the most authoritative source of information on cancer incidence and survival in the United States (About SEER, 2000). The SEER data include 11 tumor registries covering approximately 14% of the U.S. population and 12% of the U.S. population age 65 or older (NCI SEER*Stat, 2000; US Census Bureau, 2000). SEER data include population-based information on demographics, tumor types, morphology, stage at diagnosis, first course of treatment, and follow-up vital status.

To calculate the proportion of the U.S. population with each cancer type ($US_{CA(i)}$) who were 65 or older ($US_{65\_CA(i)}$) we adjusted the incidence rates from the 1997 SEER data to reflect the proportion of elderly in the nation as a whole. Formally this is expressed:

$$\frac{\left(\frac{SEER_{65\_CA(i)}}{SEER_{65}}\right)*US_{65}}{\left[\left(\frac{SEER_{65\_CA(i)}}{SEER_{65}}\right)*US_{65}+\left(\frac{SEER_{LT65\_CA(i)}}{SEER_{LT65}}\right)*US_{LT65}\right]}=\frac{US_{65\_CA(i)}}{US_{CA(i)}}$$

We first used 1998 data from the U.S. Census Bureau to determine the number of elderly ($SEER_{65}$) and the total population ($SEER_{POP}$) in all of the counties represented in the 11 SEER registries. We then, using the SEER registry data, we determined the number of new cases of cancer among the elderly ($SEER_{65\_CA(i)}$) and the population under 65 ($SEER_{LT65\_CA(i)}$) by cancer type, *i*. We divided the aggregate numbers by the respective populations in the SEER areas to yield the SEER incidence rates for each cancer in both the elderly and the non-elderly populations. To yield the number of new cases nationally for both groups, we then multiplied the incidence rates for the elderly and for the populations within SEER registry counties for each type of cancer by the number of elderly (US_{65}) and the non-elderly (US_{LT65}) within the United States. Last, we divided the national number of new cases of cancer among the elderly by the national number of new cases of cancer in the total population to calculate the estimated proportion of the total population diagnosed with cancer who were elderly. We calculated proportions for 18 specific cancer types and for all cancer types combined. We also used the SEER data to calculate the proportions of elderly who present with early and late stage cancers for each cancer type and compared these numbers to the proportion of elderly in trials for early and late stage cancers.

Statistical Analysis

Statistical analyses used Stata, v7.0 (Stata Corp.) and Excel spreadsheets (Office 2000, Microsoft Corp.) were used for data management and tables. We evaluated the distribution of elderly participants across all trials, and for trials stratified by cancer type, phase and stage. We compared the participation rates to the proportions of the elderly in the U.S. population with each cancer type using one-sample binomial tests. Two tailed p-values of 0.05 or less were considered to indicate statistical significance (Cochran, 1977).

The study submitted for publication (Lewis et al. 2003) used an ordinary least squares (OLS) regression model to examine the association between the proportion of participants in each trial who were elderly, as the dependent variable in the model, and the year the trial opened, trial phase, cancer type and stage, and protocol exclusion criteria. The model is of the form: $Y = X\beta + \varepsilon$, where Y is the (n x 1) vector of *proportions* of patients in clinical trials, X represents an (n x k) matrix of observations of independent variables, β is a vector of unknown parameter estimates, and ε a random error term with mean zero and a normal distribution (Netter et al., 1996).

We defined indicator variables for the year the trial began, trial phase, and protocol exclusion criteria by using a backward stepwise selection procedure set to retain only variables that were significant at the 0.05 level. Indicator variables for cancer type and cancer stage (late), and their interactions, were forced into the model in order to control for epidemiological differences in age and stage distribution across cancer types. Each trial was weighted in the regression by its total enrollment. We presented the results based on this model for ease of interpretation.

A potential problem with the OLS model concerns the limited dependent variable, the proportion of patients enrolled in clinical trials. The proportion can only take on values between zero and one, inclusive, but OLS can produce conditional means outside that range. To some extent this concern is decreased by the fact that the parameters of interest apply to dummy variables, and thus the OLS becomes an analysis of variance with the partial effects indicating differences in means between those trials that have particular exclusions and those that do not.

A common alternative to OLS for limited dependent variables is a logit model taking the form: $P(y \neq 0 \mid X) = \frac{\exp(X\beta + \varepsilon)}{1 + \exp(X\beta + \varepsilon)}$ (Stata 2001, Wooldridge 2000). This formulation does not work for proportions in most statistical packages, where the dependent variables are assumed to be dichotomous zero/one variables. It is possible to achieve a similar result by performing a logit transformation on the proportion and using that as the dependent variable in an OLS regression: $\text{logit}(y_i) = \ln\left[\frac{y_i}{1 - y_i}\right] = X\beta + \varepsilon_i$ (Greene 2000, p. 835). This approach, however, causes all proportions of one or zero to be set to missing. Instead, a Generalized Linear Model (GLM) with a logit link (Hardin & Hilbe, 2001) allows the range restriction $0 \leq y \leq 1$ for the dependent variable. The model becomes $\ln\left(\frac{\mu}{1-\mu}\right) = X\beta + \varepsilon,$ where $\mu = \text{E}[y]$; various assumptions can be made as to the distribution of the error term. This allows a comparison of the OLS and GLM models in terms of which criteria are significant and in terms of goodness of fit, measured by root mean squared errors and mean absolute deviations. Again, observations were weighted by total trial enrollment.

Simulations

The regression parameter estimates were then used to predict the effect that relaxing protocol exclusion criteria would be expected to have on elderly participation rates. First exclusions based on organ system functions were relaxed (by setting the value of the associated dummy variables to zero), then predicted values were generated. Similarly, functional status and life expectancy criteria were relaxed and another set of predicted values were generated.

RESULTS

Descriptive Data

Table 2.3 reports the proportion of elderly patients in phase II and phase III clinical trials for 18 cancer types. Overall, 32% of the participants in Phase II and III clinical trials combined were elderly, compared with 61% of patients with incident cancers in the U.S. population who are age 65 or older. Figure 1 shows the proportion of elderly patients in phase II and phase III clinical trials for 18 cancer types compared with the proportion of the U.S. population with each cancer type who are elderly. The elderly were significantly under-represented ($p < .05$) in Phase III myeloma trials; Phase II central nervous system (CNS), gastro-esophageal, head and neck, leukemia, and pancreatic cancer trials; and Phase II and III breast, colorectal, and lung cancer trials.

Table 2.4 reports the proportion of elderly patients in early and late stage cancer trials by cancer type compared with the proportion of the U.S. population with early and late stage cancers who are elderly. (For this analysis we excluded 62 trials that could not be classified as early or late as well as trials for leukemia or myeloma.) The elderly were less underrepresented, relative to the incidence rate, in trials for late stage cancers than in trials for early stage cancers ($p < .001$). When we combined all cancer types, 25% of participants in trials for early stage cancers and 41% of participants in trials for late stage cancers were elderly. In the U.S. population, 57% of new cases of early stage cancer and 65% of new cases of late stage cancer occur in the elderly.

Figure 2.1. Comparison of elderly participation in phase II and III trials with percent of US cancer patients who are elderly.

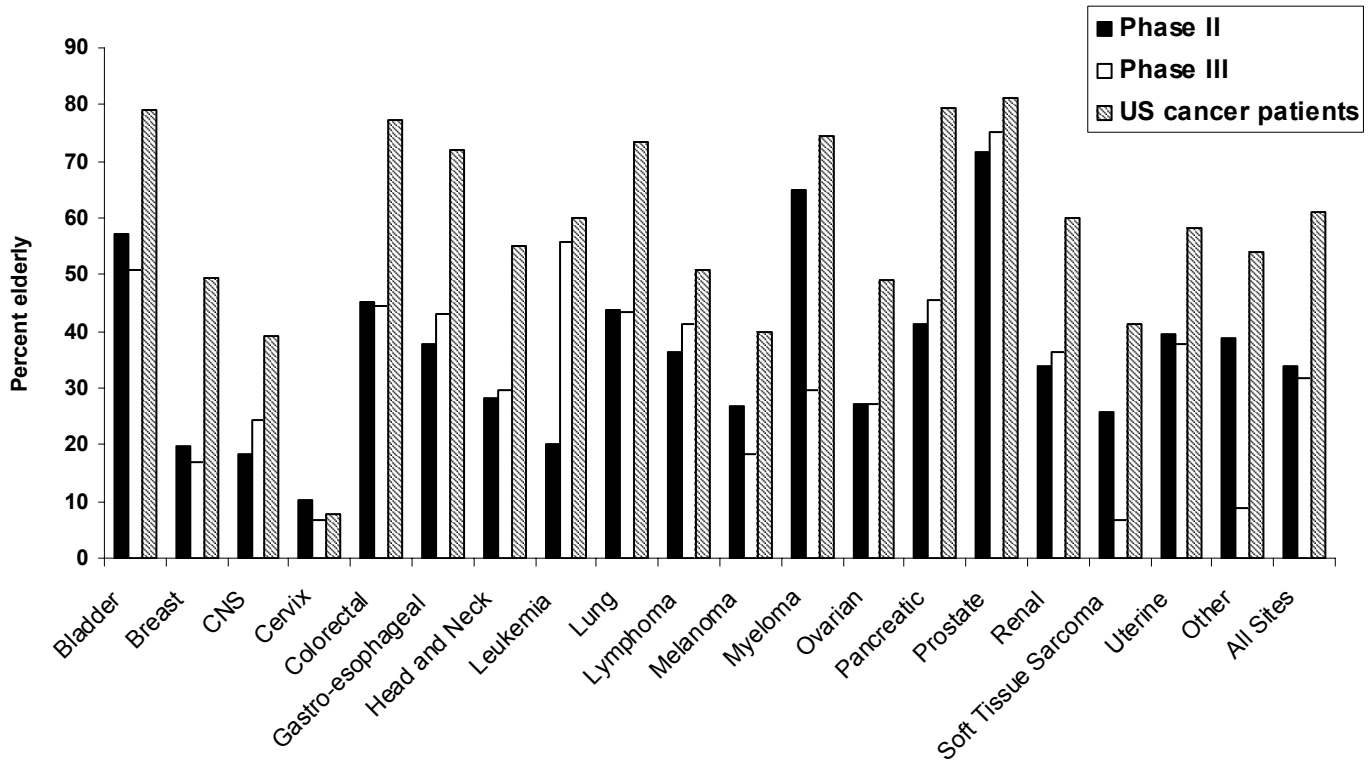


Table 2.3. Elderly Participation in NCI-Sponsored Cooperative Group Treatment Trials from 1997 through 2000 by Trial Phase.

	Phase II			Phase III		
Cancer Type	Number of Trials	Total Enrollment	Percent Elderly	Number of Trials	Total Enrollment	Percent Elderly
Bladder	7	159	57%	3	126	51%
Breast	21	1,096	20%	25	18,650	17%
CNS	34	1,396	19%	7	1,096	24%
Cervical	19	555	11%	7	780	7%
Colorectal	14	541	45%	12	5,890	44%
Gastro-esophageal	13	396	38%	3	335	43%
Head and Neck	15	793	28%	8	1,213	30%
Leukemia	27	1,030	20%	11	959	56%
Lung	41	2,516	44%	21	4,357	43%
Lymphoma	23	865	36%	9	1,147	41%
Melanoma	10	268	27%	7	1,330	18%
Myeloma	6	214	65%	5	837	30%
Ovarian	26	697	29%	10	1,952	27%
Pancreatic	10	506	41%	2	615	46%
Prostate	11	596	71%	11	3,384	75%
Renal	5	118	34%	2	44	36%
Soft Tissue Sarcoma	7	216	26%	1	30	7%
Uterine	18	538	39%	12	2,928	38%
Other	27	675	40%	5	452	9%
All Sites	334	13,175	34%	161	46,125	31%

Table 2.4. Elderly Participation in Trials by Stage at Diagnosis.*						
	Early Stage			Late Stage		
Cancer Type	Number of Trials	Percent Elderly	Incidence Rate**	Number of Trials	Percent Elderly	Incidence Rate**
Bladder	2	57%	78%	6	51%	81%
Breast	25	18%	49%	21	20%	48%
CNS	4	11%	6%	25	22%	31%
Cervical	3	4%	37%	21	9%	24%
Colorectal	4	54%	78%	16	41%	73%
Gastro-esophageal	4	42%	75%	11	39%	66%
Head and Neck	0	-	59%	22	29%	48%
Lung	13	48%	75%	46	42%	70%
Lymphoma	4	56%	48%	13	44%	51%
Melanoma	1	14%	44%	15	24%	46%
Ovarian	3	23%	31%	30	29%	59%
Pancreatic	1	40%	81%	10	45%	72%
Prostate	4	67%	82%	17	76%	73%
Renal	0	-	57%	7	35%	60%
Soft Tissue Sarcoma	0	-	41%	8	24%	40%
Uterine	5	38%	56%	23	43%	64%
Other* **	2	0%	37%	27	27%	47%
Above Sites Combined	75	25%	57%	318	41%	65%

* Excludes trials which treated patients with varying stages of cancer and therefore could not be classified as early or late and leukemia and myeloma trials for which incidence rates were unavailable.

Table 2.5 Exclusion Criteria Specified in 495 Phase II and III Trials.

Type of Exclusion	Phase II	Phase III	Aggregate
Hematological			
Strict	22%	27%	23%
Moderate	65%	48%	59%
Any	86%	75%	83%
Hepatic			
Strict	59%	61%	59%
Moderate	28%	21%	26%
Any	87%	82%	85%
Renal			
Strict	53%	43%	49%
Moderate	34%	37%	35%
Any	87%	80%	84%
Pulmonary			
Strict	1%	1%	1%
Moderate	9%	14%	11%
Any	10%	16%	12%
Psychological			
Broad	14%	20%	16%
Specific^	3%	3%	3%
Any	16%	23%	19%
Functional Status Requirements*			
Ambulatory and able to work	19%	30%	23%
Ambulatory and able to do ADLs**	71%	43%	62%
Non-ambulatory with limited self care	5%	9%	6%
Any			
Cardiac			
Congestive Heart Failure	41%	47%	43%
Coronary Artery Disease	34%	39%	35%
Conduction Disease / Arrythmia	23%	31%	26%
Hypertention	7%	8%	8%
Life Expectancy			
Life Expectancy <= 6 Months	20%	7%	16%
Life Expectancy > 6 Months	20%	25%	22%
Any	40%	33%	38%
Other			
Neurologic	16%	12%	15%
No Other Cancer	88%	91%	89%
AIDS/HIV	14%	13%	14%
Severe Disease	23%	28%	25%
Infection	41%	34%	39%

^ Specific psychiatric exclusions include organic brain syndrome, Alzheimer's Disease, and "senility".

* The protocols required individuals to function at the level detailed or better.

** Activities of Daily Living (ADLs). Requires enrollees to be capable of all self-care, but may be unable to carry out any work activities.

The majority of cancer trials prohibited participation by people with hematological, hepatic, renal, or cardiac abnormalities (Table 2.5). Over 85% of the trials required participants to be either ambulatory and capable of work or capable of carrying out their activities of daily living independently. A minority of trials excluded individuals who had specific psychiatric diseases that are more common in the elderly such as organic brain syndrome, Alzheimer's disease, or "senility." Few trials had exclusions based on pulmonary disease, but most trials excluded individuals who had a history of another cancer.

Regression Analysis

Trials with exclusions based on hypertension, cardiac, hematological or pulmonary function abnormalities enrolled lower proportions of elderly patients than trials without such exclusions (Table 2.6). For example, other things equal, the proportion of elderly patients was 7.8% lower (95% confidence interval [CI], 3.6% lower to 12.9% lower) in trials that excluded patients with cardiac abnormalities than in trials that did not exclude these patients. Similarly, trials that excluded patients with functional status limitations enrolled lower proportions of elderly patients than trials that explicitly allowed patients with impaired functional status. For instance, other things equal, the proportion of elderly patients was 22.4% lower (95% CI, 15.8% lower to 29.1% lower) in trials that excluded patients with mild functional status impairment than in trials that did not exclude these patients. Interestingly, trials that did not specify any functional status exclusions enrolled lower proportions of elderly patients than trials that explicitly allowed patients with impaired functional status. Trials that specified life expectancy requirements enrolled slightly higher proportions of elderly patients, whereas trials that specifically excluded pregnant women enrolled lower proportions of elderly patients. The

proportion of elderly patients was 18.9% higher (95% CI, 9.7% - 28.0% higher) in trials for late stage cancers. Trial phase did not effect any change in elderly participation rates.

Table 2.6. Impact of Exclusions on Participation of the Elderly in Clinical Trials. ^

Dependent Variable: Percent of Enrollment Aged 65+

	Change in Elderly Participation	(95% Confidence Interval)
Organ System		
Abnormal Cardiac function excluded	7.8% Lower	(4.5% to 11.0% Lower)
Hypertension excluded	6.4% Lower	(2.1% to 10.7% Lower)
Abnormal Hematologic function excluded	11.1% Lower	(7.4% to 14.7% Lower)
Abnormal Pulmonary function excluded	8.3% Lower	(3.6% to 12.9% Lower)
Functional Status		
Mild functional status impairment excluded	22.4% Lower	(15.8% to 29.1% Lower)
Moderate functional status impairment excluded	21.8% Lower	(15.4% to 28.2% Lower)
No Functional Status Exclusion Specified	28.4% Lower	(22.1% to 34.7% Lower)
Any Specified Life Expectancy Requirement	3.8% Higher	(1.0% to 6.7% Higher)
Late Stage Disease	18.9% Higher	(9.7% to 28.0% Higher)
Adjusted R-squared	0.653	

^ Controlling for Cancer site and stage and site-stage interactions.

The GLM regression results can be found in Appendix IV. The findings are the same with regard to which exclusion criteria are significant. Table 2.7 compares goodness of fit for the OLS regression model with the fit of the GLM regression. In terms of root-mean squared error, the fits are almost identical; OLS produced slightly higher mean absolute deviation.

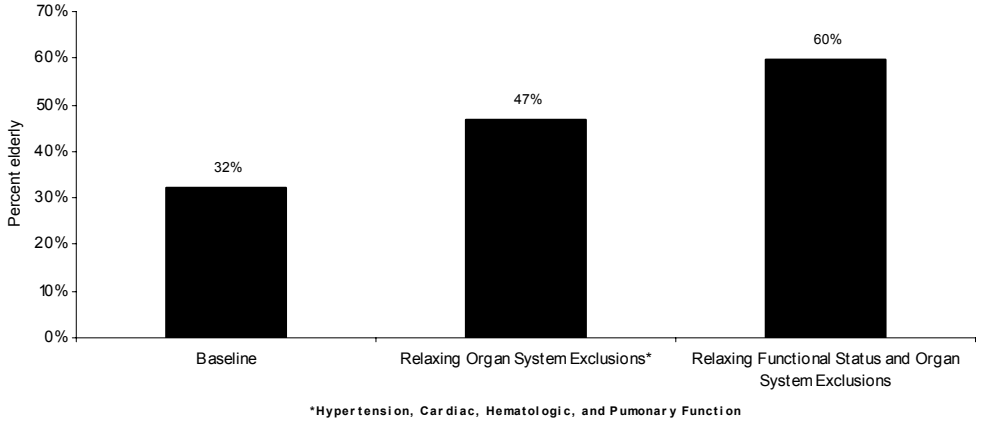
Table 2.7. Goodness of Fit Tests

	Model	
	OLS	GLM
Root Mean Squared Error	0.20238	0.20245
Mean Absolute Deviation	0.15480	0.15308

Simulations

Using the regression model, we performed simulations to predict the proportion of elderly participation that would be expected if trials did not have protocol exclusions based on organ system abnormalities or functional status limitations. When we relaxed the cardiac function, hypertension, hematological and pulmonary function exclusions, the overall predicted proportion of elderly patients rose to 47%. When we relaxed both the organ system and functional status exclusions, the overall predicted proportion of elderly patients increased to 60% (Figure 2.2).

Figure 2.2. Simulated impact on eldelry participation in cancer clinical trials of relaxing protocol exclusion criteria.



Discussion

The elderly are under-represented in cancer clinical trials relative to the proportion of patients with cancer who are elderly. Protocol exclusion criteria based on organ system abnormalities and functional status limitations are associated with lower rates of elderly participation in cancer trials and almost fully explain their observed under-representation.

Although the elderly were under-represented in these trials, they comprised a larger proportion of clinical trial participants than previously reported. Expanding the analysis of elderly participation to all cooperative groups and focusing on the past four years narrowed the gap between the 61% of the cancer population who are elderly and the previously reported 25% of cancer clinical trial participants who are elderly (Hutchins et al., 1999). Using more recent and comprehensive data, we found that 32% of patients in cancer trials were age 65 or older. Based on our simulation results, relaxing protocol exclusion criteria could result in elderly enrollment rates of up to 60%, or almost complete parity with the proportion of cancer patients 65 or older.

While Hunter et al. (1987) suggested that elderly under-representation could result from failure to meet eligibility criteria, this is the first study to summarize the protocol exclusion criteria that are used in cancer clinical trials, and to relate them to elderly participation. Empirical simulations based on the actual clinical trial data found that relaxing the protocol exclusions for hypertension and for cardiac, hematological, and pulmonary function abnormalities would be expected to increase elderly participation in cancer trials to 46%. Additionally relaxing the exclusions for functional status limitations would increase elderly participation to 59%, nearly eliminating the gap between the proportion of trial participants who are elderly and the proportion of cancer patients who are elderly. These simulations demonstrate the substantial impact of restrictive protocol exclusion criteria on elderly participation.

Of course, protocol exclusion criteria are not arbitrary. For example, it is important that participants in trials that employ nephrotoxic chemotherapies have normal renal function. Similarly, pulmonary and cardiac toxicity can be risks of cancer treatment, and it is reasonable for certain trials to require ample pulmonary or cardiac reserve in order for the patients to tolerate the therapy. Elderly patients with comorbid conditions may be more likely to die of causes other than the cancer being treated, making treatment effects more difficult to detect (Muss, 2001; Sargent et al., 2001). Nonetheless, protocol exclusion criteria based on comorbid conditions or functional status limitations disproportionately exclude older patients from clinical trials. If there are treatments that can be expected to affect elderly individuals differently, particularly the sicker elderly, then further study of outcomes, either in trials or observational studies are warranted.

As the U.S. population ages, a greater proportion of cancer patients will be elderly. Studies of many different cancers have demonstrated age-related differences in the natural history of cancer and in the effect of cancer treatment. For example, in prostate cancer, age has been found to be an independent predictor of distant metastases after treatment (Herold, Hanlon, Movsas & Hanks, 1998). In non-Hodgkin's lymphoma, age greater than 65 has been found to be a significant negative prognostic factor (Maksymiuk, 1996). Studies of leukemias have found that older patients do not tolerate intensive treatment as well as younger patients (Johnson & Liu, 1993; Ryan et al., 1992). Also, specific biologic characteristics in older patients can be associated with poor outcomes (Leith et al., 1997), and there is evidence that hematological, cardiac, gastrointestinal, and neurological toxicity related to chemotherapy may be more severe in older patients (Kimmick, Flemming, Muss & Balducci, 1997).

This study and the data have limitations. First, the data do not indicate the degree to which the protocol exclusions were followed. However, all of the trials in our sample were audited according to the CTEP guidelines. Second, the regression analyses do not demonstrate that protocol exclusion criteria are causally related to lower elderly participation; rather, they reveal associations that in some cases may have alternative explanations. For instance, we found an association between a "not pregnant" exclusion and lower elderly participation rates; which is clearly contrary to any reasonable expectation. Of note: the NCI policy since 1998 has been that patients should not be automatically excluded based on pregnancy or breast feeding (CTEP, 2001). The finding of a positive association between life expectancy requirements and higher elderly participation was also unexpected. Trials may have been actively targeting older populations and the investigators, therefore, specified life expectancy exclusions. Despite these unexpected findings, it seems likely that most of the associations we found between elderly participation and protocol exclusions based on organ system abnormalities or functional status limitations represent causal relationships.

Lastly, there remains considerable variability that remains unexplained (the R^2 values in the OLS regression was .65 and .76 in the GLM model). Not assessed are the non-clinical factors that may influence the elderly's participation in cancer trials. For example, older patients may be less likely to seek out clinical trials (Trimble et al., 1994), or more inclined to obtain treatment from community physicians rather research centers. Differences in elderly persons' preferences for trials could stem from differences in education, stronger relationships with primary care physicians, or difficulty getting to and from distant providers. The frequent visits required for aggressive cancer care or for participation in clinical trials may not be feasible for elderly persons who live alone or lack social supports. Additionally, the elderly and their families may

have preconceived notions about the potential benefits to elderly patients from participating in clinical trials or from aggressive cancer therapy.

The NCI has several initiatives in place to assess the impact of various factors that may affect the recruitment of older patients to clinical trials and to understand the effect of comorbidities on tolerance of cancer treatment (Trimble et al., 1994; Muss, Cohen & Lichtman, 2000). Future research should examine the preferences of the elderly regarding participation in trials, as well as the beliefs and behaviors of investigators regarding participation of the elderly in trials.

