DISSERTATION

RAND

Healthcare Cost-Effectiveness Analysis for Older Patients

Using Cataract Surgery and Breast Cancer Treatment Data

Arash Naeim

RAND Graduate School

DISSERTATION

RAND

Healthcare Cost-Effectiveness Analysis for Older Patients

Using Cataract Surgery and Breast Cancer Treatment Data

Arash Naeim

RAND Graduate School

This document was prepared as a dissertation in August 2002 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Policy Analysis at the RAND Graduate School of Policy Studies. The faculty committee that supervised and approved the dissertation consisted of Emmett Keeler (Chair), Ron Hays, and David Reuben. Charles Bennett was the outside reader for the dissertation.

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

RAND is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND® is a registered trademark. RAND's publications do not necessarily reflect the opinions or policies of its research sponsors.

Published 2002 by RAND
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
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

Dedicated
To

My wife, Behnaz, without whose continual support, encouragement, and grammatical skills I would never have succeeded in finishing my dissertation. You are my soulmate.

My parents, Faramarz and Ester, who provided me with all the tools I needed to succeed in both my professional and personal life. You are both the best role models.

My sister, Shiva, and brother-in-law, Robert, who provided their friendship and love.

My grandparents who have shown my how to enjoy living life even in the most difficult circumstances.

Behrooz, Edna, Behnoosh, Beata, and Bryan - the best in-laws in the world

Behrooz and Pejman – the best friends in the world

Acknowledgements

I would like to thank the Hartford Foundation and the Claude D. Pepper Center on Aging for providing funding and support for this work and without whose support I would never had the opportunity to have an extended period of protected time to pursue research focused on older patients.

I have been fortunate to have the support of the Sub-specialty Training and Advanced Research (STAR) program at UCLA. Thanks to this program and the help of Joy Frank M.D., I have had ample support and protected time from clinical activities to finish this dissertation quickly.

Very special recognition needs to be given my committee chair, Emmett Keeler Ph.D, who from the start of my graduate education always maintained an open door policy for me. His continual support, encouragement, and trust has provided my foundation for what I hope to be a successful career in health services research.

In addition to Emmett, Ron Hays Ph.D. and David Reuben M.D., have both been very thoughtful and supportive committee members. They not only spent considerable time evaluating the design, data, and analysis of the research in this dissertation, but also helped with the often painful and mundane task of assuring information is conveyed clearly and concisely.

I have been fortunate to have a very supportive outside reader, Charles Bennett M.D. Ph.D., who understanding of both the clinical and research components made him an invaluable resource throughout the dissertation process.

I have been fortunate to have strong mentors outside my dissertation committee. Carol Mangione M.D. provided me the opportunity to use her cataract surgery trial data and spent considerable time in serving mentoring me during my work on that portion of my dissertation. John Glaspy M.D. has served as a role model for a successful clinician scientist in oncology, and always made time in his busy schedule to help provide advice and guidance for my career.

Special thanks needs to be given to my colleague, Neraj Sood, without whom I would probably never have passed qualifying exams and played a special role as an outside scorer in one component of this dissertation.

Finally, I would like to thank Dean Klitgaard, Marcy Agmon, Maggie Clay, Lynn Anderson, and Margie Milrad who have all provided me with outstanding support, empathetic ears, and sturdy shoulders during my stay at the RAND Graduate School.

Summary

The percentage of elderly, over the age of 65, in the U.S. population will be changing dramatically over the next 20-30 years. The changing demographics over the next three decades will have significant consequences for Medicare and national policy. Given the economic impact of future health care expenditures by our growing elderly population, a concerted effort needs to be made to define high quality yet cost-effective medical therapy for older patients.

The purpose of this dissertation is to explore the use of both clinical trials and evidence-based decision models in performing cost-effectiveness analysis in elderly patients. For some diseases, such as cataract surgery, the majority of patients tend to be older. Therefore, previous studies focusing on younger patients do not exist. Important therapeutic and policy questions can only be addressed through a clinical trial. Other diseases, such as breast cancer, involve a wider age range of patients from early 40s to 90s. For such diseases, there is a literature of clinical trials on younger patients and the young elderly, 60 – 70. This previous literature can be used to develop decision analysis models to help define pertinent questions and areas for further research (i.e., clinical trials).

This dissertation is broken up into two main parts, demonstrating two different approaches to cost-effectiveness in an older population, a clinical trial and modeling from existing data. One part focuses on a randomized clinical trial on cataract surgery. The other part develops an evidence-based decision analysis model on the cost-effectiveness of treating older patients with early breast cancer.

The cataract surgery section has two sub-components: (a) a methodological section focusing on strategies to deal with question non-response among the older patients on the Heath Utilities Index Mark 3, HUI3, questionnaire, and (b) a cost-effectiveness analysis based on a randomized clinical trial of older patients with cataracts comparing up-front surgery versus watchful waiting in patients who have relatively good visual functioning. HUI3 analysis demonstrated a significant percentage of missing data due to "Don't Know" responses that could be handled using inspection/deduction from the pattern of completed responses. In the cost-effectiveness analysis, the use of a preoperative tool, Cataract Surgery Index, was shown to discriminate between those with high and low probability for improvement from cataract surgery, and those for whom surgery was cost-effective.

The breast cancer section focuses on an evidence-based decision analysis for older patients, >65, who have newly diagnosed early stage breast cancer. This analysis includes models taking into account longevity, aggregate comorbidity, frailty, and established preferences from quality of life literature. The uncertainty associated with treatment decision in older breast cancer patients could be mapped in this decision analysis framework. Whereas in younger estrogen positive breast cancer patients adjuvant chemotherapy was a dominant strategy, in older patients the dominant strategy was hormone therapy. In both 65 and 75 year old patients, there were scenarios for which combined hormone and chemotherapy could be considered cost-effective. Furthermore, sensitivity testing taking into account higher discount rates in older patients and different baseline quality of life states altered the cost-effectiveness of most adjuvant therapy strategies.

Many policy decisions will be made in the future pertaining to the provision of health among elderly patients. A broad set of approaches will be required to determine the cost-effectiveness of specific therapies in this population. These approaches will range from clinical trials to elaborate modeling using a combination of existing data and assumptions. The dissertation provides two examples using these approaches in performing cost-effectiveness analysis among the elderly.

Table of Contents

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 1

Introduction

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Introduction

The percentage of elderly, over the age of 65, in the U.S. population will be changing dramatically over the next 20-30 years. Currently Americans over the age of 65 constitute 12 % of the population, 32.6 million individuals [1]. More specifically, 7 % of Americans are between the ages of 65-74, 4 % between the ages of 75-84, and 1 % over the age of 85 [1]. By 2020, the Baby Boomer generation, representing the increase in births in the United States during the 1950's and 60's, will be in their age of retirement. As a result, the population age 65-74 will increase 74% between 1990-2020, while the population under 65 will increase only 24%[2]. As the Baby Boomer generation ages, it is expected that the elderly will constitute about 1 in 5 Americans. By 2030, there will be 70.2 million Americans over the age of 65 with the largest increase occurring in the sub-segment of individuals aged 75-84 [3].

The changing demographics over the next three decades will have significant consequences for Medicare and national policy. Medicare's growth in spending is outpacing its growth in revenue. Most of the revenue in Medicare is generated from a 2.9% payroll tax, which has not changed since 1986 [4]. The population increase among the elderly is projected to increase the number of Medicare beneficiaries by 77%, from the present 40 million to 70 million by 2025 [4]. Medicare spending will almost double and may cause the Medicare Trust Fund to run out of money. Furthermore, the impact of these changes is likely to also be felt by beneficiaries, for whom out of pocket expenses are likely to rise by 80% over the next several decades [4]. Among poor elderly persons, it is estimated that the proportion of income spent on health expenses will rise from 51% to 71% by 2025 [4].

Given the economic impact of future health care expenditures by our growing elderly population, a concerted effort needs to be made to define high quality yet cost-effective medical therapy for older patients. Determining appropriate care has historically been achieved using clinical trials, which when published form a collective called "evidence-based literature". Most clinical trials to date have not included a large number

of older patients. Older patients tend to have medical co-morbidities and functional disabilities. In addition, the elderly individuals may be less likely to comply with the stringency of a clinical trial protocol [5]. There is also a perception that older patients may be more susceptible to side effects or complications [6]. Since pharmaceutical companies sponsor many clinical trials seeking FDA approval, there has been a tendency to select younger, healthier populations that are easier to manage and more apt to show benefit, either in terms of life expectancy or disease free survival.

The under-representation of those over 65 in clinical trials limits the benefit of drug therapy in these populations [7]. One example pertains to the use of non-steroidal anti-inflammatory drugs, NSAIDS. Older patients who have degenerative musculoskeletal diseases commonly use NSAIDS. However, in leading trials evaluating NSAIDs, only 2 % of patients were 65 years of age or over and less than 0.1% were over 75 [8]. Even when older patients are included in clinical trials, they are generally younger, healthy and predominantly male. Frail older persons are rarely included in drug trials even though they are commonly given the drugs of interest. For example, in trials evaluating donepezil therapy for Alzheimer's disease, only patients between the ages of 65-74 were chosen. Moreover, individuals with associated co-morbid conditions common to patients over the age of 65 were excluded [9]. Considering that many patients with dementia are frail elderly individuals with multiple comorbidities, it is unclear how to extend benefits and anticipate toxicity from this type of trial for the average older patient with dementia. The external validity of any trial, particularly those involving the elderly, is very important. A trial should be designed to develop or contribute to generalizable knowledge [10].

Thus, for many diseases impacting the elderly individuals, there is a lack of "evidence-based" literature [11] [12]. To help determine appropriate care, policymakers and medical researchers have two options: (1) perform clinical studies specifically on an older population or (2) extrapolate evidence from younger patients to older patients. Each of these strategies has its disadvantages.

Performing clinical trials on the older patients is often a difficult endeavor. In many large aging trials participants are purposefully selected to reduce the risk of sub-optimal adherence and retention. This selection often involves excluding those with

barriers such as transportation needs, sensory deficits, functional dependence, major diseases limiting life expectancy, or apparent psychological distress [13]. Recruiting older patients often requires patience and utilization of a mixture of approaches including: phone, referrals, solicitations, presentations, media, mailings, and fliers [14, 15].

Many elderly patients have varying cognitive or functional impairment, which makes using traditional outcome measures difficult [16]. There is a higher rate of non-response, refusal, and loss to follow-up [17]. Furthermore, if benefit has been shown for a treatment in previous trials in younger patients, it may be unethical to repeat this trial over again in the elderly. There are also issues pertaining to appropriate informed consent among elderly patients [18, 19]. Finally, a policy of repeating all or most previous clinical trials for the elderly is costly.

On the other hand, extending data from younger patients to older patients may be inaccurate even after accounting for longevity. It is quite common for medical practitioners to extrapolate clinical trials data without realizing the inherent assumptions required for such extrapolation to be valid. Extrapolation with supporting data, termed as "off-support" has been discussed greatly among social scientists and felt to be often misleading [20]. The most common assumption in regards to extrapolation is the "invariance assumption" [20]. The invariance assumption assumes a similar behavior or outcome "off-support" as that demonstrated by the data. A sub-type of this type of assumption is temporal invariance which is routinely applied to predict the future [20].

In addition to decreased longevity, the elderly have more comorbidity and functional disability. It is not certain whether the biology of many illnesses is similar among older and younger patients. Furthermore, older patients may have a different response and toxicity profiles to treatment. In addition, older patients are likely to have different preferences for health states than younger patients, placing more weight on quality of life than longevity. The elderly are a heterogeneous population for whom age alone may be insufficient to predict benefit from treatment.

Clinical trials and evidence-based analysis based on previous work do not need to be mutually exclusive of one another. It is possible to use prior literature in a decision analysis framework to determine the upper and lower bounds of expected cost and benefit

in elderly patients and create models to take into account longevity, co-morbidity, functional status, and preferences. These analyses would identify the questions for which there is the most uncertainty and clinical importance justifying the expense of a clinical trial.

The purpose of this dissertation is to explore the use of both clinical trials and evidence-based decision models in performing cost-effectiveness analysis in elderly patients. For some diseases, such as cataract surgery, the majority of patients tend to be older. Therefore, previous studies focusing on younger patients do not exist. Important therapeutic and policy questions can only be addressed through a clinical trial. Other diseases, such as breast cancer, involve a wider age range of patients from early 40s to 90s. For such diseases, there is a literature of clinical trials on younger patients and the young elderly, 60 – 70. This previous literature can be used to develop decision analysis models to help define pertinent questions and areas for further research (i.e., clinical trials).

This dissertation is broken up into two main parts, demonstrating two different approaches to cost-effectiveness in an older population, a clinical trial and modeling from existing data. One part focuses on a randomized clinical trial on cataract surgery. The other part develops an evidence-based decision analysis model on the cost-effectiveness of treating older patients with early breast cancer. The cataract surgery section has two sub-components: (a) a methodological section focusing on strategies to deal with question non-response among the older patients on the Heath Utilities Index Mark 3, HUI3, questionnaire, and (b) a cost-effectiveness analysis based on a randomized clinical trial of older patients with cataracts comparing up-front surgery versus watchful waiting in patients who have relatively good visual functioning. The breast cancer section focuses on an evidence-based decision analysis for older patients, >65, who have newly diagnosed early stage breast cancer. This analysis will include models taking into account longevity, comorbidity, frailty, and simulated preferences.

Many policy decisions will be made in the future pertaining to the provision of health among elderly patients. A broad set of approaches will be required to determine the cost-effectiveness of specific therapies in this population. These approaches will range from clinical trials to elaborate modeling using a combination of existing data and

assumptions. The dissertation provides two examples using these approaches in performing cost-effectiveness analysis among the elderly.

1. Bureau, U.S.C., *The Older Population in the United States*. 2000, U.S. Department of Commerce:Economics and Statistics Adminstration.
2. Bureau, U.S.C., *Aging in the United States: Past, Present, and Future*. 1997, U.S. Department of Commerce:Economics and Statistics Adminstration.
3. Yanick, R., *Cancer Burden in the Aged: An Epidemiologic and Demographic Overview*. Cancer, 1997. **80**(7): p. 1273-1283.
4. Moon, M., *Growth in Medicare Spending:What will Beneficiaries Pay?* 1999, Urban Institute.
5. Cameron HJ, W.B., *Clinical trials in the elderly: Should we do more?* Drugs Aging, 1996. **9**(9): p. 307-10.
6. Tariot, P.N., *The older patient: the ongoing challenge of efficacy and tolerability.* J Clin Psychiatry, 1999. **23**: p. 29-33.
7. Paula A. Rochon, M., MPH; Philip B. Berger, MD; Michael Gordon, MD, *The evolution of clinical trials: inclusion and representation.* Canadian Medical Association Journal, 1998. **159**: p. 1373-4.
8. Rochon PA, F.P., Dear KBG, Minaker KL, Chalmers TC, *Reporting of age data in clinical trials of arthritis: deficiencies and solutions.* Arch Intern Med, 1993. **153**(2): p. 243-8.
9. Rogers SL, F.M., Doody RS, Mohs R, Friedhoff LT, Group DS., *A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease.* Neurology, 1998. **50**(1): p. 136-45.
10. Kitler, M.E., *Clinical trials in the elderly. Pivotal points.* Clin Geriatr Med, 1990. **6**(2): p. 235-55.
11. Coelho Filho, J.M., *Appraising clinical trial evidence for elderly people: special considerations.* Age Ageing, 2001. **30**(2): p. 117-9.
12. Pientka, L. and C. Friedrich, *[Evidence-based medicine--problems and application to geriatrics].* Z Gerontol Geriatr, 2000. **33**(2): p. 102-10.
13. Anderson, R.T., et al., *Issues of aging and adherence to health interventions.* Control Clin Trials, 2000. **21**(5 Suppl): p. 171S-83S.
14. Adams, J., et al., *Recruiting older adults for clinical trials.* Control Clin Trials, 1997. **18**(1): p. 14-26.
15. Boult, C., et al., *Soliciting defined populations to recruit samples of high-risk older adults.* J Gerontol A Biol Sci Med Sci, 1998. **53**(5): p. M379-84.
16. Williams, M.E. and S.M. Retchin, *Clinical geriatric research: still in adolescence.* J Am Geriatr Soc, 1984. **32**(11): p. 851-7.
17. ter Riet, G., A.G. Kessels, and P. Knipschild, *Problems in the conduct of a randomised clinical trial.* J Wound Care, 1998. **7**(5): p. 259-62.
18. Elad, P., et al., *Demented patients' participation in a clinical trial: factors affecting the caregivers' decision.* Int J Geriatr Psychiatry, 2000. **15**(4): p. 325-30.
19. Pucci, E., et al., *Information and competency for consent to pharmacologic clinical trials in Alzheimer disease: an empirical analysis in patients and family caregivers.* Alzheimer Dis Assoc Disord, 2001. **15**(3): p. 146-54.

20. Manski, C.F., *Extrapolation*, in *Indentification Problems in the Social Sciences*. 1995, Harvard University Press: Cambridge. p. 172.

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 2

Handling Missing Data in the Health Utilities Index Mark 3 (HUI3)

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Introduction

The value placed on a medical intervention has increasingly come to be based on its costs and outcomes, including both life expectancy and health related quality of life. The Health Utilities Index Mark 3 (HUI3) is one of several instruments developed to measure health-related quality of life. The HUI3 instrument is composed of 40 questions that measure eight health attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain) and one question that elicits overall health status. The attributes were selected to be structurally independent (each independently affects overall health), and the system as a whole defines 972,000 different states [1,2]. The HUI3 includes categorical information in the form of attribute levels (best score for level =1.00, score for most disabled=6.00). These attribute scores can then be converted to single attribute utility scores, ranging from 0 to 1. For example, the single attribute vision utility ranges from 1.00, representing perfect vision state, to 0.00, representing a blind state. Furthermore, a weighted-scoring algorithm is applied to combine the scores for each attribute to derive multi-attribute utility score, a value between zero and one to represent utility of the overall health state (perfect health=1.00, worst health=0.00, where worst health could conceivable be worst than a death state).

Most questions are of the "yes/no" categorical variety, and all questions allow the respondent to answer "Don't Know" or to refuse to answer. When a response is "Don't Know," the interviewer moves on to the next question in the sequence. Often the next question in sequence is not appropriate instead a skip should have taken place to the next appropriate question. The failure to follow the appropriate skip pattern eliminates the possibility of accurately scoring the attribute. Don't know responses to well defined questions asking about capabilities are very different than typical non-response missing data, for which non response is given at all. Little research has been conducted on the best way to handle Don't Know (DK) responses. One article reviewed the issue as it applied to a survey of Slovenians regarding Slovenian independence [3]: In election polls, DK means "undecided" and is valuable information to both parties. In the HUI3, DK is a valid response and may have similar utility in estimating the proportion of a survey population that cannot clearly answer a question one way or another.

In addition to DK responses, traditional non-response or data input errors causes disturbances in the skip patterns and complicate scoring (**Tables 1 and 2**). Since the HUI3 is a widely used preference measure, developing strategies to handle "Don't Know" responses and to impute non-responses so that attribute levels can be assigned would be valuable.

Types of Missing Data

There are three major categories of missing data: (1) Missing Completely At Random (MCAR), (2) Missing At Random (MAR), and (3) Non-Ignorable Non-responses (NINR) [4]. Data are considered to be MCAR if missing values represent a random sample of the entire data set or result completely by chance [5][6]. MAR means that the data are missing at random conditional on the values of some other preserved, non-missing, variables [7]. If one can identify variables that explain the reason for missing data, the problem can be modeled [5]. Finally, if the data are not missing randomly (Non-ignorable Non-responses [NINR]), the results might be biased thereby threatening the validity of the results [8]. NINR often occurs when the factors responsible for the missing data are related to values of the data and other personal characteristics of those surveyed. [9]. All forms of missing data are problematic if the number of missing items represents a large proportion of the data set, thereby affecting the power of the analysis and potentially leading to type II errors [10]. Thus, there are incentives to develop a strategy to handle missing values.

Handling missing data resulting from item non-response

In the HUI3, item non-response in interviews may be attributed to three primary factors: (a) an interviewee may refuse or be unable to answer a particular question, (b) the interviewer may fail to ask the question or to record the answer (generally rare), and (c) the recorded response may be logically inconsistent (interviewer guidelines are designed to minimize this problem). Item non-response can be dealt with in several ways. Many software packages simply delete the records with missing items (listwise deletion), which has potentially significant consequences for analysis [6]. Another option is to try to analyze records with missing items as a

separate group from records with complete data. Finally, one can impute, or fill in, plausible and consistent values for missing items.

Imputation can reduce the bias in survey estimates that results from missing data while making results easier to present and more consistent. Using imputation, analysis can be performed as if the data were complete. A good imputation procedure meets the objectives of (1) imputing values that are consistent and (2) reducing the non-response bias while preserving the variance and the relationships between items as much as possible. In addition, a good imputation procedure is one that is set up prior to data collection and evaluated in terms of its impacts on the bias and precision of estimates [11].

Non-responses to many of the questions in the HUI3 need not result in the loss of attribute level scores. In many cases, the attribute level scores are based on a pattern, and the question lacking a response does not impact the standard scoring. In most cases, analysis of all the questions involved in the attribute score will allow assignment of a appropriate score or dramatically narrow the range of possible scores, even with non-response to one or two of the questions within the attribute. This type of missing data may be termed "missing without consequence" (MWC) since its absence may have no practical consequences.

This paper attempts to present and evaluate several ways of dealing with non-response in the HUI3. Our hypothesis is that a significant portion of HUI missing data is actually MWC. This analysis will demonstrate the degree to which the range of scores can be narrowed. The use of formal techniques of imputation such as mean substitution, hot-deck and logistic regression will be compared to simpler methods. Since imputation schemes have not been formalized for this instrument, we review these options.

Methods

Data Collection

The data for this analysis are derived from the Pepper Cataract Management Trial (PCMT). The PCMT is a randomized trial that compares the benefits of cataract surgery to those of watchful waiting, using inclusion criteria defined to select older patients who might benefit only marginally from surgery. Patients older than 64 years with a diagnosis of bilateral cataracts, with and without other chronic eye diseases, candidates for first eye surgery, and a <30% probability of predicted improvement, were eligible for enrollment to the study. Additional

eligibility criteria included: English speakers, sufficient cognitive and hearing function to be able to provide informed consent and fully participate in the interview, and persons who planned to remain in the are to be seen for follow-up clinical examinations. Patients enrolled were also required to be a regular patient of the ophthalmology practices enrolled. Patients who were told by their surgeon that they should or should not have surgery were considered ineligible.

Baseline and 6-month assessments were conducted on 248 randomized patients. Measurement instruments used at these two time points included (1) the Activities of Daily Vision Scale (ADVS), (2) SF-12, and (3) HUI3. Two HUI3 attributes were chosen for non-response analysis: baseline HUI3 vision and baseline HUI3 ambulation. These two attributes were chosen because they accounted for a significant proportion of the missing data and were key outcome measures for the intervention (cataract surgery) in the randomized trial. The number of missing items per attribute and the proportion of DK responses among the patients are shown in **Table 3**. The majority of missing data are attributable to DK responses.

Analysis

The DK response to a single question for a specific attribute resulted in missing values for the question, the attribute level score, and the total weighted HUI3 score. The attribute level score was considered to be the most important component to impute, since it is used for the derivation of both single-attribute and multi-attribute utility scores. No specific imputation scheme is provided with the HUI3 for missing attribute level scores that result from specific item non-response, instead a variety of options are suggested including listwise deletion and regression imputation. However, there is no discussion on using answered internal responses to help solve the missing data issues. Our analysis takes a two-step approach. As a first step, we inspect patterns to fill in attribute values that are either independent of the value of missing data or can be assigned through the use of logical deduction. In the second step, we compare alternative imputation schemes.

Inspection, Deduction, and Model Formation

The entire set of data was inspected. For those individuals who were missing HUI3 Vision and HUI3 Ambulation scores, the pattern of response to questions within those attributes was inspected. These patterns were then compared to the scoring scheme for the particular

attribute. The patterns were considered to have "missing without consequences" (MWC) non-responses if the missing response was deemed not to alter the attribute-level score, regardless of the answer. For those patterns with relevant non-responses, an attempt was made to use the internal logic of the sequence of questions within the attribute to score the question correctly and then subsequently to provide an attribute-level score. For those patterns for which inspection and deduction could be useful two independent judges were used to determine inter-rater consistency.

In circumstances where neither inspection nor deduction allowed for definitive scoring, a model was created that consisted of the pattern of responses, both answered and missing, that was to be used in the second (imputation) part of the analysis.

Imputation Technique

The second part of the analysis concentrated on imputation techniques for HUI3 Vision attribute-level scores for the remaining missing data. We first assumed that a DK response was equivalent to missing. Two sets of imputations were performed. One set was performed prior to using inspection and deduction, while the second was performed after inspection and deduction. Four imputation techniques on the post inspection and deduction data were compared: (1) mean substitution, (2) model scoring, (3) hot deck, and (4) regression imputation.

Mean Substitution – This technique was performed using the mean of the observed data from individuals with complete data. This mean was then substituted as the HUI3 vision-level attribute score for those with missing scores. Since attribute level scores are ordinal, the mean value is difficult to interpret. Therefore, this analysis was also performed (data not shown) using single-attribute utility values and multiple-attribute utility values for perspective. The conversion of attribute levels to utility values is well described in previous literature [12].

Model Scoring –Attribute level scores ranges normally range from 1 to 5 or 6. The models developed in the first part of this analysis reduced the possible attribute level scores to between two and three choices from the five to six possibilities without the models. We then imputed a weighted-mean based on the scores of individuals with complete answers that formed patterns consistent with each of the models. Using the weighted mean is appropriate if data are missing at random, so the respondent is like the other people who answered non-missing questions similarly.

Hot-Deck – We used the HUI3 vision-attribute models to identify individuals with complete survey responses whose response pattern was similar to individuals with missing data. Pattern similarity was used to define the pool of possible donors. Attribute-level scores were then selected randomly with replacement from this pool and used as the imputed score for those with missing level scores.

Logistic Regression- This imputation technique used Activities of Daily Vision Scale (ADVS) scores as an external scale. Initial analysis was performed using an ordered logit model with ADVS as the only independent variable and HUI3 attribute vision levels as the dependent variable. This logit regression provided probabilities for each level (1-6). A random number was generated from a uniform distribution and used to assign attribute levels based on regression probabilities. In addition, this method was applied using the models created from response pattern of questions answered within the attribute by performing ordered logit regressions on only "similar" individuals.

Unweighted Mean ImputationUsing Models- Finally we looked at the implications of assuming that a DK response indicates respondents are caught in the middle of a dichotomous answer, between yes and no. In that case, it seems reasonable to impute an unweighted mean of the scores associated with yes and with no on that choice and the rest of their responses. For example, if a yes response on an item would lead to an attribute score of 2 based on the other responses and a no response would lead to an attribute score of 4, DK gets assigned a score of (2+4)/2 = 3, independently of the pattern responses and attribute level scores among people with complete data.

.

Results:

Inspection and Detection

HUI vision

The vision-attribute portion of the HUI is composed of 5 questions (2 near vision, 1 general, 2 distant vision). **Table 1** shows HUI vision questions and the scoring decision. A review of baseline HUI vision scores revealed that 39 of the 47 missing data points were a result of a DK response. With inspection, 23 cases could be assigned attribute levels with a high degree of certainty. In 19 cases, the missing data were not needed to assign appropriate attribute levels

(**Table 4, A and B**). In the cases with Pattern C, internal logical deduction was needed to assign a score (**Table 4, C**). Here, the main question was how to interpret a DK response to question 4 when the person answers "no" to question 5. Question 4 states "Have you ever been able to see well enough to recognize a friend across the street *without* glasses or contact lenses? Question 5 states "Have you ever been able to see well enough to recognize a friend across the street *with* glasses or contact lenses? Thus, it is reasonable to assume, as verified buy our two independent judges, that those who answered "no" to question 5 should answer "no" to question 4 as well.

Two cases of missing data resulted from a skip or input error (**Table 4, D and E**). In these cases, logic could also be used to assign an attribute level. For case D, the skip pattern dictated that a "no" to questions 1-3 should result in skipping questions 4 and 5. Since questions 4 and 5 were answered, it is logical to assume that the answer to question 3 was "yes," which would have forced the interviewer to ask question 4 and 5. In case E, a "yes" to question 2 required a skip to question 4. In this case, no data were entered for question 4. Thus, the "yes" response to question 3 most likely was entered in the data file incorrectly and represents the answer to question 4.

HUI Ambulation

The ambulation portion of the HUI comprises 7 related questions. **Table 2** shows the questions and scoring of this attribute**.** The initial assessment of baseline HUI ambulation-level scores revealed 24 missing values, 21 of which were due to DK and 3 to input errors. Inspection led to the immediate assignment of appropriate level scores in 16 cases, where the missing data were not relevant (**Table 5, A, B, and F**). In 5 cases, logical deduction was required (**Table 5, C, D, and E**). If an individual answered "yes" to question 20, "Have you needed mechanical support to be able to walk around the neighborhood?" then a "no" was required to questions16, 17, and 18. Therefore, all five questions were assigned the level of 3 by both of our independent judges. Three cases remained for which neither inspection nor internal deduction was sufficient to assign a level score.

Summary of Inspection and Deduction

The use of inspection and deduction dramatically reduced the missing attribute level problem for both HUI vision and ambulation. There was 100% agreement in score assignment

for missing scores resulting from DK responses between the two independent judges. The only area of agreement between occurred for the two cases resulting from skip errors or transcription error by the interviewer (**Table 4 D and E**). One judge felt comfortable assigning specific scores assuming the likely cause of the error (as described in the above text), while the other judge felt that trying to guess the etiology of the error was outside the scope of logical deduction. Nevertheless, of the 44 cases scored by inspection and deduction there was agreement on 42 representing an overall 95% inter-rater agreement. The number of missing items was reduced by almost 50% in the HUI vision attribute and almost 88% in the HUI ambulation attribute (**Table 3).** Since only 3 HUI ambulation attribute scores remained missing, the rest of the analysis focused on the remaining 24 missing scores in the HUI vision attribute.

Imputation without Prior Inspection

Without inspection, there were only 201 complete HUI3 vision attribute level scores with 47 missing values. It is possible to perform imputation neglecting the internal pattern of response. Substituting 1.960 in for the 47 missing cases provided a total sample mean of 1.960 with a standard deviation, SD, of 0.660 (**Table 6**). The artifactual variance reduction from .730 to .660 could lead to too narrow confidence intervals and overestimates of significance. Using a random draw of an attribute level from completed cases to fill in the missing cases, a simple hot deck, yielded a mean of 1.778 and SD of 0.743. Using ADVS scores via logistic regression to impute missing scores yielded a mean of 1.964 and SD of 0.734 (**Table 6**).

Imputation after Inspection

After the initial inspection and logical deduction, 24 cases of missing data remained to be analyzed in the HUI vision attribute. For the 224 complete cases (deleting cases with missing information), the mean was 1.986 with a SD of 0..734. Substituting 1.986 in the 24 cases yielded a total mean for the 248-case sample of 1.986 with SD of 0.694 (**Table 6**), thus halving the problem of variance reduction compared to mean imputation with no inspection. Logistic regression using ADVS scores as the only independent variable yielded a complete sample mean of 1.972 with a SD of 0.716. This value is a good fit to the standard deviation, but the method ignores appropriate answers to other questions within the attribute, exploited in the models below.

From the 24 missing HUI vision scores, 7 models were needed to represent the pattern of missing information from which the sets of "similar" individuals could be defined (**Table 7**). These models were used for the subsequent imputation techniques. Weighting the feasible pattern scores in each model by the distribution of the scores among those with complete answers and then substituting the weighted mean for missing values yielded a mean and a SD of 1.963 and 0.686, respectively.