Our study findings suggest that recent federal policy to expand Medicare coverage for cancer clinical trials is, by itself, unlikely to increase substantially the level at which the elderly participate in cancer treatment trials. We found that protocol exclusions based on organ system abnormalities and functional status limitations in NCI-sponsored trials disproportionately disqualify the elderly from participation, and almost fully account for elderly patients' under-representation in trials relative to their cancer burden. To raise elderly participation rates above what could be achieved by relaxing exclusion criteria, it would be necessary to actively exclude younger people from trials. In some cases that might be desirable.

As noted in the introduction, if there is reason to believe that specific treatments have differential effects in elderly individuals, it is not sufficient to see that the elderly are represented in clinical trials. It may be necessary, as has been done to conduct RCTs with age restrictions that allow only the elderly to participate. Or it may be possible to conduct observational studies using registries, administrative data, or other sources that do not involve randomization. Nonetheless, if we need to know how treatments affect specific groups, whether based on ethnicity, gender, or age, then it is necessary to conduct studies focusing on those groups.

REFERENCES

About SEER. Available at: http://www-seer. ims. nci. nih. gov/AboutSEER. html accessed September 14, 2000.

Benson AB 3d, Pregler JP, Bean JA, Rademaker AW, Eshler B, Anderson K. Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study. *J Clin Oncol*. 1991;9:2067-75.

CDC—Centers for Disease Control and Prevention, Summary Health Statistics for US Adults, National Health Interview Survey, 1997. Vital and Health Statistics, Series 10 No. 205. DHHS, 2002.

Chamberlain RM, Winter KA, Vijayakumar S, et al. Sociodemographic analysis of patients in radiation therapy oncology group clinical trials. *Int J Radiat Oncol Biol Phys*. 1998;40(1): 9-15.

Cotton P. FDA lifts ban on women in early drug tests, will require companies to look for gender differences. *JAMA*. 1993;269:2067.

CTEP Policies Guidelines and Procedures. Available at: http://ctep. info. nih. gov/Policies/PregnancyGuide. htm accessed November 28, 2001.

Fisher B, Redmond C, Poisson R, et al. Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. *N Engl J Med.* 1989;320 (13):822-8.

Food and Drug Administration. Guideline for the study and evaluation of gender differences in the clinical evaluation of drugs: notice. *Fed Regist.* 1993;58(139):39406-16.

Freedman DA, Ecological Inference and the Ecological Fallacy, in International Encyclopedia of the Social & Behavioral Sciences 6: 4027–30, Eds. Smelser NJ and Baltes PB. Technical Report No. 549, 30 October 1999.

Goodwin JS, Hunt WC, Humble CG, Key CR, Samet JM. Cancer treatment protocols. Who gets chosen? *Arch Intern Med.* 1988;148(10):2258-60.

Greene WH. *Econometric Analysis*, 4th Edition, Prentice Hall, Upper Saddle River, NJ, 2000.

Hebert-Croteau N, Brisson J, Latreille J, Blanchette C, Deschenes L. Compliance with consensus recommendations for the treatment of early stage breast carcinoma in elderly women. *Cancer* 1999;85:1104-1113.

Herold DM, Hanlon AL, Movsas B, Hanks GE. Age-related prostate cancer metastases. *Urology*. 1998; 51(6):985-90.

Hardin J & Hilbe J. *Generalized Linear Models and Their Extensions*. College Station, TX: Stata Press, 2001, pp 101-114.

Hunter CP, Frelick RW, Feldman AR, et al. Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log. *Cancer Treat Rep*. 1987;71(6):559-65.

Hutchins LF, Unger JM, Crowley JJ, Coltman CA Jr, Albain KS. Under-representation of Patients 65 Years of Age or Older in Cancer-Treatment Trials. *N Engl J Med.* 1999;341(27):2061-2067.

Johnson PRE, Liu Yin JA. Acute myeloid leukemia in the elderly: biology and treatment. *British Journal of Hematology*. 1993;83:1-6.

Kennedy P. *A Guide to Econometrics*, 4th Edition, MIT Press, Cambridge, MA, 1998.

Kimmick GG, Fleming R, Muss HB, Balducci L. Cancer chemotherapy in older adults. A tolerability perspective. *Drugs Aging.* 1997;10(1):34-49.

Klabunde CN, Springer BC, Butler B, White MS, Atkins J. Factors influencing enrollment in clinical trial for cancer treatment. *South Med J*. 1999;92(12):1189-93.

Leith CP, Kopecky KJ, Godwin J, et al. Acute Myeloid Leukemia in the Elderly: Assessment of multi-drug Resistance (MDRI) and cytogenetics distinguishes biologic subgroups with remarkably distinct responses to standard Chemotherapy: a Southwest Oncology Group study. *Blood*. 1997;89:3323-9.

Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, Housman MG, Escarce JJ. "The Participation of Patients 65 Years of Age and Older in Cancer Clinical Trials," *J Clin Oncol*, in press.

Maksymiuk AW, Haines C, Tan LK, Skinnider LF. Age-related prognostic factor analysis in non-Hodgkin's lymphoma. *Can J Oncol.* 1996;6(1):435-442.

Moore DH. Ovarian cancer in the elderly patient. *Oncology.* 1994;8(12):21-25.

Muss HB, Cohen HJ, Lichtman SM. Clinical research in the older cancer patient. *Hematol Oncol Clin North Am* 2000;14:283-291.

Muss HB. (Editorial) Older Age – Not a Barrier to Cancer Treatment. *N Engl J Med.* 2001;345(15):1128-1129.

National Cancer Institute, CancerNet, http://cancernet. nci. nih. gov/trialsrch. shtml, accessed August 31, 2001.

National Cancer Institute, SEER*Stat 3. 0 SEER Cancer Incidence Public-Use Database, 1973-1997, Issued April 2000.

National Institutes of Health Consensus Conference. Treatment of early-stage breast cancer. *JAMA*. 1991; 265(3):391-5.

National Institutes of Health Consensus Conference. Adjuvant therapy for patients with colon and rectal cancer. *JAMA*. 1990; 264(11):1444-50.

NCHS—National Center for Health Statistics. *National Vital Statistics Report*, Deaths: Preliminary Data for 1999. Vol. 49, No. 3.

Neter J, Kutner MH, Nachtscheim CJ, Wasserman W. Applied Linear Statistical Models, 4th Ed. , Chicago: Irwin, 1996, pp 348-52.

Oken, Martin M. et al. Toxicity and Response Criteria of the Eastern Cooperative Group. *Am J Clin Oncol*. 1982;5:649-655.

Oncolink. Available at: www. oncolink. upenn. edu/disease/general/faz/faz_perform_status. html accessed August 1, 2000.

Perez EA, Hesketh P, Sandbach J, et al. Comparison of Single-Dose Oral Granisetron versus Intravenous Ondansetron in the Prevention of Nausea and Vomiting Induced by Moderately Emetogenic Chemotherapy: A Multi-center, Double-Blind, Randomized Parallel Study. *JCO*. 1998;16(2):754-60.

Pinn VW. Women, Research, and the National Institutes of Health. *Am J Prev Med.* 1992;8:324-7.

Ryan DH, Kopecky KJ, Head D, Grever MR, Shiaer SM, Lipschitz DA. Analysis of Treatment Failure in Acute Non-lymphocytic Leukemia Patients over Fifty Years of Age: A Southwest Oncology Group Study. *Am J Clin Onc*. 1992;15(1):69-75.

Sargent DJ, Golberg RM, Jacobson SD, et al. A Pooled Analysis of Adjuvant Chemotherapy for Resected Colon Cancer in Elderly Patients. *N Engl J Med.* 2001;345:1091-1097.

Schmucker DL, Vesell ES. Under-representation of Women in Clinical Drug Trials. *Clin Pharmacol Ther.* 1993;54:11-5.

Tejeda HA, Green SB, Trimble EL, et al. Representation of African-Americans, Hispanics, and whites in National Cancer Institute cancer treatment trials. *J Natl Cancer Inst*. 1996;88(12):812-6.

Trimble EL, Carter CL, Cain D, Freidlin B, Ungerleider RS, Friedman MA. Representation of older patients in cancer treatment trials. *Cancer.* 1994;74:Suppl 7:2208-2214.

US Census Bureau, Annual Population Estimates by Age Group and Sex, Selected Years from 1990 to 2000. Available at: http://www. census. gov/population/www/estimates/nation2. html, accessed September 4, 2000.

WG Cochran, *Sampling Techniques*. 3rd ed. New York, NY: John Wiley and Sons 1977.

Wooldridge JM. *Introductory Econometrics: A Modern Approach*. South-Western College Publishing, 2000.

Yancik R, Ries L. A. Aging and cancer in America: demographic and epidemiologic perspectives. *Hematol Oncol Clin North Am.* 2000;14:17-23.

Appendix 2.1. Protocol Entry Criteria Exclusions for Comorbid Conditions

Hematology

Moderate Restrictions	Strict Restrictions
Adequate Hematologic Function	Normal or near normal required
WBC > = 3,500	WBC >= 4, 000
ANC >= 1,500	ANC >= 1,800
Granulocyte > = 1500	ANC >= 2,000
PLT >= 125,000	Granulocyte > = 1800
Hg > = 11 or HCT >=33	Granulocyte > = 2000
Bone Marrow Cellularity > = 30%	PLT Normal
Fibrinogen >=200mg/dl	PLT >=130,000
	Hg or HCT Normal

Hepatic

Moderate Restrictions	Strict Restrictions
Adequate Hepatic Function	Liver Function Tests Normal or near normal
No Acute Hepatitis	Bilirubin Normal
Hepatitis C status required	Bili < = 1.5 mg/dl
LFT < = 2.5* NL	Direct Bilirubin Normal
Bilirubin < 2.5 * NL of <= 5mg/dl	AST/ALT < 1.5 NL
Direct Bilirubin < .3mg above NL	AST <= 60 IU/ml
AST/ALT < 5 NL	ALT < = 56 IU/ml
GGT < 3 * NL	GGT NL
Alkaline Phosphatase < 5 * NL	AP NL
LDH < 3 * NL	AP < 1.2 * NL
Triglycerides <= 320 mg/dl	LDH NL
PT NL	
PTT NL	
Thrombin Time NL	

Renal

Moderate Restrictions	Strict Restrictions
Adequate Renal Function	Normal Renal Function
Creatinine Clearance > = 50	Creatinine Clearance > = 70
Creatinine < 2 mg/dl or < 2* NL	Creatinine < 1.8 mg/dl or < 1.3 * NL
Creatinine < .8 mg above nl	Creatinine < .3 mg above nl
Creatinine < 2 * NL	BUN < 25 or < 1.5 * NL
BUN < 33 or < 2 * NL	
Calcium < 1.2 * NL	

Pulmonary

Moderate Restrictions	Strict Restrictions
No acute respiratory infection	No History of COPD or Chronic restrictive pulmonary dx.
No active COPD	FEV1 > 80% predicted
No significant non-neoplastic pulmonary disease	DLCO > 80% predicted
Medically fit for pulmonary resection	
PFT's at least 50% predicted (unless d.t. myeloma)	
FVC >= 60% predicted	
FEV1 > 2 L or pred. Post resection > 800 mL	
FEV1 >= 60% predicted	
DLCO >= 50% predicted	
FEV1/FVC <65%	

Psychiatric

Broad psychiatric exclusions	Specific psychiatric exclusions
No condition that would preclude informed consent.	No organic brain syndrome, alzheimer's disease or altered mental status
No condition that would interfere with protocol compliance.	No senility or severe emotional instability.
No significant psychiatric disease	No hospitalizations for psychiatric illness, including depression or psychosis.
No psychoses	

Appendix 2.2 Protocol Exclusion Criteria; Performance Status Score relating the ECOG/Zubrod and the Karnofsky Scores.

Summary Exclusion Rating	ECOG/Zubrod Score	Karnofsky Score
1	0= Fully active, able to carry on all pre-disease performance without restriction. 1= Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work.	100 = Normal, no complaints; no evidence of disease 90 = Able to carry on normal activity; minor signs or symptoms of disease. 80 = Normal activity with effort, some signs or symptoms of disease. 70 = Cares for self but unable to carry on normal activity or to do active work.
2	2= Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours.	60 = Requires occasional assistance but is able to care for most of personal needs. 50 = Requires considerable assistance and frequent medical care.
3	3= Capable of only limited self-care, confined to bed or chair more than 50% of waking hours. 4= Completely disabled. Cannot carryon any self-care. Totally confined to bed or chair. 5= Deceased.	40 = Disabled; requires special care and assistance. 30 = Severely disabled; hospitalization is indicated although death not imminent. 20 = Very ill; hospitalization and active supportive care necessary. 10 = Moribund 0 = Deceased

Appendix 2.3 Protocol Entry Exclusions - Cardiac

Moderate	Strict
Congestive Heart Failure; Cardiac Function	
No condition adversely affected by sinus bradycardia	No History CHF
Adequate Cardiac Function	No cardiomegally on CXR or LVH on EKG unless LV EF > = 45%
No uncontrolled or severe cardiovascular disease	Normal MUGA or echo
No active cardiac disease that precludes doxorubicin or docetaxel	NL LVEF
No clinically evident CHF	LVEF > = 45%
No difficult to control CHF	LVEF >= 45% and 50% with ex, or LVEF >= 55%
NYHA class II required (if protocol states I or II then put here; = no class III or IV)	LVEF > = 50% on MUGA
No valvular disease with cardiac function compromise	NYHA class I required (= no NYHA class (II/III/IV)
No Pericarditis or myocarditis	
No cardiomyopathy	
Coronary Artery Disease	
No MI past 12 months	No active angina
No MI past 6 months	No MI ever
No MI past 3 months	No MI past 5 years
No MI past 6 weeks	No History Ischemic Heart Disease
No CABG past 6 months	
No unstable angina	
No angina requiring medication	
Cardiac Electro-physiology problems	
No unstable heart rhythm	No abnormal Conduction Disease
No major ventricular arrhythmia	No arrhythmia requiring treatment
No arrhythmia associated with heart failure	
No arrhythmia that is difficult to control	
No cardiac medications that alter cardiac conduction	
No symptomatic arrhythmia within past 6 months	
Conduction disease allowed if stable for 6 months	
Hypertension	
No poorly controlled hypertension	No History of Hypertension
No Systolic BP > 200 or DBP > 120	No Diastolic BP > 100 mmHG
	No Systolic BP > 160 or DBP > 100 mmHg
Other Cardiovascular	
No thromboembolic dx past 6 months	No histroy Peripheral Vascular Disease
No History DVT past 6 months	No History of stroke
	No History of TIA
	No thromboembolic disease history
	No history of chronic CVA
	No History DVT
	No History PE

Appendix 2.4 OLS Regression Output

	Parameter	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Protocol Exclusion Criteria	No Hypertension	-0.064	0.022	-2.930	0.004	-0.107	-0.021
	No Abnormal Cardiac Function	-0.078	0.017	-4.670	0.000	-0.110	-0.045
	No Hematologic Function Abnormality	-0.111	0.019	-5.910	0.000	-0.147	-0.074
	No Impaired Pulmonary Function	-0.083	0.024	-3.490	0.001	-0.129	-0.036
	Any Specified Life Expectancy	0.038	0.014	2.650	0.008	0.010	0.067
	No Mild Functional Status Impairment	-0.224	0.034	-6.660	0.000	-0.291	-0.158
	No Moderate Functional Status Impairment	-0.218	0.033	-6.670	0.000	-0.282	-0.154
	No Functional Status Criteria Specified (Omitted Variable is No Severe Functional Status Impairment)	-0.284	0.032	-8.870	0.000	-0.347	-0.221
	No Pregnancy	-0.108	0.018	-5.990	0.000	-0.143	-0.072
	Late Stage Disease	0.189	0.047	4.050	0.000	0.097	0.280
	Cancer Site						
	Bladder	0.252	0.096	2.610	0.009	0.062	0.442
	Breast	0.061	0.054	1.120	0.263	-0.046	0.167
	CNS	0.013	0.073	0.170	0.861	-0.131	0.156
	Cervical	-0.261	0.099	-2.650	0.008	-0.455	-0.068
	Colorectal	0.293	0.057	5.120	0.000	0.180	0.405
	Gastro-esophageal	0.238	0.089	2.660	0.008	0.062	0.414
	Head and Neck	0.196	0.252	0.780	0.436	-0.298	0.691
	Leukemia	-0.123	0.076	-1.620	0.105	-0.272	0.026
	Lung	0.271	0.063	4.280	0.000	0.146	0.395
	Lymphoma	0.140	0.074	1.880	0.060	-0.006	0.286
	Melanoma	-0.025	0.076	-0.330	0.740	-0.174	0.124
	Myeloma	0.495	0.152	3.270	0.001	0.197	0.794
	Ovarian	-0.254	0.088	-2.880	0.004	-0.426	-0.081
	Pancreatic	0.240	0.094	2.540	0.011	0.054	0.425
	Prostate	0.268	0.094	2.860	0.004	0.084	0.452
	Renal	-0.039	0.122	-0.320	0.749	-0.279	0.201
	Soft Tissue Sarcoma	-0.078	0.102	-0.770	0.444	-0.279	0.122
	Uterine	-0.030	0.052	-0.590	0.558	-0.132	0.071
	Site x Stage Interactions						
	Late Stage-Breast	-0.163	0.066	-2.470	0.014	-0.292	-0.033
	Late Stage-CNS	-0.129	0.076	-1.710	0.088	-0.278	0.019
	Late Stage-Cervical	-0.073	0.104	-0.700	0.486	-0.278	0.132
	Late Stage-Colorectal	-0.157	0.061	-2.560	0.011	-0.278	-0.036
	Late Stage-Gastro-esophageal	-0.162	0.116	-1.400	0.163	-0.390	0.066
	Late Stage-Head and Neck	-0.284	0.252	-1.120	0.261	-0.780	0.212
	Late Stage-Leukemia	0.128	0.082	1.570	0.118	-0.032	0.288
	Late Stage-Lung	-0.177	0.061	-2.910	0.004	-0.296	-0.058
	Late-Stage Lymphoma	-0.177	0.083	-2.120	0.035	-0.341	-0.013
	Late Stage-Melanoma	-0.073	0.090	-0.820	0.413	-0.250	0.103
	Late Stage-Myeloma	-0.486	0.156	-3.110	0.002	-0.793	-0.179
	Late Stage-Ovarian	0.138	0.087	1.580	0.115	-0.034	0.309
	Late Stage-Pancreatic	-0.188	0.104	-1.810	0.071	-0.391	0.016
	Late Stage-Prostate	-0.058	0.090	-0.650	0.519	-0.235	0.119
	Intercept	0.618	0.059	10.490	0.000	0.502	0.734

The dependent variable is the proportion of trial enrollees who were aged 65 or older
(N = 495) The Adjusted R-squared statistic was 0.653

Appendix 2.5 GLM Regression Output							
Protocol Exclusion Criteria	No Hypertension	-0.369	0.130	-2.830	0.005	-0.624	-0.113
	No Abnormal Cardiac Function	-0.400	0.085	-4.720	0.000	-0.566	-0.234
	No Hematologic Function Abnormality	-0.509	0.096	-5.290	0.000	-0.697	-0.320
	No Impaired Pulmonary Function	-0.403	0.131	-3.070	0.002	-0.660	-0.145
	Any Specified Life Expectancy	0.241	0.080	3.010	0.003	0.084	0.398
	No Mild Functional Status Impairment	-1.109	0.170	-6.510	0.000	-1.443	-0.775
	No Moderate Functional Status Impairment	-1.107	0.165	-6.710	0.000	-1.430	-0.783
	No Functional Status Criteria Specified	-1.384	0.152	-9.120	0.000	-1.681	-1.086
	No Pregnancy	-0.579	0.089	-6.500	0.000	-0.754	-0.405
	Late Stage Disease	0.894	0.240	3.730	0.000	0.424	1.364
	Bladder	1.151	0.453	2.540	0.011	0.263	2.039
	Breast	0.202	0.310	0.650	0.514	-0.405	0.809
	CNS	-0.053	0.456	-0.120	0.908	-0.946	0.841
	Cervical	-2.376	1.958	-1.210	0.225	-6.215	1.462
	Colorectal	1.395	0.311	4.480	0.000	0.785	2.006
	Gastro-esophageal	1.168	0.426	2.740	0.006	0.332	2.004
	Head and Neck	1.032	1.228	0.840	0.401	-1.375	3.438
	Leukemia	-0.861	0.472	-1.820	0.068	-1.786	0.064
	Lung	1.244	0.336	3.700	0.000	0.585	1.904
	Lymphoma	0.669	0.397	1.680	0.092	-0.109	1.447
	Melanoma	-0.219	0.476	-0.460	0.646	-1.152	0.714
	Myeloma	2.240	0.820	2.730	0.006	0.633	3.848
	Ovarian	-1.341	0.496	-2.700	0.007	-2.313	-0.368
	Pancreatic	1.186	0.442	2.680	0.007	0.320	2.053
	Prostate	1.067	0.453	2.360	0.018	0.180	1.955
	Renal	-0.158	0.566	-0.280	0.780	-1.267	0.950
	Soft Tissue Sarcoma	-0.295	0.556	-0.530	0.595	-1.386	0.795
	Uterine	-0.150	0.278	-0.540	0.589	-0.695	0.395
	Late Stage-Breast	-0.706	0.382	-1.850	0.065	-1.455	0.043
	Late Stage-CNS	-0.619	0.471	-1.320	0.188	-1.541	0.303
	Late Stage-Cervical	0.010	2.050	0.000	0.996	-4.008	4.028
	Late Stage-Colorectal	-0.646	0.295	-2.190	0.028	-1.225	-0.068
	Late Stage-Gastro-esophageal	-0.801	0.522	-1.540	0.125	-1.823	0.222
	Late Stage-Head and Neck	-1.419	1.225	-1.160	0.247	-3.820	0.983
	Late Stage-Leukemia	0.892	0.460	1.940	0.053	-0.010	1.794
	Late Stage-Lung	-0.788	0.295	-2.670	0.008	-1.366	-0.209
	Late-Stage Lymphoma	-0.851	0.406	-2.100	0.036	-1.647	-0.055
	Late Stage-Melanoma	-0.189	0.532	-0.350	0.723	-1.232	0.854
	Late Stage-Myeloma	-1.991	0.829	-2.400	0.016	-3.617	-0.366
	Late Stage-Ovarian	0.716	0.475	1.510	0.132	-0.215	1.647
	Late Stage-Pancreatic	-0.916	0.454	-2.020	0.043	-1.805	-0.027
	Late Stage-Prostate	-0.209	0.422	-0.490	0.621	-1.036	0.619
	Intercept	0.690	0.311	2.220	0.027	0.079	1.300

Appendix 2.6 Categories for Early and Late Stage Cancers

Bladder Early: Stage I, II, localized, carcinoma in situ, Stage Ta transitional.
Late: Stage III, IV, refractory, relapsed or metastatic.

Breast Early**:** Stage 0, I, II, IIIa, in situ, localized or regional by direct extension.
Late: Stage IIIb, IV, advanced or metastatic.

Cervical Early: Stage O, I, II, in situ or localized.
Late: Stage III, IV, incurable, advanced or inoperable.

CNS Early: Localized, regional by direct extension.
Late: Distant, unresectable, aggressive, poor risk, recurrent or advanced.

Colorectal Early: Stage 0,I, II, in situ, localized or completely resected.
Late: Stage III, IV, distant, metastatic or advanced.

Gastro-Esophageal
Early: In situ, localized, regional by direct extension or resectable.
Late: Regional by nodes, distant, advanced, unresectable or metastatic.

Head and Neck
Early: Stage I, II, in situ, localized, or regional by direct extension.
Late: Stage III, IV, regional by nodes, distant, advanced or metastatic.

Lung Early: Stage 0, I, II, in situ, localized or limited stage.
Late: Stage III, IV, advanced, metastatic or extensive stage.

Lymphoma; Hodgkin's and Non Hodgkin's
Early: Stage 0, I, II or localized.
Advanced: Stage III, IV, advanced or distant.

Melanoma Early: Stage O, I, II, in situ, localized, regional by direct extension.
Late: Stage III, IV, regional by nodes, distant, advanced or metastatic.

Ovarian Early: stage 0,I, II, in situ, localized or regional by direct extension.
Late: stage III, IV, regional by nodes, distant or metastatic.

Pancreatic Early: stage I, II, in situ, localized or regional by direct extension.
Late: stage III, IV, regional by nodes, distant or metastatic.

Prostate Early: Stage 0,I, II or in situ.
Late: Stage III, IV, distant or metastatic.

Renal Early: In situ, localized or regional by direct extension.
Late: Regional by nodes, distant, advanced or metastatic.

Soft Tissue Sarcoma
Early: Stage I, II, localized or regional by direct extension.
Late: Stage III, IV, regional by nodes, distant or advanced.

Uterine Early: Stage I, II or localized.
Late: Stage III, IV, regional by nodes, distant or metastatic.

Chapter 3. Comparing Data Sources for Health Services Research: Findings from the Cost of Cancer Treatment Study

When investigators are designing health service research studies the choice of data sources is among the first concerns. The type of data collected will influence how well a study can address specific aims, and the data collection is often among the most costly components of a research plan. The Cost of Cancer Treatment Study (CCTS) provides a rare opportunity to evaluate several data sources both in terms of the effort needed to acquire the data, the quality of the data developed, and agreement between different sources. This chapter presents a case study of data collection results, and while the methods described fit the context of a specific study the results illustrate strengths and weaknesses of different data source in addressing specific questions in health services research.

The Cost of Cancer Treatment Study (CCTS) selected a probability sample of adult cancer patients participating in treatment trials sponsored by the National Cancer Institute. These subjects were matched to a cohort of cancer patients who were not participating in clinical trials based on the institutions where they received treatment and disease and comorbid characteristics as detailed in trial protocols (Goldman, Adams, et al., 2000). Ultimately 1628 subjects were enrolled. The project attempted to obtain health service utilization data on these individuals from telephone surveys, medical records, provider billing records, and Medicare claims data. Thus, we have information developed from up to four sources for some individuals. This account examines the costs and quality of the data produced from each of these sources.

Medical records are generally accepted as valid sources of documentation for health services. As the delivery of health services becomes more complex and less centralized, however, complete medical records have become increasingly difficult and

expensive to obtain. Administrative data (billing records) and Medicare claims data have also seen wide use. Self-reported data on utilization rates can be subject to recall and response biases, but may be easier to obtain than are data from medical records. Indeed, for some data, such as perceptions, comprehension, and value judgments of illnesses and treatments, self-report will be the only data source.

There are two basic sources of systematic disagreement between data obtained from subject self-reports and data derived from medical records or administrative databases. Self-reported data may be subject to recall bias (or other types of response bias), and medical records or administrative data may under-report some classes of data.

The literature on the comparison of data from different sources for utilization measures is sparse. Most data on recall bias has been directed at recall of exposures or major health events (Balir and Zham 1990; Swan et al 1992; Hruska et al 2000; Tudor-Locke and Myers 2001; Cole et al 2003). However, there have been some utilization studies that provide examples of recall bias and others that provide examples of incompleteness in other data sources.

Clegg et al. (2001) compared self-reported prostate cancer treatments (i.e. data on treatment obtained from patient interviews) with medical records for a few specific treatments. They found that agreement on prostatectomy and radiation therapy was high ($\kappa > 0.8$), but modest for hormone therapy ($\kappa < 0.7$). Another study comparing physician charts and self-report for estimating the use of complementary and alternative medicine, found that such use was poorly documented in medical records (Cohen et al., 2002). Reijneveld (2000) compared survey data with heath insurance registry data for subjects in The Netherlands and found good concordance between the two sources for

hospitalization, physiotherapy, and prescriptions drug use. He also found important ethnic differences, with immigrants having much lower rates of agreement between self report and registry data.

May and Trontell (1998) compared self-reported and Medicare claims data as sources for estimates of mammography use. They found that bias can be introduced by memory telescoping that takes place when respondents misremember the dates of remote events. Burt et al. (2001) found that this effect is influenced both by the relative time of an event's occurrence and the age of respondents. Another study compared pill counts, where study personnel counted the number of pills remaining in medicine bottles, self-report and pharmacy claims data (i.e. prescriptions filled) for medication use in the elderly (Grymonpre et al.. 1998). Self-reported data agreed well with pharmacy claims, but data obtained from pill counts was found to significantly under-estimate compliance—patients reported using the prescribed medications at rates that agreed with pharmacy transaction data but rates measured by pill count were substantially lower. This was attributed to the difficulty of obtaining data using this method. Another study (West et al, 1997) comparing self-report of prescription drug use with pharmacy data found that patient education level, repetitiveness of use and type of drug all affect the probability of accurate recall.

Kvale et al. (1994) compared telephone surveys and medical records for health status assessment and found poor agreement. In contrast, Katz et al. (1996) compared comorbidity scores derived self-reported and medical records data and found high correlations in the results. It is likely that the quality of the data source varies with the type of information one is seeking.

DATA COLLECTION AND METHODS

Data collection for the CCTS involved a set of discrete tasks (Goldman et al, JCO 2001). Each of these tasks was a step in the process of obtaining data on health services utilization for cancer patients, and took place as follows:

1. Site enrollment: health service providers were approached to participate in the CCTS.
2. Subject identification: participating institutions identified eligible subjects.
3. Subject enrollment: patients identified were offered the opportunity to participate in the study.
4. Health services utilization survey: subjects were asked to identify providers and the intensity of services provided.
5. Records abstraction: medical and billing records were collected and abstracted for consenting subjects.

Ultimately, 30 out of 55 sampled sites (accounting for 66% of sampled trial participants), along with 53 affiliated institutions, agreed to participate in the study. Once an institution agreed to participate, staff contacted trial participants and asked permission for CCTS staff to contact them. They also identified cancer patients who were eligible for participation in sampled trials (i.e. patients who met protocol entry criteria) but who were not participating in any research study. These patients were then asked for permission to be contacted. After this, the remaining tasks were performed by of CCTS personnel.

Informed consent was obtained from patients who agreed to be contacted. Those who gave consent participated in a telephone interview and received $25 compensation for their time. Trained interviewers used a computer assisted questionnaire to first identify all hospitals and physicians from whom they had received care since the time of their cancer diagnosis. The interviewers then obtained information on subjects' health services utilization within the six months preceding the interview. The questionnaire also

elicited data on comorbid conditions, health status, prescription drug use, and insurance coverage, along with respondents' satisfaction with and attitudes concerning health care, and various socio-economic status and demographic characteristics. Medicare eligible respondents were asked to provide their Social Security Numbers (SSNs) and to allow CCTS staff to access their Medicare claims data.

After the interviews were completed, patients were sent consent forms to release medical and billing records for each of the providers identified. The CCTS subcontracted with the Phoenix based Health Service Advisory Group (HSAG, www.hsag.com), for the tasks of retrieving and abstracting medical and billing records. CCTS staff worked with HSAG personnel to develop computerized record tracking and abstraction tools and to verify that quality control procedures were in place. Records acquisition involved contacting providers, forwarding consent forms, receiving records, and paying for the cost of copying records.

Once received, medical records were abstracted by trained registered nurses with experience in abstracting medical records. Billing records were abstracted by trained key punch operators with extensive experience working with billing records for health insurance firms. After the training period, first a 10% and later a 5% sample of records were re-abstracted to insure an inter-rater reliability rate of at least 95%. Data entry was accomplished using abstraction tools designed using Microsoft Access (Microsoft Corporation, Redmond, WA).

Data were periodically sent to CCTS staff to check the cleanliness and credibility of the abstracted data. This allowed for the development of programs to produce analytic data files in parallel with the records abstraction process. Database construction and

management tasks were accomplished using Stata statistical applications software (Stata Corp., College Station, TX).

Acquiring Medicare claims data followed a different process. A Data Use Agreement (DUA) was prepared in consultation with the Research Data Assistance Center (ResDAC, www.resdac.org). The DUA was signed by the PI of the CCTS and by a representative of CMS and arrangements were made to purchase the data. After consent forms were obtained from survey respondents to access Medicare data, a file with SSNs was sent to the CMS programming staff. They returned a file of Health Insurance Claim (HIC) numbers to the CCTS and those HIC numbers specifically relating to claims for survey respondents was returned to CMS. This procedure is necessary since multiple beneficiaries can be associated with a single SSN, as when a beneficiary is eligible for Medicare through a spousal or dependent relationship. This file of beneficiaries was then used to query the Medicare claims database. Standard Analytic Files (SAFs) were obtained covering inpatient, hospital outpatient, Part B, home health, hospice, and durable medical equipment claims for 1998, 1999, and 2000.

We next analyze the acquisition effort expended and the relative quality of data obtained from each of the sources described above. Specific aspects of quality include the completeness of the data both in terms of response rates and in terms of coverage for various types of utilization. Another factor to consider is the accuracy and reliability of the data generated from different sources.

FINDINGS

Acquisition Efforts

Table 1 represents an ordinal ranking of the effort needed to complete the tasks required to obtain data via surveys, medical and billing records, and from Medicare claims data.

Table 1. Levels of Effort Associated with Data Sources

	Survey	Medical Records	Billing Records	Medicare Claims Data
Site Enrollment	++	++	++	?
Subject Identification	++	++	++	?
Subject Enrollment	++	++	++	?
Survey Interview	++			
Provider Identification		++	++	
Record Acquisiton		++	++	+
Record Abstraction		++	+++	
Data Entry	+	+++	++	
Programing Time	+	++	+++	+++

+ Minimal Effort; ++ Moderate Effort; +++ Maximum Effort;
? Level of Effort will vary with Study Design

Site enrollment, subject identification, and subject enrollment require comparable levels of effort regardless of the data source, with the possible exception of Medicare claims data. It is possible to obtain data from CMS for patients with specific characteristics (e.g. for specific diseases) without identifying specific individuals. However, to have conducted a study similar to the CCTS, which needed information on patients' participation in research studies, these tasks would have been required to obtain Medicare data as well.

Surveys required on average 40 minutes to complete, but in this case, a substantial fraction of the time involved identifying providers and obtaining contact information. A study solely based on survey data would not require this task, so surveys could be

shortened or more information gathered in the same period of time. A study using on medical and/or billing records would require that providers be identified in order to obtain consent and request records. Survey research does not require the tasks of record acquisition and abstraction.

The task of record acquisition is not notably more difficult for billing records, and in some health systems may be easier, than for medical records; the difference is in the return on the effort, as described below under response rates. The acquisition of Medicare claims data requires far less effort than that needed for provider records. Obtaining provider records involved contacting providers, forwarding consent forms, receiving records, and compensating providers for copying costs (on average $25 per record). In acquiring Medicare claims data one need contact only CMS. The DUA took only about a day to complete, sending the finder file and identifying the appropriate HIC numbers took an additional day of effort. At the time, the cost of three years worth of data cost about $55,000. These costs are relatively unaffected by the number of research subjects; the marginal cost of an additional record is essentially zero.

The abstraction of data from billing records presents more difficulty than for medical records. This difference derives from the relatively standard organization of medical records and the wide diversity in the type and presentation of data found in billing records. Indeed, some records were unintelligible as to the type of services provided. Conversely, data entry was more difficult for medical records. Data items had to be searched for in the medical records, and checked for duplicate entries using worksheets, and then the data were entered into the abstraction tool. Since billing records tend to follow line item formats based on dates of service, the data could be coded

directly using the abstraction tools. Survey responses do require data entry, but this task was simplified using the computer assisted questionnaire developed for the study.

Programming time here refers to the time needed to process the raw data once it has been received. Here again, survey data is relatively easy to work with. The project controlled the data generation process through the survey design, and the computer assisted questionnaire limited the opportunities for miscoded data entries. Similarly, the medical record abstraction tool contained data entry safeguards. The greater programming time primarily reflects the larger number of utilization variables that can be obtained from medical records than from self-reports. Both billing records and Medicare claims data require more time and expertise to process. Utilization measures are associated with procedure codes rather than the counts of specific procedures abstracted from the medical record. The Medicare data in particular, require careful cleaning for some variables, and require extensive knowledge of the data fields and codes. An additional cost of working with Medicare claims is the need for a 3480 or 3490E tape cartridge reader to extract the data.

Data Quality

One aspect of data quality is the completeness of the data in terms of response to requests for information. Table 2 provides the final response rates for living cancer patients who were asked to participate in the CCTS. Deceased patients were included in the study, but were handled differently in different institutions due to the variability in state laws and institutional policies regarding the treatment of medical records for deceased individuals. Of potential subjects who sites attempted to reach and ask if they

could be contacted by the CCTS staff, 62.5% agreed, the remainder either were deceased, refused permission to have their contact information released, or could not be reached. Table 2 also shows for what fraction of subjects' complete and partial records were obtained. Records were considered complete if records were obtained from the physician primarily responsible for cancer care and all inpatient records were obtained. For Medicare claims data, partial data means that the subject was either ineligible for benefits or enrolled in an HMO at some point after the cancer diagnosis date. Note that complete response rates represent a subset of partial response rates.

Table 2. Response Rates (for Living Subjects)

	Survey	Returned Consent	Medical Records	Billing Records	Medicare Claims Data	Medicare Eligibles
Response Rates						
Complete Data	91.5%	86.9%	49.5%	34.2%	18.4%	52.6%
Partial Data			81.1%	73.8%	25.8%	73.8%

Of cancer patients contacted by the CCTS, 921.5% agreed to telephone surveys; of those surveyed, 87% returned consent forms permitting CCTS staff to obtain their medical and billing records. Complete medical records were obtained for 49.5% and complete billing records for 34.2% of survey respondents. Complete Medicare claims data were obtained for 18.4% of survey respondents (52.6% of respondents who reported they were on Medicare). At least some medical and billing records data were received for 93% of subjects who consented. Regarding the Medicare data, of the 35% of subjects who indicated they were covered by Medicare, virtually all agreed to allow the CCTS to access their claims data, but 15% refused to provide SSNs, making it impossible to access their claims. As noted, the partial data represent subjects not continuously eligible for Medicare during the period and also those enrolled in Medicare HMOs.

Another measure of data quality is timeliness. Obviously, the survey response data were the first available. Surveys were conducted from September of 2000 through December of 2001. Since the database management routines were developed in parallel with the data collection, the analytic database was available as soon as the survey period closed. Since subjects were asked to recall actual resource utilization over the six months prior to the interview, this source presented the narrowest window of measurement.

Medical records acquisition took place between December of 2000 and April of 2002. Even though the Medicare claims data request was made as late as possible data, for 2001 was not available for use in the CCTS. The claims data request was forward in January 2002, and the request was filled in the next six to eight weeks. It would have required a delay until at least June of 2002 to acquire the 2001 claims data.

Another quality metric is the presence of data elements that might be of interest in different types of study. Table 3 shows what elements are available by data source.

Table 3. Data Elements Available by Source

	Survey	Medical Records	Billing Records	Medicare Claims Data
Physician Visits	✓	✓	✓	✓
Inpatient Admissions	✓	✓	✓	✓
Home Health Visits	✓	?		✓
Physical Therapy	✓	✓		✓
Prescription Drugs	✓	?		
Diagnostic Procedures		✓	?	✓
Surgical Procedures		✓	?	✓
Alternative Therapy	✓			
Cost Estimates			?	✓

✓ - data available; ? Incomplete or inconsistent availability

Subject to the time restrictions indicated above, all sources yielded information on physician visits and inpatient stays. Data on home health visits were found in the survey, medical records, and claims data, but appeared to be seriously under-estimated in the

medical records. The CCTS did not specifically target home health providers due to budget constraints. Surveys, medical records, and Medicare claims all yielded data on physical therapy. Only surveys yielded data on prescription drug use; medical records would frequently list drugs prescribed at discharge, but these lists were not always present, and provide no information on actual utilization.

Data on diagnostic and surgical procedures were available from medical records and claims data, and were also found in some billing records. Survey respondents only indicated whether hospital admissions involved surgery, not the specific procedures. Only survey responses had data on alternative therapy. Claims data and billing records should both have been sources of cost data. Of the billing records received, only 44% included data on actual payments, the rest list only charges. Therefore only Medicare claims data provided adequate information on the costs of care.

Finally, in Table 4 we are able to compare the accuracy of self-reported health utilization with Medicare claims for 245 respondents who both completed surveys and permitted us to access their Medicare billing records. The same comparison is not possible for the medical records, as the chart abstraction period extended well beyond the period for which Medicare data were available. In retrospect, if utilization data were carefully dated, it would be possible to compare comparable time periods; this would, however, add to the cost of records abstraction. The table shows three main service components that could be found in both survey responses and in Medicare claims: inpatient admissions, physician visits (including hospital outpatient visits) and home health services. The Medicare claims data were restricted to the six month recall periods prior to the interview dates for each respondent.

Subjects tended to over estimate hospital admissions ($p < 0.0001$) and days of inpatient care ($p < 0.007$), under estimated physician visits ($p < 0.0001$), but gave more accurate estimates of home health visits ($p < 0.264$). Medicare claims data are being treated as a gold standard here on the assumption that all covered services will be subject to Medicare claims for eligible persons not enrolled in an HMO. However, a regression of home health visit counts from Medicare claims on self reported counts (below) could not reject the null hypothesis that the parameter estimate on self reported visits was equal to one ($p < 0.913$) or that the intercept was equal to zero ($p < 0.289$).

Table 4. Self-Reported vs. Claims Based Utilization Rates

	(N = 245)				t-Test
	Self Report	Medicare	Difference	Proportion	p(Diff =0)
Inpatient Stays	0.171	0.082	0.090	2.10	0.000
Inpatient Days	0.971	0.465	0.506	2.09	0.007
Physician Visits	4.241	8.531	-4.290	0.50	0.000
Home Health Visits	0.502	0.714	-0.212	0.70	0.264

Regression of CMS Home Visits on Self-Reported Home Visits

```
Number of obs = 245, F(1,243)=122.48,  Prob > F = 0.0000,  R-squared =  0.3351
------------------------------------------------------------------------------
    CMS
 Home Visits |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Survey Home
  Visits     |   1.009323   .0911987    11.07   0.000     .8296819    1.188964
       _cons |   .2075644   .1952521     1.06   0.289    -.1770381     .592167
------------------------------------------------------------------------------
```

DISCUSSION

No single data source is dominant across all measures of effort and quality. Given the need to identify subjects and obtain consent, survey based methods have the highest response rates and the lowest relative collection effort. Self-report is the only one of the methods examined here to provide reliable data on prescription drug use and alternative therapies. The CCTS did not attempt to obtain pharmacy records for respondents due to the additional effort and expense that would have been required. So we cannot say what the likely response rates for these data would have been.

Two types of recall bias, recall loss and telescoping (Kalton & Schuman 1980), may account for some of the discrepancies between self-reported utilization rates and rates derived from Medicare claims data. Telescoping, or erroneously recalling major distant events as having occurred more proximately, has been reported by May and Trontell (1998) for mammography, It has also been noted that telescoping is influenced by the age of subjects (Burt et al., 2001). Recall loss has the opposite effect, as minor events are forgotten. CCTS subjects were asked to limit there responses to the previous six months so as to minimize recall loss.

We found a tendency for survey respondents to report fewer physician visits and more inpatient admissions compared with Medicare claims. For inpatient stays, we can account for the response bias by extending the cutoff period for claims data to nine months prior to the survey date. Although asked to report hospitalizations that took place in the six months prior to the interview, patient responses were more consistent with hospitalizations over a nine month period. When adjustment is made for this, the mean differences between self reported and claims based inpatient care do not significantly

differ from zero, indicating that patients telescope dates of inpatient admissions forward in memory (May and Trontell 1998; Norman et al. 2003; Prohaska et al. 1998; Carey et al. 1995; Thompson and Skowronski 1988).

Compared with surveys, all other data sources were associated with lower response rates in these data. The lowest response rate was for Medicare claims data, because only 35% of survey respondents indicated they were covered by Medicare. However, given Medicare eligibility, data were more complete for claims data than for medical records. Complete medical records data were obtained for 49.5% of all subjects, and at least partial data were available for 81.1%. Medical records data were deficient for prescription drug use and for home health care—records were not sought from home health providers.

If individual subjects need not be identified, and if restricting analysis to data on Medicare eligible individuals is acceptable, then claims data involve less expense than other sources, and can expect to be complete for covered services. Medicare claims data are limited in that prescription drugs are not covered, except for some outpatient chemotherapy drugs. If it is necessary to identify and enroll specific study subjects, claims data require lower acquisition effort and expense than data from medical or billing records. There is a caveat that the costs of acquiring the equipment and expertise to handle claims data are non-trivial. Further, Medicare data do not include outpatient prescription drug use.

Despite the variations in quality among the data sources, one source, provider billing records, was of extremely limited value. Providers were significantly less willing or able to provide billing records than medical records, and the quality of the data

provided were generally poor. Some providers expressed a reluctance to supply any financial data, and some of those who did consent required explicit reassurance that their data would not be used to compare their costs with those of other providers. In addition, providers typically have mechanisms in place to share medical records data, but it is much less common for billing records to be requested. Some institutions provided very good billing records data, and utilization rates between medical and billing records were highly convergent. So studies designed with institutions known to be able and willing to provide high quality billing records have the potential to benefit from these data.

The data developed for the CCTS were collected with a specific purpose in mind, and the data comparisons made here should be applied with some caution to other research designs. The completeness and scope of data obtained from medical records in particular could have been improved by targeting pharmacies and home health providers, and it is always possible to increase the intensity of follow-up in obtaining provider records from those who did not respond. All research efforts are subject to finite budgets, and a determination has to be made as to the costs to be allocated to data collection and the types of data likely to answer the questions being addressed.

It is often necessary to make explicit tradeoffs between data completeness, reliability, and generalizability of the findings. Medicare claims provide rich information on health care utilization and costs, but at the costs of restricting a study to Medicare eligible subjects. Similar advantages in data collection may be obtained for large health systems or insurers with uniform billing systems. The use of administrative data can greatly reduce the costs of data as well.

REFERENCES

Blair A, Zham SH. "Methodological Issues in Exposure Assessment for Case-Control Studies of Cancer and Herbicides." *Am J Industrial Med* 1990;18:285-93.

Burt CDB, Kemp S, Conway M. 'What Happens if You Retest Autobiographical Memory 10 Year On?.' Mem Cognit 2001;29:127-36.

Carey TS, Garrett J, Jackman A, Sanders L, Kalsbeek W. "Reporting of Acute Low Back Pain in a Telephone Interview: Identification of Potential Biases." *Spine* 1995;20:787-90.

Clegg, LX, Potosky AL, Harlan LC, Hankey BF, Hoffman RM, Stanford JL, Hamilton AS. 'Comparison of Self-Reported Initial Treatment with Medical Records: Results from the Prostate Cancer Outcomes Study.' Am J Epi 2001;154:582-7

Cohen RJ, Ek K, Pan CX. 'Complementary and Alternative Medicine (CAM) Use by Older Adults: a Comparison of Self-Report and Physician Chart Documentation.' J Gerontol A Biol Sci Med 2002;57:M223-7.

Cole P, Rodu B, Mathisen A. "Alcohol-Containing Mouthwash and Oropharyngeal Cancer: a Review of the Epidemiology." *J Am Dental Assn* 2003;134:1079-87.

Goldman DP, Schoenbaum ML, Potosky AL, Weeks JC, Berry SH, Escarce JJ, Weidmer B, Kilgore ML, Wagle N, Adams JL, Figlin RA, Lewis JH, Rosen MR, Cohen J, Kaplan R, McCabe M. "Measuring the Incremental Cost of Clinical Cancer Research," *Journal of Clinical Oncology*, 2001;19:105-110.

Goldman DP, Adams JL, Berry SH, Escarce JJ, Kilgore ML, Lewis JH, Rosen MR, Schoenbaum ML, Schonlau M, Wagle N, Weidmer B. "The Cost of Cancer Treatment Study's Design and Methods," RAND Health Science and Technology Policy Institute, 2000.

Grymonpre RE, Didur CD, Montgomery PR, Sitar DS. 'Pill count, Self-Report, and Pharmacy Claims Data to Measure Medication Adherence in the Elderly.' Ann Pharmacother 1998;32:749-54.

Hruska KS, Furth PA, Seifer DB, Sharara FI, Flaws JA. "Environmental Factors in Infertility." *Clin Obstetrics & Gynecol* 2000;43:821-9.

Kalton G, Schuman H, "The Effect of the Question on Survey Responses: A Review," Proceeding of the Survey Research Methods Section, ASA, 1980. http://www.amstat.org/sections/srms/proceedings/, accessed 9/5/2003.

Katz NK, Chang LC, Sangha O, Fossel AH, Bates DW. 'Can Comorbidity be Measured by Questionnaire Rather than Medical Records Review?' Med Care 1996;34:73-84.

Kvale JN, Gillanders WR, Buss TF, Gemmel D, Crenesse A, Griffiths-Marnejon J. 'Agree Between Telephone Survey and Medical Record Data fro the Elderly Patient.' Fam Practice Res J 1994;14:29-39

May DS, Trontell AE. 'Mammography Use by Elderly Women: a Methodological Comparison of Two National Data Sources.' Ann Epidemiol 1998;8:439-44.

Norman SA, Localio AR, Zhou L, Bernstein L, Coates RJ, Flagg EW, et al. "Validation of Self-Reported Screening Mammography Histories among women with and without Breast Cancer." *Am J Epidemiol* 2003;158:264-71

Posner MA, Ash AS, Freund KM, Moskowitz MA, Shwartz M. "Comparing Standard Regression, Propensity Scores Matching, and Instrumental Variables Methods for Determining the Influence of Mammography on Stage of Diagnosis," Health Serv & Outcomes Res Methodol, 2001, 2:279-90.

Prohaska V, Brown NR, Belli RF. "Forward Telescoping: the Question Matters." *Memory* 1998;6:455-65.

Reijneveld SA. "The Cross-Cultural Validity of Self-Reported Use of Health Care: A Comparison of Survey and Registration Data." J Clin Epi 2000;53:267-72.

Swan SH, Shaw GM, Schulman J. "Reporting and Selection Bias in Case-Control Studies of Congenital Malformations. *Epidemiol* 1992;1992;3:356-63.

Thompson CP. Skowronski JJ, Lee DJ. "Telescoping in Dating Naturally Occurring Events." *Memory & Cognition* 1988;16:461-8.

Tudor-Locke CE, Myers AM. "Challenges and Opportunities for Measuring Physical Activity in Sedentary Adults." *Sports Med* 2001;31:91-100.

West SL, Savitz DA, Koch G, Sheff KL, Strom BL, Guess HA, Hartzema AG. 'Demographics, Health Behaviors, and Past Drug Use as Predictors of Recall Accuracy for Previous Prescription Medication Use.' J Clin Epi 1997;50:975-80.