Using the more complex imputation approach of hot-deck provided a complete sample mean of 1.931 and a SD of 0.702 (**Table 6**). The ADVS scores and the models with internal information on responses were used to perform an ordered-logit regression across a similar population. The complete sample mean and SD were 1.935 and 0.700, respectively (**Table 6**)

If a DK response means being caught in the middle between yes and no and we substitute the unweighted mean of the possible scores for each missing pattern for the missing value, the mean and SD of the total sample were 2.274 and 0.774 respectively. The mean is higher than those with complete data, but this is appropriate if a DK response has the assumed meaning.

Discussion

Missing information is likely when using the HUI3, because survey subjects have the option to answer questions with a DK response. Thus, studies that utilize this instrument invariably need to handle DK responses and item refusals. In this paper, we present a 2-step approach for handling the missing items within an attribute on the HUI. The first step inspects patterns and uses logical deduction to fill in values that are independent of the value of missing data. The second step imputes the remaining missing values in various ways.

Pattern inspection and internal logical deduction for each attribute with missing data for specific questions proved to be valuable. In many cases, the non-missing items within the attribute determined unique attribute-level values so that the data was MWC. In many of the other cases, only two or three values were consistent with the observed responses.

When the amount of missing data is small, results may be insensitive to imputation methods, and simple methods for the second imputation step will suffice. Furthermore, disregarding the internal response patterns, from which we perform our inspection and deduction, and pursuing straight imputation from the start may be unreliable and provide incorrect attribute level scores, erroneous group means, and unwanted variance reduction.

Our hypothesis was that the missing data were not Missing At Random (MAR), but rather that they represented Non-Ignorable Non-responses, NINR. Many researchers simply assume for convenience that data are MCAR/MAR. However, for the HUI3 vision attributes, we saw evidence to suggest that our data was NINR. Patients with good to marginal vision (HUI3 levels of 1-2) had more difficulty in providing responses to specific questions than those patients with poor vision (HUI3 levels of 3-6) who provided definitive responses more easily. If missing data is NINR, methods that disregard the internal scoring pattern within attributes may yield incorrect means and standard deviations.

Our analysis showed a distinct difference between imputation with and without models based on inspection of the internal response patterns. Such differences can affect the outcomes of clinical trials. The goal of the Pepper Cataract Management Trial is to evaluate 6-month outcomes of patients with marginal vision, randomized to either cataract surgery or watchful waiting. If the missing data are NINR because those with better vision tend to answer with a DK response more often, imputing that uses hot-decking or logistic regression without considering any of the internal response patterns (a higher imputed attribute level score) will likely bias baseline scores towards poorer vision. A bias to toward (artificially) poorer baseline scores might in turn exaggerate the benefit of cataract surgery at 6 months.

More importantly, if pattern analysis and logical deduction can be used to reduce the number of missing attribute-level scores, simple imputation techniques, such as mean substitution or model pattern scoring (weighted or un-weighted), might suffice for the remaining cases. We used many complicated imputations methods, but they added little more than the use of pattern weighted and un-weighted means, because the amount of missing data after using pattern analysis was so low. Using a weighted mean value from the patterns of missing data yielded results very similar to those of mean substitution, hot-deck, and logistic regression.

After taking the pattern of responses into account, regression using ADVS scores did not contribute much to an accurate prediction of the probabilities of HUI vision-attribute scores. Outside scales for other attributes could also be considered for use in modeling. For example, the SF-12's physical components might be used in a regression imputation approach for ambulation; but these outside scales would offer little improvement to the imputation if their ability to discriminate between states did not match that of the HUI. The SF-12 might ask individuals if

they walk but not be able to discriminate whether they need help or use equipment, which is essential for imputing the HUI ambulation-attribute score.

If only a single measure of health is needed and that measure is assumed to be stable over time, using data from other points in time for the same subject might be more accurate than imputing between subjects. For our purposes, the strategy of using an individual's HUI vision or ambulation levels at 6 months post-baseline to impute that people own baseline level would not be appropriate. First, the study considered an intervention, cataract surgery, aimed at changing visual function. Therefore, 6-month HUI vision would likely be different from the baseline level. This limitation also applies to HUI ambulation scores since vision affects ambulation. Second, elderly individuals would likely undergo changes in vision or ambulation over 6 months. Using a distant time point might be more useful for an attribute such as cognition. An individual with perfect cognition 6 months from a baseline time-point may have had perfect cognition at baseline, especially if the study excluded all patients with potential confounders such as delirium. The same could be said about other attributes such as hearing that would rarely improve.

In summary, pattern inspection and logical deduction can greatly mitigate problems with DK responses and missing values in HUI3. The initial task of developing a bank of algorithms for scoring when data is missing based inspection and deduction can be time consuming, even more so than complicated imputation techniques. However, these algorithms can easily be verified and standardized using a panel of independent judges. In the long run, these algorithms will save significant amounts of time and provide for more accurate and reliable results. After an initial stage of pattern inspection and deduction, the missing data problem may become so small that simple imputation methods may suffice for the remaining missing data. This two-step strategy should simplify most HUI3 missing data cases resulting from DK responses. More work needs to be done to determine the exact meaning that should be attributed to a DK response.

References

1. **Feeny DH, Torrance GW, Furlong WJ:** Health Utilities Index. In B. S (ed): "Quality of Life and Pharmacoeconomics in Clinical Trials." Philadelphia: Lippincott-Raven Publishers, 1996:239-252.

2. **Grootendorst P, Feeny D, Furlong W**: Health Utilities Index Mark 3: evidence of construct validity for stroke and arthritis in a population health survey. Med Care 2000;38:290-299.

3. **Rubin DB, Stern HS, Vehovar V**: Handling "Don't Know" Survey Responses: The case of the Slovenian Plebisicite. Journal of the American Statistical Association 1995;90:822-828.

4. **Little R, Rubin DB**: The analysis of social science data with missing values. Sociologic Methods and Research 1989;18:292-326.

5. **Schafer JL**: "Analysis of Incomplete Multivariate Data." London: Chapman and Hall, 1997.

6. **Rubin DB**: "Multiple Imputation for Non-response in Surveys." New York: John Wiley, 1987.

7. **Rubin DB, Schenker N**: Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991;10:585-598.

8. **Brown CH**: Asymptotic comparison of missing data procedures for estimating factor loading. Psychometrika 1983;48:269-291.

9. **Gibbons RD, Hedeker D, Elkin I, et al.**: Some conceptual and statistical issues in analysis of longitudinal psychiatric data. Archives of General Psychiatry 1993;50:739-750.

10. **Ward TJ, Clark HT**: A re-examination of public versus private school achievement: the case for missing data. Journal of Educational Research 1991;84:153-163.

11. **Oshungade IO**: Some methods of handling item non-response in categorical data. The Statistician 1989;38:281-296.

12. **Furlong W, Feeny D, Torrance GW, et al.**: Mutliplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report. McMaster University Centre for Health Economics and Policy Analysis Working Paper 1998;No. 98-11.

13. **Schafer JL**: Multiple imputation: a primer. Stat Methods Med Res 1999;8:3-15.

Table 1: HUI3 Vision Attribute Questions, Response Combinations, and Attribute Level Decision Table

	Combinations									
1) During the past four weeks have you been able to see well enough to read ordinary newsprint without glasses or contact lenses ? ("Yes" -> go to question 4, "No", "Don't know", "Refused")	Y	Y	N	N	Y	N	N	N	N	N
2) Have you been able to see well enough to read ordinary newsprint With glasses or contact lenses ? ("Yes"->go to question 4, "No", "Don't know", "Refused")	--	--	Y	Y	--	Y	N	N	N	N
3) During the past four weeks, have you been able to see at all? ("Yes", "No" -> finished, "Don't Know", "Refused")	--	--	--	--	--	--	Y	Y	Y	N
4) During the past 4 weeks, have you been able to see well enough to recognize a friend on the other side of the street without glasses or contact lenses? ("Yes"->finshed, "No", "Don't know", "Refused")	Y	N	Y	N	N	N	N	Y	N	--
5) Have you been able to see well enough to recognize a friend on the other side of the street with glasses or contact lenses? ("Yes", "No", "Don't know", "Refused")	--	Y	--	-Y	N	N	Y	--	N	--
Level	1	2	2	2	3	3	4	4	5	6
# of Individuals with complete responses Total=201	44	10	89	36	4	4	4	10	0	0

Table 2: HUI3 Ambulation Attribute Questions, Response Combinations, and Attribute Level Decision Table

	Combinations							
16) During the past 4 weeks, have you been able to bend, lift, jump, and run *without difficulty and without help or equipment* of any kind? ("Yes"->finish "No", "Don't know", "Refused")	Y	N	N	N	N	N	N	N
17) Have you been able to walk around the neighborhood *without difficulty* and *without help or equipment* of any kind? ("Yes"->finish, "No", "Don't know", "Refused")	--	Y	N	N	N	N	N	N
18) Have you been able to walk around the neighborhood *with difficulty* but *without help or equipment* of any kind? ? ("Yes"->finish, "No", "Don't know", "Refused")	--	--	Y	N	N	N	N	N
19) During the past four weeks, have you been able to walk at all? ? ("Yes", "No"-> go to question 22, "Don't know", "Refused")	--	--	--	Y	Y	Y	Y	N
20) Ave you needed mechanical support, such as braces Or a cane or crutches, to be able to walk around the neighborhood? ("Yes", "No", "Don't know", "Refused")	--	--	--	Y,N	Y,N	Y,N	N	--
21) Have you needed the help of another person to walk? ("Yes", "No", "Don't know", "Refused")	--	--	--	N	N	Y	Y	--
22) Have you needed a wheelchair to get around the neighborhood? ("Yes", "No", "Don't know", "Refused")	--	--	--	N	Y	Y	Y,N	Y,N
Level	**1**	**1**	**2**	**3**	**4**	**5**	**5**	**6**

Table 3: Impact of Inspection and Deduction on Reducing Missing Data

	Before Inspection/Deduction		After Inspection/Deduction	
	Vision	Ambulation	Vision	Ambulation
Complete Responses	201	224	224	245
Missing Responses	47	24	24	3
Don't Know	*39*	*21*	*22*	*3*
Refuse/Skip Error	*8*	*3*	*2*	*0*

Table 4: HUI3 Vision Attribute Inspection and Logical Deduction

Result of "Don't Know" (21 cases)

Feasible Answers

Inspection (19 cases)

	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	Score
						N	Y		Y		2
A (18 cases)	N	Y		??	Y	N	Y		N	Y	2
						N	Y		N	N	3
B (1 case)	??	Y		N	N	Y			N	N	3

Logical Deduction (2 cases)

C (2 cases)	N	Y		?	N	N	Y		N	N	3

Result of a Skip Error (2 cases)

Logical Deduction (2 cases)

D (1 case)	N	N	0	N	N	N	N	Y	N	N	5
E (1 case)	N	Y	Y	0		N	Y		Y		2

Table 5: HUI3 Ambulation Attribute Inspection and Logical Deduction

Result of "Don't Know" (20 cases)

Inspection (15 cases)

	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Feasible Answers: Q16	Q17	Q18	Q19	Q20	Q21	Q22	Score
A (14 cases)	??	Y						Y							1
								N	Y						1
B (1 case)	N	N	N	Y	??	N	N	N	N	N	Y	Y,N	N	N	3

Logical Deduction (5 cases)

C (2 cases)	N	N	??	Y	Y	N	N								
D (2 case)	N	??	??	Y	Y	N	N	N	N	N	Y	Y	N	N	3
E (1 case)	??	N	N	Y	Y	N	N								

Result of a Skip Error (1 case)

Inspection

F (1 case)	N	N	N	N			0	N	N	N	N			Y,N	6

Table 6: HUI3 Vision Attribute Level Imputation: A comparison of imputation methods with and without prior inspection and deduction

	Mean	Standard Deviation
Complete data		
N=201 / 248 (prior to inspection)	1.960	0.730
N=224 / 248 (post inspection)	1.986	0.734
Imputation Strategy Without Step 1 Inspection		
Mean Substitution	1.960	0.660
Simple Hot Deck	1.778	0.760
Logistic Regression using ADVS	1.964	0.743
Imputation Strategy After Step 1 Inspection		
Mean Subsitution	1.986	0.694
Logistic Regression Using ADVS	1.972	0.716
Simple Model Scoring Scheme A		
Mean (weighted) Score	1.963	0.686
Hot Deck using Models	1.931	0.702
Logistic Regression Using ADVS and Models	1.935	0.700
Simple Model Scoring Scheme B		
Mean Un-weighted Score	2.274	0.774

Table 7: HUI3 Vision Attribute Imputation Patterns and Models

Model 1 (2 cases)					All possible combinations						
					Q1	*Q2*	*Q3*	*Q4*	*Q5*	*Score*	#Complete
Q1	Q2	Q3	Q4	Q5							
					Y			Y		1	44
Y			?	Y	Y			N	Y	2	10

(The random draw or regression performed on those individuals with scores who said "yes" to Q1 and not "no" to question 5, Not "no" means an answer of yes or the question was skipped)

Model 2 (9 cases)										
				N	Y		Y		2	89
				N	Y		N	Y	2	36
N	Y	?	?	N	Y		N	N	3	4

(The random draw or regression performed on those individuals who said "no" to Q1 and "yes" on Q2)

Model 3 (3 cases)									
			Y			Y		1	44
?	Y	Y	N	Y		Y		2	89

(The random draw or regression performed on those individuals who said not "no" to Q2 and "yes" to Q4)

Model 4 (3 cases)									
			Y			Y		1	44
			Y			N	Y	2	10
?	Y	?	?	N	Y		Y		2	89
				N	Y		N	Y	2	36
				Y			N	N	3	4
				N	Y		N	N	3	4

(The random draw or regression performed on those individuals who said not "no" to Q2 and not "no" to Q3)

Model 5 (2 cases)									
				Y			Y		1	44
				Y			N	Y	2	10
Y		?	?	Y			N	N	3	4

(The random draw or regression performed on those individuals who said "yes" to Q1)

Model 6 (1 case)										
				N	Y			Y		2	89
N	?	Y	Y	N	N	Y		Y		4	10

(The random draw or regression performed on those individuals who said "no" to Q1, not "no" to Q3, and "yes" to Q4)

Model 7 (2 cases)										
				Y				Y		1	44
				N	Y			Y		2	89
?	?	Y	Y	N	N	Y		Y		4	10

(The random draw or regression performed on those individuals who said not "no" to Q3 and "yes" to Q4)

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 3

Cost-effectiveness of Cataract Surgery versus Watchful Waiting: a randomized trial of patients with good visual functioning

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Introduction

Cataracts, a clouding of the eye lens, are a common disorder of aging. As a result, a significant number of elderly patients have decreased vision resulting often in difficulty performing everyday activities. Cataract surgery, in which an artificial intra-ocular lens replaces the cataract, is one of the most common outpatient surgeries performed in the United States. In the past, cataract surgery represented 5% of Medicare spending, which in 1994 represented close to $2 billion dollars [1]. The cost of cataract surgery to society will increase as the percentage of older Americans increases over the next 30 years. Even though the reimbursement for cataract surgery has been reduced, this year alone it is estimated that there will be 1.6 million cataract surgeries representing $3 billion in Medicare charges [2]. In an effort to provide high quality care that is also cost-efficient, there has been a growing demand for evidence-based support to guide the utilization of cataract surgery.

In 1993, Agency for Health Care Policy and Research, AHCPR (now the Agency for Health Care Research and Quality), asked an 18-member multidisciplinary panel of private-sector experts to perform an extensive literature review to promote appropriate management of adults with cataract-related functional impairment. The panel concluded that surgery is not justified merely because a cataract exists. One of major findings of this panel was that the appropriateness of removing a cataract depends on how much it interferes with the patient's ability to function independently, including his or her assessment of vision and lifestyle needs. Depending on the degree of functional impairment, patients, together with their eye care professionals, could then explore options such as the use of stronger eyeglasses or magnifying lenses before choosing cataract surgery [3]. The findings from the AHCPR report supporting the use of functional impairment to determine the need for cataract surgery rather than visual acuity alone have been subsequently adopted by the American Academy of Ophthalmology and American Society of Cataract and Refractive Surgery [4] [2].

Several instruments are available for assessing functional impairment related to cataract, including the Visual Function Index, VF-14, the Activities of Daily Vision Scale, ADVS, and the Visual Activities Questionnaire. Studies have provided support for the reliability and validity of these measures of patient functioning [5] [6] [7]. A study using the ADVS showed that many patients have an improvement in visual acuity without an improvement in visual functioning [8].

Even though cataract surgery is a low risk intervention after which a significant number of patients report an improvement in visual functioning, prior outcome studies have shown that a sizeable portion, between 5%-20%, of patients did not improve [9-11]. Therefore, there has been a concerted effort to distinguish between individuals who are likely to benefit or not benefit from cataract surgery. Instruments developed for this purpose would not only improve outcomes but could also save a significant amount of the Medicare budget.

Development of Prediction Rules

Several studies have attempted to develop prediction rules for cataract surgery [12] [13-15]. One study integrated the ADVS with other preoperative data to predict visual functional improvement after cataract extraction with intraocular lens implantation [15].A cataract surgery index (CSI) was developed using five preoperative clinical variables (1) age, (2) preoperative ADVS score, (3) presence or absence of a posterior sub-capsular cataract, (4) presence or absence of age-related macular degeneration, and (5) diabetes. The CSI was used in a prediction rule derived from 5 factors: for every decade over 65 years, patients receive one point; 2 points are added if there is evidence of diabetes mellitus (irregardless of the presence of retinopathy); and one point is subtracted if the patient has preoperative evidence of a posterior subcapsular cataract; and, finally, the preoperative ADVS score (range 0-100) multiplied by 0.1 is added to the total score, 2 points are added if there is evidence of macular degeneration; Patients with a predication rule score of 10 points or more are considered to have a low probability of improving (<30%) with surgery [15].

This prediction rule has been used in prior studies to successfully stratify patients into three groups in which the probabilities of substantial improvement were 85%, 34%, and 3% [15]. More recently, the CSI was evaluated in a pilot study for the current investigation whereby 182 new cataract surgery patients undergoing first-eye surgery were compared with the original sample of cataract surgery patients. The CSI was successful in discriminating patients from the new sample into two levels of improved functioning, 75% and 40%, comparable to the results when the bottom two samples of the previous study were combined.

Previous Outcome Studies

Much of the previous literature on the outcomes of cataract surgery focused on visual acuity [16]. However, more recent studies use functional visual assessment tools as well. One study that focused on 1st eye and 2nd eye surgeries [10] demonstrated that patients who underwent surgery in both eyes reported greater improvement in subjective visual function than did those who underwent surgery in one eye, indicating a benefit of binocular vision. Additional outcome studies have been performed comparing multi-focal versus mono-focal intra-ocular lens replacement [17, 18] and phaco-emulsification versus extra-capsular extraction [19]. A recent randomized trial performed on immediate second eye surgery (within 6 weeks of 1st eye surgery) versus watchful waiting (6 months) demonstrated a benefit to immediate surgery [20]. Further studies noted the benefit of 2nd eye surgery in visual acuity, visual function, and psychosocial health status, but not global and physical health status, such as ambulation [21] [22]. Most of these previous outcome studies used population samples without performing subgroup analyses. There has been some suggestion that older patients with comorbidities or macular degeneration might not benefit as much from cataract surgery [23].

Previous work on the Cost of Cataract Surgery

The cost of cataract surgery has been studied extensively over the last decade, but has also been changing due Medicare reimbursement policies. In 1991, it was felt that the typical extra-capsular cataract surgery cost $2500 constituting 3.4 billion in Medicare dollars [24]. In addition, other costs included atypical preoperative ophthalmologic tests (39 million dollars), postoperative ophthalmologic tests (7 million dollars), and perioperative medical services (18 million dollars)[24]. These costs of course dropped as cataract surgery was done more as an outpatient procedure and at facilities dedicated exclusively to performing the procedure at high volumes. Furthermore, the cost of cataract surgery varies widely between countries. An English study published in 1996 reported the cost of cataract surgery to be around $760.00 [25], whereas in Nepal the cost is below $100.00 [26].

Recent research has indicated no benefit to pre-operative medical tests for medical patients thereby reducing some of the associated costs for cataract surgery [27]. Furthermore, a

study has been performed to quantify the cost of patients waiting for cataract surgery due to additional hospitalization and home aid suggesting that direct costs for society for one year caused by 1458 patients awaiting cataract surgery with a mean waiting time of 9.8 months was approximately the same as operating 800 patients (eyes) [28]. Additionally, a recent study focused on the cost of anesthesia during cataract surgery comparing intravenous sedation with block anesthesia and an anesthesiologist present throughout the case versus oral sedation with block anesthesia and an anesthesiologist on call. Even though there was a preference among ophthalmologists for IV sedation the expected anesthesia costs per case were much greater $324 versus $42 for oral sedation and an anesthesiologist on call [29]. The most recent cost data for cataract surgery reported the average charge, payment and cost to be $3,130.00, $1,466.00, and $1,761.00 respectively [30]. The cost breakdown for the procedure was $764 for the operating room, $31 for lab tests, $5 for radiology, $249 for pharmacology, $313 for supplies, and $399 in other costs [30].

Study Goal

The goal of this study was to evaluate the cost-effectiveness of cataract surgery among patients who might be expected to benefit only marginally from surgery. Although cataract surgery is generally thought to be cost-effective, there may be a sub-group of patients for whom surgery offers only marginal benefits. Watchful waiting in this subgroup of patients may be more cost-effective. If this scenario were true, one would expect a significant role for prediction rules, based on functional assessment tools of vision, in selecting patients likely to benefit from cataract surgery. Predictive rules and models may help improve cataract surgery outcomes and save Medicare resources.

Benefit from surgery was determined using the CSI prediction rule developed from a previously published model [15]. Two hundred and fifty patients eligible for first-eye cataract surgery with low probabilities (<30%) of improvement were randomized to an intervention of watchful waiting or surgery as usual. The primary endpoint was six months post-surgery or enrollment and the primary outcome was a change in self-reported visual function as measure with the ADVS. Additionally, changes in a preference-based, utility, measure of health-related

quality of life, the Health Utility Index Mark3 (HUI3) was a secondary outcome for cost-effectiveness analysis.

Methods

Protocol

Eligibility Criteria

Patients older than 64 years with a diagnosis of bilateral cataracts, with and without other chronic eye diseases, candidates for first eye surgery, and a CSI score of 10 or more, indicating a <30% probability of predicted improvement, were eligible for enrollment to the study. Additional eligibility criteria included: English speakers, sufficient cognitive and hearing function to be able to provide informed consent and fully participate in the interview, and persons who planned to remain in the geographic area to be seen for follow-up clinical examinations. Patients enrolled were also required to be a regular patient of the ophthalmology practices enrolled. Patients who were told by their surgeon that they should or should not have surgery were considered ineligible. Those patients who were eligible for the study but refused to be randomized were asked to participate in the study as part of an observation group that would be interviewed and followed in parallel with the randomized participants.

Sample Size

The study design called for a sample size of 150 per treatment arm, large enough to detect a difference of 6 ADVS points at a 0.90 power with an alpha of 0.05. A difference of 6 ADVS points is one-third as large as the overall mean change of 16.4 (SD 15.9) points found in the original study [15], and approximately reflects a change of one line of visual acuity, the minimum measurable objective visual change. A crossover rate of 20% was included in estimating the sample size, which should have ensured 0.90 power for a minimum of 142 persons per arm. The original sample size calculation took into account the effect of clustering. However, since cataract surgery is such a simple surgery, we did not expect to see clustering at the level of the surgeon. Without clustering, the study would achieve the appropriate power as little as 115 per trial arm.

Selection of Data Collection Sites and Recruitment

Fifteen surgeons from nine private-practice clinics and one university-related ophthalmology clinic from the greater Los Angeles area were recruited to participate in the study. Offices were selected based on their volume of patients and whether adequate numbers of patients seen at these practices would meet the criteria for inclusion to the study. Surgeons identified potential eligible participants from their normal flow of patients being seen for a regular vision exam, explained the goals of the study and their treatment options, then introduced the patients to an interviewer to confirm eligibility and obtain consent.

Randomization

Once a patient gave informed consent for this trial, a computer software program was used to generate a random number, with odd and even values indicating assignment to treatment arms. The control arm for this study was surgery as usual and the intervention arm was a watchful waiting period of six months. Crossovers were defined as any surgery arm patient not having surgery initially, while crossovers in the watchful waiting arm were patients who had undergone cataract surgery within the six-month window. The study protocol required that patients assigned to the surgery arm not undergo a second-eye cataract surgery until six-months after the initial surgery. Nevertheless, the patients who did undergo second-eye cataract surgeries were not dropped from the study.

Clinical Examination

Clinical examination results include: best-corrected snellen acuity and pinhole acuity for vision worse than 20/100, the presence of posterior subcaspular cataract, any previous ocular surgeries (non-cataract), and the presence of any other ocular disease, including corneal disease, chronic uveitis, macular degeneration (dry and wet forms), glaucoma, cystoid macular edema (CME), amblyopia, diabetic retinopathy, retinal detachment, pseudoexfoliation, and any other optic nerve problem. The severity of cataract was not graded. Participants in the randomized trial were encouraged to see their ophthalmologist six and 12-months post enrollment or post surgery for a follow-up visit. Follow-up clinical information was abstracted from the medical charts.

Failure to have a follow-up visit did not constitute a protocol violation, as no reimbursement was provided to patients to cover any costs associated with the visits.

Questionnaires

The ADVS [31], HUI3 [32], SF-12 [33], and Charlson Comorbidity Scale [34] were all administered at baseline and 6 months. The 6-month follow-up interview was conducted via telephone.

The ADVS measures a persons ability to perform 20 common visual activities representing five domains: night driving, daytime driving, distance vision activities that do not require driving, near vision activities, and activities that are subject to glare. An overall total visual function score was calculated that ranges from 0-100, where 100 represents no difficulties and 0 indicates that activities are no longer performed because of visual impairment [31].

The HUI Mark3 [32] is composed of 40 questions that measure eight health attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain) and one question that measures overall self-assessed health status. The HUI3 includes categorical information in the form of attribute levels (best score for level =1.00, score for most disabled=6.00). These attribute scores can then be converted to single attribute utility scores, ranging from 0 to 1. For example, the single attribute vision utility ranges from 1.00, representing perfect vision state, to 0.00, representing a blind state. Furthermore, a weighted-scoring algorithm is applied to combine the scores for each attribute to derive a value between zero and one to represent utility of the overall health state (perfect health=1.00, deceased=0.00). Non-response to specific questions within each attribute was dealt with using an imputation strategy previously developed. Although the utility representing the overall health state is the key outcome for cost-effectiveness analysis, this outcome was compared to the single attribute utility for vision and the ADVS functional outcome score.

The SF-12 [33] is a short-form version of the SF-36 [35], one of the most commonly used measures of self-reported health status, and yields summary scores representing physical health (PCS) and mental health (MCS). Both the PCS and MCS are T-scores normed to a U.S. population sample with a mean of 50 and a standard deviation of 10.

The Charlson Comorbidity Scale [34] asks patients whether they have any of the 16 medical conditions and, if so, whether they have received treatment for it and, separately, whether the condition limits them in their activities. For the analysis presented here, the comorbidity score represents an unweighted sum across all 16 items.

Cost/Utilization Surveys

Cost data for this study was derived from several sources. Health care utilization and the direct and indirect costs from the standpoint of the patient was derived from the monthly surveys. Each monthly survey asked the following: (1) the number of doctor visits and the cost of each visit, (2) the transportation costs and amount spent on the trip, (3) lost days from work of caretakers, (4) other eye procedures performed and associated costs, (5) obtaining new glasses and cost of the glasses, (6) the use of eye medications and medication costs, and (7) the use of home care as a result of decreased vision and associated cost. An imputation method was designed to reduce the missing information in these monthly surveys. Since expenses and utilization demonstrated a clustered pattern among patients with complete records, the imputation method used for missing months was to use the value of the expenditures and utilization in the prior month. This imputation method was compared to other strategies, such as mean substitution and hot decking, and felt to be superior given that in most months there were cost and utilization was zero. Cost differences were compared between the two treatment arms.

The cost of cataract surgery will be broken up into 4 major components: (1) peri-operative medical care, (2) anesthesia care, (3) surgery, and (4) facilities. Cost data for these services will be determined using previously published articles and summaries from Medicare billing data from the Health Care Financing Administration, HCFA. Furthermore, additional cost estimates for items such as glasses, medications, and doctor visits, in addition to the cost of other eye procedures were estimated using 2001 Average Medicare Payment Information. These estimated costs are further analyzed during sensitivity analysis.

Statistical Analysis

Participants in each of the treatment arms of the trial were compared on baseline demographic variables and clinical characteristics to study-eligible persons who refused to be randomized. A special focus was paid to characteristics that comprise the cataract surgery index, CSI. Comparability between the treatment arms was assessed using the same demographic and clinical variables.

General linear models, using STATA 6.0 software, were used to assess differences between treatment arms on ADVS and HUI scores controlling for baseline CSI components (ADVS score, age, diabetes, PSC cataract, and macular degeneration), gender, number of medical comorbidities, SF-12 physical and mental health summary scores, and visual acuity in the operated eye. The CSI components were included as covariates to adjust for effects that might be due to clinical severity related to cataracts, whereas the medical comorbidities and SF-12 scores were included to adjust for general health status. Gender was included as an adjuster to control for the tendency of women to self-report slightly worse health on quality of life surveys [36]. The data were also analyzed breaking the trial into three subgroups based on CSI score (=10,11,>11). Even though the trial population was restricted to individuals with a CSI score equal to or greater than 10 representing a less than 30% overall probability of improvement with surgery, the population is still heterogeneous in that those with higher CSI scores had lower predicted probabilities of improvement. Analysis was performed using two models: (1) an intention-to-treat model with crossovers included in their original assigned study arm, and (2) treatment-received model with crossovers reassigned. Crossovers are useful in that they demonstrate the likely mixed treatment that patients would receive with any policy or guideline change. In this study, the crossover time frame is artificially set at 6 months, reflecting the primary endpoint of the study.

Cost-effectiveness Analysis

The outcome and cost data is used to derive a cost per quality of life year, Cost/QALY, ratio. In determining this ratio, a horizon for the benefit from cataract surgery of 1 year was chosen. The benefit used for cost-effectivness was the point estimate from the models adjusted

for covariates (see above). The cost-effectiveness analysis was performed on the Intention-to-treat model and the Treatment-received model. Subgroup analysis based on CSI score was performed in both models.

Sensitivity Analysis

In addition to the subgroup analysis, sensitivity analysis was performed looking at two additional outcome measures, functional visual gain and visual line of acuity gain. Functional vision gained was defined as 7 ADVS points, the minimum clinically measurable gain in prior studies [8]. Cost-effectiveness was analyzed using these two outcome measures. Further, the cost effectiveness of cataract surgery at baseline utility +/- SD and baseline cost +/- 50% was performed on both the Intention-to-treat and Treatment-received models.

A second approach in determining HUI3 total utility gain was used. Under the assumption that noise in total utility was due to attributes other than vision, a total utility score was derived holding all other attributes constant. The mean multi-attribute utility scores at baseline for all attributes other than vision were used to calculate total utility at both baseline and at 6 months. In this model, the only aspect that changed was multi-attribute visual utility scores.

A third approach to determining HUI3 utility gain by converting ADVS gain into utility was also explored. The conversion factor was determined by regressing change in utility (6 months – baseline) onto change in ADVS. This gain was then used for an alternative set of cost-effectiveness calculations.

Results

RCT Participants vs. Refusers

The division of eligible participants and participants that refused the trial (RCT Refusers) is displayed in **Figure 1**, as is the division of randomized participants into the surgery arm and the watchful waiting arm. As seen in **Figure 1**, there were 250 RCT participants and 163 refusers. Tests of differences between RCT participants and refusers were conducted on baseline measures of demographic variables, clinical variables, and outcome variables. The results are

displayed in **Tables 1,2, and 3**. Several differences were found on the demographic variables (see **Table 1**). A larger percentage of RCT participants (36%) lived with their spouses than refusers (13%). In addition, a larger percentage of RCT participants lived with relatives (10%), than refusers (2%). Furthermore, a greater percentage of RCT participants (5%) lived with non-relatives than refusers (1%). No significant differences were found between RCT participants and refusers on the following clinical variables at baseline: acuity, PSC cataract, glaucoma, macular degeneration, previous surgeries, and other diseases (see **Table 2**). No significant differences were found on baseline measures of ADVS, PCS-12, MCS-12, medical comorbitities, and CSI score (see **Table 3**).

Surgery vs. Watchful Waiting Arms

Of the 250 RCT participants, 133 were assigned to the surgery arm, and 117 were assigned to the watchful waiting arm. Tests of the differences between the trial arms were conducted on baseline measures of demographic, clinical, and outcome variables. The results are displayed in **Tables 4,5, and 6**. As seen in these tables, no significant differences were found between trial arms. It should be noted that 16 participants in the surgery arm withdrew and 17 participants in the watchful waiting arm withdrew from the trial by six months (**Figure 1**). Thus the sample sizes in the surgery arm and watchful waiting arm were 117 and 100 participants respectively at 6 months.

Crossovers

A total of 32 participants crossed over during the trial. Twenty patients crossed over from the surgery arm to the watchful waiting arm, and twelve crossed over from the watchful waiting arm into the surgery arm. Intention-to-treat analyses of differences between trial arms on key outcome variables are presented next, followed by actual treatment-received analyses of differences.

Intention-to-treat and Treatment-received Models of Outcome Variables ADVS and HUI3 at 6 months

Difference between trial arms on the ADVS score and HUI3 score at 6 months were tested using a linear regression model. In this model the dependent variable was either 6-month ADVS or HUI3 score, and the main independent variable was treatment arm. The multivariate regression included the following covariates: baseline ADVS score or HUI3 score, respectively, age, diabetes, PSC, macular degeneration, PCS-12, MCS-12, gender, and medical comorbidities. As a result of missing data, the final analysis on the ADVS and HUI3 outcome measures could only be performed on a total of 209 and 212 participants respectively.

After case-mix adjustments, those individuals in the surgery arm had a mean (M) improvement in ADVS score of 6.51 points (standard deviation (SD) =1.72) compared to those in the watchful waiting arm, which was statistically significant to a p-value of <0.0001 (**Table 7**). The mean improvement in utility, HUI3 score was 0.041 points for those in the surgery arm but this value was not statistically significant (**Table 7**). Breaking the total sample into 3 sub-groups based on CSI score (=10, =11, >11) allowed a more in-depth analysis. The group of individuals with a CSI score equal to 10 had the greatest improvement in ADVS score with surgery (M=13.04, SD=3.15) followed by with CSI scores equal to 11 (M=7.69, SD=2.50) (**Table 7**). Both these first two groups had statistically significant results. The third group with CSI scores greater than 11 did not benefit. The subgroup analysis using HUI3 as the outcome measure did not yield any significant results, however, the pattern in terms of the magnitude of mean response was identical to ADVS model.

The analysis was repeated with the some covariates with the crossovers reassigned in the Treatment-received model. In this model, individuals in the surgery arm had a mean (M) improvement in ADVS score of 6.98 points (SD=1.62 than those in the watchful waiting arm, which was statistically significant to a p-value of <0.0001 (**Table 8**). The mean improvement in utility, HUI3 score was 0.032 points greater for those in the surgery arm but this value was not statistically significant (**Table 8**). The individuals with a CSI score equal to 10 and 11 had mean improvements in ADVS score with surgery of 10.92 (SD=3.21) and 10.15, (SD=2.40) respectively (**Table 8**). Both these first two groups had statistically significant results. The third group with CSI scores greater than 11 did not benefit. The subgroup analysis in the Treatment-

received model using utility as the outcome measure did not yield any significant results, however, the pattern in terms of the magnitude of mean response was identical to ADVS model. The impact of surgery on the outcome measures was greater in the Treatment-received model compared to the Intention-to-treat Model.

Single Attribute Utility Analysis using the Intention-to-treat and Treatment-received Models

The HUI3 is composed of 8 attributes. Each of these attributes was evaluated independently using the same analytic framework and covariates except that baseline single attribute utility was used instead of baseline total utility. In the Intention-to-treat model, single vision utility had the only statistically significant improvement with surgery with a mean gain of 0.031 points (SD=0.014), while improvement in cognition approached statistical significance with a mean gain of 0.022 points (SD=0.012) (**Table 9**). In the Treatment-received model, the results were virtually identical with respect to visual utility gain (**Table 10**). There was no difference between the arms in utility gain cognition.

Cost and Resource Utilization Differences between Trial Arms over 6 months

Cost and resource utilization analysis was performed in the Intention-to-treat model for the 6-month period. The cost analysis was broken up into four groups: Non-surgery Costs, Non-Health Care Costs, Patient Co-pays, and Surgical Costs. There were two non-surgical variables with significant differences between the treatment arms: doctor visits and glasses, both excluding co-pay. Those in the surgical arm had more doctor visits with a mean cost of $207.89 (SD=16.40) and had more glasses ordered at a mean cost of $14.48 (SD=2.91) compared to those in watchful waiting who had a mean cost for doctor visits of $139.96 (SD=11.01) and mean cost for glasses of $4.64 (SD=2.28). In terms of non-healthcare costs, surgery patients traveled a greater distance for visits with a mean cost of $9.58 (SD=3.03) compared to those in watchful waiting who paid a mean total mileage cost of $2.49 (SD=0.75), see **Table 11**. In terms of patient co-pays, those in the surgical arm had a greater co-pay for glasses with a mean co-pay of $30.15 (SD=7.55) compared to those in the watchful waiting with a mean co-pay of $5.69 (SD=2.75) (**Table 11**). Surgical procedures, other than the initial 1st eye cataract surgery, were

also important. There were five 2nd eye cataract surgeries in the surgical arm, a protocol violation, and none in the watchful waiting group. Among the watchful waiting arm, there were three tear duct procedures, whereas there were none in the surgical arm. It should be noted that due to crossovers there were 12 patients in the watchful waiting arm that received surgery and 20 in the surgery arm that did not. There was statistically significant difference between the randomized arms in all four cost sub-scales (**Table 11**). The total cost of cataract surgery was $1,975 (SD=73.78) for those in the surgery arm compared to $407 (SD=76.16) in the watchful waiting arm, a difference significant to the $p<0.0001$ level and driven mainly by the surgical cost of 1st eye cataracts. The difference between the arms was $1,567 (SD=106.04).

Cost and resource utilization analysis was also performed in the Treatment-received model for the 6-month period. There were six non-surgical variables with significant differences between the treatment arms: doctor visits, visit co-pay, mileage cost, glasses cost, glasses co-pay, and number of eye medications. Those in the surgical arm had more doctor visits with a mean cost of $231.66 (SD=17,33), a greater visit co-pay with a mean of $10.81 (SD=3.66), and traveled a greater distance for visits with a mean mileage cost of $10.73 (SD=3.03) compared to those in watchful waiting who had a mean cost for doctor visit of $110.17 (SD=3.15), a mean visit co-pay of $1.17 (SD=0.54) and mean mileage cost of $1.01 (SD=0.26) respectively (**Table 12**). In addition, those in the surgical arm had more glasses ordered with a mean cost of $16.74 (SD=3.27), a greater co-pay for glasses with a mean of $33.11 (SD=7.62), and a greater mean number of eye medication costing $16.60 (SD=2.78) compared to those in the watchful waiting group with a mean cost of $1.76 (SD=1.00) for glasses, glasses co-pay of $1.81 (SD=1.61), and eye medication cost of $7.09 (SD=1.86) respectively (**Table 12**). Two surgical procedures, other than the initial 1st eye cataract surgery, were significant. There were five 2nd eye cataract surgeries in the surgical arm, a protocol violation, and none in the watchful waiting group. Among the watchful waiting arm, there were three tear duct procedures, whereas there were none in the surgical arm. There was statistically significant difference between the randomized arms in all four cost sub-scales (**Table 12**). The total cost of cataract surgery was $2,183 (SD=49.28) for those in the surgery arm compared to $136 (SD=8.93) in the watchful waiting arm, a difference significant to the $p<0.0001$ level and driven mainly by the surgical cost of 1st eye cataracts. The difference between the arms was $2,047 (SD=50.08).

Cost was also determined for subgroups based on CSI score. In the Intention-to-treat model the costs were as follows: $1,803 for a CSI=10, $1,639 for a CSI=11, and $1,284 for a CSI>11 (**Table 13, Part A**). The variation in cost difference between arms based on CSI subgroups was due largely to crossover differences between the subgroups. In the Treatment-received model the subgroup costs were as follows: $2,030 for a CSI=10, $2,142 for a CSI=11, and $1,938 for a CSI>11 (**Table 13, Part B**). The costs were broken down for these subgroups in a similar fashion and using the same methodology as in the total group analysis (data not shown).

Cost-effectiveness Analysis

Based on the gains in utility with surgery in the Intention-to treat model (**Table 7**), the gain in utility was calculated to be 0.041 utility points for the overall group over a 1-year horizon of benefit. Based on a cost difference of between the two arms of the trial as discussed in the prior section, the cost-effectiveness of cataract surgery was $38,228 per quality adjusted life years, QALY (**Table 13, Part A**). Subgroup analysis based on CSI scores demonstrated a cost-effectiveness of cataract surgery $31,638/QALY and $37,250/QALY for CSI score of 10 and 11 respectively. For CSI scores greater than 11, the cost-effectiveness was $53,500/QALY (**Table 13, Part A**). Based on the gains in ADVS outcome with surgery in the Treatment-received model (**Table 8**), the gain in utility was calculated to be 0.032 points for the overall group over a 1-year horizon of benefit. Based on a cost difference of between the two arms of the trial as discussed in the prior section, the cost-effectiveness of cataract surgery was $63,972/QALY (**Table 13, Part B**). Subgroup analysis based on CSI scores demonstrated a cost-effectiveness of cataract surgery $40,599/QUALY and $82,369/QALY for CSI score of 10 and 11 respectively. For CSI scores greater than 11, the cost-effectiveness was $138,415/QALY (**Table 13, Part B**).

Sensitivity Analysis

In addition to the subgroup analysis, sensitivity analysis was performed looking at two additional outcome measures, functional visual gain and visual line acuity gain. Functional vision gained was defined as 7 ADVS points, noted as FADVS. In the Intention-to-treat model, the overall cost/function effectiveness was $1,685/FADVS. Subgroup analysis was as follows:

$968/FADVS for CSI=10, $1,492/FADVS for CSI=11, and $6,073/FADVS for CSI>11 (**Table 14A**). In the Treatment-received model, the overall cost/function effectiveness was $2,053/FADVS. Subgroup analysis was as follows: $1,301/FADVS for CSI=10, $1,477/FADVS for CSI=11, and $8,478/FADVS for CSI>11 (**Table 14B**).

As part of this analysis, cost per line of visual acuity was also calculated. In the Intention-to-treat model the overall cost per line of visual acuity, denoted as lineVA, was $1,318/lineVA. Subgroup analysis was as follows: $985/lineVA for CSI=10, $1,490/lineVA for CSI=11, and $1,629/lineVA for CSI>11 (**Table 15A**). In the Treatment-received model the overall cost per line of visual acuity was $1,516/lineVA. Subgroup analysis was as follows: $1,187/lineVA for CSI=10, $1,610/lineVA for CSI=11, and $1,846/lineVA for CSI>11 (**Table 15B**).

Lastly, a utility and cost sensitivity analysis was performed using the initial regression point estimate (mean) and standard deviation (**Tables 7 and 8**). This analysis was performed on the overall group. Threshold values for what is acceptable cost-effectiveness for an intervention vary, usually between $50,000/QALY and $100,000/QALY [37]. For out analysis a threshold of $50,000/QALY was arbitrarily selected. In the Intention-to-treat model, cataract surgery is cost-effective at baseline, mean utility and cost value). However if cost estimated increased by 50% or utility decreased by one standard deviation, cataract surgery fell below the cost-effectiveness cut-off (**Table 16A**). In the Treatment-received model, cataract surgery was over the cut-off value at baseline, but an increase by one standard deviation in utility or a decrease of costs by 50% made cataract surgery more cost-effective (**Table 16B**). For both these models the baseline cost-effectiveness derived using the regression point estimates is similar to the cost-effectiveness using the conversion factor between ADVS and utility. Additionally, it should be noted that a decrease in utility by two standard deviations would mean negative utility, which would make cataract surgery not a viable option.

Since the models using total HUI3 utility scores were not significantly different between the two treatment groups, two separate additional analyses were performed. In the first, all attributes other than vision were held constant in the calculation of total utility. Those individuals with CSI scores of 10 and 11 were combined in one group and those with CSI scores of 12 and above were in a second group. Using this strategy in the Intention-to-Treat Model, a statistically significant gain of 0.023 utility points was seen with cataract surgery in patients with CSI scores of 10 and 11 (**Table 17A**). Based on this benefit, the cost-effectiveness of cataract surgery was

$73,286/QALY for those with CSI scores of 10 and 11. The point estimate for utility gain with surgery for those with a CSI score of 12 or greater was 0.01 and was not statistically significant. Based on this benefit, the cost-effectiveness of surgery for those with CSI of 12 and greater was $128,401/QALY.

Using this strategy in the Treatment-received Model, there was a significant gain with surgery was 0.023 for those with CSI scores of 10 and 11, and a 0.01 non-significant gain for those with CSI scores 12 or greater (**Table 17 B**). Based on this benefit, the cost-effectiveness of cataract surgery was $90,926/QALY for those with CSI scores of 10 and 11 and $193,781/QALY for those with CSI scores of 12 or greater.

A separate analysis was performed to try to correlate the gain in ADVS points with surgery to utility change. A linear regression model was used regressing the change in ADVS score (6-month value minus baseline) on the change in HUI score (6-month value minus baseline) for each individual. This model demonstrated that a 1-point gain in ADVS score translates to a mean utility gain of 0.0042 points (SD=0.0012). This relationship was statistically very significant with a p-value <0.0001.

A conversion rate of 0.0042 utility points for every 1-point in ADVS score was used for the subsequent cost-effectiveness analysis. Based on the gains in ADVS outcome with surgery in the Intention-to treat model (**Table 7**), the gain in utility was calculated to be 0.027 utility points over a 1-year horizon of benefit. Based on a cost difference of between the two arms of the trial as discussed in the prior section, the cost-effectiveness of cataract surgery was $57,323/QALY (**Table 18, Part A**). Subgroup analysis based on CSI scores demonstrated a cost-effectiveness of cataract surgery $32,928/QALY and $50,747/QALY for CSI score of 10 and 11 respectively. For CSI scores greater than 11, the cost-effectiveness was $206,565/QALY (**Table 18, Part A**). Based on the gains in ADVS outcome with surgery in the Treatment-received model (**Table 8**), the gain in utility was calculated to be 0.029 points over a 1-year horizon of benefit. Based on a cost difference of between the two arms of the trial as discussed in the prior section, the cost-effectiveness of cataract surgery was $69,829/QALY (**Table 18, Part B**). Subgroup analysis based on CSI scores demonstrated a cost-effectiveness of cataract surgery $44,261/QALY and $50,237/QALY for CSI score of 10 and 11 respectively. For CSI scores greater than 11, the cost-effectiveness was $288,365/QALY (**Table 18, Part B**).

Discussion

Cataract surgery has been shown to increase both subjective and objective visual functioning in many individuals. This paper is focused on determining if a subgroup of patients can be isolated for whom cataract surgery would not be cost-efficient. The data from this randomized trial demonstrates that the prediction rule [15], developed by Mangione et al. and based on the ADVS instrument, can successfully identify a population of patients for whom watchful waiting is relatively more cost-effective than surgery.

This study's primary outcome was change in ADVS. Since prior data ADVS data was available, determining an adequate sample size for a change in ADVS was straightforward. On the other hand, no prior data on the impact of visual change on overall utility using the HUI3 instrument was available. This trial was probably not powered to provide statistically significant results for small changes in utility. To detect a significant 0.01 change in utility between the two arms of this trial would have required a sample size between 1000-2000 individuals or ten times more than in our trial. The significant change in single attribute visual utility supports the notion that there would be an overall utility benefit from cataract surgery had this study enrolled a sufficient number of individuals. Since our direct measurement of utility was not significant, our sensitivity analysis focused on alternative approaches for additional validation of our estimated benefit. Determining utility gain by conversion of ADVS gain yielded point estimates and trends among the subgroups from the initial regression models matched the alternative method used in this paper to calculate cost-effectiveness.

The change in visual functioning and visual acuity in this study was less dramatic than reported by Javitt et al.[10]. Our study enrolled patients with relatively good baseline vision whereas prior studies focused on a more representative sample of patients undergoing cataract surgery. It should be noted that the study by Javitt et al. focused on outcomes at 1 year whereas this paper focuses on 6-month outcomes. Furthermore, a significant number of patients in the Javitt study had 2nd eye and bilateral eye cataract surgery.

In addition to the difficulty to determining the exact utility benefit, there is some uncertainty as to which cost estimates are most appropriate. The in-depth cost analysis of cataract surgery by Steinberg et al. in 1991 calculated a total cost of cataract surgery at $2500.00 [24]. The Medicare reimbursement fees change annually and there has been a significant

reduction in reimbursement for cataract surgery over the last decade. The cost estimates used in this paper reflect a cross section of outpatient centers that were surveyed in 1998. It may be the cost has decreased significantly over the last 3 years. However, our sensitivity analysis demonstrated that for the most part, the findings of this paper will hold true even with a 50% change in overall cost.

The cost-effectiveness of cataract surgery in our trial population is comparable to other interventions such as: screening mammography ever 2 years for women between the ages of 40 and 70 ($70,000/QALY), adjuvant chemotherapy for 75 year old with an node (-), estrogen (-) early stage breast cancer ($58,000/QALY), estrogen treatment in a healthy 50 year old women entering menopause ($56,000/QALY), and lovastatin for cholesterol reduction ($46,000/QALY) [38] This study though suggests that using a prediction rule can make cataract surgery more cost-effective. Furthermore, the cost to gain 1 line of visual acuity was modest ($1,320). Given that the cut-off for a valid driver's license is 20/40 vision, the benefit of cataract surgery, in the ability to drive, may far exceed the cost in a patient with 20/60 or 20/80 vision.

Although this randomized trial focused on enrolling patients deemed to have a low probability of benefit, <30%, from cataract surgery, two subgroups (patients with CSI scores 10 and 11) were identified for whom cataract surgery is cost-effective. By medically managin patients with a CSI>11 for whom cataract surgery is not cost-effective, overall outcomes can be improved and Medicare money can be saved. The subgroup of patients with CSI>11 represent about 5-10% of patients annually undergoing cataract surgery. By electing instead to have these patients undergo watchful waiting instead of surgery, Medicare would save between $10-20 million dollars annually.

An equally important question is whether screening cataract surgery patients in order to determine a CSI score is cost-effective. If one assumes that of every 100 patients only 5 will have CSI scores of 12 or greater, then the maximum saving in terms of delaying cataract surgery would be around $6,450 (5*$1,290). This maximum savings assumes that these patients will never have surgery. If a third eventually have surgery and surgery is only delayed by 6 months, the saving would be only $4,257. The cost of screening to derive a CSI score would therefore need to be less than $42.57-64.50 for it to be cost-effective. Most of the cost of screening would be to administer the ADVS questionnaire since the other elements of the CSI are already derived from a standard eye visit with an ophthalmologist. Even though the ADVS questionnaire is long,

if self-administered or administered with the aid by a dedicated individual, it is reasonable to assume the cost will be within the bounds necessary for screening to be cost-effective and perhaps save Medicare additional resources as well.

There were several limitations in this study. The decision rule used for entry into the trial was applied to patients considered eligible for cataract surgery by an ophthalmologist. However, there are many patients with cataracts who are never referred to an ophthalmologist, especially among Hispanic and African-American patients. Even though this study included a good representation of minorities it is difficult to be certain if this population is reflective of general population. Thus, this trial probably only screened only a subgroup of the potential patient pool in the community. Furthermore, it may be that patients willing to enter a randomized trial represent a special group with inherent unmeasured differences from the general population. In addition, only about 20% of the ophthalmology practices approached agreed to participate in the trial. Even though the participating practices were geographically representative, these practices may have had a more academic mentality with a deeper interest in risk stratification of patients to improve outcome.

Conclusion

This study has demonstrated that a prediction rule can be used to discriminate patients for whom cataract surgery is not likely to improve outcome and for whom cataract surgery is not cost-effective. In order to develop a more precise estimate of utility gained from cataract surgery, a larger trial may be needed. Furthermore, the cost-effectiveness of cataract surgery is largely dependent on reimbursement and surgical volume, which tend to move in opposite directions. As the proportion of elderly in the United States increases, there will be a greater need to develop methods that better allocate the diminishing Medicare resources. The strategy of developing and employing prediction rules to better select subgroups of patients for procedures will be a key element in reformulating our approach to healthcare provision for older Americans. Even though the majority of cataract surgeries provide benefit and are cost-effective, a significant amount of resources will be saved using watchful waiting in a sub-group of patients with good visual functioning.

References

1. HCFA, H.F.A., *Historical Data (1996) Green Book.* 1996.
2. AAO, A.A.o.O.A.S.P., *Cataract in theAdult Eye.* 2001, American Academy of Ophthalmology. p. 1-62.
3. Panel, C.M.G., *Cataract in Adults: Management of Functional Impairment.* AHCPR Pub. No. 93-0542, 1993.
4. Ophthalmology, A.A.o. and A.S.o.C.a.R. Surgery, *White Paper on Catarct Surgery.* 1996.
5. Steinberg, E.T.J.S.O.J.J.S.P.C.S.L.M.D.-W.M.B.E.D.A.S.D.S.A., *The VF-14. An Index of Functional Impairment in Patients with Cataract.* Arch Ophthalmol, 1994. **112**: p. 630-638.
6. ME, S., et al., *The Visual Activities Questionnaire: Developing an Instrument for Assessing Problems in Everyday Visual Tasks.* Technical Digest, Noninvasive Assessment of the Visual System, 1992. **1**: p. 26-29.
7. Mangione, C.P.R.S.J.L.M.C.E.D.R.G.L., *Development of the Activities of Daily Vision Scale - A Measure of Visual Functional Status.* Med Care, 1992. **30**: p. 1111-1126.
8. Mangione, C.M., et al., *Improved visual function and attenuation of declines in health-related quality of life after cataract extraction.* Arch Ophthalmol, 1994. **112**(11): p. 1419-25.
9. Applegate, W.B., et al., *Impact of cataract surgery with lens implantation on vision and physical function in elderly patients.* Jama, 1987. **257**(8): p. 1064-6.
10. Javitt, J.C., et al., *Outcomes of cataract surgery. Improvement in visual acuity and subjective visual function after surgery in the first, second, and both eyes.* Arch Ophthalmol, 1993. **111**(5): p. 686-91.
11. Steinberg, E.P., et al., *National study of cataract surgery outcomes. Variation in 4-month postoperative outcomes as reflected in multiple outcome measures.* Ophthalmology, 1994. **101**(6): p. 1131-40; discussion 1140-1.
12. Graney, M.J., et al., *A clinical index for predicting visual acuity after cataract surgery.* Am J Ophthalmol, 1988. **105**(5): p. 460-5.

13. Schein, O.D., et al., *Predictors of outcome in patients who underwent cataract surgery.* Ophthalmology, 1995. **102**(5): p. 817-23.

14. Tielsch, J.M., et al., *Preoperative functional expectations and postoperative outcomes among patients undergoing first eye cataract surgery.* Arch Ophthalmol, 1995. **113**(10): p. 1312-8.

15. Mangione, C.M., et al., *Prediction of visual function after cataract surgery. A prospectively validated model.* Arch Ophthalmol, 1995. **113**(10): p. 1305-11.

16. Desai, P., *The National Cataract Surgery Survey: II. Clinical outcomes.* Eye, 1993. **7**(Pt 4): p. 489-94.

17. Javitt, J.C., et al., *Outcomes of cataract extraction with multifocal intraocular lens implantation: functional status and quality of life.* Ophthalmology, 1997. **104**(4): p. 589-99.

18. Javitt, J., et al., *Cataract extraction with multifocal intraocular lens implantation: clinical, functional, and quality-of-life outcomes. Multicenter clinical trial in Germany and Austria.* J Cataract Refract Surg, 2000. **26**(9): p. 1356-66.

19. Castells, X., et al., *Clinical outcomes and costs of cataract surgery performed by planned ECCE and phacoemulsification.* Int Ophthalmol, 1998. **22**(6): p. 363-7.

20. Laidlaw, D.A., et al., *Randomised trial of effectiveness of second eye cataract surgery.* Lancet, 1998. **352**(9132): p. 925-9.

21. Castells, X., et al., *Comparison of the results of first and second cataract eye surgery.* Ophthalmology, 1999. **106**(4): p. 676-82.

22. Lundstrom, M., U. Stenevi, and W. Thorburn, *Quality of life after first- and second-eye cataract surgery(1). Five-year data collected by the Swedish National Cataract Register.* J Cataract Refract Surg, 2001. **27**(10): p. 1553-9.

23. Wong, T.Y., *Effect of increasing age on cataract surgery outcomes in very elderly patients.* Bmj, 2001. **322**(7294): p. 1104-6.

24. Steinberg, E.P., et al., *The content and cost of cataract surgery.* Arch Ophthalmol, 1993. **111**(8): p. 1041-9.

25. Cresswell, P.A., et al., *Cost effectiveness of a single-function treatment center for cataract surgery.* J Cataract Refract Surg, 1996. **22**(7): p. 940-6.

26. Marseille, E., *Cost-effectiveness of cataract surgery in a public health eye care programme in Nepal.* Bull World Health Organ, 1996. **74**(3): p. 319-24.

27. Schein, O.D., et al., *The value of routine preoperative medical testing before cataract surgery. Study of Medical Testing for Cataract Surgery.* N Engl J Med, 2000. **342**(3): p. 168-75.

28. Stenevi, U., M. Lundstrom, and W. Thorburn, *The cost of cataract patients awaiting surgery.* Acta Ophthalmol Scand, 2000. **78**(6): p. 703-5.

29. Reeves, S.W., et al., *A decision analysis of anesthesia management for cataract surgery.* Am J Ophthalmol, 2001. **132**(4): p. 528-36.

30. *Four costliest outpatient procedures.* Hosp Health Netw, 1998. **72**(18): p. 32-3, 3.

31. Mangione, C.M., et al., *Development of the 'Activities of Daily Vision Scale'. A measure of visual functional status.* Med Care, 1992. **30**(12): p. 1111-26.

32. Furlong, W.J., et al., *The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies.* Ann Med, 2001. **33**(5): p. 375-84.

33. Ware, J., Jr., M. Kosinski, and S.D. Keller, *A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity.* Med Care, 1996. **34**(3): p. 220-33.

34. Charlson, M., et al., *Validation of a combined comorbidity index.* J Clin Epidemiol, 1994. **47**(11): p. 1245-51.

35. McHorney, C.A., J.E. Ware, Jr., and A.E. Raczek, *The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs.* Med Care, 1993. **31**(3): p. 247-63.

36. Ware, J.E., Jr., et al., *SF-36 Health Survey Manual and Interpretation Guide.* 1993, Boston: Health Institute, New England Medical Center. Table 10.1.

37. Laupacis, A., et al., *How attractive does new technology have to be to warrant adoption and utililization? Tentative guidelines for using clinical and economic evaluation.* Can Med Assoc J, 1992. **146**: p. 473-81.

38. Harvard, H.C.F.R.A., *The CUA Data Base: Standardizing the Methods and Practices of Cost-Effectiveness Analysis.* 1971, Harvard School of Public Health.

Figure 1. Pepper Trial Consort Table

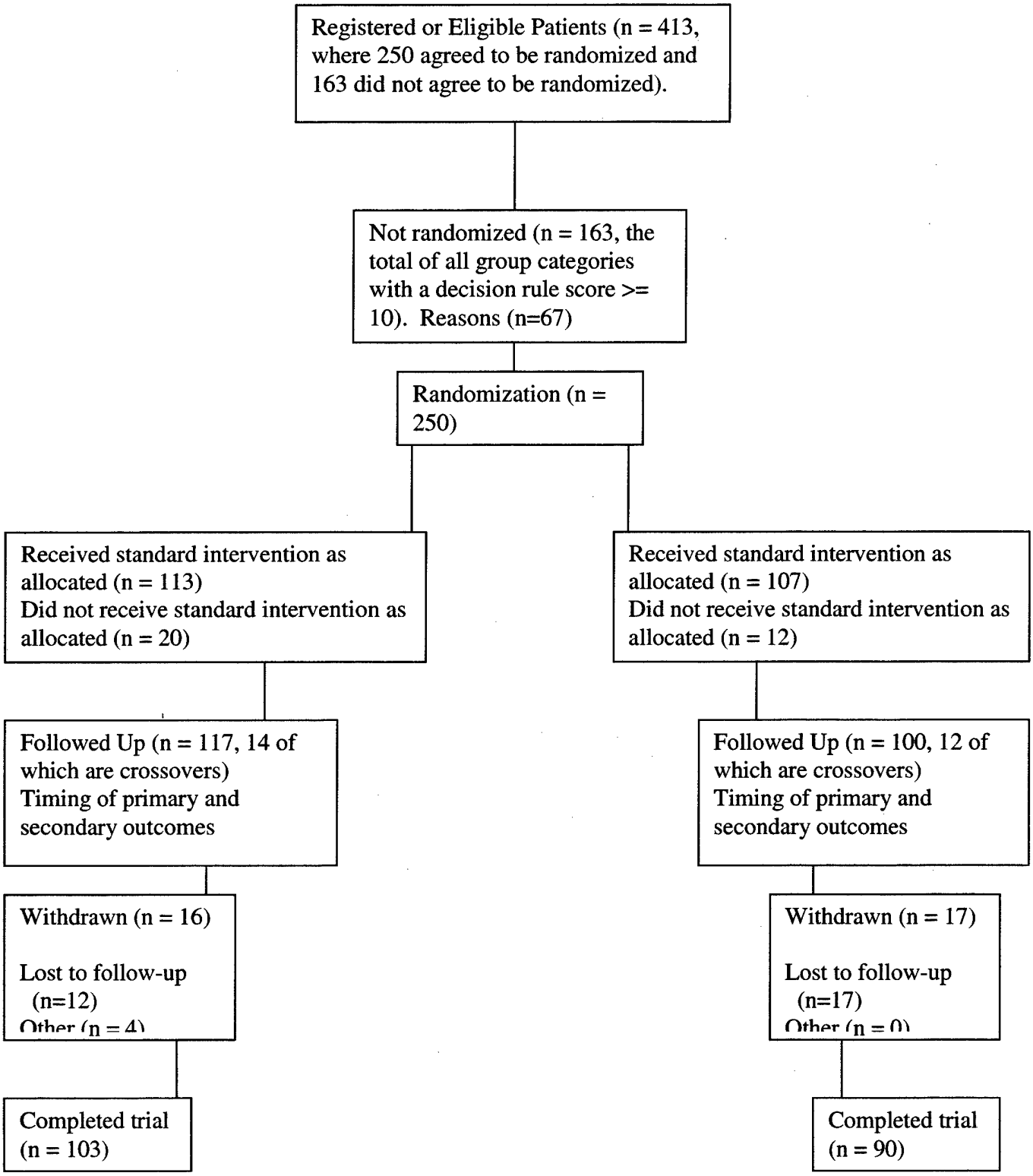


Table 1

Descriptive Statistics on Demographic Variables in RCT Participant and Refuser Groups and Significance Tests of Group Differences.