Chapter 4. Pricing Health Services Using SEER-Medicare Linked Data

The most accurate method of assessing health care costs consists of counting utilization measures, such as office visits, hospitalizations, and major procedures, then multiplying counts of the quantity of services delivered by the cost of those services. Ideally, costs should reflect the true value of the inputs used in producing services, but data on the true costs are often unavailable. Providers may be unwilling or unable to share data on their operating costs or reimbursement arrangements. Even when these data are available, they apply to specific institutions or systems and cannot readily be generalized to other settings. So it is necessary in many circumstances to estimate the costs associated with specific health services.

This chapter describes the methods used to estimate treatment costs in the Cost of Cancer Treatment Study (CCTS, Goldman et al, 2001). We first collected data on the components of care (e.g. inpatient days, office visits, tests) and then estimated the unit costs, or prices, for each component. This approach, known as micro-costing, provides the most precise estimate of health care costs for program evaluation (Drummond et al. 2000, pp 67-68). The data collected and, more importantly, the method used to derive prices for utilization measures is described below.

Estimating the costs of health services is a well known problem. Often the only data available are charges. As noted in Chapter 3, very few providers contacted by the CCTS were wiling or able to provide billing records, and the majority of those that did reported only charges, not reimbursements, for services delivered. Using provider charges as proxies for costs is problematic for two reasons (Dranove, 1995). In competitive markets, the price of a good or service can be taken to reflect the marginal cost of production; in a less competitive market, the price charged will be a function of

demand, along with political and regulatory factors. The market for health care services is distorted by numerous factors: third parties rather than consumers bear the greatest share of costs, often providers and payers have market power to set or negotiate prices, and information asymmetries abound among patients, providers, and payers. Moreover, very few providers are paid the amounts charged for services. Medicare, Medicaid, and private insurers negotiate discounts that are both large in magnitude (often 50% or more) and in variances among insures and providers.

Since charges do not reflect the true economic costs of services and often do not reflect the payments made for services, some other measure for the unit costs of services is needed. Charges may be adjusted by cost-to-charge ratios (Williams et al. 1982; Schwartz, Young, and Siegrist, 1995; Bennett et al. 2000). When such ratios are available charges can be modified to approximate average costs. An alternative to charges is the use of cost allocation systems to price services (Williams et al. 1982; Baker 1998). Both charges and accounting costs are, however, idiosyncratic to specific providers. This creates problems for generalizing findings at single institutions and for integrating cost data from different providers in multiple-site studies.

This study proposes methods for assigning prices to utilization measures using data from Medicare billing records. Here a "price" refers to an approximation of the economic costs of delivering health services. At the very least, the prices estimated reflect actual provider reimbursements, reflecting costs from the perspective of payers.

To derive prices we used Medicare claims data for cancer patients. These data include complete information on provider reimbursements, thus reflecting the costs of services to Medicare and its beneficiaries. The Resource Based Relative Value Scale

(RBRVS, Hsiao et al, 1988 & 1992) attempts to capture the intensity of resources used in providing physician services. Similarly, the prospective payment system, based on Diagnostic Related Groups (DRG), attempts to capture the costs of providing inpatient care for specific conditions, with payments periodically adjusted based on mandatory hospital cost reports. Further, prices for both inpatient and outpatient services are adjusted to account for geographic variations in the cost of providing services.

DATA AND METHODS

In this section, we first describe the data on health services utilization we collected and then the SEER-Medicare data that we used to estimate prices for our utilization measures. We then describe the methods (regression models) used to estimate the costs of care associated each of these measures.

Medical Records Abstraction

Copies of medical records were requested from all providers identified by CCTS participants. Upon receipt, they were categorized as either inpatient or outpatient records and duplicative records were culled (e.g. the same record received from more than one provider). Medical records abstraction provided counts for the types of services provided. The abstraction was performed by Registered Nurses using digital abstraction tools designed to facilitate data entry. Separate tools were used for outpatient and inpatient service providers. Inpatient records were abstracted separately for each admission. With the exception of a few relatively inexpensive service components (i.e. common laboratory tests), dates of service were listed to check the accuracy of counts. A five percent random sample of records was re-abstracted by a supervisor as a quality control check. Inter-rater reliabilities were consistently greater than 95%.

Lists of variables abstracted are provided in Table 4.1. Physician visits and consultations were classified by the specialty of the provider. Major surgical procedures were aggregated using the Berenson-Eggers Type of Service (BETOS) coding system (CMS 2002). Diagnostic procedures counted included radiology/nuclear medicine, cardiac, gastro-intestinal, and pulmonary function studies, and laboratory assays. Ancillary services included consultations for physical, occupational, and speech therapy.

Table 4.1 Variables Mapped to Outpatient Medical Record Abstracts

Variable	Label
Physicain Visits	
ervisit	ER
chir_vst	Chiropractic
gast_vst	Gastroenterology
gp_vst	General Practice
gyn_vst	OB/GYN
med_vst	Medical Specialty
np_vst	Nurse Practitioner
onc_vst	Oncology
cosm_vst	Cosmetic Surgery
psy_vst	Psychiatry
rad_vst	Radiology
srg_vst	Surgical Specialty
uro_vst	Urology
opth_vst	Ophthalmology
Surgical Procedures	
mast	Breast
colon	Colon/Rectum
chole	Cholecystectomy
turp	TURP
hyst	Hysterectomy
oth_maj	Other Major Surgery
ortho	Orthopedic
eye	Eye
minor	Minor Procedures
Radiology/Nuclear Medicine	
cxr	Chest X-ray
mammog	Mammograpy
xray	Other X-ray
barium	Barium Contrast
ct_head	Head CT Scan
ct_body	Body CT Scan
mri	MRI
angio	Angiography
bonescan	Bone Scan
nuc_med	Other Nuclear Med
us	Ultrasound
Cardiac Procedures	
cath	Cardiac Catheterization
ptca	Angioplasty
stress	Stress Test
echo	Echocardiogram
ekg	EKG
muga	Multiple Gated Cardiac Equilibrium Studies
cv	Other Cardiovascular
GI Procedures	
coloscop	Colonoscopy
ercp	Endoscopic Retrograde Cholagiopancreatography
egd	Upper GI Endoscopy
paracent	Paracentesis

Variable	Label
Path and Lab Medcine	
abg	Blood gases
chmstry	Chemistry
viro	Virology
hemat	Hematology
micro	Microbiology
cyto	Cytology
bld_bank	Blood Bank
prbc	Packed Red Cells
ffp	Plasma
pltlts	Platelets
skin_bio	Skin Biopsy
Pulmonary Procedures	
spiro	Spirometry
pft	Pulmonary Function Tests
rt	Other Respiratory Therapy
bronch	Bronchoscopy
thoracent	Thoracentesis
ctube	Chest Tube Placement
Ancillary Services	
pt	Physical Therapy
ot	Occupational Therapy
spch_tx	Speech Therapy
Line Placement	
cvp	Central Venous Line
swan	Pulmonary Artery Catheter
aline	Artelial Line
Radiation Therapy	
brachy	Brachytherapy
radtx	Other Radiation Therapy
Other Procedures	
lung_bx	Open Lung Biopsy
bm_bx	Bone Marrow Biopsy
lp	Lumbar Puncture
dialysis	Hemodialysis
chemo	Chemotherapy

The same variables were abstracted from inpatient records, with the exception of physician visits and the addition of length of stay variables.

SEER-Medicare Data

The SEER-MEDICARE linked data for breast, lung, and prostate cancer patients diagnosed from 1991 through 1996 were used to derive prices for health services (Potosky et al. 1993, Warren et al. 2002). The claims used include only individuals enrolled in both Medicare Part A and Part B, and not enrolled in a Medicare HMO. Table 4.2 details how many individuals became Medicare eligible, were first diagnosed, and died in the time frame covered. Table 4.3 shows the distribution of subjects by gender, race/ethnicity, and SEER site. There are SEER sites in every major region of the US, and these sites cover roughly 14% of the total population (SEER 2002).

Table 4.2 Count of Patients by Years of Eligibility, Diagnosis, and Death

	Year									
	<=1989	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
Year Eligible	200,235	14,061	13,449	12,959	11,893	10,660	9,374	7,946	6,873	5,826
Cumulative		214,296	227,745	240,704	252,597	263,257	272,631	280,577	287,450	293,276
Fully Eligible			**174,183**	**178,216**	**174,857**	**167,191**	**156,251**	**142,984**	**127,470**	**132,744**
(Part A & B entire year or until death)										
Year of										
1st Diagnosis	**<=1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**		
Count	17,720	2,910	49,915	52,768	48,458	44,214	42,075	39,264		
Cumulative		20,630	70,545	123,313	171,771	215,985	258,060	297,324		
Deceased		9	7,348	13,808	17,515	20,151	22,309	23,854	17,539	13,392
Cumulative			7,357	21,165	38,680	58,831	81,140	104,994	122,533	135,925

Table 4.3 Distribution of Subjects by Gender, Race, and SEER Region

Gender		**Race**			**State**		
Male	**Female**	**Unknown**	2,185	0.7%	**California**	108,115	36.4%
177,525	119,799	**White**	246,408	82.9%	**Connecticut**	33,759	11.4%
60%	40%	**Black**	27,676	9.3%	**Georgia**	15,319	5.2%
		Other	10,160	3.4%	**Hawaii**	8,665	2.9%
		Asian	6,548	2.2%	**Iowa**	31,386	10.6%
		Hispanic	4,102	1.4%	**Michigan**	44,327	14.9%
		Native Am.	245	0.1%	**New Mexico**	11,835	4.0%
					Utah	11,013	3.7%
					Washington	32,904	11.1%

Detailed distributions for physician and institutional reimbursements are provided in Appendices 4.1 through 4.4.

Price Estimation

Variables in the abstraction forms were mapped to HCFA Common Procedure Coding System (HCPCS) codes for non-institutional (formerly called carrier) provider files and revenue center codes for institutional provider files. Codes were checked using the Medicare Data Dictionary (ResDAC 1999) and the *Physicians' Current Procedural Terminology (CPT), '95* (AMA 1995), and 2001 CPT codes (Wasserman 2002). The mapping of HCPCS and revenue center codes to abstracted procedures is detailed in Appendix 4.6.

Medicare records for institutional providers include line item data on charges, but not payments. It would be feasible to estimate prices using cost to charge ratios, but this approach has drawbacks. First, cost to charge ratios tend to misallocate the cost of resources associated with specific services (Williams et al., 1982). Second, and more important, there remain a large number of service units and costs for items that were not abstracted (e.g. supplies, pharmaceuticals). These costs need to be allocated to services that were counted. This could be accomplished by regressing "other payments" on the vector of abstracted service unit counts. This method suggests an alternate approach to derive prices for utilization methods, using hedonic pricing models.

Hedonic pricing models have long been used to estimate prices and price indexes when goods possess different levels of quality and when quality changes over time (Fisher, Griliches, and Kaysen, 1949; Meullbauer, 1974). This approach has also been applied to pharmaceuticals (Berndt, Cockburn, and Griliches, 1996; Danzon and Chao, 2000) and hospital costs in Israel (Chernichovsky and Zmora, 1986). The data used for inpatient services, comprised all hospital admissions in 1995 and 1996. The total cost (in

1998 dollars) was regressed on the vector of inpatient utilization measures. For outpatient services, patient level data on all services consumed post diagnosis in the years 1994 through 1996 provided the basis for cost estimates. Total outpatient costs (aggregating both institutional and non-institutional provider files) were regressed on outpatient utilization measures along with dummy variables defining the time period for which treatment was observed post diagnosis.

One possible approach would be to use total reimbursements as the dependent variable regressed upon counts of service utilization measures. However, since Medicare reimbursements are based on a prospective payment system tied to diagnosis, this type of model would tend to miss the variances in costs that arise from differing levels of treatment intensity. We therefore used a payment to charge ratio—charges adjusted by the ratio of average total payments to total charges within each Medicare region—as the dependent variable.

Payments were converted to 1998 constant dollars using Medicare time and geographical adjustment factors for Part A and Part B. Because areas covered the SEER registries do not constitute a random or representative sample of the US population or the population of Medicare beneficiaries, failure to account for geographic variation in reimbursement rates could result in biased estimates. Geographic price adjustments for Part A were based on the Medicare Prospective Payment System (PPS) area wage index (Pope and Adamache 1993). These geographic price adjusters were combined with the Medicare PPS Hospital Input Price Index for Part A (DRI/McGraw-Hill HCC, 1995). Geographic adjusters for Part B were based on a study of actual county level differences in procedure level payments (Zuckerman et al. 1991) supplemented by the Medicare

Geographic Adjustment Factor indices for the SEER areas (Federal Register, 1991). These adjusters were extended to the time domain by using the Medicare Economic Index (MEI) Catron and Murphy, 1996). Deductibles, which do not vary geographically, were converted into 1998 dollars using the medical care component of the Consumer Price Index (Bureau of Labor Statistics, 2002).

We also desired to test the hypothesis that the intensity and mix of resource use changes with time since diagnosis, so separate regressions were run for admissions that took place within six months of diagnosis and admissions that took place thereafter. A Chow test was used to determine whether the parameter estimates from the two regressions showed statistically significant differences. Since outpatient services were aggregated for individuals, we accounted for the differences in service intensity over time by using a series of indicator variables for how long following the diagnosis utilization rates were observed.

RESULTS

The vectors of prices for inpatient are presented in Table 4.6. For each regression r-squared values in excess of 0.80 indicate a high goodness of fit, showing that the abstracted utilization measures capture the costs of care very well. Larger costs are associated with major procedures, such $8,125—8,664 for coronary artery bypass (CABG), $4,487—5,448 for angioplasty (PTCA), and $776—1,004 for magnetic resonance imaging (MRI). Some prices were negative, such as -$1,309 and -$1,142 for chest tubes, and -$879 and -$530 for mammography.

The largest fraction of the variance in inpatient costs was explained by length of stay and time spent in the intensive care unit. A separate regression (Appendix 4.5) was

performed using only variables for length of stay, length of stay squared, ICU stay, and whether any surgery was performed, with full interactions, yielded an r-squared of 0.72. The differences in prices between admissions that occurred within six months of diagnosis and after six month were statistically significant ($P < 0.000$).

Outpatient prices, including physician and outpatient institutional services, are detailed in Table 4.7. In this case there is no length of stay, but instead a series of dummy variables for length of the observation period post cancer diagnosis (mo_6—mo_36). The coefficients on these variables indicated that the amount of costs not captured by the other variables in the model increased over time from $637 for individuals observed for 6 months or less up to $1,219 for those with more than 36 or more months of data. Physician office visit costs varied by type of specialty. In the outpatient price vector, few prices were negative and these negative prices were not statistically different from zero with one exception, the cost for an office visit to a gastroenterologist of -$28.34 ($P < 0.028$). As with the inpatient data, the r-squared value was quite high (0.77). Treatments and diagnostic procedures were as important as physician visits in predicting outpatient service costs.

Table 4.6 Inpatient Hedonic Price Vectors

Regressions of Total Costs on Inpatient Utilization Measures

	Admissions within 6 Months of Diagnosis N = 40,182 Adj R-squared = 0.8329				Admissions after 6 Months of Diagnosis N = 91,048 Adj R-squared = 0.8058			
	Coef.	Std. Err.	t	P>t	Coef.	Std. Err.	t	P>t
mast	806.32	74.10	10.88	0.000	835.68	119.64	6.99	0.000
colon	2,355.72	209.01	11.27	0.000	1,901.07	167.63	11.34	0.000
chole	1,450.86	478.50	3.03	0.002	1,843.81	227.02	8.12	0.000
turp	226.72	119.23	1.90	0.057	724.53	102.13	7.09	0.000
hyst	964.05	456.97	2.11	0.035	1,225.93	244.40	5.02	0.000
oth_maj	2,117.23	54.23	39.04	0.000	1,220.78	34.38	35.50	0.000
cabg	8,664.40	386.74	22.40	0.000	8,125.81	154.76	52.50	0.000
ptca	4,487.28	320.03	14.02	0.000	5,448.86	115.03	47.37	0.000
cv	1,281.97	73.77	17.38	0.000	2,753.86	46.76	58.89	0.000
ortho	2,698.42	168.14	16.05	0.000	4,466.90	70.34	63.50	0.000
eye	868.38	686.02	1.27	0.206	980.57	233.35	4.20	0.000
minor	439.53	15.53	28.29	0.000	497.03	11.72	42.41	0.000
nuc_med	-228.14	161.19	-1.42	0.157	-153.54	95.11	-1.61	0.106
spiro	-895.75	230.13	-3.89	0.000	-569.61	166.10	-3.43	0.001
pft	-435.28	218.25	-1.99	0.046	-507.92	161.83	-3.14	0.002
rt	135.87	21.35	6.36	0.000	310.28	12.82	24.20	0.000
dialysis	415.26	39.75	10.45	0.000	619.00	21.72	28.50	0.000
abg	-862.28	279.19	-3.09	0.002	69.65	162.99	0.43	0.669
chmstry	28.28	215.00	0.13	0.895	26.99	141.01	0.19	0.848
mri	1,003.93	90.91	11.04	0.000	776.42	55.56	13.97	0.000
barium	-535.26	205.96	-2.60	0.009	-300.50	107.71	-2.79	0.005
ct_head	129.17	70.48	1.83	0.067	-128.95	37.31	-3.46	0.001
ct_body	86.86	47.60	1.82	0.068	502.12	38.90	12.91	0.000
viro	2,281.28	269.07	8.48	0.000	-310.09	195.60	-1.59	0.113
hemat	257.82	74.02	3.48	0.000	327.19	54.86	5.96	0.000
micro	60.16	630.67	0.10	0.924	-1,148.62	339.23	-3.39	0.001
skin_bio	2,035.34	591.37	3.44	0.001	-817.07	382.43	-2.14	0.033
angio	605.96	105.15	5.76	0.000	295.19	45.89	6.43	0.000
cxr	317.54	10.34	30.70	0.000	484.96	7.86	61.69	0.000
xray	200.82	26.11	7.69	0.000	115.81	14.37	8.06	0.000
bonescan	-129.88	93.18	-1.39	0.163	-204.91	75.44	-2.72	0.007
prbc	1,142.84	961.18	1.19	0.234	6,128.85	371.31	16.51	0.000
mammog	-530.29	219.58	-2.42	0.016	-879.76	180.25	-4.88	0.000
us	77.77	67.05	1.16	0.246	180.42	40.65	4.44	0.000
pt	(dropped)				1,372.48	935.63	1.47	0.142
cath	325.49	140.64	2.31	0.021	73.47	64.72	1.14	0.256
stress	-301.82	137.13	-2.20	0.028	-100.12	56.83	-1.76	0.078
echo	31.93	30.17	1.06	0.290	88.69	15.53	5.71	0.000
ercp	920.68	349.75	2.63	0.008	893.02	138.60	6.44	0.000
egd	434.68	111.23	3.91	0.000	701.59	58.97	11.90	0.000
coloscop	707.55	165.35	4.28	0.000	451.26	83.94	5.38	0.000
paracent	600.95	428.36	1.40	0.161	1,122.06	258.28	4.34	0.000
lung_bx	275.12	98.30	2.80	0.005	314.79	114.61	2.75	0.006
bronch	627.26	74.99	8.36	0.000	576.04	92.24	6.24	0.000
thoracent	-109.86	113.63	-0.97	0.334	-560.91	95.46	-5.88	0.000
ctube	-1,309.14	148.60	-8.81	0.000	-1,142.18	150.14	-7.61	0.000
cvp	1,051.17	94.15	11.16	0.000	1,364.04	69.50	19.63	0.000
swan	2,277.42	182.44	12.48	0.000	2,062.00	112.95	18.26	0.000
lungscan	1,019.16	263.33	3.87	0.000	312.75	159.86	1.96	0.050
muga	777.41	318.51	2.44	0.015	295.15	208.95	1.41	0.158
ekg	2,053.45	1,309.28	1.57	0.117	1,307.26	507.60	2.58	0.010
bm_bx	191.63	24.55	7.81	0.000	280.10	25.36	11.05	0.000
lp	-307.00	365.15	-0.84	0.400	1,169.60	200.86	5.82	0.000
brachy	1,691.84	244.96	6.91	0.000	846.78	269.92	3.14	0.002
radtx	389.32	38.51	10.11	0.000	409.53	44.08	9.29	0.000
chemo	224.38	275.18	0.82	0.415	171.48	258.74	0.66	0.507
los	715.72	7.62	93.87	0.000	792.88	3.55	223.62	0.000
los1	302.91	82.41	3.68	0.000	981.75	59.32	16.55	0.000
los_sq	2.21	0.12	18.47	0.000	-1.25	0.02	-60.85	0.000
los1_icu	-703.56	239.03	-2.94	0.003	-571.45	105.55	-5.41	0.000
icudays	1,167.97	10.03	116.46	0.000	1,037.46	6.32	164.22	0.000
_cons	1,144.29	51.22	22.34	0.000	-11.98	27.68	-0.43	0.665

Table 4.7 Outpatient Hedonic Price Vector

Regression of Total Costs on Outpatient Utilization
N = 60995 Adj R-squared = 0.7732

	Coef.	Std. Err.	t	P>t		Coef.	Std. Err.	t	P>t
ervisit	324.31	9.67	33.52	0.000	prbc	307.34	60.78	5.06	0.000
chir_vst	23.61	22.90	1.03	0.303	ffp	865.71	524.16	1.65	0.099
gast_vst	-28.34	12.92	-2.19	0.028	pltlts	2,937.89	730.42	4.02	0.000
gp_vst	16.43	2.22	7.41	0.000	mammog	125.87	13.47	9.35	0.000
gyn_vst	44.51	19.82	2.25	0.025	us	163.07	11.35	14.36	0.000
med_vst	20.01	2.73	7.32	0.000	pt	43.47	5.52	7.88	0.000
np_vst	-65.44	114.04	-0.57	0.566	ot	1,080.11	54.04	19.99	0.000
onc_vst	81.56	3.55	22.95	0.000	spch_tx	730.69	72.83	10.03	0.000
cosm_vst	18.08	43.47	0.42	0.677	cath	409.01	115.23	3.55	0.000
psy_vst	145.67	19.55	7.45	0.000	stress	99.74	27.23	3.66	0.000
rad_vst	595.73	17.24	34.56	0.000	echo	155.64	12.44	12.51	0.000
srg_vst	-1.77	5.58	-0.32	0.751	ercp	1,329.56	154.33	8.62	0.000
uro_vst	174.29	5.50	31.69	0.000	egd	224.22	29.75	7.54	0.000
opth_vst	39.46	9.88	3.99	0.000	coloscop	403.72	35.70	11.31	0.000
mast	1,641.43	73.42	22.36	0.000	paracent	40.94	123.53	0.33	0.740
colon	-1,520.04	936.92	-1.62	0.105	lung_bx	306.58	64.26	4.77	0.000
chole	637.65	1,008.13	0.63	0.527	bronch	305.48	82.39	3.71	0.000
turp	1,611.47	163.30	9.87	0.000	thoracent	139.64	66.34	2.10	0.035
hyst	-714.91	1,008.24	-0.71	0.478	ctube	422.39	429.16	0.98	0.325
oth_maj	710.14	32.69	21.72	0.000	cvp	2,099.90	175.91	11.94	0.000
ptca	2,017.81	605.81	3.33	0.001	swan	-796.20	689.61	-1.15	0.248
cv	1,588.40	72.73	21.84	0.000	aline	240.57	297.43	0.81	0.419
ortho	1,297.30	130.89	9.91	0.000	muga	793.27	70.61	11.23	0.000
eye	1,392.59	27.95	49.82	0.000	ekg	116.48	9.93	11.72	0.000
minor	103.64	2.21	46.91	0.000	bm_bx	36.67	10.07	3.64	0.000
nuc_med	233.57	10.87	21.49	0.000	lp	532.63	374.27	1.42	0.155
spiro	92.36	14.54	6.35	0.000	brachy	1,989.44	77.74	25.59	0.000
pft	15.35	26.22	0.59	0.558	radtx	345.39	2.33	148.32	0.000
rt	0.69	6.32	0.11	0.913	chemo	334.00	4.88	68.45	0.000
dialysis	2,055.62	20.92	98.24	0.000	mo_6	637.16	101.05	6.31	0.000
abg	38.79	45.21	0.86	0.391	mo_8	861.52	98.02	8.79	0.000
chmstry	32.16	1.67	19.22	0.000	mo_10	824.30	96.75	8.52	0.000
mri	635.75	25.38	25.05	0.000	mo_12	867.17	94.14	9.21	0.000
barium	-22.56	42.95	-0.53	0.599	mo_14	922.51	115.19	8.01	0.000
ct_head	471.25	28.35	16.62	0.000	mo_16	1,049.28	112.99	9.29	0.000
ct_body	696.31	10.58	65.80	0.000	mo_18	1,043.42	112.42	9.28	0.000
viro	14.05	7.22	1.95	0.052	mo_20	1,118.43	106.81	10.47	0.000
hemat	39.98	1.72	23.19	0.000	mo_22	1,098.80	108.85	10.09	0.000
micro	0.57	4.32	0.13	0.896	mo_24	1,202.08	107.99	11.13	0.000
cyto	397.09	13.87	28.62	0.000	mo_26	1,043.31	121.01	8.62	0.000
bld_bank	214.25	14.24	15.05	0.000	mo_28	1,030.42	118.48	8.70	0.000
skin_bio	54.43	27.75	1.96	0.050	mo_30	1,229.24	119.05	10.33	0.000
angio	294.79	48.09	6.13	0.000	mo_32	1,200.08	115.93	10.35	0.000
cxr	11.68	6.76	1.73	0.084	mo_34	1,157.91	113.76	10.18	0.000
xray	76.55	6.36	12.03	0.000	mo_36	1,219.04	112.24	10.86	0.000
bonescan	-217.47	223.42	-0.97	0.330	_cons	368.38	64.04	5.75	0.000

Abbreviations: see Table 4.1; mo_6—Outpt data available for 6 months post diagnosis; mo_8—data available for 6-8 months post diagnosis; ect.