Variable	RCT Participants	Refusers	Test Statistic	p
Age	M = 78.8 (5.9)	M = 78.7 (5.5)	t(409)= .07	.95
Gender				
Male	96 (38%)	56 (35%)	χ^2 (1) = .55	.46
Female	154 (62%)	105 (65%)		
Race				
White	193 (77%)	132 (81%)		
African American	30 (12%)	18 (11%)		
Asian/Pacific Islander	10 (4%)	7 (4%)		
Hispanic	14 (6%)	4 (2%)	χ^2 (5) = 4.40	.49
Multi-Racial	3 (1%)	1 (1%)		
American Indian	0 (0%)	1 (1%)		
Education				
8th grade or less	13 (7%)	4 (9%)		
Some HS	27 (14%)	3 (7%)		
HS graduate/GED	46 (24%)	13 (30%)	χ^2 (5) = 6.53	.26
1-3 years college	50 (26%)	8 (18%)		
4-year college grad	26 (14%)	11 (25%)		
>4-year college grad	30 (16%)	5 (11%)		
Living alone				
Yes	80 (42%)	19 (44%)	χ^2 (1) = .08	.78
No	111 (58%)	24 (56%)		
Living with spouse				
Selected	91 (36%)	22 (13%)	χ^2(1)=26.04	<.0
Not selected	159 (64%)	141 (87%)		01
Living with relatives				
Selected	26 (10%)	4 (2%)		
Not selected	224 (90%)	159 (98%)	χ^2(1)=9.25	.002
Living with non-relatives				
Selected	12 (5%)	1 (1%)	χ^2 (1) = 5.67	
Not selected	238 (95%)	162 (99%)		.02

Table 1 (continued)

Descriptive Statistics on Demographic Variables in RCT Participant and Refuser Groups and Significance Tests of Group Differences.

Variable	RCT Participants	Refusers	Test Statistic	p
Volunteer work				
Selected	36 (14%)	1 (1%)	χ^2 (1)=22.99	<.0001
Not selected	214 (86%)	162 (99%)		
Work full or part time?				
Selected	33 (13%)	4 (2%)	χ^2(1)=13.97	.0002
Not selected	217 (87%)	159 (98%)		
Retired?				
Selected	150 (60%)	39 (24%)	χ^2(1)=51.73	<.0001
Not selected	100 (40%)	124 (76%)		
Keep house?				
Selected	61 (24%)	14 (9%)	χ^2 (1) =16.60	<.0001
Not selected	189 (76%)	149 (91%)		
Have regular medicare?				
Selected	78 (31%)	29 (18%)	χ^2(1)=9.24	.002
Not selected	172 (69%)	134 (82%)		
Medicare with HMO coverage?				
Selected	103 (41%)	13 (8%)	χ^2(1)=54.30	<.0001
Not selected	146 (59%)	150 (92%)		
Have medicaid/medi-cal?				
Selected	10 (4%)	1 (1%)	χ^2(1)=4.36	.03
Not selected	240 (96%)	162 (99%)		
Insurance plan with vision coverage?				
Yes	163 (95%)	26 (81%)	χ^2(1)=7.24	.007
No	9 (5%)	6 (19%)		

Descriptive Statistics on Clinical Variables in RCT Participant and Refuser Groups and Significance Tests of Group Differences.

Variable	RCT Participants	Refusers	Test Statistic	p
Acuity – operated eye (log-transformed)	M = .49 (.26)	M = .43 (.19)	t(289)=1.63	.10
Diabetes				
Yes	48 (19%)	26 (16%)	χ^2 (1) = .71	.40
No	202 (81%)	137 (84%)		
PSC present - operated eye				
Yes	25 (10%)	25 (15%)	χ^2 (1) = 2.64	.10
No	225 (90%)	138 (85%)		
Glaucoma present – operated eye				
Yes	24 (10%)	6 (12%)	χ^2 (1) = .19	.67
No	217 (90%)	44 (88%)		
Macular degeneration present – operated eye				
Yes	52 (21%)	20 (12%)	χ^2 (1) = 4.99	.03
No	198 (79%)	143 (88%)		
Dry macular degeneration present – operated eye				
Yes	34 (14%)	6 (12%)	χ^2 (1) = .16	.69
No	207 (86%)	44 (88%)		
Previous surgery-operated eye				
Yes	7 (10%)	2 (12%)	χ^2 (1) = .02	.88
No	60 (90%)	15 (88%)		
Other disease-operated eye				
Yes	100 (43%)	24 (50%)	χ^2 (1) = .77	.38
No	132 (57%)	24 (50%)		

Table 3

Descriptive Statistics on Outcome Variables for RCT Participant and Refuser Groups and Significance Tests of Group Differences.

Variable	RCT Partcipants	Refusers	Test Statistic	p
ADVS	M = 84.9 (11.0)	M = 86.8 (9.5)	t(410)= -1.79	.07
PCS 12	M = 45.4 (10.5)	M = 46.4 (9.4)	t(294)= -.64	.53
MCS 12	M = 54.6 (8.8)	M = 54.3 (8.6)	t(294)= .18	.86
Medical Comorbidities	M = 2.8 (1.6)	M = 2.9 (1.6)	t(296)= -.38	.71
CSI Score	M = 11.2 (1.2)	M = 11.3 (1.3)	t(411)= -.34	.73

Descriptive Statistics on Demographic Variables in Surgery and Watchful Waiting Trial Arms and Significance Tests of Group Differences.

Variable	Surgery N = 133	Watchful Waiting N = 117	Test Statistic	p
Age	M = 78.7 (6.0)	M = 78.8 (5.7)	t(247)= .17	.87
Gender				
Male	58 (44%)	38 (32%)	$\chi^2(1)=3.26$	.07
Female	75 (56%)	79 (68%)		
Race				
White	106 (80%)	87 (74%)		
African American	13 (10%)	17 (15%)		
Asian/Pacific Islander	5 (4%)	5 (4%)	$\chi^2(4) = 1.72$	.79
Hispanic	7 (5%)	7 (6%)		
Multi-Racial	2 (2%)	1 (1%)		
Education				
8th grade or less	9 (9%)	4 (4%)		
Some HS	15 (15%)	12 (13%)		
HS graduate/GED	24 (24%)	22 (24%)	$\chi^2(5) = 2.22$	.82
1-3 years college	24 (24%)	26 (29%)		
4-year college grad	13 (13%)	13 (14%)		
>4-year college grad	17 (17%)	13 (14%)		
Living alone				
Yes	36 (36%)	44 (49%)	$\chi^2(1) = 3.43$	.06
No	65 (64%)	46 (51%)		
Living with spouse				
Selected	52 (39%)	39 (33%)	$\chi^2(1)=.89$	.34
Not selected	81 (61%)	78 (67%)		
Living with relatives				
Selected	16 (12%)	10 (9%)	$\chi^2(1)=.81$	.37
Not selected	117 (88%)	107 (91%)		
Living with non-relatives				
Selected	5 (4%)	7 (6%)	$\chi^2(1) = .67$	.41
Not selected	128 (96%)	110 (94%)		

Table 4 (continued)

Descriptive Statistics on Demographic Variables in Surgery and Watchful Waiting Trial Arms and Significance Tests of Group Differences.

Variable	Surgery N = 133	Watchful waiting N = 117	Test Statistic	p
Volunteer work				
Selected	14 (11%)	22 (19%)	χ^2 (1)=3.46	.06
Not selected	119 (89%)	95 (81%)		
Work full or part time?				
Selected	18 (14%)	15 (13%)	χ^2(1)=.03	.87
Not selected	115 (86%)	102 (87%)		
Retired?				
Selected	77 (58%)	73 (62%)	χ^2(1)=.52	.47
Not selected	56 (42%)	44 (38%)		
Keep house?				
Selected	30 (23%)	31 (27%)	χ^2 (1) =.52	.47
Not selected	103 (77%)	86 (73%)		
Have regular medicare?				
Selected	44 (33%)	34 (29%)	χ^2(1)=.47	.49
Not selected	89 (67%)	83 (71%)		
Medicare with HMO coverage?				
Selected	56 (42%)	47 (40%)	χ^2(1)=.13	.72
Not selected	76 (58%)	70 (60%)		
Have medicaid/medi-cal?				
Selected	3 (2%)	7 (6%)	χ^2(1)=2.25	.13
Not selected	130 (98%)	110 (94%)		
Insurance plan with vision coverage?				
Yes	86 (95%)	77 (95%)	χ^2(1)=.03	.87
No	5 (5%)	4 (5%)		

Table 5

Descriptive Statistics on Clinical Variables in Surgery and Watchful Waiting Trial Arms and Significance Tests of Group Differences.

Variable	Surgery	Watchful Waiting	Test Statistic	p
Acuity – operated eye (log-transformed)	M = .48 (.25)	M = .51 (.27)	t(239)=1.06	.29
Diabetes				
Yes	22 (17%)	26 (22%)	χ^2 (1) = 1.29	.26
No	111 (83%)	91 (78%)		
PSC present - operated eye				
Yes	12 (9%)	13 (11%)	χ^2 (1) = .30	.58
No	121 (91%)	104 (89%)		
Glaucoma present – operated eye				
Yes	10 (8%)	14 (13%)	χ^2 (1) = 1.51	.22
No	119 (92%)	98 (87%)		
Macular degeneration present – operated eye				
Yes	28 (21%)	24 (21%)	χ^2 (1) = .01	.92
No	105 (79%)	93 (79%)		
Dry macular degeneration present – operated eye				
Yes	20 (16%)	14 (13%)	χ^2 (1) = .45	.50
No	109 (84%)	98 (87%)		
Previous surgery-operated eye				
Yes	4 (11%)	3 (9%)	χ^2 (1) = .08	.78
No	31 (89%)	29 (91%)		
Other disease-operated eye				
Yes	50 (41%)	50 (46%)	χ^2 (1) = .64	.42
No	73 (59%)	59 (54%)		

Table 6

Descriptive Statistics on Outcome Variables for Surgery and Watchful Waiting Trial Arms and Significance Tests of Group Differences.

Variable	Surgery	Watchful Waiting	Test Statistic	p
ADVS	M = 84.4 (11.0)	M = 85.6 (10.9)	t(248)= .85	.40
PCS 12	M = 45.8 (10.6)	M = 44.9 (10.5)	t(245)= -.69	.49
MCS 12	M = 55.0 (8.5)	M = 54.1 (9.1)	t(245)= -.77	.44
Medical Comorbidities	M = 2.8 (1.6)	M = 2.7 (1.6)	t(246)= -.48	.63
CSI Score	M = 11.1 (1.1)	M = 11.4 (1.1)	t(248)= 1.85	.07

Table 7

Results from Intention-to-Treat Analysis Multivariate Regression Models Testing for Differences Between Trial Arms on Six Month Endpoint Measures

Variable	Impact of Surgery	p
ADVS (N=209)	M=6.51 (1.64)	<0.0001
ADVS (CSI=10) (N=59)	M=13.04 (3.15)	<0.0001
ADVS (CSI=11) (N=76)	M=7.69 (2.50)	0.003
ADVS (CSI>11) (N=72)	M=1.48 (2.95)	0.617
HUI (N=212)	M=0.041 (0.029)	0.156
HUI (CSI=10) (N=60)	M=0.057 (0.056)	0.312
HUI (CSI=11) (N=78)	M=0.044 (0.042)	0.302
HUI (CSI>11) (N=74)	M=0.024 (0.053)	0.657

Note: Control variables in the ADVS analysis are the following: Baselines ADVS, age diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities. Control varibales in the HUI analysis are the following: Baseline HUI, age, diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities.

Table 8

Results from Intention-to-Treat Analysis Multivariate Regression Models Testing for Differences Between Trial Arms on Six Month Endpoint Measures
with Crossovers Reassigned

Variable	Impact of Surgery	p
ADVS (N=209)	M=6.98 (1.62)	<.0001
ADVS (CSI=10) (N=59)	M=10.92 (3.21)	0.001
ADVS (CSI=11) (N=76)	M=10.15 (2.40)	<0.001
ADVS (CSI>11) (N=72)	M=1.60 (2.96)	0.591
HUI (N=212)	M=0.032 (0.029)	0.262
HUI (CSI=10) (N=60)	M=0.050 (0.055)	0.361
HUI (CSI=11) (N=78)	M=0.026 (0.043)	0.542
HUI (CSI>11) (N=74)	M=0.014 (0.054)	0.792

Note: Control variables in the ADVS analysis are the following: Baselines ADVS, age diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities. Control varibales in the HUI analysis are the following: Baseline HUI, age, diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities.

Table 9

Results from Intention-to-Treat Analysis Multivariate Regression Models Testing for Differences Between Trial Arms on Six Month HUI Specific Attributes (N=212)

Variable	Impact of Surgery	p
Vision	M=0.031 (0.014)	0.035
Hearing	M=0.000 (0.009)	0.979
Speech	M=0.005 (0.003)	0.1
Ambulation	M=0.012 (0.026)	0.65
Dexterity	M=0.014 (0.016)	0.383
Emotion	M=0.003 (0.016)	0.842
Cognition	M=0.022 (0.012)	0.062
Pain	M=0.023 (0.035)	0.522

Note: Control variables in the above analysis are the following: Baselines Single Attribute Utility, age diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities.

Table 10

Results from Intention-to-Treat Analysis Multivariate Regression Models Testing for Differences Between Trial Arms on Six Month HUI Specific Attributes (N=212)
With Crossovers Reassigned

Variable	Impact of Surgery	p
Vision	M=0.033 (0.014)	0.025
Hearing	M=0.005 (0.009)	0.538
Speech	M=0.005 (0.003)	0.96
Ambulation	M= - 0.020 (0.026)	0.428
Dexterity	M=0.018 (0.016)	0.26
Emotion	M=0.009 (0.016)	0.583
Cognition	M=0.015 (0.012)	0.215
Pain	M= - 0.007 (0.035)	0.848

Note: Control variables in the above analysis are the following: Baselines single attribute utility, age diabetes, PSC, AMD, gender, baseline PCS 12, baseline MCS 12, medical comorbidities.

Table 11
Results from Intention-to-Treat Analysis of Utilization and Cost

	WATCHFUL WAITING				SURGERY						
Non-Surgery Costs (excluding Co-pay)	**Mean**	**Price**	**Mean Cost**	**SD**	**Mean**	**Price**	**Mean Cost**	**SD**	**p value**		
New Doctor Visit	1.00	91.00	$91.00		1.00	91.00	$91.00				
Follow-up Doctor Visits	1.03	47.44	$48.96		2.46	47.44	$116.89				
Total Doctor Visits	2.03		***$139.96***	11.01	3.46		***$207.89***	16.40	0.001	107.3	175.15
Glasses	0.08	55.05	***$4.64***	2.28	0.26	55.05	***$14.48***	2.91	0.01	22.23	31.12
Medications	0.37	30.30	***$11.15***	2.57	0.44	30.30	***$13.21***	2.45	0.55	25.02	26.14
Total Utilization Cost			**$155.74**	14.18			**$235.58**	19.35	0.0015	138.24	206.55
Non-Health Care Costs											
Mileage	8.29	0.30	***$2.49***	0.75	31.95	0.30	***$9.58***	3.03	0.037	7.27	32.38
Other Travel Costs	0.77		***$0.77***	0.57	1.08		***$1.08***	0.47	0.67	5.55	5.05
Lost Work Days	0.01	50.00	***$0.25***	0.26	0.08	50.00	***$3.95***	3.11	0.28	2.56	33.19
Total Non-Health Care Costs			**$3.51**				**$14.61**				
All Patient Co-Pays											
Visit Co-pay	4.71		***$4.71***	1.94	7.95		***$7.95***	3.40	0.43	18.95	36.33
Surgery Co-pay	0.58		***$1.18***	0.58	1.18		***$1.18***	0.71	0.52	5.64	7.59
Glasses Co-pay	5.70		***$5.70***	2.75	30.15		***$30.15***	7.55	0.005	26.82	80.66
Medication Co-pay	6.13		***$6.13***	2.30	4.28		***$4.28***	1.31	0.46	22.38	13.99
Total Co-pays			**$17.72**	6.18			**$43.56**	10.95	0.005	60.27	116.89
Surgery Cost (excluding Co-pay)											
1st eye cataract	0.13	1761.00	***$221.89***	60.34	0.90	1761.00	$1,591.94	48.91	<0.0001	588.12	522.25
2nd eye cataracy	0.00		***$0.00***	0.00	0.04	1761.00	$77.48	33.92	0.04		362.21
Tear duct surgery	0.03	260.54	***$8.34***	4.70	0.00	261.00	$0.00	0.00	0.056	45.8	0
Glaucoma Surgery	0.00		***$0.00***	0.00	0.01	720.02	$6.31	6.32	0.36		67.43
Retained Lens Material	0.00		***$0.00***	0.00	0.01	573.11	$5.03	5.03	0.36		53.68
Total Surgical Cost			**$230.22**	60.20			**$1,680.77**	62.12	<0.0001	586.75	663.26
Total Cost			***$407***	76.16			***$1,975***	73.78	<0.0001	742.31	787.77

Difference in cost (Surgery-Watchful Waiting)		
	Mean	**$1,567**
	SD	**$106**

Table 12
Results from Treatmen Received Analysis of Utilization and Cost Accounting for Crossovers

	WATCHFUL WAITING				SURGERY				
Non-Surgery Costs (excluding Co-pay)	**Mean**	**Price**	**Mean Cost**	**SD**	**Mean**	**Price**	**Mean Cost**	**SD**	**p value**
New Doctor Visit	1.00	91.00	$91.00		1.000	91.00	$91.00		
Follow-up Doctor Visits	0.40	47.44	$19.17		2.965	47.44	$140.66		
Total Doctor Visits	1.40		***$110.17***	3.15	3.965		***$231.66***	17.33	<0.0001
Glasses	0.03	55.05	***$1.76***	1.00	0.304	55.05	***$16.74***	3.27	0.0001
Medications	0.23	30.30	***$7.09***	1.86	0.548	30.30	***$16.60***	2.78	0.007
Total Utilization Cost			**$119.02**	4.93			**$265.00**	20.73	<0.0001
Non-Health Care Costs									
Mileage	3.35	0.30	***$1.01***	0.26	35.780	0.30	***$10.73***	3.03	0.004
Other Travel Costs	0.34		***$0.34***	0.24	1.430		***$1.43***	0.63	0.14
Lost Work Days	0.01	50.00	***$0.25***	0.27	0.078	50.00	***$3.90***	3.08	0.28
Total Non-Health Care Costs			**$1.60**				**$16.06**		
All Patient Co-Pays									
Visit Co-pay	1.17		***$1.17***	0.54	10.810		***$10.81***	3.66	0.02
Surgery Co-pay	0.00		***$0.00***	0.00	1.650		***$1.65***	0.85	0.08
Glasses Co-pay	1.81		***$1.81***	1.61	33.110		***$33.11***	7.62	0.0003
Medication Co-pay	3.70		***$3.70***	1.75	6.280		***$6.28***	1.80	0.31
Total Co-pays			**$6.68**	3.24			**$51.85**	11.33	<0.0001
Surgery Cost (excluding Co-pay)									
1st eye cataract	0.00	1761.00	***$0.00***	0.00	1.000	1761.00	***$1,761.00***	0.00	
2nd eye cataracy	0.00		***$0.00***	0.00	0.044	1761.00	***$77.48***	33.63	0.04
Tear duct surgery	0.03	260.54	***$8.34***	4.75	0.000	261.00	***$0.00***	0.00	0.054
Glaucoma Surgery	0.00		***$0.00***	0.00	0.009	720.02	***$6.31***	4.98	0.36
Retained Lens Material	0.00		***$0.00***	0.00	0.009	573.11	***$5.03***	6.26	0.36
Total Surgical Cost			**$8.34**	4.75			**$1,849.82**	34.35	<0.0001
Total Cost			***$136***	8.93			***$2,183***	49.28	<0.0001

Difference in cost (Surgery-Watchful Waiting)	**Mean**	**$2,047**
	SD	**$50**

Table 13

Part A

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting based on Intention to Treat

Group	Utility Gain	Cost	Cost/QALY
Entire Sample	0.041	$1,567	$38,228
CSI = 10	0.057	$1,803	$31,638
CSI = 11	0.044	$1,639	$37,250
CSI => 12	0.024	$1,284	$53,500

Part B

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting with Crossovers Reassigned

Group	Utility Gain	Cost	Cost/QALY
Entire Sample	0.032	$2,047	$63,972
CSI=10	0.050	$2,030	$40,599
CSI=11	0.026	$2,142	$82,369
CSI=>12	0.014	$1,938	$138,415

Table 14

Part A

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting
Analyzed from the standpoint of functionally relevant improvement of 7 ADVS points based on Intention to Treat

Group	ADVS Score Gain	Cost	Cost/FADVS
Entire Sample	6.51	$1,567	$1,685
CSI=10	13.04	$1,803	$968
CSI=11	7.69	$1,639	$1,492
CSI=>12	1.48	$1,284	$6,073

Part B

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting
Analyzed from the standpoint of functionally relevant improvement of 7 ADVS points with Crossovers Reassigned

Group	ADVS Score Gain	Cost	Cost/FADVS
Entire Sample	6.98	$2,047	$2,053
CSI=10	10.92	$2,030	$1,301
CSI=11	10.15	$2,142	$1,477
CSI=>12	1.60	$1,938	$8,478

FADVS= 7 points ADVS points, equal to the minimum functionally appreciable benefit

Table 15

Part A

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting Analyzed from the standpoint of one line of visual acuity based on Intention to Treat

Group	Number of VA lines	Cost	Cost/line(VA)
Entire Sample	1.19	$1,567	$1,318
CSI=10	1.83	$1,803	$985
CSI=11	1.10	$1,639	$1,490
CSI=>12	0.79	$1,284	$1,629

Part B

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting Analyzed from the standpoint of one line of visual acuity with Crossovers Reassigned

Group	ADVS Score Gain	Cost	Cost/line(VA)
Entire Sample	1.35	$2,047	$1,516
CSI=10	1.71	$2,030	$1,187
CSI=11	1.33	$2,142	$1,610
CSI=>12	1.05	$1,938	$1,846

Table 16

Utility and Cost Sensitivity Analysis

A. Intention to Treat (Cost-effectiveness of surgery) units are in Cost/QALY

		Utility	
Cost	(-)1SD	Baseline	(+)1SD
(+)50%	$195,917	$57,341	**$33,586**
Baseline	$130,611	**$38,228**	**$22,390**
(-)50%	$65,306	**$19,114**	**$11,195**

B. Treatment Received (Cost-effectiveness of surgery) units are in Cost/QALY

		Utility	
Cost	(-)1SD	Baseline	(+)1SD
(+)50%	$852,960	$79,965	**$41,949**
Baseline	$682,370	$63,972	**$33,559**
(-)50%	$341,190	**$31,987**	**$16,780**

Regression Data from Table 7 and 8 were used for this sensitivity analysis. For Part A, the Baseline Utility was 0.041 (SD=0.029) and Baseline Cost was $1,567. For Part B, the Baseline Utility was 0.032 (SD=0.029) and Baseline Cost was $2,047.

Note: Two standard deviations negative of the Baseline Utility, there would be no benefit from surgery.

Table 17

Part A

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting based on Intention to Treat

Group	Utility Gain*	Cost	Cost/QALY
CSI = 10+11	0.023	$1,686	$73,286
CSI => 12	0.010	$1,284	$128,401

Part B

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting with Crossovers Reassigned

Group	Utility Gain*	Cost	Cost/QALY
CSI=10+11	0.023	$2,091	$90,926
CSI=>12	0.010	$1,938	$193,781

*** Utility Gain is calculated here using mutli-attribute weights but holding all attributes except vision constant (mean baseline attribute utility scores used). In this way changes in vision were mapped to changes in overall total utility while minimizing noise from other attributes.**

Table 18

Part A

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting Analyzed by converting ADVS Gain into Utility based on Intention to Treat

Group	ADVS Score Gain	Utility Gain	Cost	Cost/QALY
Entire Sample	6.51	0.027	$1,567	$57,323
CSI=10	13.04	0.055	$1,803	$32,928
CSI=11	7.69	0.032	$1,639	$50,747
CSI=>12	1.48	0.006	$1,284	$206,565

Part B

Cost-Effectiveness of Cataract Surgery Over Watchful Waiting Analyzed by converting ADVS Gain into Utility with Crossovers Reassigned

Group	ADVS Score Gain	Utility Gain	Cost	Cost/QALY
Entire Sample	6.98	0.029	$2,047	$69,829
CSI=10	10.92	0.046	$2,030	$44,260
CSI=11	10.15	0.043	$2,142	$50,237
CSI=>12	1.60	0.007	$1,938	$288,365

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 4

Clinical Trials in Older Cancer Patients: An overview of obstacles in generating evidence-based data.

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Introduction

The randomized trial comparing cataract surgery versus watchful waiting discussed in the previous chapters illustrates that clinical trials are feasible in older populations. Few alternatives other than clinical trials exist to determine treatment benefit in disease such as cataract surgery where the majority of patients are elderly. Even though the cataract trial was feasible, there were numerous issues that complicated its implementation and analysis. First, it was a laborious effort to recruit sufficient patients onto the trial, which limited the power of the clinical trial. Moreover, since the effectiveness outcome was quality-adjusted life years, generic instruments were used to determine subjective valuations of patient's health. Using subjective responses in an elderly population required limiting the trial to English speaking patients with adequate hearing and mental functioning. These exclusion criteria may limit the generalizability of the study findings to a community population. Even with these inherent difficulties, valuable information was derived from the trial. Is the experience from this cataract surgery trial similar to other areas of medicine? A good focus area for comparison is cancer treatment.

Cancer incidence increases with age and cancer clinical trials form the largest proportion of clinical trials performed in the U.S. A review of cancer trials between 1993 and 1996 performed by the Southwest Oncology Group (SWOG) demonstrated that the percentage of patients over the age of 65 in clinical trials, 25%, underrepresented the percentage of elderly U.S. patients, 63%, with cancer [1]. Elderly patients were underrepresented in all categories of cancer clinic trials to varying degrees. For example, in prostate cancer, elderly patients represented 64% of cancer trial participants but 77% of all U.S. patients with prostate cancer. Differential participation by age was more prominent in colorectal cancer (40% clinical trials versus 72% U.S.) and lung cancer (39% clinical trials versus 66% U.S.). This under-representation was most dramatic in breast cancer where only 9% of trial patients were over 65 even though 49% of U.S. breast cancer patients were over 65 [1] (See **Figure 1**).

The SWOG data are only from one cooperative group so some clinicians argue they are not representative of clinical trials in general. Furthermore, the above study did

not distinguish participation of elderly patients by type of trial (Phase I, II, or III) or by stage at presentation. A more recent study examined the participation of older patients in National Institute of Cancer (NCI) sponsored cancer trials active from 1997 through 2000 using data from multiple cooperative groups. The results were very similar to the earlier SWOG data showing that older patients were under-represented in all clinical trials in oncology [2]. One of the largest discrepancies in participation was in breast cancer trials were even though the elderly represented 49% of U.S. breast cancer patients; they represented less than 20% of clinical trial patients [2]. In this study, the percentage of older patients were underrepresented but participation did not differ based Phase of clinical trial or stage of cancer.

There are many reasons that might explain why older women were not a part of breast cancer trials. Three major categories are (a) lack of referral by physicians, (b) lack of willingness to participate by patients, and (c) exclusion criteria inherent in the structure of the trials. Although these categories are listed separately, significant overlap exists between them.

Attitudes and Perceptions of Clinical Trials (Physicians and Patients)

Reluctance by oncologists and surgeons to refer and recruit breast cancer patients to clinical trials was recently studied. Factors were identified as to (1) why physicians are generally reluctant to participate in clinical trials and (2) why participating physicians refer only a small percentage of their patients [3]. Physicians were more likely to refer patients to a clinical trial when they knew which trials the patient was eligible for or when patients were more involved in decision-making. Attitudinal factors that were important in determining likelihood of referral to a clinical trial included: (a) comfort in explaining trials to patients, (b) perceived level of patient interest in a clinical trial, (c) perception that patients would remain in the local community and followed closely, (d) whether paper-work associated with a clinical trial was considered to be too time-consuming, or (e) if there was a belief that trial entry requirements were too stringent [3]. Interestingly, the only patient characteristic that predicted referral to a clinical trial was the patient's age, with the older patients being less likely to be referred.

Equally important as physician attitudes towards clinical trials in breast cancer are the attitudes of breast cancer patients. One in-depth study using focus groups and surveys on 60 consecutive patients, the majority of which had breast cancer, at an outpatient cancer clinic found that patient knowledge about randomized trials was not high. The three most important factors in willingness to join a clinic trial were (1) patients perception, favorable or unfavorable, of their physician, (2) their personal attitude towards experimentation and uncertainty, and (3) whether their was a perception that clinical trial participation would be inconvenient or represent a loss of control [4]. A larger study on patient's attitudes using cross-sectional surveys of women attending a breast clinic demonstrated that women who would consider participating in a randomized clinical trial were younger, more likely to want an active role in decision-making, and more knowledgeable about randomized clinical trials [5].

Only recently has research focused specifically on barriers to participation of older women with breast cancer in clinical trials. A recent study evaluated 77 pairs of younger (mean age 50.4 years) and older (mean age 76.5 years) women with breast cancer matched on physician type (surgeon, medical oncologist, or radiation oncologist), stage (early or late) who were eligible for at least on open trial in their institution and treated within one year of the study date. Both physicians and patients were surveyed [6]. A significantly greater number of patients in the younger group, 51%, were offered a clinical trial option compared to the older group, 35%. Patients with a higher number of comorbidities were less likely to be offered clinical trial participation [6]. After controlling the number of comorbidities and functional status*, age still significantly predicted whether patients were offered a clinical trial. Factors that influenced physician referral for older patients included: (a) the presence of comorbid conditions not excluded by the clinical trial but that the referring physician felt would have affected the patient's response to treatment, and (b) a perception that the clinical trial regiment was too toxic [6].

Surprisingly, patient's difficulty in understanding and costs of the clinical trial, transportation issues, and shorter life expectancy did not influence physician's clinical

* Functional Status refers to the ability of an individual to perform required daily tasks at home and at work. It reflects the level of physical strength, mobility, and energy level of an individual.

trial referral decisions for either younger or older patients. Of patients offered a clinical trial, there was no difference among younger and older patients in the percent who consented, 56% versus 50%, respectively [6]. Therefore, the failure of clinicians to offer a clinical trial to eligible older patients is a significant barrier to enrollment.

Exclusion Criteria: Age, Comorbidity, and Functional Status

As a result of demographic changes in the U.S. population that have and will further increase the proportion of older patients seeking cancer treatment, the U.S. Food and Drug Administration (FDA) published "Guidelines for the Study of Drugs Likely to be Used in the Elderly" in 1989 [7]. These guidelines state that the population studies should reflect the population likely to be treated. Nevertheless, there is a significant discordance between the "study" patient and the "ordinary" patient when it comes to clinical trials [8]. This discordance increases with age since most cancer clinical trials set an upper age bound at 70, and exclude patients with multiple comorbidites or functional impairment. Since pharmaceutical companies, whose primary goal is to demonstrate efficacy of their drug to the FDA, sponsor many clinical trials, selection criteria for trials reflect a desire for a younger healthier population in order to provide clear and convincing proof of drug benefit.

A study reviewing the protocol exclusion criteria in over 500 Phase II and III NCI sponsored cooperative group trials demonstrated that over 80% had exclusion criteria based on hematological, hepatic, and renal functioning [2]. In addition, over 90% had an exclusion based on a minimal level of functional status with over 60% requiring patients to be able to perform all Activities of Daily Living (ADLS) [see subsequent discussion] [2]. It is therefore worthwhile to examine the relationship of age, comorbidity, functional status, and treatment outcome.

Comorbidity increases with age. One study on colon cancer patients demonstrated that patients aged 55-64 had an average 3 comorbid conditions with an addition of one comorbid condition on average per additional decade of age [9]. Patients 75 and over, therefore, had a mean of 5 comorbid conditions (See **Figure 2**). Comorbidity can influence treatment in two ways: (1) by decreasing the resilience that patients have to the

toxic effects of treatment or (2) by affecting treatment selection [10]. Furthermore, comorbidity can be split into "covert" comorbidity which is not recognized by the physician and "overt" comorbidity, which is recognized usually integrated into clinical trial exclusion criteria [10].

Baseline comorbidity has been shown to be predicitive of mortality in a population-based study [11]. In cancer patients, there is a relationship between overall survival and total comorbidity, with mortality risk ratios ranging from 1.33-1.85 in those with 3 or more comorbidities [9]. In a study of older patients with early breast cancer patients, only 51% of deaths were due to breast cancer showing the mortality effect of multiple competing illnesses in patients [12].

Complications caused by comorbid condition during treatment may exceed the expected benefit from the treatment itself. It is often difficult to separate patients who will recover from the morbidity of treatment and those whose health will decline and eventually die. Furthermore, patients with comorbid conditions get less optimal treatment. Among patients with early breast cancer, patients over 70 were significantly less likely to receive therapy consistent with the National Institute of Health (NIH) consensus statement for the treatment of breast cancer [12]. Therefore, it is difficult to use observational data to determine whether comorbidity or treatment selection leads to poor outcome. The association between comorbid condition and less treatment selected may be biased since other factors such as age bias may be the driving force in their connection [13].

A separate yet intertwined criteria in evaluating a patient's acceptability for a clinical trial is functional status or impairment. Functional impairment is the inability to perform daily life activities normally. Several scales have been used to measure functional impairment.

Geriatricians, physicians involved in the general medical care of older patients, break functional status into: (a) activities of daily living (ADLs) and (b) instrumental activities of daily living (IADLs). ADLs are a measure of 6 basic functions: bathing, dressing, toileting, continence, transferring, and feeding [14]. IADLs are a measure of eight higher level functions: using the telephone, traveling, shopping, preparing meals, laundry, doing housework, taking medicine, and managing money [15]. The number of

impairments in ADLs and IADLs increases with age (See **Figure 3**). A recent survey indicated that 10-13% of older persons between the age of 65-69 have difficulty getting out of bed and between 6-10% need help with routine care. With age this need increases, with 24-29% of those over the age of 80 requiring help getting out of bed, and 29-42% requiring help with routine care [16].

Oncologists use alternative functional assessment instruments such as the Karnofosky Performance Scale (KPS) or the Eastern Cooperative Oncology Group (ECOG) Performance Status Scale. The ECOG scale has 6 possible scores: 0 (fully active), 1 (restricted in physically strenuous activity, but ambulatory and able to carry out light work), 2 (ambulatory and capable of all self-care) but unable to carry out work activities, 3 (capable of only limited self care, confined to bed or chair >50% of waking hours), 4 (completely disabled), and 5 (dead) [17]. The ECOG scale correlates well with ADLs but any deficiency in ADLs makes a 3 the highest sore possible on the ECOG. Most clinical trials limit patients to those with ECOG scores of 2 or lower. As a result many older patients are excluded from clinical trials due to functional limitations.

Functional status is also related to treatment outcome. Medical conditions may present first (or only) as a functional disturbance. In addition, functional loss affects the quality of life** of older patients. Furthermore, functional losses may lead to further disability and institutionalization. Lastly, functional impairment is a predictor of morbidity and mortality. In a study of 189 older individuals living in a rural community, the relative death risk from going from independence to dependence in on or more ADLs, adjusted for other illnesses, was 6.5 [18]. There was an interaction between age and functional status with the relative risk of death being 12.02 and 13.60 in those aged 80-84 and over 84 respectively. Patients with two impairments in ADL had a relative risk of death of 13.66 [18]. This data demonstrates that independent of age and health condition, functional status is a predictive factor for short-term mortality in non-institutionalized older individuals.

** Quality of Life is a subjective measure of one's overall enjoyment of life. It includes some of the components of functioning, similar to Functional Status, but also includes other physical (pain, nausea,...), social, and emotional components. Therefore, even though an individual may objectively have a poor functional status, they may rate their overall quality of life high, or vice versa.

Conclusion

Unlike cataract surgery where any clinical data requires the enrollment of older individuals, the treatment guidelines for cancer, and more specifically breast cancer, are largely derived from trials of patients under the age of 70. There may be many reasons for the lack of adequate representation of older patients in clinical trials including: the lack of referral from physicians or the lack of willingness to participate from patients. In most cases, though, clinical trials have criteria that exclude patients over the age of 70 and patients with significant comorbidities or functional impairment. As a result, the "study" patients are often not reflective of the typical older breast cancer patient, and data derived from these trials are not easily generalizable to community practice involving older cancer patients. In the future, clinical trials will need to be designed to include older patients and a broader range of baseline health conditions/status. Until these trials are designed and data is generated, alternative approaches are needed in determining the care of older cancer patients. Modeling involving different assumptions of treatment benefits and costs while incorporating the impact of age, comorbidity, and functional status maybe helpful in guiding care in older breast cancer patients and may even help direct where clinical trials are most critical. An example of this type of approach is outlined in the next chapter.

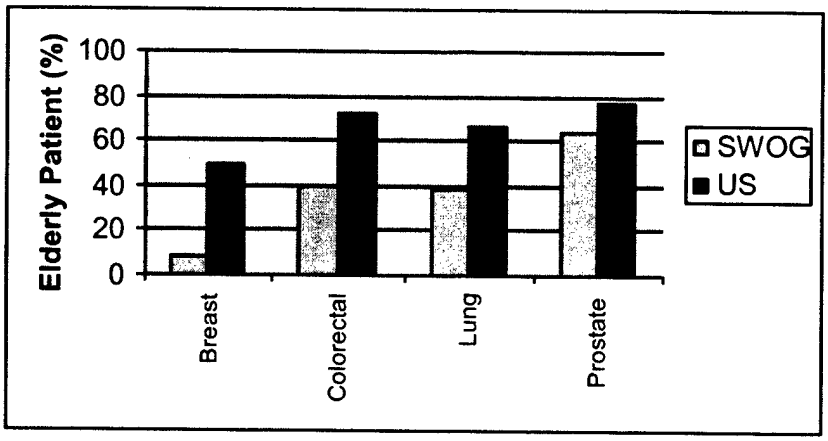


Figure 1: Older Persons Under-represented in clinical trials; SWOG – Clinical trials of Southwestern Oncology Group; US – U.S. patients with specific cancer.

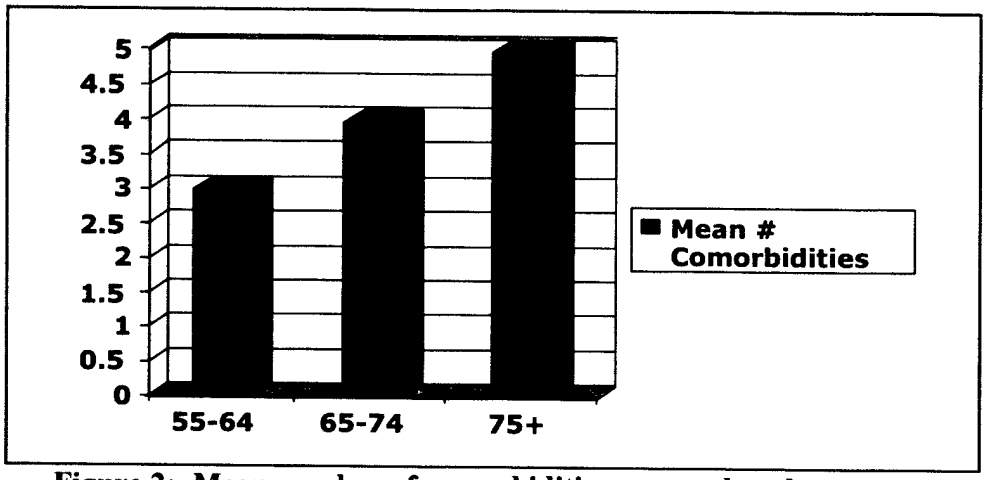


Figure 2: Mean number of comorbidities per age bracket: 55-64, 65-74, 75+.

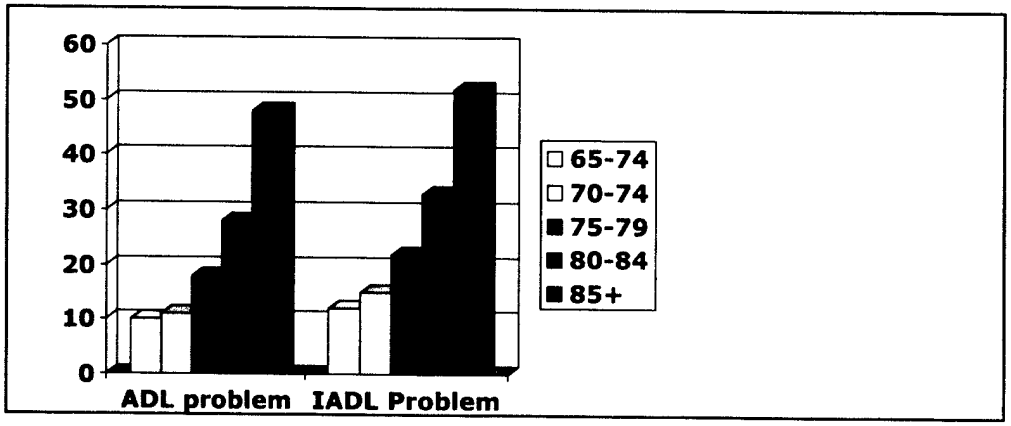


Figure 3: Percentage of individuals with functional disabilities in ADL (Activities of Daily Living) or IADLS (Instrumental Activities of Daily Living).

References

1. Hutchins LF, Unger JM, Crowley JJ, Coltman CA, Jr., Albain KS: Underrepresentation of patients 65 years of age or older in cancer-treatment trials. *N Engl J Med* 341:2061-2067., 1999
2. Kilgore M: The participation of patients 65 years of age and older in cancer clinical trials, Personal Communication to Naeim A, Santa Monicia, 2002
3. Siminoff LA, Zhang A, Colabianchi N, Saunders Sturm CM, Shen Q: Factors that predict the referral of breast cancer patients onto clinical trials by their surgeons and medical oncologists. *J Clin Oncol* 18:1203-1211, 2000
4. Ellis PM, Dowsett SM, Butow PN, Tattersall MH: Attitudes to randomized clinical trials amongst out-patients attending a medical oncology clinic. *Health Expect* 2:33-43., 1999
5. Ellis PM, Butow PN, Tattersall MH, Dunn SM, Houssami N: Randomized clinical trials in oncology: understanding and attitudes predict willingness to participate. *J Clin Oncol* 19:3554-3561., 2001
6. Kemeny M, Muss H, Kornblith A, Peterson B, Wheeler J, Cohen H, CALGB: Barriers to Participation of Older Women with Breast Cancer in Clinical Trials, in *American Society of Clinical Oncology* (vol 19), 2000, p 602a Abstract 2371
7. Administration. UFaD: Guidelines for the Study of Drugs Likely to be Used in the Elderly, in, Rockville, MD, Food and Drug Adminstrattion/Center for Drug Evaluation and Research, 1989
8. Dowd R, Recker RR, Heaney RP: Study subjects and ordinary patients. *Osteoporos Int* 11:533-536, 2000
9. Yancik R, Wesley MN, Ries LA, Havlik RJ, Long S, Edwards BK, Yates JW: Comorbidity and age as predictors of risk for early mortality of male and female colon carcinoma patients: a population-based study. *Cancer* 82:2123-2134., 1998
10. Balducci L, Yates J: General guidelines for the management of older patients with cancer. *Oncology (Huntingt)* 14:221-227., 2000

11. Seeman TE, Guralnik JM, Kaplan GA, Knudson L, Cohen R: The health consequences of multiple morbidity in the elderly. The Alameda County study. *J Aging Health* 1:50-66., 1989

12. Yancik R, Wesley MN, Ries LA, Havlik RJ, Edwards BK, Yates JW: Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older. *Jama* 285:885-892., 2001

13. Schneeweiss S, Maclure M: Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol* 29:891-898., 2000

14. Katz S, Downs T, Cash H, al. e: Progress in development of the index of ADL. *Gerontologist* 10:20-30, 1970

15. Avlund K, Schultz-Larsen K, Kreiner S: the measurement of instrumental ADL: Content validity and construct validity. *Aging (Milano)* 5:371-383, 1993

16. Ostechega Y, Harris T, Hirsch R, al. e: The prevalence of functional limitations and disability in older persons in the U.S.: Data from the National Health and Nutrition Examination Survey III. *J Am Geriatr Soc* 48:1132-1135, 2000

17. Oken M, Creech R, Tormey D, al. e: Toxicity and Response criteria of the Eastern Cooperative Oncology Group. *J Clin Oncol* 5:649-655, 1982

18. Benitez-Rosario MA, Hernandez-Estevez P, Aguirre-Jaime A, Gonzalez-Freire G, Asensio-Fraile A: Functional status and mortality in community-dwelling older people. *J Am Geriatr Soc* 49:1009-1010., 2001

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 5

Decision Analysis Modeling of the Adjuvant Treatment of Older Patients with Breast Cancer

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Introduction

The proportion of older individuals in our society is increasing rapidly with 20% of the population expected to be over the age of 65 by 2030 [1]. The biggest population growth among the elderly will be in two older age groups, those over 75 and those over 85, expected to increase by 1.5 to 2-fold [1]. The majority of all cancers and cancer deaths occur in the elderly. Determining cost-effective care for older cancer patients is often hindered by a lack of data, since many clinical trials either excluded older patients or were not powered for such sub-group analysis. In addition, the impact of comorbid conditions and functional status, the ability to perform activities of daily living, on an older individual's quality* and quantity of life requires special attention. These issues, combined with potential communication barriers with older patients, such as cognitive impairment and hearing loss, make treatment and research in geriatric-oncology challenging.

Of the cancers that affect the elderly, breast cancer is the most common in women. Age-adjusted rates reported by the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Program Data reveal that women, 65 years or older, have an incidence rate of 440.6 per 100,000 as compared to 74.5 per 100,000 for women under 65 years of age. The peak breast cancer incidence rate of 483.3 per 100,000 is in the age group 75-79 years. The annual breast cancer mortality rate for women under 65 years of age is 14.7 per 100,000 as compared to the rate for women 65 years and older which is 122.9 per 100,000 and 200.5 per 100,000 for women 85 years and older [2].

Although the therapy of older patients in the early stages of breast cancer may be similar to that of younger patients, chronological and physiologic differences affect the use of adjuvant treatment. Previous research indicates that the elderly are a heterogeneous group, and classifying them solely by chronological age would be misleading [3-5]. Furthermore, there has been an increasing effort to find predictors of treatment prognosis other than age [6] .

---

* Quality of Life is a subjective measure of one's overall enjoyment of life. It includes some of the components of functioning, similar to Functional Status, but also includes other physical (pain, nausea,...), social, and emotional components. Therefore, even though an individual may objectively have a poor functional status, they may rate their overall quality of life high, or vice versa.

In the early half of the 20th century, breast cancer was considered primarily a loco-regional disease and surgery was the mainstay of treatment [7]. During this period there were reports comparing the natural history of untreated breast cancer to surgical treatment of breast cancer [8]. One of the more rigorous studies involved a series of 250 patients showing 68% 10-year survival in untreated patients compared to 84% in those treated with surgery [9]. The natural history of node positive and negative cancer from 1927-1987 has been examined with 42% of patients with node-positive disease and 23% of patients with node-negative disease dying at the 10-year mark [10] [11].

It wasn't until the 1960s that animal models demonstrated that breast cancer quickly becomes a systemic disease [12]. As a result many trials were established that looked at adjuvant therapy, treatment given after the primary treatment (surgery +/- radiation) to increase the chances of a cure. Adjuvant therapy includes hormone therapy, chemotherapy (monotherapy and polytherapy), and combined hormone-chemotherapy [13]. The data from many of these trials conflicted, and as a result the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) was formed. The EBCTCG has undertaken systemic overviews of randomized trials every five years since 1984-85, using rigorous methods for trial identification, data checking, and meta-analysis [14]. Since the EBCTCG consolidated data from randomized trials it allowed for sufficient number of patients to be analyzed, overcoming the problems of the individual smaller, insufficiently powered trials [14].

Adjuvant Hormonal Therapy

There are currently several options for hormonal therapy for breast cancer. Three main categories include: (1) selective estrogen-receptor modulators, such as tamoxifen (Nolvadex), toremifene (Fareston), and (2) aromatase inhibitors such as exemestane (Aromasin), letrozole (Femara), anastrozole (Arimidex), and (3) estrogen receptor antagonists, such as fulvesrant (Faslodex). Of these agents, tamoxifen has been studied the most extensively.

The EBCTCG met-analysis of tamoxifen for early beast cancer showed an overall reduction of 26% in recurrence rate and 14% in mortality rate for post-menopausal women [15]. A comparison of tamoxifen use in patients with estrogen receptor positive (ER+) tumors demonstrated increased reduction of death rates (11%, 14%, and 23%) with

increased duration of tamoxifen use (1 year, 2 years, and 5 years), respectively [15]. Patients with ER (-) tumors did not benefit from tamoxifen. Additionally, there was no benefit of a higher dose of tamoxifen (40mg) over the current standard 20 mg dose.

Subgroup analysis by age categories revealed that 1 year of tamoxifen is substandard, less survival gain from treatment, across all age groups. Even though 2 years of tamoxifen is substandard for those below the age of 70, the impact of 2 and 5 years of tamoxifen on mortality was similar in those aged 70 and above. The data did reflect a slightly higher rate of recurrence though in this population with just 2 years of tamoxifen [15]. Of course, it is not clear whether this data applies to patients over the age of 80 since the majority of the trials involved in the meta-analysis had an upper age limit of 70 for enrollment. Although trials [15] [17] are still ongoing to determine if there is additional benefit from extending tamoxifen beyond 5 years, data from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14 trial demonstrated no advantage of 10 years versus 5 years of tamoxifen in ER (+) lymph node (-) women [18].

Of the newer hormonal agents available, toremifene (Fareston) has been shown to have a similar efficacy and toxicity profile to tamoxifen [19]. Interim data was presented recently from the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial that evaluates, in a randomized double-blind design, anastrozole (Arimidex), alone or in combination with tamoxifen, relative to tamoxifen alone as 5-year adjuvant treatment for post-menopausal women with early breast cancer. After a median 33 months of follow-up, anastozole alone was found significantly better in prolonging disease-free survival than tamoxifen alone [20]. Many of the newer hormonal agents have already shown efficacy in metastatic disease resistant to tamoxifen. As a result of these trials and numerous additional ongoing trials, it is conceivable that other agents will replace tamoxifen as the standard for adjuvant therapy in the future.

Adjuvant Chemotherapy

Chemotherapy has shown significant benefits in reducing both recurrence (24% reduction) and death (15% reduction) when compared to no chemotherapy [21]. Evidence has shown combination chemotherapy is more effective than single agent therapy [22]. The EBCTCG explored the efficacy of adjuvant polychemotherapy in a recent meta-

analysis [23]. The benefits of adjuvant chemotherapy in terms of recurrence and survival decreased with increasing age of the patient. The reduction in 10-year mortality with chemotherapy was 27%, 14%, and 8% for patients aged <50, 50-59, and 60-69, respectively [23]. The impact of chemotherapy on those over the age of 70 was not analyzed since this subgroup represented such a small percentage (3%) of the overall patients enrolled in clinical trials. However, the data from this small pool of 609 patients did not support a benefit to chemotherapy in those older than 70 [23].

The most frequently tested chemotherapy regimen uses cyclophosphamide, methotrexate, and 5-fluorouracil (5-FU) collectively abbreviated as CMF. CMF therapy has been compared extensively with anthracycline-containing regimens of which the most recognized uses doxorubicin (Adriamycin) and cyclophosphamide (Cytoxan) collectively abbreviated as AC. The NASBP B-15 trial compared 4 cycles of AC with 6 cycles of CMF and found that they were equivalent [24]. Other studies, however demonstrated a benefit in reducing recurrences and mortality with AC. The EBCTG reported reductions of 12% in recurrence and 11% in mortality with the use of anthracycline based chemotherapy over CMF [23], although some of the trials treated for longer lengths than the NASBP B-15 trial.

The additional benefit of anthracycline therapy over CMF probably occurs in only special circumstances. A recent study looked at a proto-oncogene encodes a Human Epidermal Growth Factor Receptor-2, HER2/neu, that is overexpressed in 20-30% of metastatic breast cancers, of patients involved in the South-Western Oncology Group (SWOG) 8814 trial [25]. This over-expression of HER2/neu is associated with decreased survival and decreased relapse-free periods. In the SWOG-8814 trial, postmenopausal women who had positive lymph node involvement and ER (+) were randomized to tamoxifen alone or tamoxifen with cyclophosphamide, adriamycin, and 5-FU (CAF). Only those patients with high Her2/neu expression benefited from chemotherapy [25]. Therefore, anthracycline-based chemotherapy may be advantageous in those patients with high Her2/Neu expression. It should be noted however that these studies did not involve many patients over the age of 70, and the EBCTCG report of anthracycline therapy did provide age-related subgroup analysis. A recent retrospective review of protocols containing AC chemotherapy at MD Anderson Cancer Center in Texas demonstrated that

of 1,898 patients enrolled only 260 patients were between 60-69 and only 40 patients were >69, resulting in low power for analysis of difference in response [26].

To further improve upon these chemotherapy regimens, researchers have used two strategies: (1) adding other drugs to already active regimens and (2) increasing the dose of chemotherapy in order to overcome the resistance of the tumor. As a result, there has been interest in combining AC therapy with paclitaxel (Taxol), abbreviated as T. Two trials recently compared AC x 4 cycles with AC x 4 cycles plus T x 4 cycles in patients with lymph node positive disease. Preliminary analysis of data from the Cancer and Leukemia Group B (CALGB) 9344 study demonstrated a 26% reduction in mortality at a median follow-up of 21 months [27]. Unfortunately, at 52 months of follow-up the results between the two arms where no longer statistically significant [13]. The NSABP B-28 trial also found no statistical improvement by adding paclitaxel during their interim analysis at a median follow-up of 34 months [13]. Again patients over the age of 70 were underrepresented in these trials, and based on the current data there is little to support sequential paclitaxel after AC therapy. There is no convincing evidence that the use of high-dose chemotherapy with peripheral stem cell transplant is superior to standard chemotherapy for adjuvant treatment even in high-risk women. Considering the data available and the treatment-related morbidity and mortality associated with this therapy, older patients should not be considered for high dose chemotherapy unless within the context of a clinical trial [28].

Combined Hormone and Chemotherapy

The EBCTCG also looked at trials combining CMF chemotherapy and tamoxifen. In patients under the age of 50, the addition of CMF to tamoxifen reduced recurrence by 21% and mortality by 25% over tamoxifen alone. In patients between the ages of 50-69, the addition of CMF reduced recurrence by 19% and mortality by 11% over tamoxifen alone [23]. In patients with ER (+) tumors, it is unclear whether the combination of an adriamcyin-based chemotherapy regiment, such as AC + tamoxifen, would improve the results seen with CMF + tamoxifen. Furthermore, it is not clear whether the benefit of combination therapy would extend to those over the age of 70. For those with ER (-) disease, the NSABP B-23 trial, a four-armed study comparing ACx4 to CMFx6 to

ACx4+tamoxifen to CMFx6+tamoxifen, found no significant benefit of combined hormone-chemotherapy in patients with lymph node (-) disease at 5 year follow-up [29].

Herceptin

Herceptin is a monoclonal antibody that works by binding to and blocking the HER2 cell surface receptors produced by the HER2/neu gene. Some women have more than two copies of the HER2/neu gene thereby producing too much HER2 protein, in turn causing the cells to reproduce uncontrollably [30]. About 25 to 30 percent of all breast cancers are HER2/neu positive, and Herceptin has made these tumors significantly easier to treat. Women with HER2-positive tumors that have metastasized are now routinely offered Herceptin in combination with Taxol (Paclitaxel), shown to reduce the risk of death by 20%, as the first line of treatment [31]. Currently, women with HER2-positive tumors are not put on Herceptin in the adjuvant setting. There are several trials examining the role of Herceptin for adjuvant treatment of breast cancer, including: (1) the North Central Cancer Treatment Group (NCCTG)-N9831 trial, (2) the NSABP-B31 trial, (3) the Herceptin Adjuvant Trial (HERA), and (4) the Breast Cancer International Research Group (BCIRG)-006 trial [30]. In the future Herceptin may play a significant role in the adjuvant setting.

Breast Cancer Care in Older Patients

Several studies have noted variations in the care of older patients with breast cancer. On retrospective study of older breast cancer patients demonstrated that women age 80 years and older were 70% less likely (odds ratio = 0.3; 95% CI, 0.1-0.8) to receive chemotherapy than women ages 67-79 years, controlling for comorbidity, functional status, and other covariates [32]. This variation may be secondary to a greater total illness burden in older patients. One conceptual model for total illness burden has chronic health conditions, severity of illness, and functional status/disability all affecting the life expectancy [33]. A study evaluating different measures of illness burden and treatment in older cancer patients demonstrated that life expectancy, estimated using a Declining Exponential Approximation of Life Expectancy (DEALE), had the largest and most consistent effect on treatment, and that the relationship between age and treatment was independent of burden of illness [34].

The life expectancy of older patients is at least in part a reflection of their underlying comorbidities. Comorbidity, the incidence of multiple conditions, does increase with age. Data from the National Institute on Disability and Rehabilitation Research shows that of individuals with chronic conditions 29%, 51%, and 69% of those in the age groups 18-44, 45-64, and >65 respectively have more than one chronic condition [35]. Furthermore among women in their 60s, 70s, and 80s, the percentage of those with more than one chronic condition is 45%, 61%, and 70%, respectively [35]. Comorbidity in patients with breast cancer is a strong predictor of survival independent of breast cancer stage. One study demonstrated that in breast cancer patients, aged 40-84, with 3 or more significant comorbid conditions (myocardial infarction, other heart disease, diabetes, other forms of cancer, and respiratory, gallbladder, and liver conditions) had a 20-fold increase in non-breast cancer mortality rate and a 4-fold increase in all-cause mortality compared to individuals without co-morbid conditions [36]. Several investigators have examined the effect of age and comorbidity in the treatment of elderly women with breast cancer with the conclusion that aggregate comorbidity does not adequately explain age-related patterns in the initial treatment of elderly patients with breast cancer [35] [38]. Age, 75 years or more, rather than the comorbidity level was found by some researchers to be the most significant risk of receiving non-standard treatment [39]. This may be partly due to the focus of clinicians on the patient's chronological age rather than patient's physiologic age.

Perhaps the variability in the treatment of older breast cancer patients is related to underlying difference in the biology of breast cancer. Elderly women (over 70 years of age) usually have hormone receptor positive (ER+) breast cancer, perhaps reflective of a more indolent tumor pattern, and a higher likelihood of response to hormonal therapy [21]. An analysis of patients >75 showed that 89% were ER (+), 78% were node (-), and 61% had a slow growing cancer (low S-phase) [40]. Among patients who did not receive systemic adjuvant therapy, there was no difference in 5-year survival among patients with node (-) breast cancer and the general population [40]. Only those with node (+) breast cancer had reduced 5-year survival (52%) compared to the normal population (67%) [40]. This data is supported by analysis of Surveillance, Epidemiology, and End Results (SEER) data showing no difference in 5-year survival rates by stage among patients <65, 65-74, and >75, even though older women were generally less aggressively treated [41].

The Frail Patient

There is no universal definition of frailty. Some clinicians define a frail elderly individual as someone dependent in their activities of daily living (ADLS) with three or more comorbid conditions and one or more geriatric related health problems, or syndromes. [42]. Other researchers and clinicians use alternative criteria. On alternative definition for frailty is 3 or more of the following: weight loss (>10 lbs/year), self-reported exhaustion, weakness (grip strength), slow walking speed, and low physical activity. Although there is overlap between comorbidity, functional decline, and frailty, comorbidity is usually a risk factor for frailty whereas disability is on outcome resulting from frailty [43] [44].

Frailty, regardless of definition, is not equivalent to near death. The average life expectancy of a frail person is in excess of 2 years, even though about 50% of frail individuals die within 2 years [45]. Although frail patients may derive palliative benefit from mild chemotherapy for metastatic breast cancer, there is very little data on the benefit of chemotherapy in primary breast cancer in the frail patient.

Quality of Life

Quality of life can be defined as one's overall enjoyment of life. The overall, self-rated, quality of life of women after breast cancer treatment is very good and comparable to age-matched women without breast cancer [46]. However, patients' quality of life is adversely affected by treatment side effects, which are temporary and resolved after treatment is completed [46-50]. Previously reported preference literature has been used to assign values to health states associated with adjuvant chemotherapy for breast cancer. Previous models have assigned a well health-state the best value of 1.0, minor toxicity with chemotherapy a value of 0.9, and major toxicity with chemotherapy a value of 0.8 [51, 52].

The impact of adjuvant therapy on the quality of life of older patients is unclear since they have been underrepresented in past studies. Recent long-term follow-up of disease free survivors of breast cancer have demonstrated poorer sexual, physical and social functioning in patients who received adjuvant chemotherapy [53]. This analysis though also demonstrated the strong association of mental health, social support, and the

overall number of medical conditions with quality of life, all key issues in older patients [53].

A health state is a description of all the medical conditions a person may have in a given situation (example healthy except for non-insulin dependent diabetes and a left leg limp due to a motor vehicle accident). Combining health states with quality of life assignments allows the preference rating of different health states. Unfortunately, preference rating, valuation, procedures often do not take into account underlying comorbidities in measuring the impact of a specific disease, such as breast cancer, on quality of life. Ideally, one would want to measure both the patient's health state valuation without breast cancer but with the patient's baseline comorbidities and the health state valuation that include both breast cancer and baseline comorbidities [54]. In older patients, who generally have more comorbidities, it may be that instead of baseline health state assigned as 1.0, patients have lower baselines perhaps 0.9 or 0.8 [54]. On the other hand, older patients may be more stoic and assign higher health state values, reflecting better quality of life, to disease processes impact younger patients more negatively.

Cost-Effectiveness

Traditionally, cost-effectiveness analysis in breast cancer has focused on a single specific decision point along the treatment path. Most of the published work in this area previous to 1998 has focused on the use of adjuvant therapy in early breast cancer [55]. One study focused on 45 and 60 year-old women with node-negative breast cancer receiving adjuvant chemotherapy estimated a lifetime benefit from chemotherapy of 5.1 and 4.0 quality-adjusted months respectively at a cost of $15,400 and $18,800 per quality-adjusted life year (QALY) respectively [51]. More in-depth analysis using EBCTCG results demonstrated that combined therapy (hormone+chemotherapy) was beneficial and cost-effective in estrogen receptor-positive cancer [52].

More recently, cost-effective analysis has broadened to include a more diverse set of treatments and patients at a later clinical stage. In patients with more advanced disease, research suggests cost-effectiveness of 2nd line [56, 57] and 3rd line chemotherapy [58]. Furthermore, a recent study compared the cost-effectiveness of different hormone therapies in advanced breast cancer [59].

Researchers from RAND evaluated the cost-effectiveness of breast cancer treatment in women with early breast cancer evaluating multiple types of treatments (surgery, radiation therapy, adjuvant therapy, bone marrow transplantation, and reconstruction) using an in-depth evidence-based model with information updated to 1999 [60]. This work focused specifically on patients under the age of 65 who had early stage breast cancer and relied heavily on meta-analyses conducted by the EBCTCG. This analysis strongly supported the following: (1) lymph node dissection with either mastectomy or lumpectomy, (2) radiation therapy with all lumpectomies and for those with larger tumors or positive lymph nodes, (3) adjuvant chemotherapy with Adriamycin and Cytoxan for patients with a greater than 10% risk of dying from breast cancer, (4) additional chemotherapy (e.g. Taxol) for women with lymph-node-positive breast cancer, and (5) five years of hormone therapy, Tamoxifen, in patients with ER+ tumors.

A few cost-effectiveness analyses in older breast cancer patients have been performed. One analysis looking at node (-) ER (-) older breast cancer patients determined that chemotherapy in prolongs survival and that the cost of this benefit, $28,200-44,200/QALY, is high but within the range of commonly reimbursed procedures [52]. More recent work focused on adjuvant chemotherapy in node-negative patients age 60-80 demonstrated QALY benefits of 2.8 and 1.8 months and cost per QALY gained was $31,300 and $44,400 for 65 and 75 year old individuals, respectively [61]. This study distinguished between normal and active life expectancy. Active life expectancy is defined as the percentage of patient's life he/she can perform routine activities of daily living. Therefore, using a patient's active life expectancy as a proxy, one can model functional status and comorbidity [61]. Another approach to incorporating comorbidity in outcomes analysis used 3 levels of comorbidity (better than average, normal, and worse than average) to determine threshold 10-year risk of relapse on mortality demonstrating that even though reduction in relapse is similar between older and younger patients, there is a marked divergence on the effect of chemotherapy on mortality [62].