DISCUSSION

We have demonstrated a method for estimating the costs associated with discrete measures of health care utilization. The prices for services developed here are at best proxies for the actual costs of care in the strict economic sense (i.e., the opportunity costs of resources used in delivery of health care services), though it can be argued that these prices are reasonable proxies for costs.

Most of the prices generated by this procedure seem reasonable on inspection. The negative prices could be interpreted as substitution effects—some procedures result in cost savings. Moreover, an approach to pricing that restricted inclusion only to statistically significant values would eliminate most of the prices with negative signs, especially for the outpatient data. In that case, some utilization counts would simply not be considered in cost calculations.

One important finding is the limited extent to which detailed information on the use of diagnostic procedures and therapies add information on the cost of hospital stays. It seems, based on the regression results in Appendix 4.5 that having very limited data on hospital stays—length of stay, type of admission, ICU stay—provide quite adequate predictors of the cost of care. Thus, it may be unnecessary for most purposes to collect detailed data on inputs to inpatient care.

An important limitation to our approach concerns our reliance on Medicare data. These data include significant costs associated with medical education subsidies, especially for inpatient care, along with disproportionate share reimbursements for hospitals providing high levels of indigent care. The extent to which these factors bias cost estimates is unclear. Furthermore, it may be the case that the costs of delivering

services to Medicare beneficiaries differ from the costs of providing the same types of service to the general population. The advantage to this approach remains, however, in that it provides a consistent set of weights for valuing different measures of health service utilization that can be readily generalized nationally or to specific regions of the country.

In future research claims data from large private insurers could be used to determine whether costs of specific services are different for different age groups. If that turns out to be the case, studies could draw on appropriate price vectors according to the population of interest. If the differences were not significant or of negligible magnitude, then the method presented here could suffice for most broadly designed cost studies.

This approach also allows prices to be developed for both very coarse and more detailed measures of service utilization. If all that is known is the number and length of inpatient stays, rough prices can be assigned to these measures that capture the average cost of tests and procedures. When more detailed data on services provided are available, it is possible to produce a vector of prices reflecting the change in overall health care costs associated with a change in each type of service utilization.

REFERENCES

American Medical Association (AMA). *Physicians' Current Procedural Terminology: CPT '95*. Chicago, IL: AMA, 1994.

Baker JJ. *Activity-Based Costing and Activity-Based Management for Health Care*. Gaithersburg, MD: Aspen Publishing, 1998.

Bennet CL, Stinson TJ, Vogel V, Robertson L, Leedy D, O'Brien P, Hobbs J, Sutton T, Ruckdeschel JC, Chirikos TN, Weiner RS, Ramsey MM, Wicha MS 'Evaluating the Financial Impact of Clinical Trials in Oncology: Results from a Pilot Study from the Association of American Cancer Institutes/Northwestern University Clinical Trials Costs and Charges Project.' *J Clin Oncol* 2000;18:2805-10

Berndt ER, Cockburn IM, Grivilches Z. Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs. *Brookings Papers on Economic Activity. Microeconomics*, Vol. 1996. (1996), pp. 133-188.

Bureau of Labor Statistics, http://www.bls.gov/cpi/home.htm, Accessed June 2002.

Catron B, Murphy B. Hospital price inflation: what does the new PPI tell us? *Monthly Labor Review* 1996(July):24-31.

Centers for Medicare and Medicaid Services (CMS). "Berenson-Eggers Type of Service (BETOS) Codes." CMS 2002. http://cms.hhs.gov/data/betos/default.asp, accessed 9/19/03.

Chernichovsky D, Zmora I. A hedonic prices approach to hospitalization costs. The case of Israel. *J Health Econ*. 1986 Jun;5(2):179-91.

Danzon PM, Chao LW. Cross-national price differences for pharmaceuticals: how large, and why? *J Health Econ*. 2000 Mar;19(2):159-95.

DRI/McGraw-Hill HCC, 4th Qtr 1995, released by HCFA, Office of National Health Statistics, Baltimore, Maryland.

Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic evaluation of Health Care Programmes*, 2nd Ed., NY: Oxford University Press, 2000.

Federal Register, pages 59785-59790, November 25, 1991.

Feenstra RC. Exact Hedonic Price Indexes. *The Review of Economics and Statistics*, Vol. 77, No. 4. (Nov., 1995), pp. 634-653.

Fisher FM, Griviliches Z, Kaysen C. The Costs of Automobile Model Changes since 1949. *The Journal of Political Economy*, Vol. 70, No. 5. (Oct., 1962), pp. 433-451.

Goldman DP, Schoenbaum ML, Potosky AL, Weeks JC, Berry SH, Escarce JJ, Weidmer B, Kilgore ML, Wagle N, Adams JL, Figlin RA, Lewis JH, Rosen MR, Cohen J, Kaplan R, McCabe M. "Measuring the Incremental Cost of Clinical Cancer Research," *Journal of Clinical Oncology*, 2001;19:105-110.

Muellbauer J. “Household Production Theory, Quality, and the "Hedonic Technique". *AER*, Vol. 64, No. 6. (Dec., 1974), pp. 977-994.

Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. 'Potential for Cancer Related Health Services Research Using a Linked Medicare-Tumor Registry Database.' *Med Care* 1993;31:732-48

Pope GC, Adamache KW, Center for Health Economics Research. Hospital Wages and the Prospective Payment System Wage Index. Final Report (Cooperative Agreement No. 17-C-99500/1-01). Baltimore, HCFA, 1993.

ResDAC. *SEER-Medicare Data Dictionary* http://www.resdac.umn.edu/ddvh/, accessed June 2002.

Surveillance, Epidemiology, and End Results (SEER) Program, National Cancer Institute. http://seer.cancer.gov/about/. Accessed June 2002.

Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. “Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population.” *Med Care*. 2002 Aug;40(8 Suppl):IV-3-18.

Wasserman Y. *RBRVS-EZ-Fees*™ [computer product]. Medical Publishers Ltd. 2002.

Zuckerman S, Holahan J, Popkin J, Beasley P, Achenbach V. Measuring Prices of Medicare Physician Services: 1985-1988. Computer Product (Cooperative Agreement No. 17-C-99473/3-01). Baltimore: HCFA, 1991.

Appendix: Price Imputations

4.1. Physician Office Visits

Number of Patients	122,522
Mean Visits/Patient	9.15
Number of Visits	1,120,728
Total Allowed Office Charges	$ 214,281,233
Laboratory Charges	$ 10,841,158
Mean Charges per Visit	$ 191
Lab Charges per Visit	$ 10
Allowed Charges per Visit	$ 201

4.2. Home Health Visits

Number of Patients	18,070
RN Visits per Patient	23.15
Total RN Visits	418,395
Total Home Health Charges	$ 93,370,840
Total Home Health Payments (all sources)	$ 70,664,125
Payment/Charge Ratio	0.757
Fully Burdened Charges per RN Visit	$ 223
Estimated Payments per RN Visit	$ 169

4.3. Hospice Care

Number of Patients	4,958
Total Hospice Charges	$ 27,378,300
Total Receipts	$ 25,283,456
Payment/Charge Ratio	0.92
Home Hospice Visits per Patient	46
Total Home Visits	229,424
Home Hospice Charges	$ 23,290,021
Charge per Visit	$ 101
Est. Payment/Visit	$ 94
Inpatient Hospice Day per Patient	2
Total Inpatient Days	9,983
Total Inpatient Charges	$ 4,854,698
Charges per Day	$ 486
Est. Payment/Day	$ 449

4.4. Outpatient Hospital Services

Number of Patients	97,894
Total Charges for Outpatient Services	$ 363,943,160
Total Outpatient Hospital Receipts	$ 189,657,627
Payment/Charge Ratio	0.52
Emergency Room Visits per Patient	.43
Total ER Visits	42,238
ER Share of Charges	$ 28,174,952
Independent Ambulance Allowed Charges Linked to ER Visits	$ 2,783,070
Independent Physician Allowed Charges for ER Services	$ 3,001,804
Outpatient Hospital Charges per ER Visit	$ 667
Est. Payments per Visit	$ 348
Ambulance Charges per Visit	$ 66
Physician Charges per Visit	$ 71
Total Imputed Cost per ER Visit	$ 485
Non-Emergency Outpatient Visits per Patient	5.2
Total Non-ER Visits	512,227
Non-ER Share of Charges	$ 318,934,673
Physician Allowed Charges for Outpatient Hospital Services	$ 31,582,810
Hospital Charges per Visit	$ 623
Est. Payments per Visit	$ 324
Allowed Physician Charges per Visit	$ 62
Total Imputed Cost per Non-ER Visit	$ 386
Aggregate of MD Office & Outpatient Visits	$259
Physical/Occupational Tx Payments per Visit	$ 59

Appendix 4.5 Regressions of Inpatient Cost on Survey Response Variables

Dependent Variable: Charges Adjusted by Payment/Charge Ratios

Source	SS	df	MS	Number of obs = 101603
				F(15,101587) =17377.65
Model	1.1702e+13	15	7.8014e+11	Prob > F = 0.0000
Residual	4.5606e+12	101587	44893137.0	R-squared = 0.7196
				Adj R-squared = 0.7195
Total	1.6263e+13	101602	160061995	Root MSE = 6700.2

inptcost	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
LOS	630.8464	14.76668	42.72	0.000	601.9039	659.7889
LOS = 1	-376.5385	131.9444	-2.85	0.004	-635.1479	-117.9291
LOS-squared	2.836039	.3565107	7.95	0.000	2.137283	3.534795
ICU Flag	-735.621	177.1503	-4.15	0.000	-1082.833	-388.4087
ICU x LOS	602.7564	34.6482	17.40	0.000	534.8463	670.6664
ICU x LOS = 1	737.6531	251.5162	2.93	0.003	244.6845	1230.622
ICU x LOS-squared	-11.70035	.9005228	-12.99	0.000	-13.46536	-9.935334
Surgery Flag	1931.098	97.07128	19.89	0.000	1740.84	2121.357
Surg x LOS	240.7863	16.19729	14.87	0.000	209.0398	272.5328
Surg x LOS = 1	343.5678	170.0454	2.02	0.043	10.28094	676.8546
Surg x LOS-squared	-3.409791	.3607966	-9.45	0.000	-4.116948	-2.702635
Surg x ICU	-1627.989	200.3831	-8.12	0.000	-2020.738	-1235.241
Surg x ICU x LOS	697.8695	36.00668	19.38	0.000	627.2969	768.4422
Surg x ICU x LOS = 1	2575.99	318.0229	8.10	0.000	1952.669	3199.311
Surg x ICU x LOS-squared	8.743419	.9049889	9.66	0.000	6.969652	10.51719
_cons	1462.405	79.88724	18.31	0.000	1305.827	1618.983

4.6. Revenue Center and Procedure Codes Mapped to Service Utilization Measures

		Medicare Codes		
Utilization Measures		HCPCS	Revenue Center	Specialty Code
Physician Visits, by Specialty				
Variable Name	Description			
ervisit	ER	99281-99285	0450-0459	93
chir_vst	Chiropractic	-	-	35
gast_vst	Gastroenterology	99201-99205 99211-99215	-	10
gp_vst	General Practice	(as above)	-	01,08,11
gyn_vst	OB/GYN	(as above)	-	16
med_vst	Medical Specialty	(as above)	-	03,06,13,39,44, 46,98
onc_vst	Oncology	(as above)	-	83,90
cosm_vst	Cosmetic Surgery	(as above)	-	24
psy_vst	Psychiatry	(as above)	-	26,86
rad_vst	Radiology	(as above)	-	31,32
srg_vst	Surgical Specialty	(as above)	-	02,04,14,20,28, 33,77,78,91
uro_vst	Urology	(as above)	-	34
opth_vst	Ophthalmology	(as above)	-	18
np_vst	Nurse Practitioner	(as above)	-	50,97
Ancillary Services				
pt	Physical Therapy	97000-97799	0421	65
ot	Occupational Therapy	-	0431	67
spch_tx	Speech Therapy	92506-92508	0441	-
Cardiac Procedures				
cath	Cardiac Catheterization	93531-93562	0481	
stress	Stress Test	93015-93024	0482	
echo	Echocardiogram	93307-93350	0483	
ekg	EKG	93000-93010	0730-0739	
muga	Multiple Gated Cardiac Equilibrium Studies	78470-78473	-	
Surgical Procedures		(BETOS)*		
mast	Breast	P1A		
colon	Colon/Rectum	P1B		
chole	Cholecystectomy	P1C		
turp	TURP	P1D		
hyst	Hysterectomy	P1E		
oth_maj	Other Major Surgery	P1G		
ptca	Angioplasty	P2D		
cabg	CABG	P2A		
cv	Other Cardiovascular	P2B,P2C,P2E,P2F		
ortho	Orthopedic	P3A-P3D		
eye	Eye	P4A-P4D		
minor	Minor Procedures	P6A-P6D		

Appendix 4.5 (Continued)

Utilization Measures			Medicare Codes		
			HCPCS	Revenue Center	Specialty Code
Radiology/Nuclear Medicine Procedures					
cxr	Chest X-ray		71010-71035	0324	
mammog	Mammograpy		76090-76092	0401	
barium	Barium Contrast		74246-74249 74270-74283	-	
ct_head	Head CT Scan		70450-70498	0351	
ct_body	Body CT Scan		71250-71275 72120-72133 72191-72194 73200-73206 76070-76085 76355-76380	0350,0352, 0359	
mri	MRI		70540-70553 71550-71555 72141-72159 72195-72198 73218-73225 73718-73725 74181-74185 75552-75556 76390-76400	0610-0619	
angio	Angiography		75600-75893	0321	
bonescan	Bone Scan		78300-78320	0341	
nuc_med	Other Nuclear Med		78070-78099 78199,78299, 78399,78499, 78599,78660, 78699,78799, 78807,78890, 78891,78999, 79100-79999	0340, 0342-0349 0974	
us	Ultrasound		76506-76999	0402	
xray	Other X-ray		70010-74775	0320,0329	
		(and not in any above)			
Pulmonary Procedures					
spiro	Spirometry		94010-94016	-	
pft	Pulmonary Function Tests		94160-94200	0460-0469	
rt	Other Respiratory Therapy		94060-94799	0410-0419	
bronch	Bronchoscopy		31620-31626	-	
thoracent	Thoracentesis		32000-32002	-	
ctube	Chest Tube Placement		32020	-	

Appendix 4.5 (Continued)

Utilization Measures			Medicare Codes		
			HCPCS	Revenue Center	Specialty Code
Path and Lab Medcine Assays					
abg	Blood gases		82800-82810	-	
chmstry	Chemistry		80002-80019 82000-84999	0301,0309	
viro	Virology		86000-86800	0302	
hemat	Hematology		85002-85999	0305	
micro	Microbiology		87001-87999	0306	
cyto	Cytology		88230-88299	0310-0319	
bld_bank	Blood Bank		86850-86922	-	
prbc	Packed Red Cells		36430-36431	0381	
ffp	Plasma		36430-36431	0383	
pltlts	Platelets		36430-36431	0384	
skin_bio	Skin Biopsy		11100-11101	0314	
	GI Procedures				
coloscop	Colonoscopy		45330-45385	0750	
ercp	Endoscopic Retrograde Cholagiopancreatography		43260-43269	-	
egd	Upper GI Endoscopy		43200-43272	-	
paracent	Paracentesis		49080-49081	-	
	Line Placement				
cvp	Central Venous Line		36488-36491	-	
swan	Pulmonary Artery Catheter		93503-93503	-	
aline	Arterial Line		36120-36140	-	
Radiation Therapy					
brachy	Brachytherapy		77750-77799	-	
radtx	Other Radiation Therapy		77261-77499	-	
Other Procedures					
lung_bx	Open Lung Biopsy		31625-31629	-	
bm_bx	Bone Marrow Biopsy		85102-88305	-	
lp	Lumbar Puncture		62270-62272		
dialysis	Hemodialysis		90935-90940	0820-0829	
chemo	Chemotherapy		96400-96549	0331,0332, 0335	
*Surgical procedures coded using Berenson-Eggers Type of Service codes.					

Chapter 5. The Effect of Clinical Trial Participation on Prescription Drug Utilization [1]

[1] To be submitted for publication with Dana Goldman as coauthor.

Introduction

The financing of care for patients in the context of clinical trials has been the subject of considerable scrutiny. Several studies have been undertaken to ascertain what effect clinical trial participation has on health services utilization and treatment costs (Wagner *et al.*, 1999; Fireman *et al.*, 2000; Bennet *et al.*, 2000). Each has found a small increase in treatment costs for trial participants. While these studies have shed light on the issue, each represents only one or a few institutions, and their findings are not readily generalizable to the national population of clinical trial participants. Further, these studies did not address the use of outpatient prescription medications (except for chemotherapy), which is a growing concern, both for patients and for Congress considering adding a prescription drug benefit to Medicare.

Obtaining data on the cost of outpatient prescription drug use can be difficult and expensive. For example, the Medical Expenditure Panel Survey provides a comprehensive estimation of total health care costs, including prescription drugs, but had a budget of over $40 million in 2001 (MEPS, 2002). Due to the effort and expense associated with collecting the data many studies of health care costs omit the cost of outpatient drugs altogether, even though these costs could be substantial. Whereas physicians and hospitals are used to sharing information both to guide treatment and assist in research, pharmacies are not. When patients obtain prescriptions from large chains or from discount department stores, it may not even be clear who in the organization would have the authority to release data on pharmaceutical purchases. So even when it is possible to identify prescription drug suppliers for study participants,

adding considerable expense to a study, these efforts are unlikely to result in reliable and complete data on prescription drug use.

An estimate of the impact of clinical trial participation on utilization rates and costs of prescription drugs should be of interest to health policy makers and also to patients and physicians deciding whether or not to join research studies. Many patients bear a greater fraction of the costs for prescription drugs than for other types of health services. Therefore, if higher drug costs are associated with clinical trial participation, that is something patients and their physicians need to know to make informed choices. Third party payers are more likely to be concerned with total treatment costs, especially if prescription drug use is a substitute for other types of health care.

To obtain an accurate assessment of the costs of clinical trial participation, the National Cancer Institute selected RAND to conduct the Costs of Cancer Treatment Study (CCTS) (Goldman *et al.*, 2000). The study enrolled a national probability sample of cancer clinical trial participants, and a matched cohort of cancer patients who did not enroll in any research study, but received treatment in the same institutions and met the protocol entry criteria of the same clinical trials. CCTS participants received an extensive telephone interview regarding their health services and prescription drug utilization, and were asked to allow the study to access medical and billing records from all their health service providers from the time they were diagnosed with cancer.

This paper proposes a new method for estimating the cost of prescription drug consumption that does not require access to pharmacy transaction data linked to research subjects. We then use this method to estimate the impact of participation in cancer treatment trials on prescription drug costs. The remainder of the paper is organized as

follows. The data and methods section describes how CCTS participants were selected, how surveys were conducted to elicit data on prescription drug use, how prescription drug costs were estimated, and how the analysis was conducted. The results section first reports the main findings on the effect of trial participation on prescription drug utilization, costs, and patient out-of-pocket spending.

The two key variables we are interested are drug costs and out-of-pocket expenses. We assume that patients are able to identify the prescription drugs they have used recently and their out-of-pocket expenditures, but that they will usually not be aware of the total costs of drugs, particularly those covered by health insurance or Medicare supplemental policies. We therefore use self-reported out-of-pocket expenditures directly in our analysis. For total drug costs we use self-reported prescription drug use as a basis for estimating drug costs.

DATA AND METHODS

The data sources used in this analysis include data on prescription drug use obtained from surveys of CCTS participants and a data on prescription drug costs obtained from a database of pharmacy transactions. The essential idea was to link data from patients on which prescription drugs they used, to costs derived from averages for a large number of persons using those drugs. This allows cost estimates to incorporate factors such as compliance and differential prices for drugs.

Sampling Methods

The CCTS selected a sample of patients drawn from all Phase III cancer treatment trials conducted by NCI-sponsored Cooperative Groups at all participating institutions in the United States. The sampling design is described at length in Adams *et al.* (2001). Thirty-five cancer treatment trials were selected with probabilities proportionate to their accrual, and then fifty-five institutions were selected with probabilities proportional to their accrual of patients in the selected trials. These institutions included academic health centers, community hospitals and clinics, and physician group practices participating in NCI's Community Clinical Oncology Program. Chapter 3 describes response rates for institutions and individuals approached to participate in the study.

The CCTS enrolled 923 clinical trial participants and another 693 individuals who met the matching criteria for clinical trials, but were not enrolled in research studies. Interviews were completed on 781 clinical trial participants, referred to hereafter as "cases," and 595 non-participants, referred to as "controls" for our purposes. The remaining 142 cases and 98 controls died before they could be interviewed, but did contribute medical and/or billing records. For those individuals, however, data on prescription drug use was unavailable. Tables 5.1 and 5.2 compare the interviewed cases and controls based on health status, demographics, and insurance coverage; Table 5.3 summarizes provider characteristics.

Interviews on Pharmaceutical Utilization

Computer assisted telephone interviews were conducted by trained interviewers in RAND's Survey Research Group. There is evidence that survey respondents tend to

under-report prescription drug utilization and costs (Berk *et al.*, 1990; Grootendorst, 1995). To compensate for this tendency, CCTS participants were asked to describe their utilization only for the six months preceding the interview, and subjects were sent reminder cards prior to being interviewed listing the 86 drugs most frequently used by cancer patients. The interviewer asked, using both the trade and generic drug names, whether the subject used each drug in the preceding six months. The drug list is included in appendix 5.1. Participants were also asked about their out-of-pocket expenditures for prescription drugs and other health care.

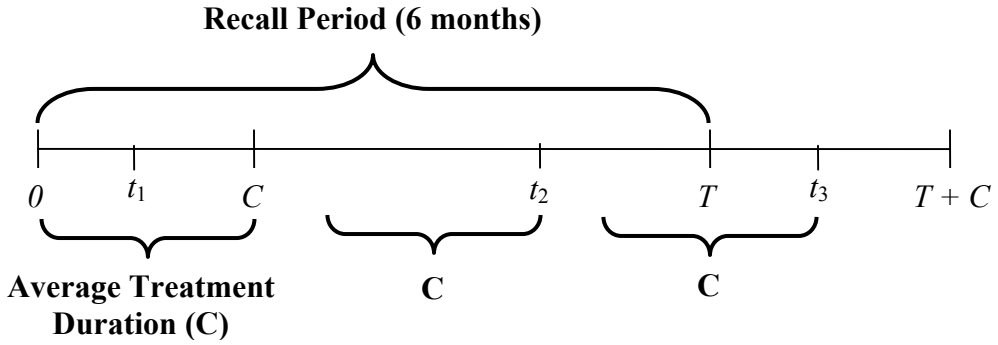
Respondents were asked to report their out-of-pocket expenditures for medications during the six months preceding the interview. Those who unable to provide a precise estimate were asked to bracket their medication expenditures within ranges of 0-100, 100-150, 150-250, 250-500, and greater than 500 dollars. The level of expenditure was then imputed using the average spending for individuals who reported estimates within those ranges.

Costs of Treatment for Survey Respondents

To estimate the expected costs per course of prescription drug treatment we used a national database covering approximately 1.8 million beneficiaries of employer group health insurance plans (Ingenix, New Haven Connecticut). These data include information on pharmacy transactions, including the total amount paid for the prescription, the number of days for which drugs are to be taken, and whether the prescription is a refill. Where cost of treatment estimates were available from both data sources, it is possible to compare those estimates and determine which seems to best

reflect expected utilization and costs. It is also possible to identify drugs that are typically not prescribed, or not taken, according to package insert recommendations.

Applying the typical course of treatment to the survey responses, however, would tend to over-estimate the treatment costs for the six months preceding the interview. The degree of potential bias correlates with the duration of treatment. Consider the timeline below. *C* represents the average duration of a course of treatment for a specific drug. A subject answering yes to a survey question indicates she used the drug within the time frame from zero to *T*; here *T* is the six month recall period. The subject could thus have *concluded* a course of treatment at any point between 0 and $T + C$ and some or all of the treatment course would fall within period *T*.



If an individual concluded a course of treatment at time t_1, between zero and *C*, then t_1 days of treatment would fall within the recall period. Treatment completed between time *C* and *T* (time t_2) would have the entire course of treatment fall within the period, and a treatment course completed at t_3 would be ongoing at the time of the interview and thus fall within the recall period for $T + C - t_3$ days. The expected duration of the treatment, E[*Y*], occurring within the time frame can thus be expressed:

$$E[Y] = \int_0^C t \cdot f(t)dt + \int_C^T C \cdot f(t)dt + \int_T^{T+C} (T + C - t) f(t)dt$$

where $f(t)$ is some probability density function on t, the endpoint of a course of treatment. If we assume a uniform distribution for t the expression becomes:

$$E[Y] = \int_0^C \frac{t}{T+C}\,dt + \int_C^T \frac{C}{T+C}\,dt + \int_T^{T+C} \frac{(T+C-t)}{T+C}\,dt = \frac{C \cdot T}{C+T}.$$

The estimated costs of treatment are then estimated as $\frac{E[Y]}{C}G$, where G is the cost of a full course of treatment. For drugs used to treat chronic conditions, subjects were assumed to be on the drug throughout the six-month period.

Statistical Analysis

Sampling weights for CCTS participants are the reciprocals of their selection probabilities based on the trial and institution pair in which they were recruited. These probabilities were calculated using simulations (Adams *et al.*, 2001).

Cases and controls are not randomly assigned to become trial participants or non-participants; trial participation was the result of choices made by patients and providers, introducing a potential selection bias. Some bias was eliminated by requiring that controls meet the protocol entry criteria in order to be eligible for the CCTS. Nevertheless, there are observable differences between the two groups (Tables 5.1, 5.2 and 5.3), and these differences could affect both trial participation and the utilization of prescription drugs. We addressed this issue with an additional weighting factor derived from propensity scores (Posner et al., 2001; Hirano, Imbens, and Ridder, 2000; Rosenbaum and Rubin, 1983 & 1984), as discussed in Chapter 1. Briefly, propensity scores were derived using logit regression to predict the probability of trial participation.

Weights for controls were calculated as the reciprocal of the probability of trail participation, and for cases as the reciprocal of the probability's complement.

We present descriptive comparisons of the number and types of prescription medications used by cases and controls, along with weighted OLS models of drug costs run to control for covariates. Robust standard errors were computed to account for the clustering of subjects within trials-institution pairs. When then explore the potential effects of interactions between trial participation and type of insurance coverage. This allows us to test the hypothesis that trial participation has differential effects depending on participants insurance coverage. A separate regression is presented with out-of-pocket expenditures as the dependent variable.

Alternative Model Specifications

OLS results are presented for their ease of interpretation. We did, however, explore the results derived from Two Part Models, with and without log-transformation of drug costs, and Generalized Linear Models. There are large numbers of zero-cost observations; 24% of those surveyed reported no prescription drug use during the previous six months. This potential problem was dealt with by using a two-part regression model (Mullahy, 1998; Newhouse, 1994). First, a logit regression was used to estimate the probability of having non-zero drug costs. A second linear regression of costs or log-transformed costs on predictor variables was run conditionally for respondents with non-zero costs. Expected costs become $\Pr(Cost > 0 \mid \mathbf{x}) * \mathrm{E}[Costs \mid Costs > 0; \mathbf{x}]$, the probability of non-zero expenditures times the expected expenditures, conditional on non-zero values and a vector, $\mathbf{x}$, of explanatory

variables. When a log transformation was made to compensate for the skewed distribution of costs for subjects who had costs greater than zero. Expected costs were calculated using a variation the smearing estimate proposed by Duan (1983):

$$E[Costs] = \exp(\log[Cost(X)]) \cdot S$$

where the smearing estimate, $S = \frac{1}{N}\sum_i \exp[e_i]$, where e_i indexes the vector of residuals from the log-transformed regression and N is the number of observations. The variation involves correcting the smearing estimate for heteroscedasticity in the error terms (Mullahy 1998, Manning 1998) such that $S_t = \frac{1}{N_t}\sum_i \exp[e_{it}]$, where t indexes subgroups of the data. Here the subgroups are defined by six percentile partitions in the range of fitted values (0-10%, 10-25%, 25-50%, 50-75%, 75-90%, and 90-100%).

As an alternative to OLS regression using a log transformation, we also used a Generalized Linear Model (GLM) with a log link function. The link function internalizes the log transformation by in effect transforming predictors rather than the dependent variable (Hardin and Hilbe, 2001, p 59). The resulting model specification takes the functional form: $Y = e^{X\beta} + \varepsilon; \quad \varepsilon \square N[0, \sigma^2]$. The log-likelihood function can be expressed:

$$\Lambda = \sum_{i=1}^{n}\left[\frac{y_i \ln(\mathbf{x}_i\beta - \{\ln(x_i\beta\}^2/2}{\sigma^2} - \frac{y_i 2}{2\sigma^2} - \frac{1}{2}\ln\left(2\pi\sigma^2\right)\right].$$

We used the parameter estimates from each of the models to simulate the effect of trial enrollment on prescription drug expenditures. This is accomplished by predicting mean costs when the dummy variable for case is set to one for all observations and comparing this to mean costs when the dummy is set to zero for all observations. In the

simple OLS case, the difference is the same as the parameter estimate for the dummy variable's partial effect. Finally, we repeat the entire analytic procedure using self-reported out-of-pocket expenditures as the dependent variable.

Each of these models has been used in cost estimations, but there is ongoing discussion as to what specification is "best" in a specific instance. Therefore we compared goodness of fits for the models according to a pre-selected set of validation criteria. We chose three criteria defined here:

$$\text{Root Mean Squared Error (RMSE)} \quad = \left\{ \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})^2 \right\}^{1/2}$$

$$\text{Mean Absolute Deviation (MAD)} \quad = \frac{1}{N} \sum_{i=1}^{N} | y_i - \hat{y} |$$

$$\text{Average Prediction Error (APE)} \quad = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})$$

Smaller values for RMSE and MAD indicate greater efficiency of the estimates. Larger absolute values of APE indicate bias, noting APE must equal zero for OLS regression.

RESULTS

Table 5.4 compares utilization rates among respondents, by case/control status, for various types of prescription drugs. For all types of drugs, utilization was higher among cases. The differences were significant ($p < 0.05$) for antibiotics, antidepressants, and anxiolytics, and marginally significant ($p < 0.10$) for erythropoietics and chemotherapy agents.

Table 5.5 provides the weighted least squares regression of prescription drug costs. Trial participation is associated with a $131 increase in drug costs ($p < 0.012$). The strongest predictor of drug costs was self-reported general health status. The category "poor" health was omitted, and "fair", "good", "very good", and "excellent" health responses were associated with decreases in drug costs of $656, $753, $860, and $894, respectively ($p < 0.001$). Weight loss was associated with higher costs, and treatment in an NCI designated cancer center was associated lower costs for prescription drugs. Respondents who indicated a preference for "home remedies" over prescription drugs had lower costs, but those who indicated they did not feel the need for help from medical professionals had higher drug costs.

Table 5.6 shows the effects of interacting trial participation with insurance coverage along with the main effects associated with insurance status; the omitted group includes all those not covered by Medicare or private insurance. None of the interactions of trial participation with insurance status yielded significant differences. Only the main effect of Medicare coverage (without supplemental insurance) was significant. Persons enrolled in Medicare without supplemental coverage had lower drug costs ($p < 0.02$)

independent of trial participation. The full regression results are presented in Appendix 5.6.

Table 5.7 presents results for the regression of out-of-pocket drug expenditures on the same set of predictor variables. In this case a backward stepwise regression was run to retain only variables significant at the $p < 0.10$ level of confidence. The dummy for trial participation was forced into the model, and was not found to differ significantly from zero ($p < 0.84$). As in the previous regression, the strongest effects were associated with health status. Respondents who had Medicare supplemental insurance reported higher out-of-pocket drug expenditures, and did those with breast cancer, diabetes, and hypertension. Complications of diabetes, alcohol abuse, and treatment in teaching hospitals or hospitals in the West or Midwest were associated with lower expenditures.

Alternative Model Designs

Table 5.8 shows the incremental differences in predicted drug costs for clinical trial participation estimated using selected models. In each case the difference shown is derived from a simulation in which the costs predicted if all subjects were enrolled in trials subtracted from the predicted costs if none were enrolled. The full results of each of the regression models are appended. The logistic regression of a dummy variable for non-zero drug costs on the listed predictor variable found that cases were more likely to have non-zero drug costs, but the effect was only marginally significant (odds-ratio: 1.50, 95% CI 1.04-2.17), the weighted regression with log-transformed costs found that, conditional on having non-zero costs, cases had higher prescription drug costs than did controls. The magnitude of the difference in drug costs associated with clinical trial

participation ranges runs from a low of $43 using a Generalized Linear Model to a high of $130 when a weighted OLS model in linear costs is used. Percentage differences range from 45% with the GLM model up to 50% for the two-part model using untransformed costs as the conditional dependent variable.

Table 5.9 provides statistics for comparing goodness of fit among the models. Statistics include RMSE, MAD, and APE for raw weighted means for comparison. No single model structure dominates across all measures of fit. The GLM model produced the lowest RMSE and MAD, but the worst absolute predictive error, indicating a negative bias in the estimator. That is to say, the expected values derived from the model results do not equal the observed mean value of prescription drug costs. The OLS model and the two-part model with log costs as the dependent variable produce the least bias, with OLS yielding a lower RMSE and the TPM a lower MAD. OLS and the TPM with log costs yield nearly identical results in the parameter of interest—cost differences of $130 (47%) and $124 (44%), respectively.

DISCUSSION

The results from a variety of models indicate that participation in cancer treatment trials is associated with higher rates of prescription drug utilization and costs, but that these higher costs do not translate into higher out-of-pocket expenditures for patients. These findings are robust to different model specifications. While the increase in drug costs is significant, the magnitude of the cost difference is small in relation to total cancer treatment costs.

The interaction effects suggest that there is no difference in the effect of trial participation for individuals with different types of insurance coverage. Trial participation did not exhibit differential effects for individuals with different types of insurance coverage, although Medicare beneficiaries without supplemental coverage had lower drug costs, as expected.

We are able to compare alternative models, both in terms of goodness of fit and in terms of how the cost of trial participation is conceptualized. As noted, no model stands out as dominant in measures of goodness of fit. There appears to be a tradeoff between bias and MAD/RMSE in the estimators. OLS estimates the effect of interest as a constant, as opposed to proportional, difference in average drug costs between trial participants and non-participants. This implicitly assumes that the effect of trial participation is a constant, regardless of baseline expenditures. This may be a reasonable assumption for third party payers making decisions about coverage, but may be less informative for researchers or trial participants.

One solution to this would be to estimate log effects (Appendix 5.5); this model suggests that trial participation is associated with a 34% increase in costs; thus the absolute magnitude of the difference varies with the baseline expected costs for trial participants. A limitation of this model is that the log transformation sets zero values to missing, and a substantial number of respondents (24%) reported no prescription drug use during the recall period.

Two-part models allow us to accommodate subjects with zero expenditures. The skewness of non-zero cost observations, and the resulting heteroscedasticity in the regression residuals can be addressed with a log-transformation on prescription drug

costs. This two-part model estimated a $125 or 46% increase in drug costs over a six-month period for clinical trial participants, cases had a higher likelihood of incurring costs and also had higher costs, conditional on non-zero costs. The problem with the two-part model is that, while it is possible to estimate incremental effects, there is no straightforward way to combine the parameter estimates from each part to arrive at the goal of estimating the proportional effect originally sought from the log-transformed model.

The solution here is to estimate the log effect using a Generalized Linear Model, as describe in the methods section. This allows us to obtain an estimate of proportionate changes in drug cots for trial participants without ignoring those subjects with zero drug use. The regression results are presented in Appendix 5.7, and we are unable to reject the hypothesis no proportional effects of trial participation on baseline drug costs.

There are limitations and caveats to consider in evaluating these results. Perhaps the strongest caveat would be that cancer treatment trial participants have already made the decision to pursue aggressive treatment rather than primarily palliative care. Non-participants could have decided either way. This could introduce a bias toward finding higher treatment costs for clinical trial participants compared with others who might follow dissimilar courses of treatment. To the extent that responses to questions about the patients perceived health locus of control, insurance status, and other observed variables impact both the decision to pursue aggressive treatment and trial participation, the use of propensity score weights can serve to mitigate selection bias that may be present. At any rate, the results reported here likely represent at least an upper bound on the effect of trial participation.

From the perspective of third party payers, the increase in drug costs for clinical trial participants may or may not be of concern. If prescription drug utilization substitutes for more costly inpatient or outpatient services, then overall costs could be reduced. If, on the other hand, utilization rates are higher for all types of services, then prescription drugs are simply one more factor in the economic burden of trial participation. From the perspective of potential trials participants, there is no evidence that trial participation imposes an increased burden in costs for prescription drugs.

Table 5.1 Basic Demographic Information, SES, Insurance Coverage

	Cases	Controls	
N	781	595	
Mean Age at Interview	57.9	60.5	***
Married	70%	69%	
Female	76%	77%	
Non-white	11.7%	7.4%	***
Income	$55,692	$62,588	*
Household Wealth	$330,633	$404,997	***
Highest Education			
HS Graduate	27%	28%	
Some College	22%	20%	
College Graduate	40%	42%	
Insurance	(not mutually exclusive)		
Private Insurance	67%	64%	
Medicare	32%	39%	***
Medicaid	5.6%	4.9%	
No Insurance	3.8%	2.5%	
Self-Reported Health Status			
Excellent	17%	20%	
Very Good	35%	35%	
Good	31%	30%	
Fair	13%	10%	
Poor	4%	4%	
Cancer Site			
Breast	46%	52%	***
Colo-Rectal	16%	16%	
Gynecologic	14%	13%	
Hematologic	7%	3%	***
Lung	2%	1%	
Prostate	7%	10%	***
Other	8%	4%	***
Comorbid Conditions			
Myocardial Infarction	4%	4%	
Congestive Heart Failure	2%	2%	
Stroke	5%	4%	
Emplysema	4%	5%	
Ulcer	9%	8%	
Diabetes Mellitis	13%	9%	*
Diabetic Complications	2%	1%	
End Stage Renal Disease	0%	1%	
Impaired Renal Function	2%	2%	
Arthritis	38%	40%	
Liver Cirrhosis	1%	2%	
Other Cancer	9%	13%	**
Hypertension	32%	34%	
Alcohol Abuse	1%	1%	
Phlebitis	2%	2%	
Deep Vein Thrombosis	5%	4%	
Weight Loss	17%	13%	

Difference significant at *p < .10; **p < .05; ***p < .01

Table 5.2 Reponses to Health Locus of Control Questions

Response to Locus of Contol Questions	Cases	Controls
1. I can overcome most illnesses without help from medically trained professionals.		
Strongly Disagree	0.44	0.47
Somewhat Disagree	0.23	0.21
Neutral	0.05	0.04
Somewhat Agree	0.18	0.18
Strongly Agree	0.10	0.10
2. Home remedies are often better than drugs prescribed by a doctor.		
Strongly Disagree	0.45	0.40 *
Somewhat Disagree	0.30	0.29
Neutral	0.05	0.08 **
Somewhat Agree	0.15	0.20 **
Strongly Agree	0.04	0.03
3. If I get sick, it is my own behavior which determines how soon I get well again.		
Strongly Disagree	0.16	0.18
Somewhat Disagree	0.15	0.16
Neutral	0.05	0.06
Somewhat Agree	0.37	0.34
Strongly Agree	0.27	0.26

Difference significant at *p < .10; **p < .05; ***p < .01

Table 5.3 Provider Characteristics

	Cases	Controls	
Type of Facility			
Academic Health Center	0.44	0.40	
Community Clinical Oncology Program	0.44	0.46	
NCI Designated Cancer Center	0.28	0.31	
Region			
Northeast	0.07	0.05	
Midwest	0.56	0.54	
South	0.20	0.12	***
West	0.17	0.28	***
Distance (Miles) from Patient's Home to:			
Nearest Hospital	5	6	
Nearest Teaching Hospital	56	76	***
Nearest Cancer Center	101	98	

Difference significant at *p < .10; **p < .05; ***p < .01

Table 5.4 Average Number of Prescription Drugs Used by Patient Type

	Cases	Controls	
Analgesic	0.591	0.523	
Antibiotic	0.039	0.036	**
Antidepressant	0.243	0.191	**
Antiemetic	0.275	0.201	
Anxiolytic	0.164	0.112	**
Appetite	0.236	0.213	
Chemo	0.066	0.040	*
Erythropoietic	0.291	0.222	*
Hypnotic	0.573	0.513	

Difference significant at *p < .10; **p < .05; ***p < .01

Table 5.5 Weighted Least Squares Regression
Dependent Variable--Prescription Drug Costs

Number of Obeservations = 1282; R-squared = 0.2193

Variable	Coefficient	Robust Standard Error	t	P > |t|
Case	**130.62**	**51.46**	**2.54**	**0.012**
Male	73.94	78.91	0.94	0.350
Married	-2.01	47.71	-0.04	0.966
Age at Diagnosis	-1.71	3.89	-0.44	0.662
Education				
High School	14.79	80.74	0.18	0.855
Some College	86.72	80.11	1.08	0.280
College Graduate	97.17	84.37	1.15	0.251
Comorbidities				
Myocardial Infarction	35.77	121.05	0.30	0.768
Congestive Heart Failure	-182.45	129.63	-1.41	0.161
Stroke	84.99	93.50	0.91	0.364
Emphysema	35.09	71.01	0.49	0.622
Gastric Ulcer	218.15	134.82	1.62	0.107
Diabetes	101.62	105.39	0.96	0.336
Diabetic Complications	163.22	261.99	0.62	0.534
End Stage Renal Disease	-453.74	144.78	-3.13	0.002
Chronic Renal Disease	-70.15	212.36	-0.33	0.741
Arthritis	34.16	54.33	0.63	0.530
Liver Cirrhosis	169.96	226.10	0.75	0.453
Other Cancer	76.77	64.47	1.19	0.235
Hypertension	-75.36	47.28	-1.59	0.112
Alcohol Abuse	-183.54	181.94	-1.01	0.314
Phlebitis	-150.23	145.29	-1.03	0.302
Deep Vein Thrombosis	77.59	110.22	0.70	0.482
Weight Loss	**201.28**	**64.00**	**3.14**	**0.002**
Type of Cancer				
Breast	25.53	87.30	0.29	0.770
Lung	208.17	301.02	0.69	0.490
Gynecological	172.50	111.10	1.55	0.122
Colorectal	-163.33	92.12	-1.77	0.077
Prostate	284.22	153.28	1.85	0.065
Bone Marrow Transplant	267.47	209.16	1.28	0.202
General Health Status (Omitted Value "Poor")				
Excellent	**-894.18**	**179.55**	**-4.98**	**0.000**
Very Good	**-860.40**	**176.88**	**-4.86**	**0.000**
Good	**-753.16**	**173.99**	**-4.33**	**0.000**
Fair	**-655.73**	**191.13**	**-3.43**	**0.001**
Insurance Coverage				
Private Insurance	16.85	99.78	0.17	0.866
Medicare	-135.64	124.45	-1.09	0.277
Medigap Policy	101.78	71.09	1.43	0.154

Table 5.5 (Continued)

Treating Institution				
Academic Health System	235.77	143.21	1.65	0.101
Community Clinical Oncology Program	39.42	50.27	0.78	0.434
NCI Designated Cancer Center	**-369.78**	**149.31**	**-2.48**	**0.014**
South	-144.53	130.14	-1.11	0.268
West	3.55	129.01	0.03	0.978
Midwest	81.50	124.19	0.66	0.512
Distance of Patient Home to Nearest:				
Hospital	-3.07	2.82	-1.09	0.277
Teaching Hospital	0.56	0.44	1.27	0.205
Cancer Center	-0.49	0.42	-1.18	0.240
Does not need help from medical professionals.				
Strongly Disagree	151.56	123.85	1.22	0.222
Somewhat Disagree	108.66	136.54	0.80	0.427
Somewhat Agree	**308.48**	**145.77**	**2.12**	**0.035**
Strongly Agree	235.46	187.19	1.26	0.210
Home remedies are better than prescription drugs.				
Strongly Disagree	-262.41	146.38	-1.79	0.074
Somewhat Disagree	-248.93	156.22	-1.59	0.112
Somewhat Agree	**-341.54**	**155.94**	**-2.19**	**0.029**
Strongly Agree	-245.21	175.02	-1.40	0.163
My own behavior determines how soon I will get well.				
Strongly Disagree	-30.47	151.57	-0.20	0.841
Somewhat Disagree	-169.19	149.65	-1.13	0.259
Somewhat Agree	-151.96	158.11	-0.96	0.337
Strongly Agree	-74.24	158.43	-0.47	0.640
Constant	1,174.51	425.26	2.76	0.006

Table 5.6 Interaction Effects

	Coefficient	Robust Std. Err.	t	P>t
Medicare	**-230.47**	**98.69**	**-2.34**	**0.020**
MC Interaction	153.78	129.22	1.19	0.235
Private Insurance	-36.42	108.58	-0.34	0.738
Private Interaction	70.77	63.80	1.11	0.268
Medigap	28.93	70.22	0.41	0.681
Medigap Interaction	157.82	148.57	1.06	0.289

Table 5.7 Stepwise Regression Results
Dependent Variable: Out-of-Pocket Drug Expenses

	Coef.	Robust Std. Err.	t	P>|t|
Case	-4.48	22.49	-0.20	0.842
Medigap	160.34	33.32	4.81	0.000
Breast Cancer	68.05	23.97	2.84	0.005
Diabetes	167.00	54.43	3.07	0.002
DM Complications	-169.84	100.71	-1.69	0.093
Alcohol Abuse	-99.05	41.69	-2.38	0.018
Hypertension	68.20	25.92	2.63	0.009
Health Status				
Fair	-228.85	129.76	-1.76	0.079
Good	-328.12	128.13	-2.56	0.011
Very Good	-372.70	127.06	-2.93	0.004
Excellent	-373.02	130.31	-2.86	0.005
Teaching Hospital	-48.31	25.58	-1.89	0.060
Midwest	-99.22	41.29	-2.40	0.017
West	-85.20	42.45	-2.01	0.046
_cons	519.15	140.94	3.68	0.000

Table 5.8 Comparing Simulation Results from Different Models

Dependant Variable—Rx Drug Costs

	Expected Costs			
	Cases	Controls	Difference	(%)
Ordinary Least Squares	408	278	130	(47%)**
Two Part Model, Linear Costs	350	234	116	(50%)**
Two Part Model, Log-Transformed Costs	400	276	124	(45%)***
Genralized Linear Model (GLM)	128	86	42	(49%)

Difference significant at *p < .10; **p < .05; ***p < .01

Table 5.8 Goodness of Fit Measures

Model	Predicted Mean	Root Mean Squared Error	Mean Absolute Deviation	Average Prediction Error
Raw Weighted Mean	373	652	354	0
OLS	373	600	349	0
TPM-Untransformed	298	601	316	-75
TPM-Log Transformed	341	619	314	-32
GLM-Log Link	254	578	290	-119

Appendix 5.1. List of Specific Drugs used in Patient Interviews

Pain Medications	Anxiolytics, Sleeping Pills	Antidepressants
Codeine	Ativan	Zoloft
Demerol	Xanax	Paxil
Dilaudid	Valium	Prozac
Darvocet	Librium	Luvox
Darvon	Klonopin	Elavil
Duragesic	Tranxene	Anafranil
Levo-Dromoran	Paxipam	Sinequan
Roxanol (Morphine)	Centrax	Tofranil
MS Contin	Doral	Norpramin
Roxicodone	Halcion	Aventyl/Pamelor
Oxycontin	Dalmane	Effexor
Percodan	Restoril	Wellbutrin
Percocet	Prosom	Serzone
TC #3 or 4	Ativan	Desryel
Vicodin	Ambien	Remeron
Tegretol	Benadryl	
Neurontin		Chemotherapy Agents
Elavil	Heme-Rescue Drugs	Uracil
Tofranil	GCSF/Neupogen	Leucovorin
	GMCSF/Leukine	Tamoxifen
Anti-emetics /	Procrit/Epogen	Premarin
Appetite Stimulants		Megace
Megace	Antibiotics	Depo-Provera
Prednisone	Cipro	Cytoxan
Marinol	Bactrim	Prednisone
Zofran	Diflucan	Bicalutamide
Kytril	Sporanox	Interferon
Anzemet	Mycelex	Interleukin-2
Reglan	Nizoral	Goserelin
Compazine	Mycostatin	
Decadron	Zovirax	
Ativan	Ganciclovir	
Dramamine	Gamciclovir	
Marinol	Valtrex	
Phenergan	Foscavir	
Tigan		
Torecan/Norzine		

Appendix 5.2. Variable Names and Descriptions

Variable	Description
case	Trial Participant
male	Male
married	Married
agedx	Age at Diagnosis
	Highest Education
hs_grad	High School
somecoll2	Some College
college2	College Graduate
	Comorbid Conditions
mi	Myocardial Infarction
chf	Congestive Heart Failure
cva	Stroke
emphys	Emphysema
ulcer	Gastric Ulcer
dm	Diabetes
dm_comp	Diabetic Complications
esrd	End Stage Renal Disease
ren_dis	Chronic Renal Disease
arthrit	Arthritis
cirrhosi	Liver Cirrhosis
oth_ca	Other Cancer
htn	Hypertension
etoh	Alcohol Abuse
phleb	Phlebitis
dvt	Deep Vein Thrombosis
wt_loss	Weight Loss
	Cancer Type
breast	Breast
lung	Lung
gyn	Gynecological
colorect	Colorectal
prostate	Prostate
bmt	Bone Marrow Transplant
	Genral Health Status
gh_excl	Excellent
gh_vgood	Very Good
gh_good	Good
gh_fair	Fair
	Insurance Coverage
pvt_ins	Private Insurance
medicare	Medicare
medigap	Medicare Supplemental Insurance