Research Objective

The research objective of this study is to determine the appropriate treatment for breast cancer in elderly women using an evidence-based analytic model. The most recent data from the EBCTCG meta-analyses in combination with multiple models taking a

wide range of values for the impact of therapy in areas without current data is used to help develop a treatment map comparing 65, 75, and 85 year-old women to 45 year-old women with early stage breast cancer. These data in conjunction with cost of adjuvant therapy are used for both cost-effectiveness and incremental cost-effectiveness of hormone therapy, chemotherapy, and combined therapy.

Methods

Model

Life tables were created for 45, 65, 75, and 85-year olds based on 1999 mortality data of women from the National Center for Health Statistics [63]. To derive non-breast cancer mortality rates these life tables were modified by subtracting breast-cancer specific mortality from total mortality. This provided baseline non-breast cancer life expectancy of 36.41, 19.27, 12.29, and 6.94 for a 45, 65, 75, and 85 year-old woman respectively. Mortality rates for breast cancer were then added to the model (see below), and 6 treatment options were considered: (a) Tamoxifen (HRT) – 2 years, (b) Tamoxifen (HRT)- 5 years, (c) CMFx6 chemotherapy, (d) ACx4 chemotherapy, (e) Tamoxifen (HRT)-CMF, and (f) Tamoxifen (HRT)-AC.

Data Sources

Our primary data sources on the benefits of adjuvant therapy were meta-analyses conducted by the EBCTCG. EBCTCG summaries of the studies of breast cancer treatment are considered the "gold standard" since they include the individual data on the thousands of women enrolled in breast cancer treatment trials over the past two decades. These studies were supplemented with focused reviews evaluating the impact of duration of Tamoxifen use on breast cancer recurrence and survival. In addition, the 1999, 2000, 2001, and 2002 Proceedings of the American Society of Clinical Oncology online abstracts were searched for randomized controlled trials pertinent to our analysis. For data on quality of life during and after breast cancer, previous reviews [60, 64] were utilized.

Measuring Benefits of Treatment

The primary goal of adjuvant breast cancer treatment is the extension of life. This analysis is based on the 10-year mortality following various treatments, as reported in systematic overviews of clinical trials. Since excellent information exists on most early breast cancer treatments, it is possible to calculate the expected incremental benefit from more powerful treatments. These gains are measured in additional years of life, and vary by age and the initial probability of dying.

In the meta-analyses of the many clinical trials of each treatment, the effects of adjuvant therapy are generally modeled in terms of reductions in the odds ratio. The odds of dying = p/(1-p), where p is the probability of dying. Conversely the probability of dying = O/(1+O). So, for example, a 20 % probability of dying corresponds to odd of 0.2/0.8 = 0.25. Suppose a woman given the base case treatment has a 20% probability of dying. If a more powerful treatment reduces the odds of 10-year mortality by 40%, the new odds are (1-0.4) x 0.25 = 0.15. The post-treatment odds correspond to a probability of dying of 0.15/1.15 = 13%. If 100 such women get the more powerful treatment, then 100 x (0.2-0.13) = 7 more will survive for 10 years. This way of modeling treatment effects reflects the generally greater payoff in cases where the initial probability of dying is higher.

This analysis used initial 10-year mortality rates of 20% and 40%, which roughly correspond to the natural history of women with node-negative and node-positive disease respectively. The odds reductions in 10-year mortality from some adjuvant therapies are greater in younger women and decrease with age. For adjuvant therapies, especially in women over the age of 70, data on survival benefits were not readily available. Therefore a range of three possible values was used (high, medium, and low). The high values reflected an impact of the same magnitude as a younger cohort of women. The medium values reflected a continuing linear trend of less benefit with increasing age. The low values represented minimal benefit. Furthermore, hormone therapy has only been shown to work for women with estrogen-receptor positive breast cancer. As a result, we report the preferred strategy for 16 hypothetical women: all combinations of (1) 45, 65, 75, and 85 year old, (2) estrogen receptor positive or negative, and (3) node positive or negative. In addition, selected analyses were performed for the "frail" patient defined as an individual with a 2-year mortality of 50%.

Since breast cancer treatment has negative side effects, a measure of its health cost, negative benefit, is included as part of the calculation of the effects of treatment. The health cost is estimated using previously reported aggregate health state valuations based on a time tradeoff of the number of days of healthy life a patient would give up to avoid the side effects of adjuvant therapy. The baseline health state assigned was 1.0 representing perfect health on a 0-1 scale. Previously derived treatment health states valuations of 0.99 of baseline for hormone therapy, 0.90 of baseline for minor toxicity with chemotherapy and 0.80 of baseline for major toxicity with chemotherapy were used in this analysis [52] [65]. The health cost of living 100 days with the minor toxicities of chemotherapy would therefore be 100*(1-0.90*1) = 10 days. A course of chemotherapy over 6 months (168 days based on 28 day cycles) where a patient would be getting chemotherapy one day a month, major toxicity, but having minor toxicity for the rest of the time would have a health cost = (0.97*168)*(1-0.9*1)+(0.03*168)*(1-0.8*1)=17.3 days.

The added years of life minus the health costs of treatment in days of life represents the added quality-adjusted life years we used as our health outcome for cost-effectiveness analysis. Sensitivity analysis using baseline health states of 0.8 and 0.6 was also performed. Since people usually value immediate health more than future health, it is standard to discount health gains as well as health costs. Following standard practice, future life years were discounted at 3% similar to our cost analysis. Sensitivity analysis was performed using 0%, 5%, 10%, and 20%. The higher values in sensitivity analysis were used to determine if very high discounting of future life would alter treatment strategies.

Measuring Costs

Costs of treatment include direct medical costs, as well as indirect costs, such as costs incurred by the patient for transportation or lost wages. Since this analysis is focused on the cost to Medicare of providing high quality care cost-effective to the elderly, only direct costs of initial treatment were included in the analyses. Cost estimates were based on the cost of health services in 2001 Medicare-allowed charges. For drug costs, 2001 Average Wholesale Prices (AWP) were used, but sensitivity analysis was also conducted using the generally lower Public Health Service (PHS) price that

pharmaceutical manufactures offer to facilities with a disproportionate share of indigent patients. The cost of adjuvant therapy includes the cost of the initial consultation, chemotherapy administration, the drugs, biweekly laboratory testing, and biweekly follow-up visits. Costs were discount using a standard 3% discount rate, but sensitivity analysis was also performed using discount rates of 0% and 5%. Our costs estimates were compared to published values in other cost-effectiveness analyses and to previous analysis performed by RAND researchers in 1999 focused on patients under the age of 65. The assumptions behind the cost estimates are described at the end of the Methods section.

Cost-Effectiveness Analysis

In order to better describe the treatment interventions, each with a different potency and expense, plot of the costs and benefits were constructed. These plots were then used not only for standard cost-effectiveness analysis, but for incremental cost-effectiveness analysis. We were also able to use these plots to determine the dominant therapies for each age group. For older patients for whom the therapeutic benefit was uncertain, scenarios using the high, medium, and low range of possible responses were used to determine incremental cost-effectiveness.

Assumptions

1. Based upon information from published clinical trials on the toxicities of chemotherapy, it was estimated that 10% of patients would require daily Granulocyte Colony Stimulating factor (G-CSF) for 10 days each cycle after the 1st cycle of chemotherapy to treat low white blood counts [66] [67].
2. In addition, it was estimated that 3% would require hospitalization for fever and neutropenia (low blood counts).
3. For models on Tamoxifen therapy, the percentage patients dying in the first five years on therapy in our life expectancy models was used in calculating both the costs and disutility of treatment.
4. Tamoxifen and chemotherapy were assumed to have no impact on non-cancer mortality rate.

5. Node-negative and node-positive patients were assumed to have a 20% and 40% chance of dying respectively after surgical resection for breast cancer.
6. It was assumed that there were no long-term effects of adjuvant therapy and therefore only adjustments were made for toxicity during therapy.
7. The 10-survival benefit from chemotherapy was divided equally over all the years.

Results

This section presents the results of out analysis of treatment costs, treatment benefits, effects of treatment on quality of life, the cost-effectiveness and the incremental cost-effectiveness of different treatments.

Costs

The first two columns of **Table 1** show estimates of the costs of treatment using two sets of drug prices: the 2001 Average Wholesale Price (AWP) and the Public Health Service Price (PHS). The cost estimates discussed in this analysis are based on AWP, which is taken as the base case, unless otherwise noted. The PHS cost for treatment is 40-60% less than the AWP cost. The cost trends for adjuvant treatment between AWP and PHS are similar, except that AC chemotherapy is less expensive than CMF in the PHS pricing system. Hormone therapy (HRT) is listed as a class of therapy but in reality reflects Tamoxifen pricing. Since hormone therapy can last either 2 or 5 years in the models presented in this analysis, the price is reflective not only of a discount rate of 3% annually for future costs but also the proportion of individuals expected alive in the future. The proportion of individuals alive is dependent on both breast cancer hazard, which is greater in Node (+) patients than Node (-) patients, and the effect of breast cancer treatment. The costs listed in **Table 1** are reflective of a 45 year old with Node (-) breast cancer unless otherwise noted.

The health costs of treatment in **Table 1** reflect the days lost due to the decreased quality of life of patients during treatment. Since CMF for 6 cycles occurs over a longer time period than AC for 4 cycles, it is not surprising that the health costs for CMF are greater, 17 days, than AC, 11 days. Five years of tamoxifen therapy has approximately

the same health costs, 17 days, as CMF x 6 cycles. The use of 2 years of tamoxifen rather than 5 years reduces both financial and health costs. Combination adjuvant therapy has greater financial and health costs than either hormone or chemotherapy alone.

Benefits of Treatment

Table 2 summarizes the estimates used for the benefits of treatment in terms of a 10-year odds reduction in mortality. The benefits of adjuvant therapy vary with patient age and the type of adjuvant therapy used. For hormone therapy, tamoxifen use for five years is not age-sensitive, with odds reduction of mortality ranging from 0.32-0.34 for those between the age of 45-75. The use of tamoxifen for two years has much less impact in younger patients aged 45 and 65. However, for patients over 75, data suggests equivalent or better effect with two years of tamoxifen (0.36) compared to five (0.34). For adjuvant chemotherapy, there is a benefit in younger women 45-65 at risk for recurrent breast cancer, but the benefit is substantially greater for younger women. The use of AC x 4 cycles is a little more effective than CMF x 6 cycles in women under 50. Chemotherapy is much less efficacious in older women. For a 65-year old woman, hormone therapy is much more efficacious than chemotherapy. There is an advantage in treating with combined chemotherapy and hormone therapy in younger women.

Since there are no good outcomes data on the use of hormone therapy in patients age 85 and chemotherapy in patients age 75 and 85, a range of values were considered. The high bound used assumes that older patients, 75 and 85, have a response equal in magnitude to a 65 year old. The middle bound assumes that affect of chemotherapy decreases with age almost linearly. The low bound assumes almost no benefit from chemotherapy.

The life expectancy without breast cancer is 36.41, 19.27, 12.29, and 6.94 years for a 45, 65, 75 and 85 year old, respectively, **Tables 3-7.** With breast cancer prior to adjuvant therapy (BCPATH), the life expectancy for a 45 year old is 16.27 years in node (-) disease and 12.06 years in node (+) disease, **Table 3**. The BCPATH life expectancy for a 65 year old is 11.78 and 9.58 years for node (-) and node (+) disease, respectively (**Table 4**). The BCPATH life expectancy for a 75 year old is 8.75 and 7.52 years for node (-) and node (+) disease, respectively (**Table 5**). The BCPATH life expectancy for a 85

year old is 5.63 and 5.11 years for node (-) and node (+) disease, respectively (**Table 6,7**).

Adjuvant therapy always adds more years of life to those with Node (+) breast cancer than in those with Node (-) breast cancer. In a 45 year old, **Table 3,** with Node (+) ER (+) disease there is more added years of life with chemotherapy (AC – 2.09 years, CMF – 1.80 years) than with 5 years of hormone therapy 1.67 years. The greatest benefit (4.19 years) is with a combination AC x 4 cycles and hormone therapy for 5 years in a 45 year old with node (+) ER (+) breast cancer. In a 65 year old with Node (+) ER (+) breast cancer, the added years of life from even two years of hormone therapy (0.30 years) is better than chemotherapy (AC 0.26, CMF 0.16, **Table 4**). Nevertheless, the combination of hormone and chemotherapy is still superior to the use of hormone or chemotherapy alone.

In a 75 year old, the benefit of chemotherapy is less uncertain. In a 75 year old with Node (-) breast cancer, chemotherapy primarily adds benefit if one assumes the same benefit seen in a 65 year old, the high bound (**Table 5**). If lower levels of benefit are used, the health effects outweigh the benefits and negative years of added life result. For those with ER (+) breast cancer, it is very difficult to surpass the benefit of either 2 or 5 years of hormone therapy. The high level values of combined therapy are superior to hormone therapy. On the other hand, the mid level values of combined therapy are comparable to the use of hormone therapy alone since the benefit of chemotherapy is balanced by its negative health cost. The greatest benefit would be derived if two years of hormone therapy, instead of five years, can be used in combination with chemotherapy (high bound) with an equivalent effect.

In an 85 year old, the benefit of both hormone and chemotherapy is uncertain. If one assumes that hormone therapy is not age sensitive (high bound), then the addition of chemotherapy adds very little in patients with ER (+) breast cancer, **Table 6**. The high-bound for combined treatment is comparable to hormone therapy alone since the health costs again offset the health benefit. The benefit of combined therapy increases if it can be achieved with only two years of hormone therapy since the health costs of hormone therapy decrease. For an 85 year old with node (-) breast cancer, only the high bound of AC therapy added any years of life (0.02), see **Table 7**. In node (+) breast cancer, the high bound of both AC and CMF had overall positive benefits.

The cost-effectiveness of adjuvant therapy is listed in **Tables 8-11.** In a 45 year-old woman, the cost-effectiveness of most adjuvant strategies, with the exception of 2 years of hormone therapy, is comparable ranging from $3,217 - $4,906/ QALY in node (-) disease and from $2,538 -$ 3,718/QALY in node (+) disease, **Table 8.** Chemotherapy and combined therapy is more cost-effective than hormone therapy for five years alone. Since hormone therapy for 2 years provides sub-optimal benefit, it is most cost-ineffective at $7,293/QALY of the options. In a 65 year-old woman, hormone and combined strategies are much more cost-effective, $10,194 - $13,842/QALY in node (-) disease and $6,520 - $9,060/QALY in node (+) disease, than using adjuvant chemotherapy alone, $30,451-$33,137/QALY in node (-) disease and $22,941-$28,547 in node (+) disease, **Table 9.**

The cost-effectiveness analysis in 75 and 85 year old becomes more complex due to the uncertainties associated with benefits from treatment. In a 75 year old, hormone therapy of either 2 or 5 years is very cost-effective, $7,584/QALY or $19,530/QALY respectively in node (-) and $4,503/QALY or $10,965/QALY in node (+) disease, **Table 10.** Chemotherapy alone in those with ER (-) disease is still relatively cost-effective in those with node (+) disease, $42,605 - $65, 251/QALY, if one assumes a high bound for its effect. Chemotherapy alone is less cost-effective in node (-) disease and cost-ineffective if the mid and low bounds for benefit are used. On the other hand, combined therapy still is valuable. Using the high bound for benefit and 5 years of hormone therapy, the cost-effectiveness of combined therapy is about $28,000/QALY for node (-) disease and $15,000/QALY for node (+) disease. If one can gain the same benefit from combined therapy with just 2 years of hormone therapy instead of 5 years, the cost-effectiveness improves even further at about $18-19,000/QALY in node (-) disease and $11-13,000/QALY in Node (+) disease, using the high bound. The mid level cost-effectiveness of combined treatment with 2 years of hormone therapy is comparable to the high bound of combined treatment with 5 years of hormone therapy.

In an 85 year old, very little is cost-effective except for hormone therapy and the high bound for combined therapy. Assuming that the benefits of hormone therapy are age-insensitive, the cost-effectiveness of either 2 or 5 years of hormone therapy is $18,206/QALY or $55,085/QALY, respectively, in node (-) disease and $10,011/QALY or $26,463/QALY, respectively, in node (+) disease, **Table 11.** Chemotherapy alone is

not cost-effective (>$100,000/QALY). Combined therapy is really only cost-effective if the high bound for benefit is assumed, with the most-effectiveness occurring in those with node (+) disease about $42,000/QALY if 5 years of hormone therapy are needed or $27-28,000/QALY if 2 years of hormone therapy is sufficient. It should be noted that combined therapy in node (+) positive disease using only 2 years of hormone therapy if using the mid level values for benefit is still generally cost-effective at $60-72,000/QALY.

A model run for a frail patient, with a 50% 2 year survival, demonstrated that even using 2 years of hormone therapy for palliation was not cost-effective (data not shown).

In describing, interventions of various potency and expense, a plots of their costs and benefits, **Figures 1-6**, are useful to understand the incremental cost-effectiveness of one adjuvant treatment compared to another, **Tables 12-15**. For example, in a 45-year old with node (-) and node (+) ER (+) breast cancer, the plot of adjuvant therapies is seen in **Figure 1a and 1b.** The plot of the costs of each therapy on the horizontal axis, and the expected added years of life on the vertical axis. The higher the point is on the graph the more effective the treatment, and the farther the point to the right the more expensive the treatment. Treatments that are better (i.e. higher) and cheaper (i.e. to the left) are preferred. Treatment strategies that form the solid line connecting the points lying left and upward are the economically rational subset of choices. Points lying beneath the line represent treatment strategies that are not as effective for any given amount of money as a point lying on the line and are said to be "dominated" strategies. The slope between any two points represents the inverse of the incremental cost-effectiveness ratio (i.e. the steeper the slope of line the more cost-effective the incremental addition of therapy). A flatter slope connecting two treatment choices reflects a very high incremental cost-effectiveness with decreased returns in terms of effectiveness per expenditure. When point lie close in terms of benefit and cost, they should be considered reasonable alternatives. This is especially important when cost differences are relatively insignificant since non-economic reasons, such as patient or physician acceptability, play a much larger role in treatment selection.

The graph shows that in a 45 year-old woman, we would not want to give hormone therapy (HRT) x 5 years alone. Since chemotherapy with CMF x 6 is cheaper

and more effective that HRT x 5 years, HRT x 5 years is "dominated" by CMF and should not be used (from a cost-effectiveness standpoint). To buy as much health as possible for a given budget (ex: Medicare budget), and a given patient, decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point. So with initial funds for adjuvant therapy in 45 year old women, adjuvant therapy would start by giving CMF. Once all women receive CMF, then combined therapy with CMF can be given (HRT-CMF). The health gains per dollar for funding HRT-CMF must be measure relative to the alternative uses of the money. Treatment such as HRT x 2 years and AC which lie below the line are not cost-effective, since more health can be bought using a strategy of giving some people CMF and some people HRT-CMF, than giving AC to everyone. The incremental cost-effectiveness of CMF and combined therapy in a 45 year old with ER (+) breast cancer is <$10,000/QALY, **Table 12.** In a 45-year old with ER (-) disease both CMF and AC are options, and both have incremental cost-effectiveness again <$10,000/QALY.

In a 65-year old woman with either node (-) or node (+) ER (+) disease, 5 years of hormone therapy dominates both CMF and AC chemotherapy, **Figures 2a-2b**. Combined therapy is also on the incremental cost-effectiveness line. In node (-) breast cancer HRT-AC dominates HRT-CMF, whereas in node (+) breast cancer both HRT-CMF and HRT-AC are viable options. The incremental cost-effectiveness of adding AC to hormone therapy is $22,220/QALY in node (-) disease, and $12,890/QALY when CMF is added in node (+) disease, **Table 13**. The incremental cost-effectiveness of HRT-AC over HRT-CMF in node (+) patients is $13,972/QALY. On the other hand in patients who are node (-) ER (-), the incremental cost effectiveness of chemotherapy is $30,451/QALY for CMF and $46,572/QALY for AC, whereas in node (+) disease it is $28,547/QALY for CMF and $13, 972/QALY for AC.

The incremental cost effectiveness analysis in a 75 year old is a little more complex due to uncertainty of the benefit of adjuvant chemotherapy and combined therapy in this age range. In a 75 year-old with either node (-) or node (+) disease ER (+) disease, two years of hormone therapy "dominates" all other adjuvant therapy except combined therapy at the high bound, **Figure 3a-3b.** Assuming the EBCTCG data on 2 years of HRT therapy and the high bound for combined (5 years of hormone therapy HRT5) therapy labeled *Scenario 1*, the incremental cost-effectiveness of adding

chemotherapy to hormone therapy only makes sense in node (+) patients at $60,925/QALY using HRT5-AC, **Table 13**. If on the other hand, one rejects the EBCTG data and believes 5 years of hormone therapy is required for maximal benefit, labeled *Scenario 2,* it might be incrementally cost-effective to add chemotherapy in both node (-), $54,530/QALY, and node (+) patients, $27,406/QALY using HRT5-AC. Clearly, adding chemotherapy to hormone therapy at the mid-level effect for combined therapy is not incrementally cost-effective, *Scenario 3.* Alternatively, the analysis might be different if two years of hormone therapy (HRT2) is just as good as five years in combined therapy, **Figure 4a-4b**. In this case, HRT2 dominates all other adjuvant therapy except for combined therapy at the high bound. If one believes the high bound can be achieved in a 75 year old with only HRT2, *Scenario 4,* then chemotherapy may be incrementally cost-effective for both node (-) and node (+) breast cancer patients at $66,308/QALY and $33,174/QALY respectively for HRT2-AC, **Table 13**. If the high bound for combined therapy HRT2-AC, *Scenario 5,* then adding chemotherapy is not incrementally cost-effective. In a patient who is ER (-) , AC therapy dominates CMF and is incrementally cost-effective if the high bound for benefit is assumed, $75,559/QALY in node (-) disease and $42,605/QALY in node (+) disease, **Table 13** *Scenario 6 and 7.*

Given the fact that the benefit of both hormone therapy chemotherapy is unclear in 85 year old patients, it is not surprising that more scenarios need to be evaluated, **Figures 5a-5d**. *Scenario 1* is that hormone therapy is age-insensitive and that two years of hormone therapy works best when used alone but five years is needed for the high bound of combined therapy. In this case, for an 85 year-old node (-) ER (+) patient (**Figure 5a**), HRT2 is dominant over all other adjuvant therapy, and even though in node (+) patients (**Figure 5c**) there is some benefit to combined therapy (high bound), it is not incrementally cost-effective, **Table 14**. Even if HRT2 really is not as effective as HRT5, *Scenario 2,* only two years of hormone therapy is incrementally most cost effective in patients with node (-) breast cancer since it costs $45,418/QALY for HRT2 and $66,686/QALY for HRT5. On the other hand a case could be made for using the full five years in a node (+) patients since it costs $24,460/QALY for HRT2 and an additional $28,465/QALY for HRT5. Adding chemotherapy to either *Scenario 1 or 2* is not incrementally cost-effective, **Table 14.**

If hormone therapy is really age-sensitive at older ages, then one would expect reduced benefit (mid values), **Figures 5b and 5d**. Still this reduced benefit may be equivalent if either 2 years or 5 years of hormone therapy is used, *Scenario 3*. If this is the cases, then combined therapy (high bound) is incrementally cost-effective only in node (+) disease at $53,823/QALY, **Table 14.** If on the other hand, only 5 years of hormone therapy provides this benefit, *Scenario 4*, treating a node (-) patient is not cost-effective. In a node (+) patient, HRT5 and combined therapy (HRT5-CMF, HRT5-AC) all have incremental cost-effectiveness from $41,000-54,000/QALY, **Table 14**. Alternatively, it may be that 2 years of hormone therapy is optimal both alone and in combination with chemotherapy (high bound), *Scenario 5*. If this scenario is true, it might be beneficial to add chemotherapy to node (-), $59,730/QALY or node (+), $29,850/QALY, patients. *Scenario 6* represents the unusual circumstance where two years of hormone therapy is not sufficient alone, but sufficient in combined therapy and where 5 years of hormone therapy alone provides optimal benefit. In this scenario for a node (-) 85 year-old HRT5 is dominated, and the incremental cost-effectiveness of adding AC to HRT2 is $59,730. In a node (+) 85 year-old, neither HRT5 nor HRT2-CMF (high) are dominated, and the incremental cost-effectiveness ranges from $24,000-$35,000/QALY for all non-dominated adjuvant therapies, **Table 14**. It should be noted that for an 85 year-old patient, combined therapy was found to be incrementally cost-effective in selected instances only when the high bound was assumed. The mid level and low bound were always dominated or demonstrated negative benefit. Chemotherapy was not cost-effective in 85 year-old patients with ER (-) breast cancer, **Table 14,** *Scenario 7.*

Finally an analysis on the benefits of hormone therapy in the frail breast cancer patient was performed. The best-case scenario where maximal benefit would be derived from only two years of hormone therapy was used. Under this scenario hormone therapy, the cost-effectiveness of hormone therapy was $391,198/QALY, **Table 15.**

Sensitivity analysis was performed to determine the degree of efficacy needed for chemotherapy and combined therapy to be cost-effective. Two cut-off points for cost-effectiveness were used $50,000/QALY (A) and $100,000/QALY (B). The reference efficacy used was the benefit of the selected treatment in a 65 year old (i.e. the high bound). In a 75 year old with node (+) breast cancer AC chemotherapy would need to be as 86% and 50% as efficacious as the high bound in order to be cost-effective at the A

and B cut-offs see **Table 16**. CMF therapy would not meet cut-off A even at 100% of the high bound, but could cut-off B at 94% the efficacy of the high bound. The incremental benefit of combined therapy, HRT-AC, was cost-effective only if one could assume the treatment was either 89% or 83% as efficacious as the high bound for cut-off A and B respectively, **Table 16**. For a 75 year old with node (-) breast cancer, cut-off A could not be reached for chemotherapy and combined therapy even if efficacy was identical to the high bound. However, cut-off B could be reached if efficacy for AC and HRT-AC was 82% and 89% of the high bound respectively.

Additional sensitivity analysis was performed using a variety of discount rates. This was little difference in the rank order of treatment choice using 0%, 5%, and 10% discount rates. However, at 25% discount rate, several interesting results were obtained. In a 65 year old with Node (+) ER (+) breast cancer, the benefit of combined therapy added at most 0.04 added years of life costing about $150,000/QALY, see **Table 17**. In a 65 year old with Node (+) ER (-) breast cancer, the use of chemotherapy added at most 0.02 added years of life costing about $300,000/QALY. These benefits decreased substantially as with drops in the baseline QALY. The benefit of combined therapy dropped to 0.02 with baseline QALY of either 0.8 or 0.6, and no benefit was seen in chemotherapy alone at these reduced baseline QALYs, **Table 17**.

A similar analysis could be performed on in 75 and 85-year old patients with Node (+) ER (+) breast cancer looking at the benefit of hormone therapy. The maximum benefit at a discount rate of 25% is 0.08 and 0.05 added years of life for a 75-year old and 85-year old, respectively. For an 85 year old this would be slightly over $100,000/QALY. These benefits drop with decreases in baseline QALY. Chemotherapy ceases being cost-effective in a 75 year old when baseline QALY is 0.06, **Table 18**.

Discussion

In comparing the benefits of adjuvant therapy in early breast cancer between different age groups of women, a starting reference point should be the maximum amount of life expectancy that can be gained with treatment. For example, in a patient with node (+) breast cancer the maximum life expectancy that could be gained by adjuvant therapy would be 24 years in a 45 year-old woman, but a little less than 10 years in a 65 year-old

woman. In older node (+) breast cancer patients, the life expectance drops significantly with the maximum gain being less than 5 years for a 75 year-old woman and less than 2 years for an 85 year-old woman. These numbers reflect the maximum gain meaning the existence of a magic pill that would cure all patients without toxic side effects.

In reality, the adjuvant treatments (hormone, chemotherapy, or combined modalities), which are available, allow the gain of only a fraction of the "maximum" life expectancy. In the models presented in this paper, the highest gains were 4.2, 1.4, 0.75, 0.28 years, respectively, for 45, 65, 75, and 85 year-old women with node (+) breast cancer. In older women aged 75 and 85, these "highest" gains reflect the most optimistic assumptions of treatment efficacy, a response equivalent to that of a 65 year-old. Although most of the available data supports the notion that hormone therapy is age-insensitive, data for chemotherapy suggests marked age-sensitivity. Even though age in these contexts most often refers to chronological age, physicians need to assess the older patient carefully to derive an estimate of a patient's physiological age (i.e. 75 year-old but healthier than a typical 65 year-old or a frail 75 year-old with an illness burden greater than a typical 85 year-old).

The Average Wholesale Price (AWP) costs of hormone therapy (Tamoxifen) and AC chemotherapy were relatively close differing only by $600.00. The difference in AWP cost between AC and CMF chemotherapy was also only about $600.00. The major difference in AWP costs, $3,600-8,000, occurred between combination hormone-chemotherapy and single-modality treatment. All costs were lower under Public Health Service pricing. Given the fact that the costs and benefits of AC and CMF are relatively close, the decision as to the appropriate treatment must include non-economic factors, such as physician and patient preference. For example, AC is not the best choice in an 85 year-old with poor cardiac function. Furthermore, from a policy standpoint, the recommended therapy must take into account societal cut-offs for health expenditures. It may be that combination therapy always adds more benefit than single modality therapy but at a very high cost. Whether the cut-off for appropriate cost-effective therapy is $50,000/QALY or $100,000/QALY or some other number is a complicated social and political decision.

A simplified summary of the data presented in this paper is shown in **Figure 7 and 8**. All modalities of adjuvant therapy are cost-effective in women at the age of 45.

Chemotherapy though is more cost-effective than hormone therapy and combination therapy is incrementally as cost-effective as chemotherapy alone, especially in node (+) women, **Figure 7 Panel A**. In 65 year-old women, hormone therapy is more cost-effective than chemotherapy, and combination therapy is incrementally still cost-effective, **Figure 7 Panel B**. The gap in cost-effectiveness between hormone therapy and chemotherapy continue to increase with age, even when one assumes high efficacy from chemotherapy, **Figure 7 Panels C and D**. Nevertheless combination chemotherapy is still incrementally cost-effective ($50,000-$100,000/QALY) in 75 year-old node (-) and node (+) patients and 85 year-old node (+) patients provided one ***assumes*** that chemotherapy will have the same efficacy in these older populations as seen in a 65 year-old. If instead, there is lower efficacy from chemotherapy in older breast cancer patients, chemotherapy is not cost-effective in either node (-) or node (+) patients, **Figure 8 Panels B and D**. Interestingly, combination therapy may still be incrementally cost-effective ($120,000/QALY) in a node (+) assuming lower efficacy from chemotherapy and ***if one also assumes*** that 5 years of hormone therapy are needed for maximum benefit, **Figure 8 Panel B**. If maximal benefit can be obtained with only two years of chemotherapy, then combination therapy is no longer a cost-effective option, **Figure 7 Panel D**.

In order to begin to formulate policy on cost-effective therapy in early breast cancer patients, it is useful to model benefit using a wide range of expected benefits, and to use scenarios that represent different views on the likely benefit of treating on older individual with breast cancer. This approach allows physicians, patients, and policy makers participating in the debate on appropriate treatment in the older breast cancer patient to see their viewpoint with respect to other prevailing viewpoints. This type of analysis can point to areas where more information is most critical for which clinical trials would be essential.

The analysis in this paper suggests two areas for which clinical trial data would be very important. First, the optimal duration of hormone therapy in older, greater than age 70, ER (+) breast cancer patients needs to be better defined, whether it is 2 years, 5 years, or some intermediate duration. The answer to this question will not only improve older patient's quality of life be reducing side effects and complications from hormone therapy, but will also help in addressing the overall incremental benefits and cost-effectiveness

from combination therapy. Second, there is a subgroup of older breast cancer (node +, Her2-Neu +) patients for whom hormone therapy may be less efficacious and chemotherapy, especially anthracyclines, may be more efficacious. Our analysis shows that chemotherapy and combination therapy can still be cost-effective in older patients under certain conditions. Clinical trials focused the use of chemotherapy or combination therapy on older patients with poor prognostic features might help resolve some of these uncertainties.