Appendix 5.2 Continued

Treating Institution

ahc	Academic Health System
ccop	Community Clinical Oncology Program
can_ctr	NCI Designated Cancer Center
south	South
west	West
midwest	Midwest
hospdist	Distance of Patient Home to Nearest Hospital
ahcdist	Distance to Nearest Teaching Hospital
ccdist	Distance to Nearest Cancer Center

Health Locus of Control Responses

I do not need help from medical professionals.

selfcur1	Strongly Disagree
selfcur2	Somewhat Disagree
selfcur4	Somewhat Agree
selfcur5	Strongly Agree

Home remedies are better than prescription drugs.

homecur1	Strongly Disagree
homecur2	Somewhat Disagree
homecur4	Somewhat Agree
homecur5	Strongly Agree

My own behavior determines how soon I will get well.

behave1	Strongly Disagree
behave2	Somewhat Disagree
behave4	Somewhat Agree
behave5	Strongly Agree

Appendix 5.3 Logit Regression—Dependent Variable: Positive Drug Costs

Number of obs = 1282 Log pseudo-likelihood = -589.62528 Pseudo R2 = 0.1994

Indicator: Cost >0	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
case	.4088455	.1880318	2.17	0.030	.0403098	.7773811
male	-.3993893	.276526	-1.44	0.149	-.9413703	.1425917
married	.2022355	.2114884	0.96	0.339	-.2122741	.6167451
agedx	.0040962	.011869	0.35	0.730	-.0191666	.027359
hs_grad	.0273111	.3335547	0.08	0.935	-.6264441	.6810664
somecoll2	.0300555	.3539395	0.08	0.932	-.6636531	.7237642
college2	.5035216	.3479951	1.45	0.148	-.1785362	1.18558
mi	-.0324492	.4755621	-0.07	0.946	-.9645337	.8996354
chf	-.4220771	.7161084	-0.59	0.556	-1.825624	.9814695
cva	.3186803	.4813001	0.66	0.508	-.6246505	1.262011
emphys	.9437012	.6048608	1.56	0.119	-.2418042	2.129207
ulcer	.5839882	.3628887	1.61	0.108	-.1272607	1.295237
dm	-.0895673	.2534929	-0.35	0.724	-.5864042	.4072697
dm_comp	1.052222	.8395159	1.25	0.210	-.5931986	2.697643
esrd	-1.301266	.8891085	-1.46	0.143	-3.043887	.4413542
ren_dis	1.341061	.781156	1.72	0.086	-.1899763	2.872099
arthrit	.1056585	.2110045	0.50	0.617	-.3079028	.5192198
cirrhosi	.2904108	1.152432	0.25	0.801	-1.968314	2.549135
htn	-.0453303	.178496	-0.25	0.800	-.395176	.3045155
etoh	-.3201203	.8727373	-0.37	0.714	-2.030654	1.390413
phleb	.2765307	.585348	0.47	0.637	-.8707302	1.423792
dvt	-.3114874	.4247715	-0.73	0.463	-1.144024	.5210494
wt_loss	.3280878	.2385012	1.38	0.169	-.1393661	.7955416
breast	1.575126	.3520485	4.47	0.000	.8851232	2.265128
lung	-.8578981	.5499182	-1.56	0.119	-1.935718	.2199219
gyn	.4158079	.3942123	1.05	0.292	-.356834	1.18845
colorect	-.8129586	.2872339	-2.83	0.005	-1.375927	-.2499904
prostate	-.1002772	.3831952	-0.26	0.794	-.8513261	.6507716
bmt	-.2964772	.5844966	-0.51	0.612	-1.442069	.849115
gh_excl	-3.112192	.8162784	-3.81	0.000	-4.712068	-1.512315
gh_vgood	-3.171351	.8019005	-3.95	0.000	-4.743047	-1.599655
gh_good	-2.586808	.8136248	-3.18	0.001	-4.181483	-.9921322
gh_fair	-1.905883	.839513	-2.27	0.023	-3.551298	-.2604676
medigap	-.1170453	.4029249	-0.29	0.771	-.9067636	.6726731
pvt_ins	.1069057	.309447	0.35	0.730	-.4995993	.7134108
medicare	.4973619	.4312745	1.15	0.249	-.3479207	1.342644
ahc	-.3006799	.5001112	-0.60	0.548	-1.28088	.6795201
ccop	-.3443668	.2735729	-1.26	0.208	-.8805598	.1918263
can_ctr	.0047849	.3895324	0.01	0.990	-.7586845	.7682544
south	-.7002843	.5760346	-1.22	0.224	-1.829291	.4287228
west	-.458995	.6258959	-0.73	0.463	-1.685728	.7677384
midwest	-.1760137	.5960719	-0.30	0.768	-1.344293	.9922657
hospdist	.014738	.0146863	1.00	0.316	-.0140466	.0435227
ahcdist	-.0011785	.0018284	-0.64	0.519	-.0047622	.0024051
ccdist	.0010377	.0018165	0.57	0.568	-.0025226	.004598
selfcur1	-.314597	.4791182	-0.66	0.511	-1.253651	.6244574
selfcur2	-.3044573	.4939954	-0.62	0.538	-1.272671	.663756
selfcur4	-.451242	.5136647	-0.88	0.380	-1.458006	.5555224
selfcur5	-.1467725	.5041185	-0.29	0.771	-1.134826	.8412815
homecur1	-.0779305	.5089309	-0.15	0.878	-1.075417	.9195557
homecur2	-.3477583	.5187305	-0.67	0.503	-1.364451	.6689348
homecur4	-.2159429	.4892883	-0.44	0.659	-1.17493	.7430445
homecur5	-.3428081	.6951627	-0.49	0.622	-1.705302	1.019686
behave1	.154382	.4551925	0.34	0.734	-.737779	1.046543
behave2	-.0396595	.4760589	-0.08	0.934	-.9727177	.8933987
behave4	.1467811	.4181872	0.35	0.726	-.6728508	.966413
behave5	.2057933	.43599	0.47	0.637	-.6487315	1.060318
_cons	3.340733	1.469154	2.27	0.023	.4612442	6.220222

Appendix 5.4 Weighted OLS Regression: Non-Zero Rx Drug Costs

Number of obs = 978 F(58,209) = 4.79 Prob > F = 0.000 R^2 = 0.2674

rx_cost	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
case	94.47828	59.30883	1.59	0.113	-22.44192	211.3985
male	171.7461	129.0292	1.33	0.185	-82.61935	426.1115
married	2.937533	59.21439	0.05	0.960	-113.7965	119.6716
agedx	-2.532129	4.604725	-0.55	0.583	-11.60979	6.545531
hs_grad	-24.19699	94.10531	-0.26	0.797	-209.7143	161.3203
somecoll2	85.61221	97.32304	0.88	0.380	-106.2484	277.4728
college2	52.00451	102.0579	0.51	0.611	-149.1903	253.1993
mi	61.86958	155.0966	0.40	0.690	-243.8846	367.6237
chf	-235.8003	183.5228	-1.28	0.200	-597.5934	125.9928
cva	92.64041	101.0779	0.92	0.360	-106.6226	291.9034
emphys	-27.42228	83.13186	-0.33	0.742	-191.3067	136.4622
ulcer	219.8331	150.8215	1.46	0.146	-77.4932	517.1595
dm	143.5264	136.9742	1.05	0.296	-126.5018	413.5546
dm_comp	80.91803	273.8606	0.30	0.768	-458.9651	620.8012
esrd	-477.728	341.0708	-1.40	0.163	-1150.108	194.6521
ren_dis	-219.1406	257.3813	-0.85	0.396	-726.5368	288.2555
arthrit	34.61738	60.8467	0.57	0.570	-85.33455	154.5693
cirrhosi	230.1445	234.5296	0.98	0.328	-232.2024	692.4913
oth_ca	98.75975	85.45612	1.16	0.249	-69.70669	267.2262
htn	-90.71012	64.66133	-1.40	0.162	-218.1821	36.7619
etoh	-171.0461	227.946	-0.75	0.454	-620.4141	278.3219
phleb	-240.1661	242.9564	-0.99	0.324	-719.1252	238.7931
dvt	126.4262	134.9799	0.94	0.350	-139.6703	392.5227
wt_loss	**236.2978**	**85.18365**	**2.77**	**0.006**	**68.36855**	**404.2271**
breast	-13.62159	109.3671	-0.12	0.901	-229.2257	201.9825
lung	418.0778	362.4964	1.15	0.250	-296.5402	1132.696
gyn	236.9944	140.3161	1.69	0.093	-39.62181	513.6106
colorect	-133.216	114.8823	-1.16	0.248	-359.6925	93.26055
prostate	489.7609	196.7872	2.49	0.014	101.8187	877.7032
bmt	322.876	216.7015	1.49	0.138	-104.3249	750.077
gh_excl	**-793.3435**	**189.7265**	**-4.18**	**0.000**	**-1167.366**	**-419.3206**
gh_vgood	**-744.791**	**182.7913**	**-4.07**	**0.000**	**-1105.142**	**-384.44**
gh_good	**-644.843**	**180.6102**	**-3.57**	**0.000**	**-1000.894**	**-288.7918**
gh_fair	**-571.255**	**195.0693**	**-2.93**	**0.004**	**-955.8107**	**-186.6993**
medigap	136.3137	97.52222	1.40	0.164	-55.93962	328.567
pvt_ins	49.50546	122.0355	0.41	0.685	-191.0728	290.0837
medicare	-208.5534	158.0012	-1.32	0.188	-520.0338	102.927
ahc	**341.1231**	**162.4488**	**2.10**	**0.037**	**20.87478**	**661.3714**
ccop	88.77669	64.74764	1.37	0.172	-38.86549	216.4189
can_ctr	**-465.1772**	**163.4666**	**-2.85**	**0.005**	**-787.432**	**-142.9225**
south	-77.17344	160.1662	-0.48	0.630	-392.9218	238.5749
west	51.99899	149.959	0.35	0.729	-243.6271	347.6251
midwest	99.12202	142.4327	0.70	0.487	-181.667	379.911
hospdist	-5.187447	3.669984	-1.41	0.159	-12.42238	2.047484
ahcdist	.667286	.5216974	1.28	0.202	-.3611775	1.695749
ccdist	-.5648223	.4856622	-1.16	0.246	-1.522247	.3926022
selfcur1	263.1554	163.6179	1.61	0.109	-59.39768	585.7084
selfcur2	199.5069	178.4907	1.12	0.265	-152.3659	551.3798
selfcur4	**488.5825**	**204.928**	**2.38**	**0.018**	**84.59171**	**892.5732**
selfcur5	343.7125	240.3256	1.43	0.154	-130.0603	817.4854
homecur1	**-284.7264**	**158.2385**	**-1.80**	**0.073**	**-596.6746**	**27.2217**
homecur2	-251.5042	168.9249	-1.49	0.138	-584.5192	81.51083
homecur4	**-433.3251**	**180.0746**	**-2.41**	**0.017**	**-788.3205**	**-78.32973**
homecur5	-264.2663	199.4665	-1.32	0.187	-657.4904	128.9578
behave1	-79.2229	178.898	-0.44	0.658	-431.8988	273.453
behave2	-195.5359	177.0099	-1.10	0.271	-544.4896	153.4178
behave4	-198.3549	185.3738	-1.07	0.286	-563.797	167.0871
behave5	-118.1022	188.9847	-0.62	0.533	-490.6626	254.4583
_cons	1114.228	453.863	2.45	0.015	219.492	2008.964

Appendix 5.5 Weighted Regression of Log-Transformed Rx Costs

Number of obs = 978 F(58,209)= 4.12 Prob > F = 0.000 R^2 = 0.1828

lcost	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
case	**.3045925**	**.1037682**	**2.94**	**0.004**	**.100026**	**.5091591**
male	.2401103	.2301157	1.04	0.298	-.2135351	.6937556
married	-.0663922	.09997	-0.66	0.507	-.2634711	.1306867
agedx	-.0012293	.0066478	-0.18	0.853	-.0143347	.0118761
hs_grad	-.0795146	.1774125	-0.45	0.654	-.4292619	.2702327
somecoll2	.1679097	.1627237	1.03	0.303	-.1528804	.4886999
college2	.0717386	.179455	0.40	0.690	-.2820353	.4255125
mi	.0256816	.3029408	0.08	0.933	-.5715297	.6228929
chf	-.3918068	.398233	-0.98	0.326	-1.176875	.3932615
cva	-.0922053	.2689671	-0.34	0.732	-.6224414	.4380308
emphys	-.070097	.2399093	-0.29	0.770	-.5430493	.4028552
ulcer	.0493595	.1672039	0.30	0.768	-.2802628	.3789817
dm	.1864962	.1637995	1.14	0.256	-.1364148	.5094073
dm_comp	.2753136	.4032057	0.68	0.495	-.5195578	1.070185
esrd	-.2052096	.6329177	-0.32	0.746	-1.45293	1.042511
ren_dis	-.4371537	.3609804	-1.21	0.227	-1.148783	.2744756
arthrit	.1189786	.1056689	1.13	0.261	-.089335	.3272921
cirrhosi	**.8910068**	**.3352335**	**2.66**	**0.008**	**.2301344**	**1.551879**
oth_ca	.1042437	.1367679	0.76	0.447	-.1653777	.373865
htn	-.1729622	.124327	-1.39	0.166	-.418058	.0721336
etoh	**-.8039823**	**.3776279**	**-2.13**	**0.034**	**-1.54843**	**-.0595344**
phleb	-.7140933	.804719	-0.89	0.376	-2.3005	.8723132
dvt	**.5306526**	**.2267906**	**2.34**	**0.020**	**.0835623**	**.977743**
wt_loss	**.3017348**	**.1471858**	**2.05**	**0.042**	**.0115757**	**.5918939**
breast	.1601437	.205186	0.78	0.436	-.2443558	.5646432
lung	-.6232426	.7049347	-0.88	0.378	-2.012936	.7664512
gyn	.0132523	.2609913	0.05	0.960	-.5012607	.5277652
colorect	**-.4806737**	**.2279836**	**-2.11**	**0.036**	**-.9301158**	**-.0312316**
prostate	.2782502	.3952499	0.70	0.482	-.5009374	1.057438
bmt	.2968129	.2320171	1.28	0.202	-.1605808	.7542066
gh_excl	**-1.140647**	**.2047485**	**-5.57**	**0.000**	**-1.544283**	**-.7370096**
gh_vgood	**-1.08898**	**.2022895**	**-5.38**	**0.000**	**-1.48777**	**-.6901907**
gh_good	**-.9917476**	**.195908**	**-5.06**	**0.000**	**-1.377957**	**-.6055386**
gh_fair	**-.9839195**	**.2397832**	**-4.10**	**0.000**	**-1.456623**	**-.5112158**
medigap	-.045796	.1925879	-0.24	0.812	-.4254598	.3338678
pvt_ins	.0629083	.1708146	0.37	0.713	-.2738321	.3996487
medicare	-.158661	.230718	-0.69	0.492	-.6134937	.2961717
ahc	.2072282	.2696219	0.77	0.443	-.3242989	.7387553
ccop	.1204538	.1050985	1.15	0.253	-.0867351	.3276428
can_ctr	-.3190392	.2733971	-1.17	0.245	-.8580087	.2199302
south	.053833	.2402391	0.22	0.823	-.4197694	.5274354
west	-.0880366	.2252887	-0.39	0.696	-.5321662	.356093
midwest	.1013915	.2172995	0.47	0.641	-.3269883	.5297713
hospdist	-.0116404	.0072809	-1.60	0.111	-.0259939	.0027131
ahcdist	.0005578	.0009036	0.62	0.538	-.0012235	.0023392
ccdist	-.0004668	.0008257	-0.57	0.572	-.0020946	.001161
selfcur1	.2603922	.2361464	1.10	0.271	-.2051419	.7259263
selfcur2	.160747	.2578598	0.62	0.534	-.3475925	.6690865
selfcur4	.3575042	.2953058	1.21	0.227	-.2246556	.9396639
selfcur5	.3538693	.2621112	1.35	0.178	-.1628513	.8705899
homecur1	-.2578254	.1935448	-1.33	0.184	-.6393755	.1237248
homecur2	-.2679223	.2067865	-1.30	0.197	-.675577	.1397324
homecur4	**-.3665251**	**.2102756**	**-1.74**	**0.083**	**-.7810581**	**.048008**
homecur5	-.1419502	.2717661	-0.52	0.602	-.6777043	.3938039
behave1	-.0390078	.2613831	-0.15	0.882	-.554293	.4762774
behave2	-.234488	.2451985	-0.96	0.340	-.7178672	.2488913
behave4	-.3107893	.2370843	-1.31	0.191	-.7781723	.1565937
behave5	-.0811819	.2303771	-0.35	0.725	-.5353426	.3729787
_cons	6.432155	.5785207	11.12	0.000	5.291671	7.572639

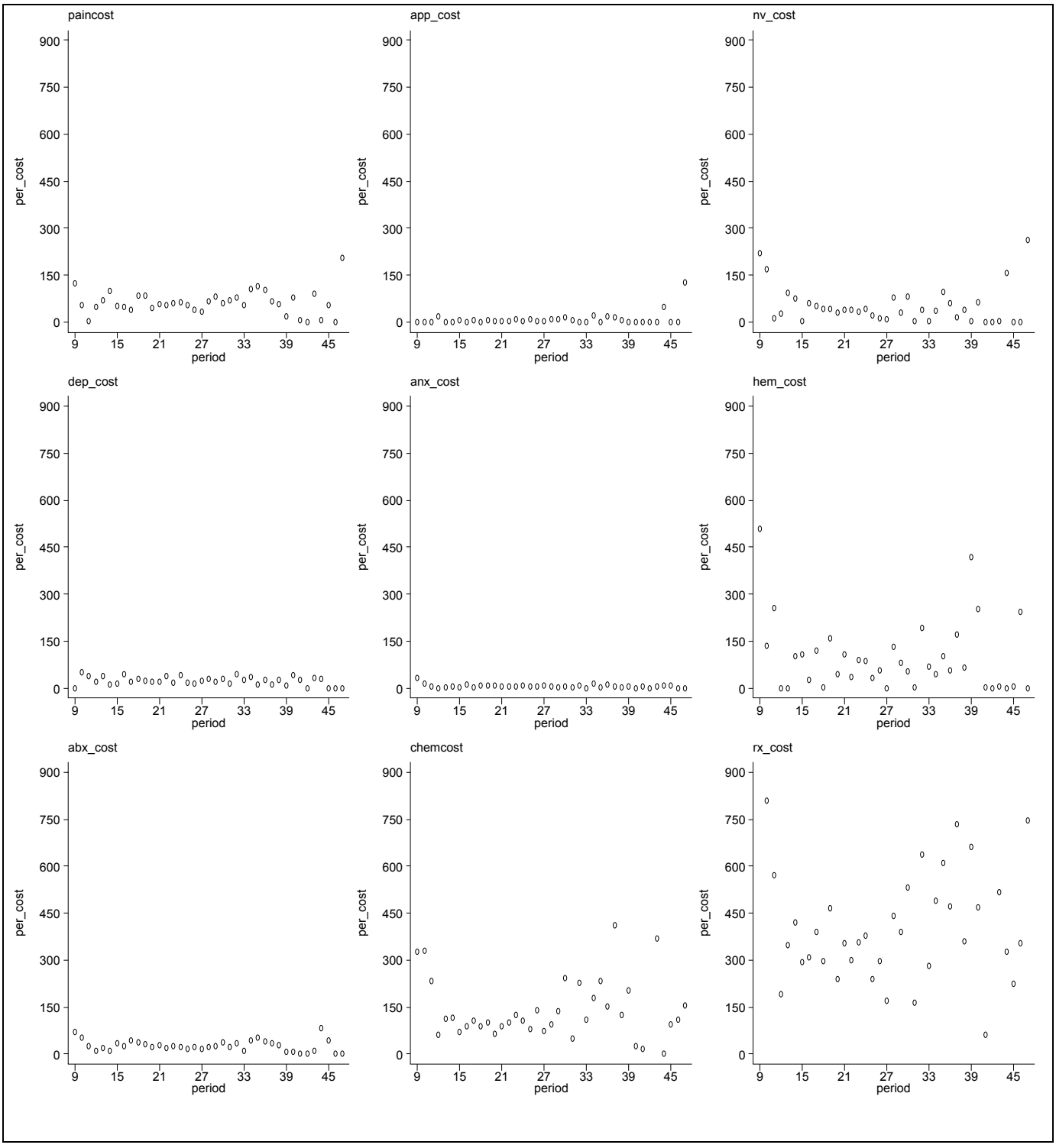
Appendix 5.6 OLS Regression with Insurance/Participant Interaction Terms

rx_cost	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
mc_type	153.7811	129.2176	1.19	0.235	-100.7643	408.3264
pvt_type	70.76548	63.79745	1.11	0.268	-54.90897	196.4399
gap_type	157.8209	148.5699	1.06	0.289	-134.8466	450.4885
male	75.57477	78.92293	0.96	0.339	-79.89531	231.0449
married	-4.312837	47.8592	-0.09	0.928	-98.59056	89.96489
agedx	-1.553062	3.920122	-0.40	0.692	-9.2753	6.169176
hs_grad	22.96773	80.96443	0.28	0.777	-136.5239	182.4594
somecoll2	92.94069	80.0956	1.16	0.247	-64.83944	250.7208
college2	103.4431	84.69271	1.22	0.223	-63.39283	270.2791
mi	31.71867	118.2136	0.27	0.789	-201.1499	264.5873
chf	-178.4925	130.1037	-1.37	0.171	-434.7834	77.7984
cva	83.73266	91.7576	0.91	0.362	-97.02041	264.4857
emphys	35.4359	67.63824	0.52	0.601	-97.8045	168.6763
ulcer	207.969	135.1436	1.54	0.125	-58.25018	474.1881
dm	107.7131	106.2829	1.01	0.312	-101.6534	317.0796
dm_comp	131.4055	263.9067	0.50	0.619	-388.4636	651.2746
esrd	**-515.7849**	**155.634**	**-3.31**	**0.001**	**-822.368**	**-209.2018**
ren_dis	-40.7909	217.6116	-0.19	0.851	-469.4634	387.8816
cirrhosi	194.7555	222.2529	0.88	0.382	-243.0599	632.571
htn	-73.57258	46.46883	-1.58	0.115	-165.1114	17.96624
etoh	-185.2535	179.4029	-1.03	0.303	-538.6588	168.1517
phleb	-160.712	146.853	-1.09	0.275	-449.9973	128.5733
dvt	70.78365	113.9884	0.62	0.535	-153.7618	295.3291
wt_loss	205.847	63.53103	3.24	0.001	80.69742	330.9967
breast	34.52712	88.02241	0.39	0.695	-138.868	207.9223
lung	203.6841	286.9105	0.71	0.478	-361.5001	768.8684
gyn	170.4117	110.2195	1.55	0.123	-46.70944	387.5328
colorect	**-161.5389**	**94.11716**	**-1.72**	**0.087**	**-346.9401**	**23.86225**
prostate	**270.8726**	**150.2275**	**1.80**	**0.073**	**-25.06025**	**566.8054**
bmt	272.3925	206.937	1.32	0.189	-135.2522	680.0372
gh_excl	**-894.9807**	**180.2493**	**-4.97**	**0.000**	**-1250.053**	**-539.908**
gh_vgood	**-863.3482**	**177.8845**	**-4.85**	**0.000**	**-1213.763**	**-512.9339**
gh_good	**-751.3472**	**175.3064**	**-4.29**	**0.000**	**-1096.683**	**-406.0115**
gh_fair	**-652.8845**	**191.0484**	**-3.42**	**0.001**	**-1029.23**	**-276.5387**
medigap	28.93409	70.22412	0.41	0.681	-109.4002	167.2684
pvt_ins	-36.42133	108.5811	-0.34	0.738	-250.315	177.4723
medicare	**-230.4736**	**98.69072**	**-2.34**	**0.020**	**-424.8842**	**-36.06296**
ahc	231.4138	141.9055	1.63	0.104	-48.1254	510.953
ccop	45.59407	49.64957	0.92	0.359	-52.2105	143.3986
can_ctr	**-364.6586**	**147.9402**	**-2.46**	**0.014**	**-656.0856**	**-73.2317**
south	-172.7854	129.5357	-1.33	0.184	-427.9575	82.38658
west	-14.95881	128.9136	-0.12	0.908	-268.9053	238.9877
midwest	54.48798	124.4874	0.44	0.662	-190.7395	299.7154
hospdist	-2.932929	2.843204	-1.03	0.303	-8.533751	2.667892
ahcdist	.5642593	.4352747	1.30	0.196	-.2931873	1.421706
ccdist	-.5012322	.4150184	-1.21	0.228	-1.318776	.3163115
selfcur1	136.8551	125.1835	1.09	0.275	-109.7436	383.4538
selfcur2	93.27071	135.3622	0.69	0.491	-173.3789	359.9203
selfcur4	284.8025	146.2707	1.95	0.053	-3.335874	572.9409
selfcur5	218.2933	184.4848	1.18	0.238	-145.1229	581.7095
homecur1	-256.4908	145.3238	-1.76	0.079	-542.7639	29.7822
homecur2	-236.6626	154.6565	-1.53	0.127	-541.3201	67.99484
homecur4	-324.7382	153.1534	-2.12	0.035	-626.4347	-23.04175
homecur5	-233.7809	174.4194	-1.34	0.181	-577.3693	109.8075
behave1	-33.85547	149.1399	-0.23	0.821	-327.6458	259.9349
behave2	-170.5463	150.4107	-1.13	0.258	-466.8399	125.7473
behave4	-149.6879	157.8819	-0.95	0.344	-460.6991	161.3233
behave5	-77.15477	158.9887	-0.49	0.628	-390.3462	236.0367
_cons	1265.598	414.5373	3.05	0.003	449.0015	2082.194

Appendix 5.7 Results of GLM, Log Link

rx cost	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
case	.0753842	1.742382	0.04	0.965	-3.339622	3.49039
male	-.1682377	1.345195	-0.13	0.900	-2.804771	2.468295
married	-.0942431	.645471	-0.15	0.884	-1.359343	1.170857
agedx	.0067275	.1414484	0.05	0.962	-.2705063	.2839613
hs_grad	-.1508903	10.52024	-0.01	0.989	-20.77018	20.4684
somecoll2	1.112065	5.699827	0.20	0.845	-10.05939	12.28352
college2	1.158995	5.405636	0.21	0.830	-9.435856	11.75385
mi	.4566865	3.167972	0.14	0.885	-5.752424	6.665797
chf	-.2197728	3.435506	-0.06	0.949	-6.95324	6.513694
cva	-.6722801	6.158122	-0.11	0.913	-12.74198	11.39742
emphys	-1.34164	.	.	.	.	.
ulcer	.8058581	.5220794	1.54	0.123	-.2173987	1.829115
dm	.4351477	2.109611	0.21	0.837	-3.699614	4.569909
dm_comp	.3790262	2.565668	0.15	0.883	-4.649591	5.407643
esrd	-4.855368	12.70937	-0.38	0.702	-29.76528	20.05454
ren_dis	.8181729	1.681875	0.49	0.627	-2.478241	4.114587
arthrit	-.2960227	3.486875	-0.08	0.932	-7.130172	6.538126
cirrhosi	-.1783089	4.655902	-0.04	0.969	-9.303709	8.947091
oth_ca	-.7043059	3.227121	-0.22	0.827	-7.029347	5.620736
htn	-.104406	1.570674	-0.07	0.947	-3.18287	2.974058
etoh	1.815404	5.545251	0.33	0.743	-9.053088	12.6839
phleb	-5.290545	6.814252	-0.78	0.438	-18.64623	8.065145
dvt	-.4077595	5.552191	-0.07	0.941	-11.28985	10.47434
wt_loss	.9043255	1.107586	0.82	0.414	-1.266502	3.075153
breast	.0261514	1.588842	0.02	0.987	-3.087921	3.140224
lung	2.239373	3.378349	0.66	0.507	-4.38207	8.860815
gyn	1.14409	5.323234	0.21	0.830	-9.289256	11.57744
colorect	-.1873887	3.53839	-0.05	0.958	-7.122506	6.747729
prostate	2.103928	.8337041	2.52	0.012	.4698977	3.737958
bmt	1.188974	1.686259	0.71	0.481	-2.116032	4.493981
gh_excl	-2.39163	3.327977	-0.72	0.472	-8.914345	4.131086
gh_vgood	-2.572405	3.900834	-0.66	0.510	-10.2179	5.073089
gh_good	-1.571133	2.636865	-0.60	0.551	-6.739294	3.597028
gh_fair	-1.456336	2.15775	-0.67	0.500	-5.685449	2.772777
medigap	-.4249298	3.186355	-0.13	0.894	-6.67007	5.82021
pvt_ins	-.0292476	1.486518	-0.02	0.984	-2.942769	2.884273
medicare	.202684	6.38513	0.03	0.975	-12.31194	12.71731
ahc	1.0105	1.256831	0.80	0.421	-1.452844	3.473844
ccop	.2589444	1.183848	0.22	0.827	-2.061356	2.579244
can_ctr	-1.966134	1.436031	-1.37	0.171	-4.780703	.8484343
south	-2.333441	.	.	.	.	.
west	-.1333904	.7957466	-0.17	0.867	-1.693025	1.426244
midwest	.6346679	1.924795	0.33	0.742	-3.137862	4.407198
hospdist	-.0211474	.1131613	-0.19	0.852	-.2429394	.2006446
ahcdist	.0068408	.0077005	0.89	0.374	-.0082518	.0219335
ccdist	-.0064412	.0079498	-0.81	0.418	-.0220226	.0091403
selfcur1	.6246026	1.610671	0.39	0.698	-2.532255	3.78146
selfcur2	.2798841	3.039295	0.09	0.927	-5.677025	6.236793
selfcur4	1.420632	1.700946	0.84	0.404	-1.91316	4.754424
selfcur5	.7974202	2.038028	0.39	0.696	-3.19704	4.791881
homecur1	-1.21151	1.735552	-0.70	0.485	-4.613131	2.19011
homecur2	-.3009483	.9503505	-0.32	0.751	-2.163601	1.561704
homecur4	-1.518429	4.166982	-0.36	0.716	-9.685563	6.648706
homecur5	-.9512821	5.693163	-0.17	0.867	-12.10968	10.20711
behave1	-.6078091	4.540459	-0.13	0.894	-9.506946	8.291328
behave2	-.5348363	.8464597	-0.63	0.527	-2.193867	1.124194
behave4	-.2697298	1.142491	-0.24	0.813	-2.508972	1.969512
behave5	-.0754344	3.953534	-0.02	0.985	-7.82422	7.673351
_cons	5.597264	11.02438	0.51	0.612	-16.01012	27.20465

Appendix 5.8 Patterns of Drug Costs over Time


paincost
app_cost
nv_cost
dep_cost
anx_cost
hem_cost
abx_cost
chemcost
rx_cost
per_cost
period
0
150
300
450
600
750
900
9
15
21
27
33
39
45

References

Adams JL, Goldman DP, Kilgore ML, Schonlau M, Schoenbaum MS, Escarce JJ. "Multi-Stage, Complex, Non-Nested Samples: Sampling Patients Within and Across Health Care Providers," *Health Serv Res* 2001, *submitted*.

Bennet CL, Stinson TJ, Vogel V, Robertson L, Leedy D, O'Brien P, Hobbs J, Sutton T, Ruckdeschel JC, Chirikos TN, Weiner RS, Ramsey MM, Wicha MS 'Evaluating the Financial Impact of Clinical Trials in Oncology: Results from a Pilot Study from the Association of American Cancer Institutes/Northwestern University Clinical Trials Costs and Charges Project.' J Clin Oncol 2000;18:2805-10.

Berk ML, Schur CL, Mohr P. "Using Survey Data to Estimate Prescription Drug Costs," *Health Affairs* 1990;9:146-56.

Charlson ME, Pompei P, Ales KL, MacKensie CR. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation," *J Chron Dis* 1987;40:373-83.

Fireman BH, Fehrenbacher L, Gruskin EP, Ray GT. "Cost of care for patients in cancer clinical trials," *J Natl Cancer Inst* 2000:92:136-42.

Goldman DP, Schoenbaum ML, Potosky AL, Weeks JC, Berry SH, Escarce JJ, Weidmer B, Kilgore ML, Wagle N, Adams JL, Figlin RA, Lewis JH, Rosen MR, Cohen J, Kaplan R, McCabe M. "Measuring the Incremental Cost of Clinical Cancer Research," *Journal of Clinical Oncology*, 2001;19:105-110.

Goldman DP, Adams JL, Berry SH, Escarce JJ, Kilgore ML, Lewis JH, Rosen MR, Schoenbaum ML, Schonlau M, Wagle N, Weidmer B. "The Cost of Cancer Treatment Study's Design and Methods," RAND Health Science and Technology Policy Institute, 2000.

Grootendorst PV. "A Comparison of Alternative Models of Prescription Drug Utilization," Health Econ 1995 May-Jun;4(3):183-98.

Hirano K, Imbens GW, Ridder G. "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score," NBER Technical Working Paper 251, http://www.nber.org/papers/T0251, accessed 2/3/2002.

Lipscomb J, Ancukiewicz M, Parmagiani G, Hasselblad V, Samsa G, Matchar DB. 'Predicting the Cost of Illness: A comparison of Alternative Models Applied to Stroke.' Med Decis Making 1998;18:S39-S56.

Manning WG. 'The Logged Dependent Variable, Heteroscedasticity, and the Retransformation Problem.' J Health Econ 1998;17:283-95.

Medical Expenditure Panel Survey (MEPS). From: Justification for Budget Estimates for Appropriations Committees, Fiscal Year 2003. February 2002. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/about/cj2003/meps03.htm

Mosby's GenRx 2000, 10th ed. CD-ROM. St. Louis: Harcourt Health Sciences Co.

Mullahy J. "Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics," J Health Econ 1998;17:247-82.

Rosenbaum PR, Rubin DB. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 1983, 70:41-55.

Rosenbaum PR, Rubin DB. "Reducing Bias in Observational Studies using Sub-classification on the Propensity Score," *JASA* 1984, 79:516-24.

Sullivan SD, Liljas B, Buxton M, Lamm CJ, O'Byrne P, Tan WC, Weiss KB; START Steering Committee. "Design and analytic considerations in determining the cost-effectiveness of early intervention in asthma from a multinational clinical trial." Control Clin Trials 2001 Aug;22(4):420-37.

Newhouse JP. *Free for All: Lessons form the RAND Health Insurance Experiment*. 1994, Boston: Harvard Univ Press.

Wagner JL, Alberts SR, Sloan JA, Cha S, Killian J, O'Connell MJ, Grevenhof PV, Lindman J, Chute CG. "Incremental costs of enrolling cancer patients in clinical trials: a population-based study." *J Natl Cancer Inst* 1999;91:847-53.

Hsiao WC, Braun P, Yntema D, Becker ER. Estimating physicians' work for a resource-based relative-value scale. N Engl J Med. 1988 Sep 29;319(13):835-41.

Hsiao WC, Braun P, Dunn D, Becker ER, DeNicola M, Ketcham TR. Results and policy implications of the resource-based relative-value study. N Engl J Med. 1988 Sep 29;319(13):881-8.

Hsiao WC, Braun P, Kelly NL, Becker ER. Results, potential effects, and implementation issues of the Resource-Based Relative Value Scale. JAMA. 1988 Oct 28;260(16):2429-38.

Hsiao WC, Braun P, Dunn D, Becker ER. Resource-based relative values. An overview. JAMA. 1988 Oct 28;260(16):2347-53.

Hsiao WC, Braun P, Becker ER, Dunn DL, Kelly N, Causino N, McCabe MD, Rodriguez E. Results and impacts of the Resource-Based Relative Value Scale. Medical Care 1992 Nov;30(11 Supp):NS61-79.

Centers for Medicare and Medicaid Services. *Healthcare Cost Report Information System (HCRIS) DATASET - HOSPITAL Release 1.0, November 1, 2002.* ***http://cms.hhs.gov/data/download/readmore_sept02_hcris.asp***.

Chapter 6. CONCLUSION

This chapter provides of review of the material covered in the dissertation, along with a discussion of implications for both policy and future research. We first summarize the key theoretical issues addressed, then discuss the significance of findings in the studies of trial participation rates for older cancer patients, the strengths and weaknesses of data sources for health services research, estimating the economic costs of health services, and the effect of trial participation on prescription drug costs.

Theoretical Findings

The two principal issues in theory of concern from the first chapter that are employed at various points subsequently concern representation of trial subjects in relation to generalizability and selection bias arising when non-randomized study designs are used in research. The policy implications relate to the interpretation of research results assessing the extent to which the findings of a specific study may be informative for decisions in different contexts. How this plays out in practice very much depends on the context of the question one is interested in and the quantity and quality of information available to inform decision making. The more interesting general points arising from this project concern the issue of representativeness for the design of research.

Taking the simplest possible case as an illustration, an appropriately applied t-Test for differences in means, even substantial differences in treatment effectiveness between subgroups in a research study would not be detectable without multiplying the sample size several times (incurring proportionately higher study costs). This was the case even if only one sub-population of interest were involved. Further stratification, for example by gender and race or ethnicity, would compound the problem exponentially. This calls into question an insistence on proportional representation of specific subgroups

in the design of clinical trials, particularly for groups that make up relatively small fractions of the general population. This does not imply that the inclusion of specific populations in clinical trials is undesirable, but rather that simple "representativeness" (i.e. proportional representation) is unlikely to provide usable data on outcomes for minority populations. Instead, where prior evidence indicates that there may be substantial differences in treatment effects for specific groups, trials need to be designed to focus on them and not on the general population. A current example of this problem has arisen with respect to the use of selective serotonin reuptake inhibitors (SSRIs) for the treatment of depression in children. The literature on the subject yields mixed results (Mitka 2003; Olfson et al. 2003; Wagner et al. 2003). Initial evaluations of SSRIs included only adults, but pediatric psychiatrists have subsequently used then in treating children and adolescents. Anecdotal evidence and at least one large observational study (Olfson 2003) suggest that SSRIs may pose an increased risk of suicide in children.

This example illustrates a number of theoretical issues related to the design of trials. Including a small number of children in the original studies would not have identified the problem. Indeed, two studies focused on children failed to detect any increased risk of suicide. The problem is that suicide attempts in children are extremely rare events. The question remains open, although there is evidence that children taking SSRIs have higher suicide rates than those who do not, it has not been possible to establish a causal relationship—does the effect arise from the drugs or from the disease the drugs are supposed to treat? Which leads to the question of drawing inference from studies other than randomized controlled trials (RCTs).

In contrast to RCTs, observational studies lack control over assignment to treatment or exposure. In the Cost of Cancer Treatment Study is an example; trial participants were compared to other cancer patients who were not participants, and no randomized design was feasible. The lack of random assignment can produce biased estimates of treatment effects. Three basic modeling approaches have been used to address these problems are difference-of-differences (DoD, including fixed effects models), propensity scores, and instrumental variables (IV). DoD methods generally involve panel data, with repeated observations of the same units over time, and have not been explored here. IV models produce results that can be considered comparable to those obtained from randomized studies, but depend on the availability of valid instruments—variables that effect outcomes only through their influence on intermediate variables of interest. Propensity scores, by contrast, have a lesser ability to overcome problems related to confounding effects, but can be implemented wherever sufficient rich covariates are available.

It is interesting to contrast the clinical and economic literatures on IV and propensity score models. A MEDLINE (US Library of Medicine 2004) search of the clinical literature since 1990 yielded 587 citations referring to propensity scores but only 81 citations for IV. Furthermore, the overwhelming majority of citations referenced statistical or methods oriented publications and only one paper was published in a major general interest clinical journal. Most citations referencing propensity scores were published in general or subspecialty clinical journals. A search of the JSTOR® database (Journal Storage, Inc. 2004) for citations in economic journals yielded 1194 citations for IV and only 11 for propensity scores. Although IV models represent a substantial

improvement in the validity of inferences drawn from non-randomized analyses, it would appear that the rarity of valid instruments presents a barrier to their use. Propensity scores, in contrast, may provide a less powerful but more practical set of tools, and this was the approach taken in the Cost of Cancer Treatment Study and in the investigation of prescription drug costs presented in Chapter 5.

The theoretical issues discussed above, however, did not constitute the central subject of this dissertation, but were rather pursued to clarify issues relevant to the analysis of clinical trial design and evaluation. The remaining sections summarize the key findings of this investigation and some of their policy implications.

Older Patients in Clinical Trials

As noted in Chapter 2, numerous studies have noted the lack of trial participation among older adults in comparison with the incidence of cancer for different age groups. Contrary to the discussion of representativeness for relatively small minorities within the population, individuals 65 or older represent the majority of adult cancer patients. Studies that fail to include older adults effectively exclude the apparent population of interest in assessing cancer treatment, and there has been considerable speculation about barriers to entry into trials for older adults.

Two of our principal findings bear directly on these questions. The first is that, when we examine a census of NCI-sponsored clinical trials, the degree of under-representation for older adults is less than previously reported. We found that 32% of adult trial participants were 65 or older, in comparison with proportions of 25% or less reported elsewhere (Hutchins et al. 1999). However, 32% is still considerably lower than the proportion (61%) of newly diagnosed cancer patients who are 65 or older. Our second

and more crucial finding is that it is possible to account for the disparity between cancer incidence and trial participation for older patients by taking protocol eligibility criteria into account. It is apparent that age in itself is not the issue, but rather health status—trials are often restricted to relatively healthy individuals, and may well include healthy older adults in proportion to their numbers in the population of cancer patients.

A primary policy implication of these findings is that research designs should be careful to avoid arbitrary exclusion criteria. If treatments are expected to be harmful to persons with specific comorbid conditions then exclusion criteria are obligatory. Arbitrary exclusion criteria, on the other hand, can impose serious limitations on the generalizability of trial results, so it is incumbent upon investigators and reviewers to insure the trial designs are appropriate with regard to the potential risks and benefits of specific experimental treatments.

In a more technical vein, in modeling the effects of trial design on participation rates, we had to consider the appropriate statistical methods for use with rates and proportions, where the range of possible values is restricted to between 1 and 0, inclusive. In this instance, the ordinary least squares model yielded the same results as did the "better" generalized linear model. This is likely due to the fact that the parameters of interest all attached to binary variables for the presence of protocol exclusion criteria and thus concerned simple differences in means. It is generally advisable to adjust modeling approaches to conform to the nature of the data being analyzed, and practical tools are now widely available to do so (Fleiss, Levin an Paik, 2003).

Data Sources for Health Services Research

In an effort to achieve the clearest possible picture of the effects of clinical trial participation on treatment costs the CCTS collected data from patient interviews, medical records abstraction, provider billing records, and Medicare claims. This design provided an opportunity to compare a variety of data sources for use in health services research and health economics. The results of these comparison have implications for the design of future studies.

The most striking finding is that great care should be taken before implementing a research design intended to use provider billing records as a primary data source. In the CCTS we found that relatively few providers, whether individual physicians, practice groups, or institutions, were willing to provide any financial data at all and that most of the data provided listed only charges, not actual reimbursements. At the same time a few providers, particularly those in closely integrated health systems, provided quite detailed billing records including detailed data on services and procedures, charges, and payments from various sources. An earlier study of cancer treatment costs in the context of clinical trials within the Northern California Kaiser Permanente health system (Fireman et al. 2000). Where such data is known to be available, it is quite useful and may be easily obtained. As a general rule, however, attempts to obtain billing records may be prohibitively expensive and/or produce data of dubious quality.

In contrast to provider billing records, Medicare claims can provide a valuable source of data on health services utilization and costs. Medicare records contain data on all covered services, including provider charges, cost-to-charge ratios (for institutional providers), and reimbursements from Medicare and from beneficiaries. The costs of obtaining these data are less than from other sources of comparable quality and the

marginal costs are negligible—adding individual beneficiaries does not affect the costs of obtaining the data. The primary limitation of the Medicare data is obvious—Medicare for the most part covers only people 65 or older or people with kidney disease. Further, Medicare claims data are missing for individuals enrolled in managed care plan. Finally, Medicare has not, with few exceptions, covered outpatient drugs, which make up a substantial fraction of health care costs.

One class of providers the CCTS did not pursue were pharmacists, instead we obtained data on prescription drug utilization and expenditures from surveys. While there are acknowledged problems with the reliability of self-reported utilization data, there were steps taken to mitigate response bias, and more to the point, no better option was truly available. Previous experience had shown that attempting to obtain data from pharmacists and retailers on prescription drugs is prohibitively expensive and subject to considerable non-response rates. And while medical records do contain data on prescription drugs, these data generally constitute second hand self-reports from patients and may not include information on compliance. Thus, unless research is focused on groups of subjects all participating in centrally administered plans that cover prescription drugs, survey responses may be the best source for drug data.

Medical records abstraction has a long history in health services research. Expertise in collecting and abstracting records is readily available and quality control methods have been developed to ensure the integrity of abstracted data. For most types of health services utilization, especially when claims data are unavailable or unreliable, medical records provides accessible data rich in details of what services and procedures were used to treat study subjects. The key problems with medical records involve the

expense of collecting and abstracting data and the procedures required to safeguard the confidentiality of the data.

To summarize, it is necessary to consider what types of data are needed and what sources are likely to provide the data best suited to the specific aims of particular studies. No single source dominates the others. As a general rule, large administrative databases, such as Medicare claims data or records from other health systems, provide a convenient and economical source for data on covered services. The utility of such databases, though, is limited by the types of services covered and the individuals included in the health plan.

Pricing Health Services

We provide an example for deriving "prices" for health services using hedonic regressions. A few points in the model design are worth emphasizing. First, the use of Medicare reimbursements presents a reasonable proxy for actual costs. These are the costs from the CMS perspective, and the various payment scales are attempts to relate payment levels for services to the actual costs of providing them. Second, adjustment factors are available to smooth out differences in costs across different geographic regions at different points in time, allowing the prices derive to reflect constant dollar costs. Finally, the costs derived for cancer services were obtained using a large sample of cancer patients, so the impact of utilization measures on costs can be expected to reflect the specific population under investigation.

Hedonic regressions allowed us to apply prices to utilization that reflected their impact on total costs, not limited to the cost of inputs for those specific services. This

allows the prices for measured health services to reflect the cost of materials and supplies or ancillary services that it was impractical to measure directly.

The main findings were the price vectors used in subsequent analyses. However, one general finding my have broader implications. We found that obtaining data on lengths of stay, types of admissions, and intensive care use provided almost as much information for pricing inpatient services as did very detailed inventories of tests and procedures performed during admissions. This means that the expense of detailed medical records abstraction may not be necessary for many studies concerned with inpatient care costs.

Trial Participation and Prescription Drug Use

In the examination of the effects of trial participation on prescription drug use and costs, several modeling issues had to be addressed. First, as noted earlier, CCTS subjects were not randomly assigned to participate in trials or to refrain from participating; they chose, presumably in consultation with their physicians, whether or not to enroll. It is likely that there could be considerable differences between the two groups that influenced their decisions. The CCTS design sought to reduce potential selection bias in three ways:

1) Controls for the study received cancer treatment from the same providers as cases.

2) Controls had to meet the relevant protocol entry criteria as did cases, and thus had similar disease characteristics and health profiles.

3) Propensity score weights were used to adjust for observed differences between the two groups.

These measures may not have completely addressed all possible biases, but did insure that the comparison group was selected ad weighted to resemble the group of trial participants as closely as possible. Finally, the most important likely differences between

cases and controls (in terms of treatment costs) was thought to concern their attitudes toward cancer treatment. If some controls chose not to participate in trials because they had decided not to pursue aggressive cancer treatment, that would have obvious implications for differing costs of care between the two groups. This bias would, however, produce findings that would be of concern only if substantial higher costs were found to be associated with trial participation.

The other modeling issue concerns whether average treatment costs are the chief concern, or whether it might be of more interest to determine whether costs differences might be increasing as a function of baseline costs for non-participants. One typical approach to estimate such a non-linear cost function is to use a log transformation on the cost variable. This approach does not work when substantial numbers of study participants report zero costs, as was the case in the CCTS. While two-part models were explored, the results are difficult to interpret in terms of marginal effects. The use of a generalized linear model testing for the presence of a log-linear relationship of costs to trial participation allowed this issue to be addressed directly.

The principal findings were that trial participation is associated with a small but statistically significant increase in prescription drug costs, but that the magnitude of the costs was trivial relative to other treatment costs and did not translate into higher out-of-pocket costs for trial participants. In terms of policy then, the conduct of clinical trials is unlikely to pose an undue economic burden on either third party payers or on study participants. The incremental costs are trivial in comparison with the potential improvements in treatments for cancer.

References

Fireman BH, Fehrenbacher L, Gruskin EP, Ray GT. "Cost of care for patients in cancer clinical trials," *J Natl Cancer Inst* 2000:92:136-42.

Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*, 3rd Ed. 2003, Hoboken, NJ: Wiley-Interscience.

Hutchins LF, Unger JM, Crowley JJ, Coltman CA Jr, Albain KS. Under-representation of Patients 65 Years of Age or Older in Cancer-Treatment Trials. *N Engl J Med.* 1999;341(27):2061-2067.

Journal Storage, A Scholarly Archive, JSTOR, Inc. http://www.jstor.org/, accessed 2/6/04.

Mitka M. FDA Alert on Antidepressants for Youth. *JAMA*. 2003;290:2534.

Olfson M, Shaffer D, Marcus SC, Greenberg T. Relationship between antidepressant medication treatment and suicide in adolescents. *Arch Gen Psychiatry* 2003;60(10):978-82.

US National Library of Medicine and National Institutes of Health Database, http://medlineplus.gov/, accessed 2/6/04.

Wagner KD, Ambrosini P, Rynn M, Wohlberg C, Yang R, Greenbaum MS, Childress A, Donnelly C, Deas D; Sertraline Pediatric Depression Study Group. Efficacy of sertraline in the treatment of children and adolescents with major depressive disorder: two randomized controlled trials. *JAMA* 2003;290(8):1033-41.