Chemotherapy drugs are often available at discount prices. To evaluate the impact of lower drug costs on the cost-effectiveness of adjuvant therapy, a sensitivity analysis using Public Health Service discounted pricing for drugs (the price that pharmaceutical companies will sell to hospitals with a disproportionate share of indigent patients) was performed (data not shown). While all adjuvant therapies became more cost-effective when the price of the drug was discounted, there were minor differences in the rank order of the different strategies. AC always dominated CMF in this pricing system and combined therapy with AC always dominated combined therapy with CMF. This is a result of AC being less expensive than CMF under the PHS pricing system. Thus, obtaining drugs at the PHS price generally exaggerated the results of our underlying model.

The cost of treating metastatic breast cancer or of providing hospice care to patients who die are downstream costs that are in part affected by current decisions, Previous analysis (Malin and Keeler, 2000) has shown that including the effects of averted downstream costs in our model would slightly increase the cost-effectiveness of more-powerful adjuvant therapies. Patients who recur will have greater expenses than will those who do not. Insofar as one adjuvant therapy strategy decreases recurrences and death from breast cancer more than another, it will have lower downstream costs. These reductions can offset some of the initial treatment costs. Each 1% decrease in the probability of dying at 10 years was shown to reduce the discounted costs of treating downstream metastatic cancer by about $450 for a 45 year-old woman and $250 for a 60 year-old woman in prior analysis [60]. Given the fact that the older a patient becomes, the greater the likelihood of dying from competing causes rather than breast cancer, the overall costs of treating downstream metastatic cancer are likely to be even smaller for a 75 year-old and 85 year-old woman. Therefore, these costs were not incorporated into the

analysis presented since their effect is fairly small compared to the costs of initial treatment, and would probably not alter any of our recommendations.

Typically low discount rates, 0-5%, are used in cost-effectiveness analysis. However, there has been published work that in certain populations (such as gamblers, those engaged in high-risk sex or drug use, or depressed and suicidal individuals) that a high discount rate is internalized by individuals. These high discount rates may also apply to older patients diagnosed with cancer and may help provide a rational explanation as to why these patients may decline treatment. Often it helps to translate discount rates into an intertemporal trade-off for illustration. For example, a discount rate of 3 % means one would be willing to give up 7 days of life today for 9.4 days of life 10 years from now, whereas a 15% discount rate means one would be willing to give up a week of life now for slightly less than 1 month of life 10 years from now. In our analysis we looked at a 25% discount rate, which translates to giving up a week now for slightly over 2 months of life 10 years from now. Furthermore, these higher discount rates may be internalized by those patients with a worse baseline quality of life due to comorbidity. As a result of these high discount rates and lower baseline QALY, some younger patients, age 65, may decline chemotherapy or combined therapy and some older patients may, age 75 and 85, might decline hormone therapy.

There are many limitations to this type of study. Medicare-allowed charges were used to estimate costs. While these are not the actual resource costs of providing medical care, they may be good surrogates, since they are determined in a manner that attempts to incorporate the variation in intensity of resource use. In addition, the system of Medicare charges is now frequently used as a benchmark in negotiating managed-care contracts.

In using decision-analysis modeling, many assumptions were used to limit the number of models and scenarios used in the analysis. The same probability of neutropenia and resulting hospitalization was used for all age groups. A strong argument can be made that these neutropenic events increase with age due to decreasing bone marrow reserve. As a result older patients may have increased costs due to the use of G-CSF, antibiotics, and hospital stays. An even bigger assumption is that adjuvant therapy, whether chemotherapy or hormone therapy, has no impact on non-cancer mortality rate. Chemotherapy can complicate underlying comorbidities (such cardiac disease, diabetes, or hypertension) and accelerate functional decline. It may be that the age-sensitivity of

chemotherapy is a reflection not only of the impact of competing underlying causes of mortality but also the interaction between the treatment and comorbidities. Finally, the preference weights used to estimate the health cost associated with the side effects of chemotherapy are based on aggregate data derived from limited studies. In other cancers, such as ovarian cancer, studies have used preference states for chemotherapy showing greater impact on quality of life than the analysis in this study used (weights of 0.65 for chemotherapy state instead of 0.80) [68]. This issue is particularly important in terms of hormone therapy since the side effects, such as hot flashes, were minimally weighted (0.99 weight for hormone therapy). It may be that older women are in fact more stoic about the side effects from hormone therapy than those who are perimenopausal, but more research on the quality of life effect of hormone therapy in older women needs to be performed for better estimates of negative health effect to be generated.

Cost-effectiveness analyses are only as good as the available data on costs and benefits. It is fortunate that there are excellent meta-analyses on standard adjuvant therapy for breast cancer in younger women. In younger women the uncertainty in estimated benefits is small. The standard deviations of the estimates in mortality reductions in the trials vary from 1/2 to 1/10 of the mean estimated used in this analysis. However, for older women over 70, traditionally under-represented in clinical trials, there is a high degree of uncertainty in expected benefit. In part, this uncertainty is due to decreased longevity, but also compounded by heterogeneity in older individuals in terms of comorbidity and functional status not to mention biologic differences in the breast cancer itself. The use of decision-analysis models can help provide a framework to examine the areas, magnitude, and consequence of treatment uncertainty. Visuals maps of alternative possibilities, or scenarios, can foster debate, which will hopefully in turn motivate well-developed and strategic clinical trials to help address the most important issues related to the treatment of older women with early breast cancer. Focusing analytic and clinical research on older cancer patients will help ensure that high-quality cost-effective care can be provided in the next several decades, during which our countries finances may become strained.

References

1. Hobbs FB: The Elderly Population, in, U.S. Census Bureau, January 2001
2. SEER: SEER 1978-1998 Public Use Data, in, National Cancer Institute, 1999
3. Frame PS: Preventive care for the elderly: getting by in the absence of evidence. *Am Fam Physician* 59:1747-1748, 1750., 1999
4. Dwyer JT, Gallo JJ, Reichel W: Assessing nutritional status in elderly patients. *Am Fam Physician* 47:613-620., 1993
5. MacDonald P, Johnstone D, Rockwood K: Coronary artery bypass surgery for elderly patients: is our practice based on evidence or faith? *Cmaj* 162:1005-1006., 2000
6. Conte HA, Chen YT, Mehal W, Scinto JD, Quagliarello VJ: A prognostic rule for elderly patients admitted with community-acquired pneumonia. *Am J Med* 106:20-28., 1999
7. Fisher B: The evolution of paradigms for the management of breast cancer: a personal perspective. *Cancer Res* 52:2371-2383, 1992
8. Johnstone P, Notron M, Riffenburgh R: Survival of Patients With Untreated Breast Cancer. *Journal of Surgical Oncology* 73:273-277, 2000
9. Bloom H, Richardson W, Harries E: Natural History of Untreated Breast Cancer (1805-1933). *British Medical Journal*:213-221, 1962
10. Quiet CA, Ferguson DJ, Weichselbaum RR, Hellman S: Natural history of node-negative breast cancer: a study of 826 patients with long-term follow-up. *J Clin Oncol* 13:1144-1151., 1995
11. Quiet CA, Ferguson DJ, Weichselbaum RR, Hellman S: Natural history of node-positive breast cancer: the curability of small cancers with a limited number of positive nodes. *J Clin Oncol* 14:3105-3111., 1996
12. Fisher B: Systemic chemotherapy as an adjuvant to surgery in the treatment of breast cancer. *Cancer* 24:1286-1289, 1969
13. Aapro MS: Adjuvant therapy of primary breast cancer: a review of key findings from the 7th international conference, St. Gallen, February 2001. *Oncologist* 6:376-385., 2001

14. EBCTCG: Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. Early Breast Cancer Trialists' Collaborative Group. *N Engl J Med* 319:1681-1692., 1988

15. EBCTCG: Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 351:1451-1467., 1998

16. Stewart H: The Scottish trial of adjuvant tamoxifen in node negative breast cancer. Scootish Cancer Trials Breast Group. *J Natl Cancer Inst Monogr*:117-120, 1992

17. Tormey D, Gray R, Falkson H: Postchemotherapy adjuvant tamoxifen bejond 5 years in patients with lymph node-positive breast cacncer. Easter Cooperative Oncology Group. *J Natl Cancer Inst Monogr*:117-120, 1996

18. Fisher B, Constantino J, Redmond CK, al. e: Endometrial cancer in tamoxifen-related breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project NASBP B-14. *JNCI* 86:527-537, 1994

19. Holli K, Valavaara R, Blanco G, al. e: Safety and efficacy results of a randomized trial comparing toremifene and tamoxifen in postmenapausal patients with node-positive breast cancer. *J. Clin. Oncol.* 18:3487-3494, 2000

20. Baum M, Group AT: The ATAC (Arimidex, Tamoxifen, Alone or in Combination) adjuvant breast cancer trial in post-menopausal (PM) women., in *The 24th San Antonio Breast Cancer Symposium*, 2001

21. NIH Consensus Statement on Adjuvant Therapy for Breast Cancer, in *2000 NIH Consensus Conference on Adjuvant Breast cancer* (vol 17), 2000, pp 1-35

22. EBCTCG: Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* 339:71-85, 1992

23. EBCTCG: Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 352:930-942., 1998

24. Fisher B, Brown A, Dimitrov N, Poisson R, Redmond C, Margolese R, Bowman D, Wolmark N, Wickerham D, Kardinal C, al. e: Two months of doxorubicin-

cyclophosphamide with and without interval reinduction therapy compared with 6 months of cyclophosphamide, methotrexate, and fluorouracil in positive-node breast cancer patients with tamoxifen-nonresponsive tumors: results from the National Surgical Adjuvant Breast and Bowel Project B-15. *J Clin Oncol* 8:1483-1496, 1990

25. Albain K, Green S, Osborne K, al. e: Tamoxifen (T) versus cyclophosphamide, Adriamycin, and 5-FU plus either concrueent or sequential T in postmenopausal, receptor(+), node (+) breast cancer: SWOG-8814, in *American Society of Clinical Oncology* (vol A450), 1997

26. Valero V, Buzdar AU, Asmar L, al. e: Benefits of Adjuvant Doxorubicin/Cyclophosphamide-containing therapy in women with breast cancer by age (<60, 60-69, >69): The University of Texas Cancer Center Experience, in *American Society of Clinical Oncology* (vol 2233), 1999

27. Henderson IC, Berry DA, Demetri C, al. e: Improved disease-free survival (DFS) and overall survival (OS) from the addition of sequential paclitaxel (T) but not from the escalation of doxorubicin (A) dose in the adjuvant chemotherapy of patients with node positive primary breast cancer., in *American Society of Clinical Oncology* (vol 17), 1998, p 101a

28. Viens P, Maranchi D: High-dose chemotherapy in advanced breast cancer. *Critical Reviews in Oncology Hematology* 41:141-149, 2002

29. Fisher B, Anderson S, Tan-Chiu E, Wolmark N, Wickerham DL, Fisher ER, Dimitrov NV, Atkins JN, Abramson N, Merajver S, Romond EH, Kardinal CG, Shibata HR, Margolese RG, Farrar WB: Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. *J Clin Oncol* 19:931-942., 2001

30. New Uses for Herceptin, in, www.susanlovemd.com, 2001

31. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 344:783-792., 2001

32. Mandelblatt JS, Hadley J, Kerner JF, Schulman KA, Gold K, Dunmore-Griffith J, Edge S, Guadagnoli E, Lynch JJ, Meropol NJ, Weeks JC, Winn R: Patterns of breast carcinoma treatment in older women: patient preference and clinical and physical influences. *Cancer* 89:561-573., 2000

33. Buschner D, Wagner E: "Preventing Frail Health". *Clinical Geriatric Medicine* 8:1-17, 1992

34. Mandelblatt JS, Bierman AS, Gold K, Zhang Y, Ng JH, Maserejan N, Hwang YT, Meropol NJ, Hadley J, Silliman RA: Constructs of burden of illness in older patients with breast cancer: a comparison of measurement methods. *Health Serv Res* 36:1085-1107., 2001

35. Jans L, Stoddard S: Chartbook on Women and Disability in the United States. *U.S. National Institute on Disability and Rehabilitation Research*:45-46, 1999

36. Satariano WA, Ragland DR: The effect of comorbidity on 3-year survival of women with primary breast cancer. *Ann Intern Med* 120:104-110., 1994

37. Newschaffer CJ, Penberthy L, Desch CE, Retchin SM, Whittemore M: The effect of age and comorbidity in the treatment of elderly women with nonmetastatic breast cancer. *Arch Intern Med* 156:85-90., 1996

38. Greenfield S, Aronow HU, Elashoff RM, Watanabe D: Flaws in mortality data. The hazards of ignoring comorbid disease. *Jama* 260:2253-2255., 1988

39. Bergman L, Dekker G, Kerkhoof E, Peterse H, Dongen JA, Leeuwen F: Influence of age and comorbidity on treatment choice and survival in elderly patients with breast cancer. *Breast Cancer Research and Treatment* 18:189-198, 1991

40. Diab SG, Elledge RM, Clark GM: Favorable Biologic Characteristics and Clinical Outcome in Elderly Patients with Invasive Ductal Carcinoma of the Breast, in *American Society of Clinical Oncology* (vol 261), 1999

41. Kant A, Glover C, Horm J, Schatzkin A, Harris T: Does cancer survival differ for older patients. *Cancer* 70:2734-2740, 1992

42. Balducci L, Extermann M: Cancer in the Elderly, in, Clinical Geriatrics, 2001

43. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA: Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 56:M146-156, 2001

44. Fried LP: Conference on the physiologic basis of frailty. April 28, 1992, Baltimore, Maryland, U.S.A. Introduction. *Aging* 4:251-252, 1992

45. Balducci L, Stanta G: Cancer in the frail paitent: a coming epdiemic. *Hematol Oncol Clin North Am* 14:235-250, 2000

46. Ganz PA, Rowland JH, Meyerowitz BE, Desmond KA: Impact of different adjuvant therapy strategies on quality of life in breast cancer survivors. *Recent Results Cancer Res* 152:396-411., 1998

47. Swain SM, Rowland J, Weinfurt K, Berg C, Lippman ME, Walton L, Egan E, King D, Spertus I, Honig SF: Intensive outpatient adjuvant therapy for breast cancer: results of dose escalation and quality of life. *J Clin Oncol* 14:1565-1572., 1996

48. Macquart-Moulin G, Viens P, Bouscary ML, Genre D, Resbeut M, Gravis G, Camerlo J, Maraninchi D, Moatti JP: Discordance between physicians' estimations and breast cancer patients' self-assessment of side-effects of chemotherapy: an issue for quality of care. *Br J Cancer* 76:1640-1645., 1997

49. Fetting JH, Gray R, Fairclough DL, Smith TJ, Margolin KA, Citron ML, Grove-Conrad M, Cella D, Pandya K, Robert N, Henderson IC, Osborne CK, Abeloff MD: Sixteen-week multidrug regimen versus cyclophosphamide, doxorubicin, and fluorouracil as adjuvant therapy for node-positive, receptor-negative breast cancer: an Intergroup study. *J Clin Oncol* 16:2382-2391., 1998

50. Lindley C, Vasa S, Sawyer WT, Winer EP: Quality of life and preferences for treatment following systemic adjuvant therapy for early-stage breast cancer. *J Clin Oncol* 16:1380-1387., 1998

51. Hillner BE, Smith TJ: Efficacy and cost effectiveness of adjuvant chemotherapy in women with node-negative breast cancer. A decision-analysis model. *N Engl J Med* 324:160-168., 1991

52. Hillner BE, Smith TJ, Desch CE: Assessing the cost effectiveness of adjuvant therapies in early breast cancer using a decision analysis model. *Breast Cancer Res Treat* 25:97-105., 1993

53. Ganz PA, Desmond KA, Leedham B, Rowland JH, Meyerowitz BE, Belin TR: Quality of life in long-term, disease-free survivors of breast cancer: a follow-up study. *J Natl Cancer Inst* 94:39-49., 2002

54. Harris RA, Nease RF, Jr.: The importance of patient preferences for comorbidities in cost-effectiveness analyses. *J Health Econ* 16:113-119., 1997

55. Desch CE, Hillner BE, Smith TJ, Retchin SM: Should the elderly receive chemotherapy for node-negative breast cancer? A cost-effectiveness analysis examining total and active life-expectancy outcomes. *J Clin Oncol* 11:777-782., 1993

56. Li N, van Agthoven M, Willemse P, Uyl-de Groot C: A cost--utility analysis comparing second-line chemotherapy schemes in patients with metastatic breast cancer. *Anticancer Drugs* 12:533-540., 2001

57. Lister-Sharp D, McDonagh MS, Khan KS, Kleijnen J: A rapid and sytematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast and ovarian cancer. *Health Technol Assess* 4:1-113., 2000

58. McLachlan SA, Pintilie M, Tannock IF: Third line chemotherapy in patients with metastatic breast cancer: an evaluation of quality of life and cost. *Breast Cancer Res Treat* 54:213-223., 1999

59. Hillner BE, Radice D: Cost-effectiveness analysis of exemestane compared with megestrol in patients with advanced breast carcinoma. *Cancer* 91:484-489., 2001

60. Malin J, Keeler E: Cost-effectiveness and Health Benefit of Breast Cancer in Women With Early Breast Cancer, in, Los Angeles, RAND, 2000, p 28

61. Desch CE, Hillner BE, Smith TJ, Retchin SM: Should the elderly receive chemotherapy for node-negative breast cancer? A cost-effectiveness analysis examining total and active life-expectancy outcomes. *J Clin Oncol* 11:777-782, 1993

62. Extermann M, Balducci L, Lyman GH: What threshold for adjuvant therapy in older breast cancer patients? *J Clin Oncol* 18:1709-1717, 2000

63. NCHS: United States Life Tables, 1999, in, National Center for Health Statistics, 1999

64. Kattlove H, Liberati A, Keeler E, Brook RH: Benefits and costs of screening and treatment for early breast cancer. Development of a basic benefit package. *Jama* 273:142-148., 1995

65. Hillner BE, Smith TJ: Cost effectiveness and other assessments of adjuvant therapies for early breast cancer. *Oncology (Huntingt)* 9:129-134., 1995

66. Silber JH, Fridman M, DiPaola RS, Erder MH, Pauly MV, Fox KR: First-cycle blood counts and subsequent neutropenia, dose reduction, or delay in early-stage breast cancer therapy. *J Clin Oncol* 16:2392-2400, 1998

67. Fisher B, Brown AM, Dimitrov NV, Poisson R, Redmond C, Margolese RG, Bowman D, Wolmark N, Wickerham DL, Kardinal CG, et al.: Two months of doxorubicin-cyclophosphamide with and without interval reinduction therapy compared with 6 months of cyclophosphamide, methotrexate, and fluorouracil in positive-node breast cancer patients with tamoxifen-nonresponsive tumors: results from the National Surgical Adjuvant Breast and Bowel Project B-15. *J Clin Oncol* 8:1483-1496, 1990

68. Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang YT, Gold K, Barter J, Shah K: Benefits and costs of using HPV testing to screen for cervical cancer. *Jama* 287:2372-2381, 2002

Table 1 Financial and Health Costs of Adjuvant Treatment

Treatment	Medicare (AWP)*	Medicare (PHS)*	Health Cost in Days &
CMF x 6	$4,567.58	$2,833.45	17
AC x 4	$5,964.73	$2,317.70	11
HRT x 2 years#	$2,771.38	$1,468.83	7
HRT x 5 years#	$6,475.84	$3,432.18	17
HRT (2) +CMF^	$7,300.95	$4,282.13	25
HRT (2) +AC^	$8,698.10	$3,766.38	18
HRT (5) +CMF#	$11,130.08	$6,311.56	35
HRT (5) + AC#	$12,527.23	$5,795.81	28

* Future costs discounted at 3%
HRT=Hormone therapy which for this analysis was Tamoxifen
# Cost of HRT is for a 45 year old with Node (-) breast cancer and is adjusted for difference in mortality in the first 5 years. Cost is lower in older ages and with Node (+) disease
It is assumed that there is no difference in the cost of chemotherapy with age or disease characteristics
^Combined adjuvant therapy with only two years of tamoxifen modeled only in the 75 and 85 year old models
& Rounded to the nearest full day. Health cost are calculated as follows:
For chemotherapy prior weighted health states used are 0.9 for minor chemotherapy toxicty and 0.8 for major toxicity of chemotherapy. For 6 cycles of chemotherapy lasting 6x28days=168days the health cost is 0.2 (1-0.8) on the one day a cycle when chemotherapy is given and 0.1 (1-0.9) for the rest of the cycle. Health cost = 0.97*168*0.1+0.03*168*0.2=17.3 days.
The health state weight used for HRT is 0.99, but health costs for HRT over 2 or 5 years were also subject to discounting.

Table 2: Odds Reduction in 10 year mortality*

Adjuvant Therapy	Age	Odds Reduction	Standard Deviation	Number of Patients in Meta-analyses	Low Bound	Mid-Range	High Bound
HRT- 5 years	45	0.32	0.10	1327			
	65	0.33	0.06	3174			
	75	0.34	0.13	390			
	85	??			0.20	0.25	0.36
HRT - 2 years	45	0.1	0.06	4153			
	65	0.12	0.05	4466			
	75	0.36	0.07	1476			
	85	??			0.10	0.18	0.38
CMF IV	45	0.34	0.07	1660			
	65	0.08	0.04	6161			
	75	??			0.00	0.02	0.08
	85	??			0.00	0.02	0.08
AC IV	45	0.38	0.09	6950			
	65	0.11	0.06	6950			
	75	??			0.00	0.033	0.11
	85	??			0.00	0.033	0.11
HRT-CMF	45	0.57	0.17	640			
	65	0.44	0.11	9192			
	75	??			0.30	0.36	0.44
	85	??			0.20	0.25	0.44
HRT-AC #	45	0.60	N/A	N/A			
	65	0.46	N/A	N/A			
	75	??			0.30	0.38	0.46
	85	??			0.20	0.27	0.46

HRT= Adjuvant Hormone Therapy which for this analysis was Tamoxifen.
IV= Intravenous
N/A - Not Available since an estimate from a combination of studies but a standard deviation is at least as great as HRT-CMF and likely greater.
*Source of data was primarily meta-analyses performed by the Early Breast Cancer Trialists Collaborative Group (EBCTCG)

This calculation is based on only patients with Her2-Neu + benefitting more from AC than CMF, Her2-Neu expression is present only in 25% of patients, decreasing with age.

Since there are no good outcomes data on the use of hormone therapy in patients age 85 and chemotherapy in patients age 75 and 85, a range of values were considered. The high bound used assumes that older patients, 75 and 85, have a response equal in magnitude to a 65 year old. The middle bound assumes that affect of chemotherapy decreases with age almost linearly. The low bound assumes almost no benefit from chemotherapy.

Table 3: Added Years of Life, QALYS, with Adjuvant Treatment for a 45 year old woman with breast cancer

45 year old with ER+ Breast Cancer	LE	Node (-)	Node (+)
Life Expectancy without Breast Cancer	***36.41***		
Life Expectancy with Breast Cancer without treatment		16.27	12.06
With Treatment Added Years of Life			
Hormone (HRT) 2 years		0.38	0.44
HRT 5 years		1.32	1.67
CMF IV*		1.42	1.80
AC IV*		1.63	2.09
HRT-CMF		2.75	3.86
HRT-AC		2.94	4.19
45 year old with ER- Breast Cancer			
Life Expectancy without treatment		16.27	12.06
With Treatment Added Years of Life			
CMF IV*		1.42	1.80
AC IV*		1.63	2.09

*An assumption for this analysis was that the impact of chemotherapy was similar among ER (-) and ER(+) patients.

Table 4: Added QALYS with Adjuvant Treatment for a 65 year old woman with breast cancer

65 year old with ER+ Breast Cancer	**LE**	**Node (-)**	**Node (+)**
Life Expectancy without Breast Cancer	***19.27***		
Life Expectancy with Breast Cancer without treatment		11.78	9.58
With Treatment Added Years of Life			
Hormone (HRT) 2 years		0.21	0.30
HRT 5 years		0.62	0.93
CMF IV		0.15	0.16
AC IV		0.18	0.26
HRT-CMF		0.82	1.29
HRT-AC		0.89	1.39
65 year old with ER- Breast Cancer			
Life Expectancy without treatment		11.78	9.58
With Treatment Added Years of Life			
CMF IV		0.15	0.16
AC IV		0.18	0.26

*An assumption for this analysis was that the impact of chemotherapy was similar among ER (-) and ER(+) patients.

Table 5: Added QALYS with Adjuvant Treatment for a 75 year old woman with breast cancer

75 year old with ER+ Breast Cancer	**LE**	**Node (-)**	**Node (+)**
Life Expectancy without Breast Cancer			
Life Expectancy with Breast Cancer without treatment	***12.29***		
With Treatment Added Years of Life		8.75	7.52
Hormone (HRT) 2 years			
HRT 5 years		0.36	0.6
CMF IV (high)		0.31	0.53
CMF IV (mid)		0.03	0.07
CMF IV (low)		-0.03	-0.02
AC IV (high)		-0.05	-0.05
AC IV (mid)		0.08	0.14
AC IV (low)		0	0.02
HRT5-CMF (high)		-0.03	-0.03
HRT5-CMF (mid)		0.38	0.69
HRT5-CMF (low)		0.29	0.52
HRT5-AC (high)		0.11	0.23
HRT5-AC (mid)		0.42	0.75
HRT5-AC (low)		0.33	0.58
HRT2-CMF (high)		0.13	0.25
HRT2-CMF (mid)		0.41	0.72
HRT2-CMF (low)		0.31	0.55
HRT2-AC (high)		0.13	0.25
HRT2-AC (mid)		0.45	0.78
HRT2-AC (low)		0.36	0.61
75 year old with ER- Breast Cancer		0.15	0.27
Life Expectancy without treatment			
With Treatment Added Years of Life		8.75	7.52
CMF IV (high)			
CMF IV (mid)		0.03	0.07
CMF IV (low)		-0.03	-0.02
AC IV (high)		-0.05	-0.05
AC IV (mid)		0.08	0.14
AC IV (low)		0	0.02
		-0.03	-0.03

Table 6: Added Years of Life with Adjuvant Treatment for a 85 year old woman with breast cancer

85 year old with ER+ Breast Cancer	LE	Node (-)	Node (+)
Life Expectancy without Breast Cancer	*6.94*		
Life Expectancy with Breast Cancer without treatment		5.63	5.11
With Treatment Added Years of Life			
Hormone (HRT) 2 years (high)		0.15	0.27
HRT 2 years (mid)		0.06	0.11
HRt 2 years (low)		0.03	0.05
HRT 5 years (high)		0.11	0.22
HRT 5 years (mid)		0.06	0.14
HRT 5 years (low)		0.04	0.10
CMF IV (high)		-0.01	0.01
CMF IV (mid)		-0.04	-0.03
CMF IV (low)		-0.05	-0.05
AC IV (high)		0.02	0.05
AC IV (mid)		-0.01	0.00
AC IV (low)		-0.03	-0.03
HRT5-CMF (high)		0.10	0.25
HRT5-CMF (mid)		0.02	0.09
HRT5-CMF (low)		-0.01	0.05
HRT5-AC (high)		0.13	0.28
HRT5-AC (mid)		0.03	0.09
HRT5-AC (low)		0.01	0.07
HRT2-CMF (high)		0.13	0.27
HRT2-CMF (mid)		0.05	0.12
HRT2-CMF (low)		0.02	0.07
HRT2-AC (high)		0.16	0.31
HRT2-AC (mid)		0.05	0.12
HRT2-AC (low)		0.04	0.09
85 year old with ER- Breast Cancer			
CMF IV (high)		5.62	5.12
CMF IV (mid)		5.59	5.08
CMF IV (low)		5.58	5.06
AC IV (high)		5.65	5.16
AC IV (mid)		5.62	5.11
AC IV (low)		5.60	5.08

Table 7A: Added Years of Life with Adjuvant Treatment (health cost of treatment subtracted out) for a 85 year old woman with breast cancer

85 year old with ER- Breast Cancer	LE	Node (-)	Node (+)
Life Expectancy without Breast Cancer	***6.94***		
Life Expectancy with Breast Cancer without treatment		5.63	5.11
With Treatment Added Years of Life			
CMF IV (high)		-0.01	0.01
CMF IV (mid)		-0.04	-0.03
CMF IV (low)		-0.05	-0.05
AC IV (high)		0.02	0.05
AC IV (mid)		-0.01	0
AC IV (low)		-0.03	-0.03

Table 7B - Cost per QALY for Adjuvant Therapy for Breast Cancer

Treatment	Cost per QALY	Cost per QALY
For a 45 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years	$7,293.00	$6,215.00
CMF	$3,217.00	$2,538.00
ACX4	$3,659.00	$2,854.00
HRT-5 Years	$4,906.00	$3,718.00
HRT+CMF	$4,047.00	$2,840.00
HRT+AC	$4,261.00	$2,950.00
For a 45 Year-old woman with ER- Breast Cancer		
CMF	$3,217.00	$2,538.00
ACX4	$3,659.00	$2,854.00

Table 8 - Cost per QALY for Adjuvant Therapy for Breast Cancer

	Cost per QALY	Cost per QALY
For a 65 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years	$13,116	$9,060
CMF	$30,451	$28,547
ACX4	$33,137	$22,941
HRT-5 Years	$10,194	$6,520
HRT+CMF	$13,320	$8,297
HRT+AC	$13,842	$8,706
For a 65 Year-old woman with ER- Breast Cancer		
CMF	$30,451	$28,547
ACX4	$33,137	$22,941

Table 9 - Cost per QALY for Adjuvant Therapy for Breast Cancer

	Cost per QALY	Cost per QALY
For a 75 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years	$7,584	$4,503
CMF (low)	XXX	XXX
CMF (mid)	XXX	XXX
CMF (high)	$152,253	$65,251
ACX4 (low)	XXX	XXX
ACX4 (mid)	XXX	$298,237
ACX4 (high)	$75,559	$42,605
HRT-5 Years	$19,530	$10,965
HRT5+CMF (low)	$96,249	$29,037
HRT5+CMF (mid)	$36,662	$20,000
HRT5+CMF (high)	$28,033	$15,135
HRT5+AC (low)	$92,188	$32,061
HRT5+AC (mid)	$36,452	$20,340
HRT5+AC (high)	$28,690	$15,787
HRT2+CMF (low)	$56,102	$29,038
HRT2+CMF (mid)	$23,544	$13,219
HRT2+CMF (high)	$17,807	$10,106
HRT2+AC (low)	$57,936	$32,061
HRT2+AC (mid)	$24,155	$14,210
HRT2+AC (high)	$19,329	$11,119
For a 75 Year-old woman with ER- Breast Cancer		
CMF (low)	XXX	XXX
ACX4 (low)	XXX	XXX
CMF (mid)	XXX	XXX
ACX4 (mid)	XXX	$298,237
CMF (high)	$152,253	$65,251
ACX4 (high)	$75,559	$42,605

Table 10 - Cost per QALY for Adjuvant Therapy for Breast Cancer

	Cost per QALY	Cost per QALY
For a 85 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years (L)	$90,761.00	$53,719.00
HRT-2 Years (M)	$45,418.00	$24,460.00
HRT-2 Years (H)	$18,206.00	$10,011.00
CMF (low)	XXX	XXX
CMF (mid)	XXX	XXX
CMF (high)	XXX	$456,758.00
ACX4 (low)	XXX	XXX
ACX4 (mid)	XXX	XXX
ACX4 (high)	$298,237.00	$119,295.00
HRT-5 Years (L)	$150,494.00	$57,413.00
HRT-5 Years (M)	$100,533.00	$41,184.00
HRT-5 Years (H)	$55,085.00	$26,463.00
HRT5+CMF (low)	XXX	$206,178.00
HRT5+CMF (mid)	$529,733.00	$114,705.00
HRT5+CMF (high)	$106,526.00	$41,774.00
HRT5+AC (low)	$1,198,449.00	$167,229.00
HRT5+AC (mid)	$399,727.00	$130,229.00
HRT5+AC (high)	$92,690.00	$42,288.00
HRT2+CMF (low)	$364,663.00	$103,706.00
HRT2+CMF (mid)	$145,882.00	$60,510.00
HRT2+CMF (high)	$56,161.00	$26,948.00
HRT2+AC (low)	$217,260.00	$96,184.00
HRT2+AC (mid)	$173,825.00	$72,153.00
HRT2+AC (high)	$54,363.00	$27,978.00
For a 85 Year-old woman with ER- Breast Cancer		
CMF (low)	XXX	XXX
ACX4 (low)	XXX	XXX
CMF (mid)	XXX	$456,758.00
ACX4 (mid)	XXX	XXX
CMF (high)	XXX	XXX
ACX4 (high)	$298,237.00	$119,295.00

Table 11 - Incremental Cost per QALY for Adjuvant Therapy for Breast Cancer

Treatment	Cost per QALY	Cost per QALY
For a 45 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years	XXX	XXX
CMF	$3,217	$2,538
ACX4	XXX	XXX
HRT-5 Years	XXX	XXX
HRT+CMF	$4,934	$3,105
HRT+AC	$7,353	$4,234
For a 45 Year-old woman with ER- Breast Cancer		
CMF	$3,217	$2,538
ACX4	$6,653	$4,818

Table 12 - Incremental Cost per QALY for Adjuvant Therapy for Breast Cancer

	Cost per QALY	Cost per QALY
For a 65 Year-old woman with ER+ Breast Cancer	***Negative Lymph Nodes***	***Positive Lymph Nodes***
HRT-2 Years	XXX	XXX
CMF	XXX	XXX
ACX4	XXX	XXX
HRT-5 Years	$10,194	$6,520
HRT+CMF	XXX	$12,890
HRT+AC	$22,220	$13,972
For a 65 Year-old woman with ER- Breast Cancer		
CMF	$30,451	$28,547
ACX4	$46,572	$13,972

Table 13 - Incremental Cost per Additional Life Year for Adjuvant Therapy for Breast Cancer

		Cost per QALY *Negative Lymph Nodes*	Cost per QALY *Positive Lymph Nodes*
For a 75 Year-old woman with ER+ Breast Cancer	**Scenario 1: HRT2-true, HRT5-Chemo (H) True**		
	HRT-2 Years	$7,584.00	$4,503.00
	CMF (high)	XXX	XXX
	ACX4 (high)	XXX	XXX
	HRT-5 Years	XXX	XXX
	HRT5+CMF (high)	XXX	XXX
	HRT5+AC (high)	$155,322.00	$60,925.00
	Scenario 2: HRT2-false,HRT5-true HRT5-Chemo (H) True		
	CMF (high)	XXX	XXX
	ACX4 (high)	XXX	XXX
	HRT-5 Years	$19,530.00	$10,965.00
	HRT5+CMF (high)	XXX	XXX
	HRT5+AC (high)	$54,530.00	$27,406.00
	Scenario 3: HRT2-false,HRT5-true HRT5-Chemo (H) False, HRT5-Chemo (M) True		
	CMF (high)	XXX	XXX
	ACX4 (high)	XXX	XXX
	HRT-5 Years	$19,530.00	$10,965.00
	HRT5+CMF (mid)	XXX	XXX
	HRT5+AC (mid)	$298,742.00	$119,715.00
	Scenario 4: HRT2-true, HRT2-Chemo (H) True		
	HRT-2 Years	$7,584.00	$4,503.00
	CMF (high)	XXX	XXX
	ACX4 (high)	XXX	XXX
	HRT-5 Years	XXX	XXX
	HRT2+CMF (high)	XXX	XXX
	HRT2+AC (high)	$66,308.00	$33,174.00
	Scenario 5: HRT2-true, HRT2-Chemo (H) False, HRT2-Chemo (M) True		
	HRT-2 Years	$7,584.00	$4,503.00
	HRT2+AC (mid)	XXX	$596,602.00
For a 75 Year-old woman with ER- Breast Cancer	**Scenario 1: Chemo (H) True**		
	CMF (high)	XXX	XXX
	ACX4 (high)	$75,559	$42,605
	Scenario 2: Chemo (H) False, Chemo (M) True		
	CMF (mid)	XXX	XXX
	ACX4 (mid)	XXX	$298,237

Table 14 - Incremental Cost per QALY for Adjuvant Therapy for Breast Cancer

		Cost per QALY	Cost per QALY
		Negative Lymph Nodes	*Positive Lymph Nodes*
For a 85 Year-old woman with ER+ Breast Cancer	**Scenario 1:**	**HRT2 (high)-true, HRT5-Chemo (H) True**	
HRT-2 Years (high)		$18,206.00	$10,011.00
ACX4 (high)		XXX	XXX
HRT-5 Years (high)		XXX	XXX
HRT5+CMF (high)		XXX	XXX
HRT5+AC (high)		XXX	$913,749.00
	Scenario 2:	**HRT2 (mid)-True, HRT5(high)-True, HRT5-Chemo (H) True**	
HRT-2 Years (mid)		$45,418.00	$24,460.00
ACX4 (high)		XXX	XXX
HRT-5 Years (high)		$66,686.00	$28,465.00
HRT5+CMF (high)		XXX	XXX
HRT5+AC (high)		$299,517.00	$100,313.00
	Scenario 3:	**HRT2 (mid)-True, HRT5(high)-False, HRT5-Chemo (H) True**	
HRT-2 Years (mid)		$45,418.00	$24,460.00
ACX4 (high)		XXX	XXX
HRT-5 Years (mid)		XXX	XXX
HRT5+CMF (high)		XXX	XXX
HRT5+AC (high)		$133,209.00	$53,823.00
	Scenario 4:	**HRT2 (mid)-False, HRT5(mid)-True, HRT5-Chemo (H) True**	
ACX4 (high)		XXX	XXX
HRT-5 Years (mid)		$100,533.00	$41,184.00
HRT5+CMF (high)		XXX	$42,015.00
HRT5+AC (high)		$133,209.00	$53,823.00
	Scenario 5:	**HRT2 (high)-True, HRT2-Chemo (H) True**	
HRT-2 Years (high)		$18,206.00	$10,011.00
ACX4 (high)		XXX	XXX
HRT-5 Years (high)		XXX	XXX
HRT2+CMF (high)		XXX	XXX
HRT2+AC (high)		$596,715.00	$149,251.00
	Scenario 6:	**HRT2 (mid)-True, HRT5(high)-True, HRT2-Chemo (H) True**	
HRT-2 Years (mid)		$45,418.00	$24,460.00
ACX4 (high)		XXX	XXX
HRT-5 Years (high)		XXX	$28,465.00
HRT2+CMF (high)		XXX	$29,084.00
HRT2+AC (high)		$59,730.00	$34,929.00
For a 85 Year-old woman with ER- Breast Cancer	**Scenario 1:**	**Chemo (H) True**	
CMF (high)		XXX	XXX
ACX4 (high)		$298,237.00	$119,295.00

Table 15 - Cost per QALY for Hormone Therapy for Breast Cancer in Frail Woman.

	Cost per QALY
For a frail woman with ER+ Breast Cancer	
HRT - 2 years	$391,198.00

Table 16 - Percent Activity (Compared to activity of agent in a 65 year old) Needed in a 75 year old woman for Treatment to be Cost-Effective

	Cost-Effectiveness Cut-Off	
Treatment	**$50,000/QALY**	**$100,000/QALY**
Node (+)		
AC	86%	50%
CMF	XXX	94%
HRT+AC	89%	83%
Node (-)		
AC	XXX	82%
HRT+AC	XXX	89%

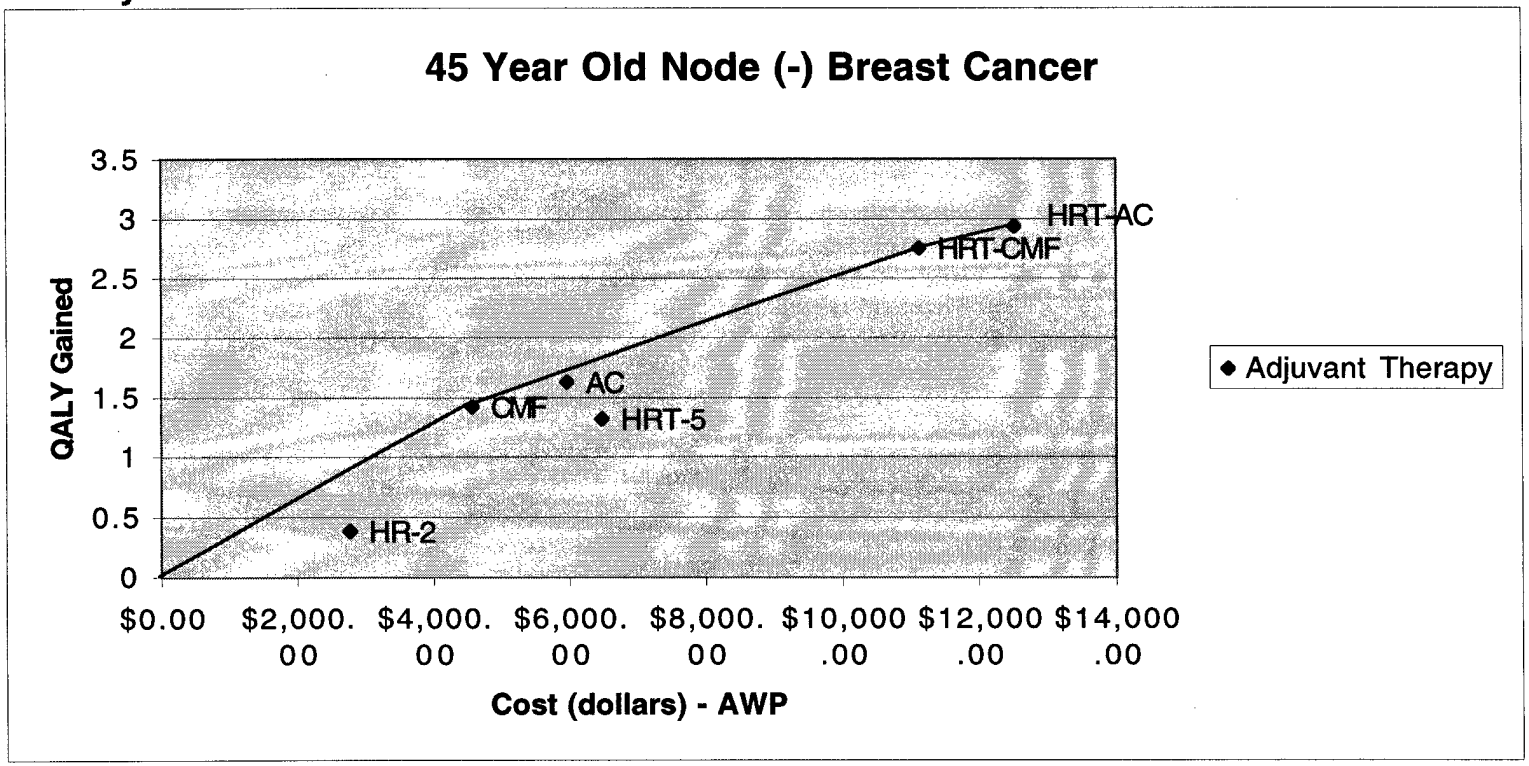
Table 17: Added QALY with Adjuvant Treatment (health cost of treatment subtracted out) for a 65 year old woman with breast cancer (using a 20% discount rate)

	Baseline Health State		
	1	*0.8*	*0.6*
65 year old with ER+ Breast Cancer	**Node (+)**	**Node (+)**	**Node (+)**
Life Expectancy with Breast Cancer without treatment	4.07	3.26	2.44
With Treatment Added Years of Life			
HRT 5 years	0.11	0.08	0.05
CMF IV	-0.02	-0.03	-0.03
AC IV	0.02	0	0
HRT-CMF	0.12	0.08	0.04
HRT-AC	0.15	0.1	0.07
65 year old with ER- Breast Cancer			
Life Expectancy without treatment	4.07	3.26	2.44
With Treatment Added Years of Life			
CMF IV	-0.02	-0.03	-0.03
AC IV	0.02	0	0

Table 18: Added QALY with Hormone Therapy (health cost of treatment subtracted out) for a 75 and 85 year old woman with breast cancer (using a 20% discount rate)

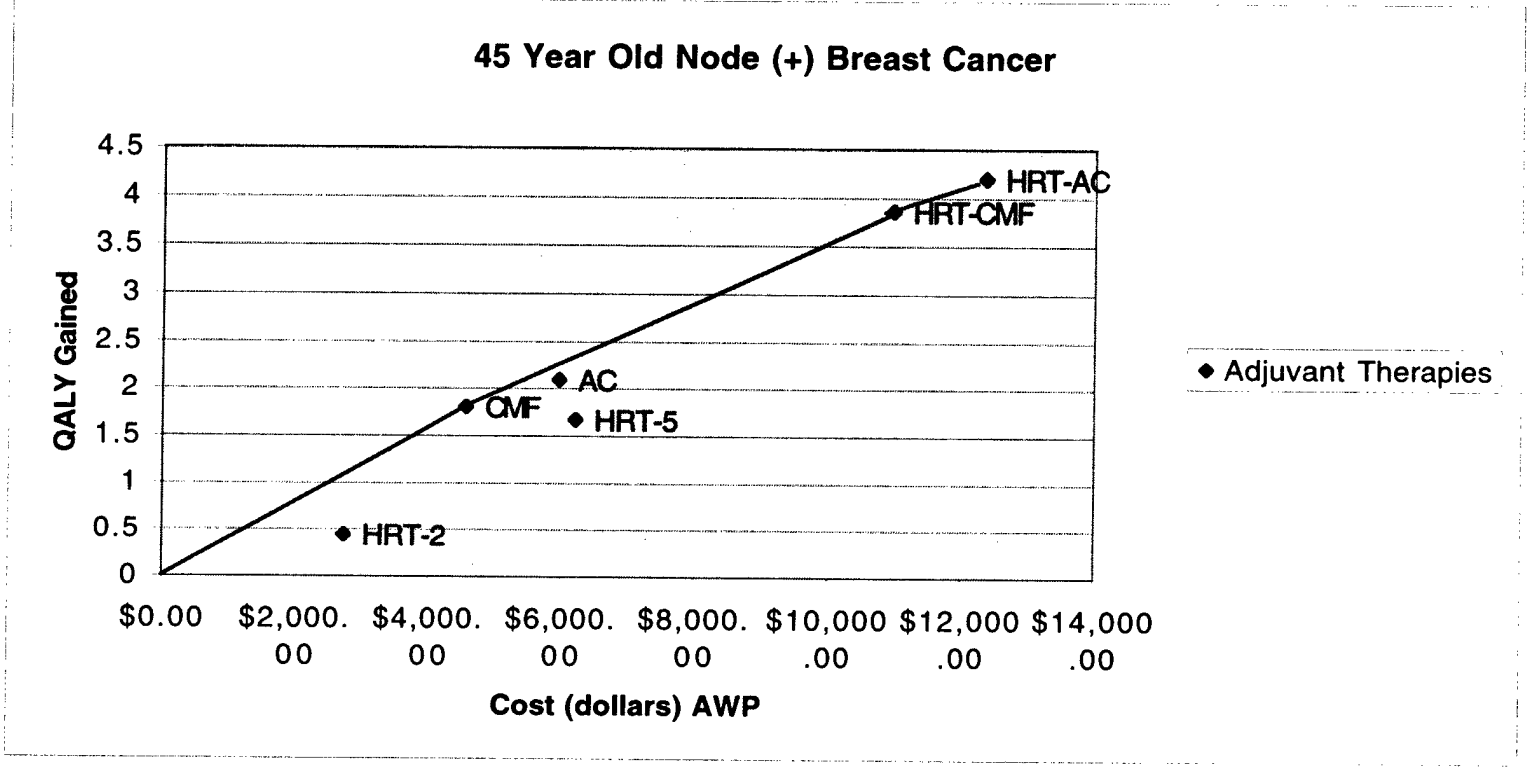
	Baseline Health State		
	1	*0.8*	*0.6*
75 year old with ER+ Breast Cancer	**Node (+)**	**Node (+)**	**Node (+)**
Life Expectancy with Breast Cancer without treatment	3.77	3.01	2.26
With HRT Added Years of Life	0.08	0.06	0.04
85 year old with ER+ Breast Cancer			
Life Expectancy with Breast Cancer without treatment	3.14	2.51	1.89
With HRT Added Years of Life	0.05	0.03	0.01

Figure 1a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 45 year old with breast cancer



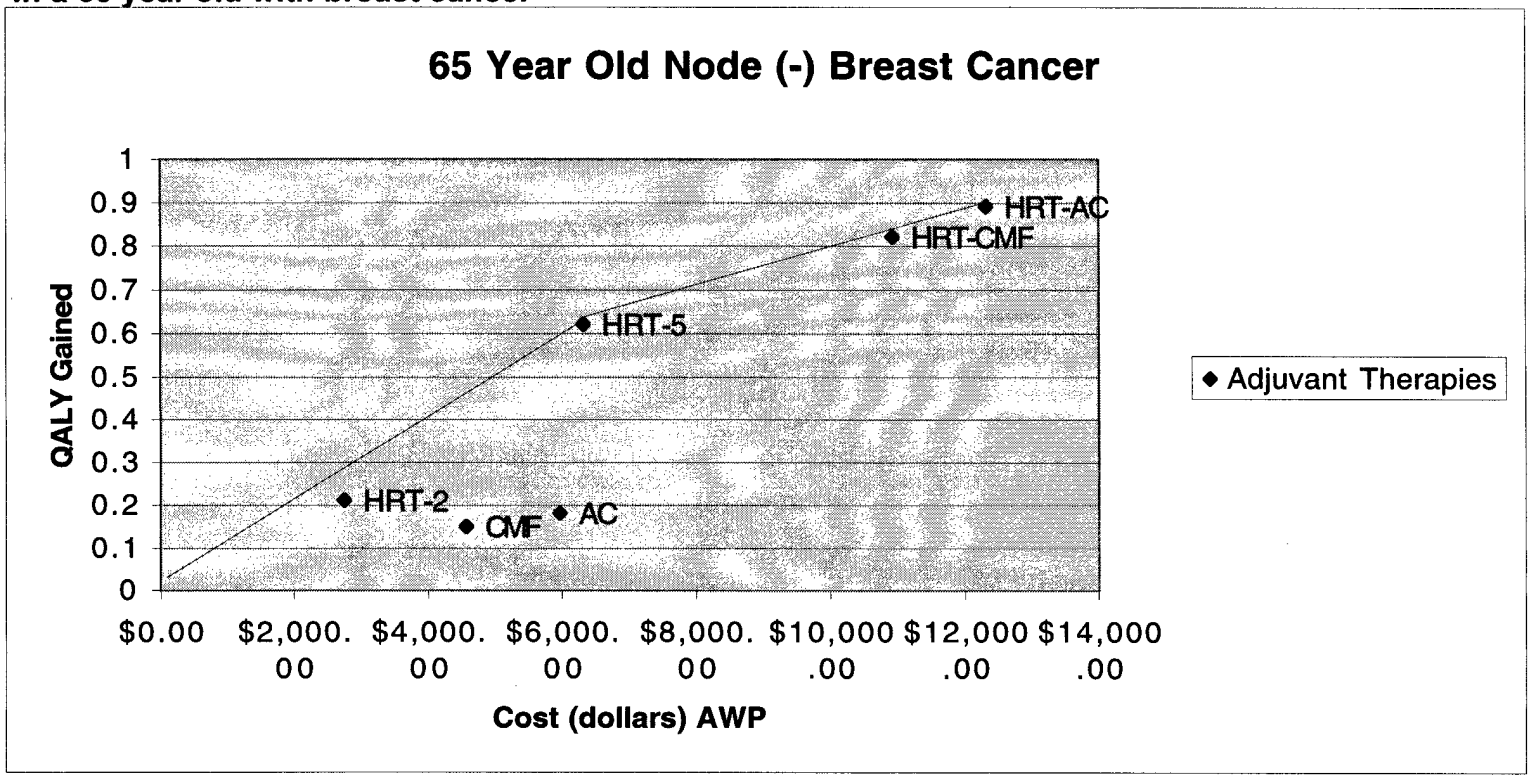
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 1b: Cost and Benefits of Diiferent Adjuvant Therapies In a 45 year old with breast cancer



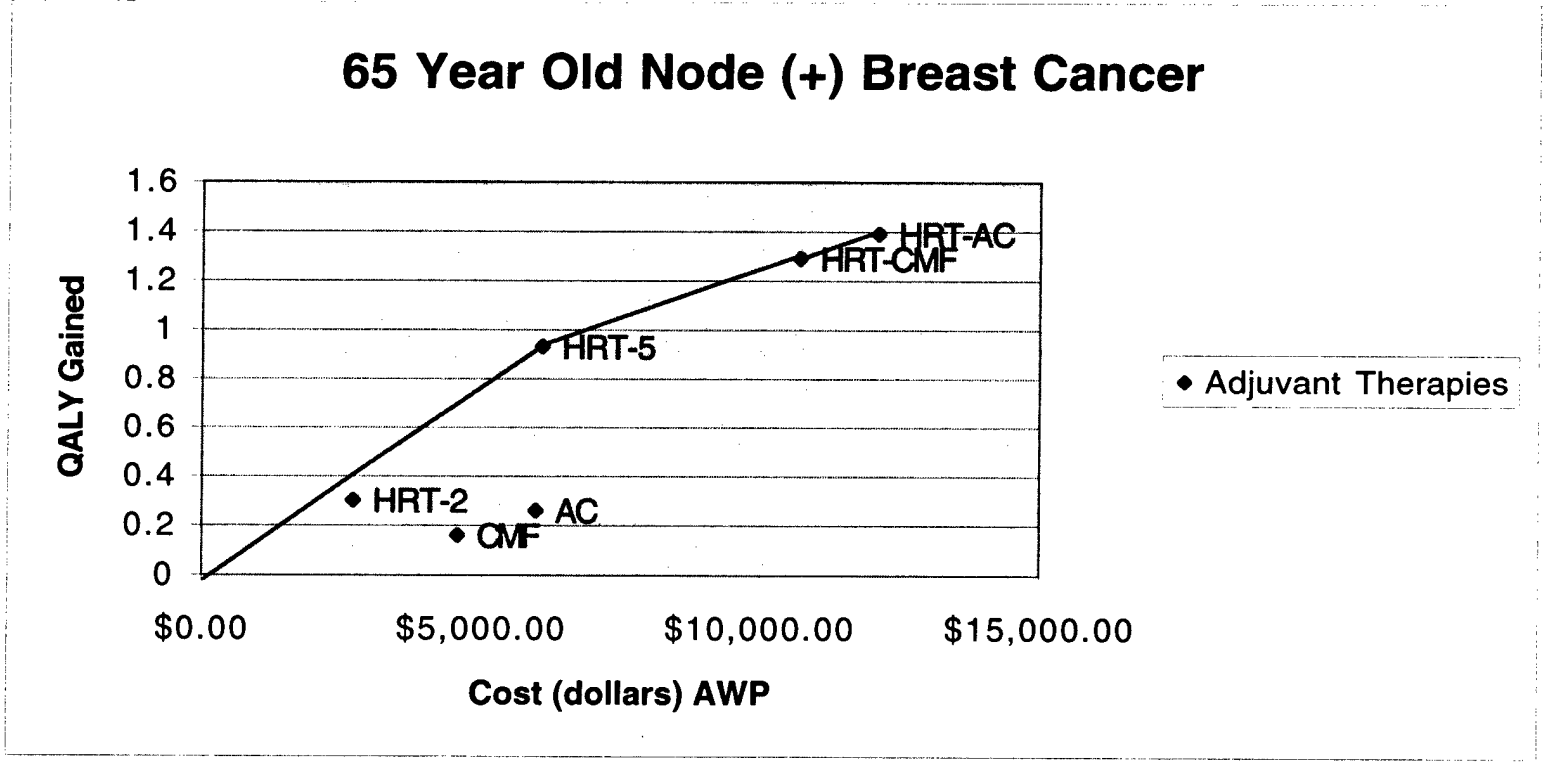
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 2a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 65 year old with breast cancer



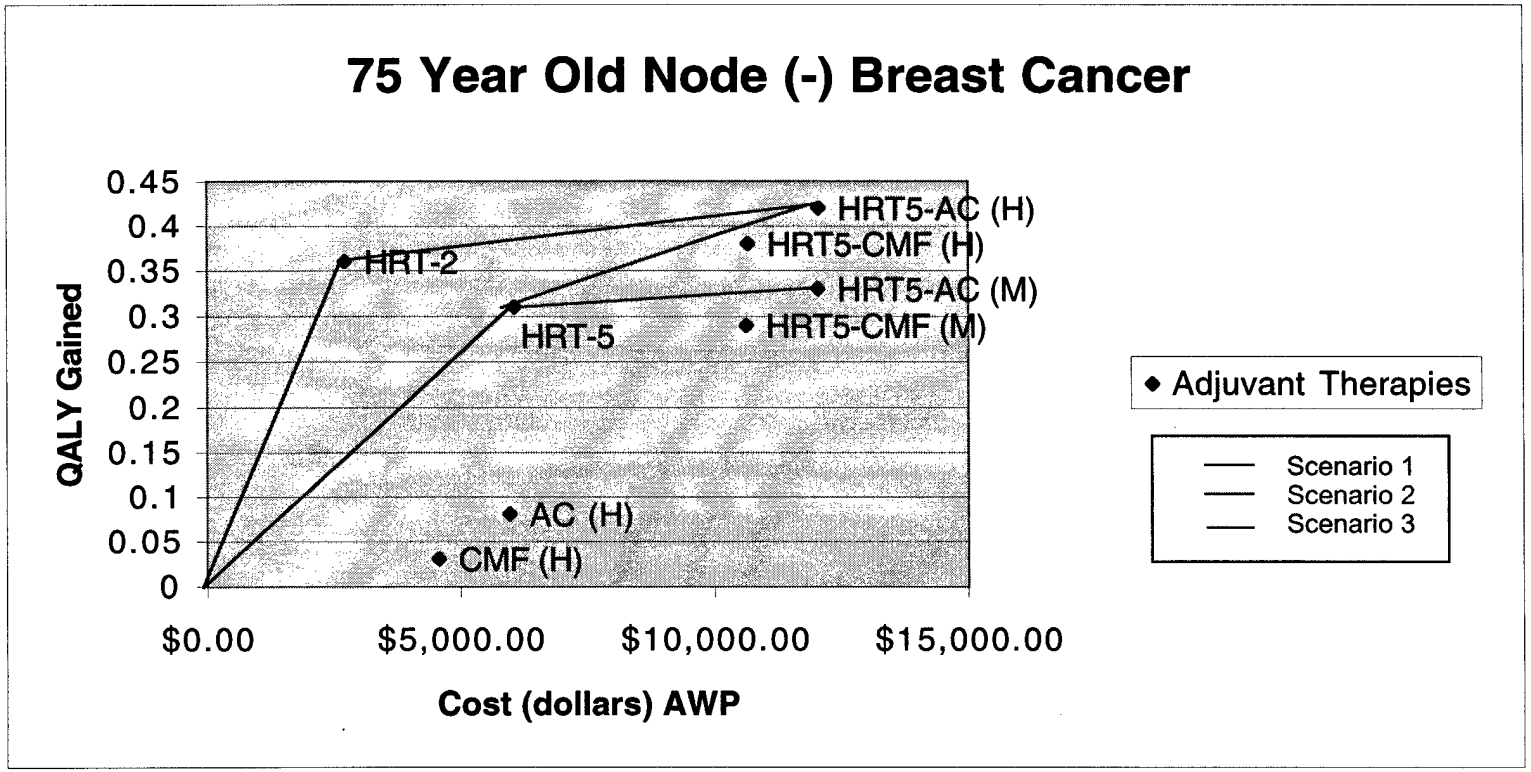
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 2b: Cost and Benefits of Diiferent Adjuvant Therapies In a 65 year old with breast cancer



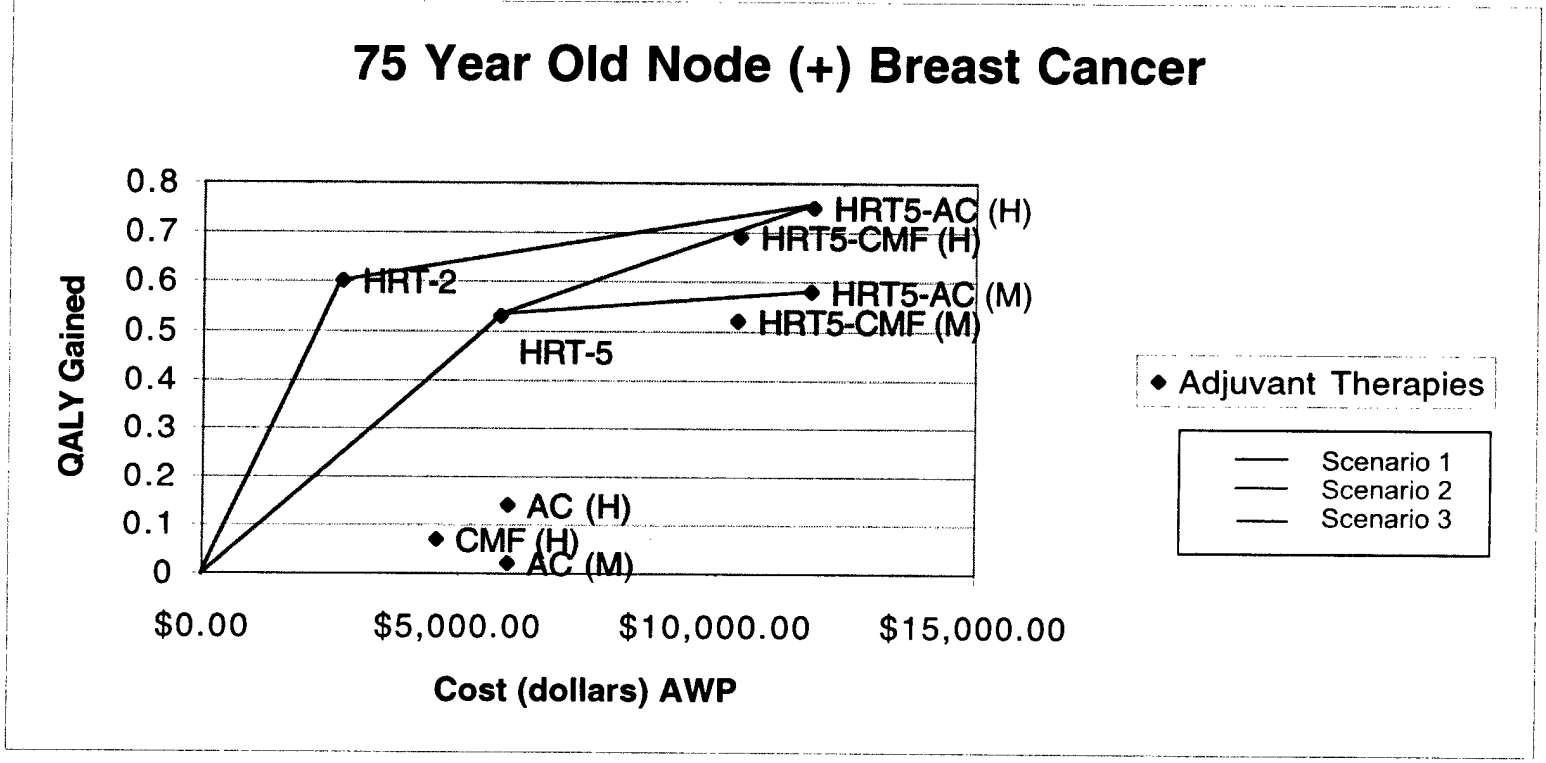
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 3a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 75 year old with breast cancer



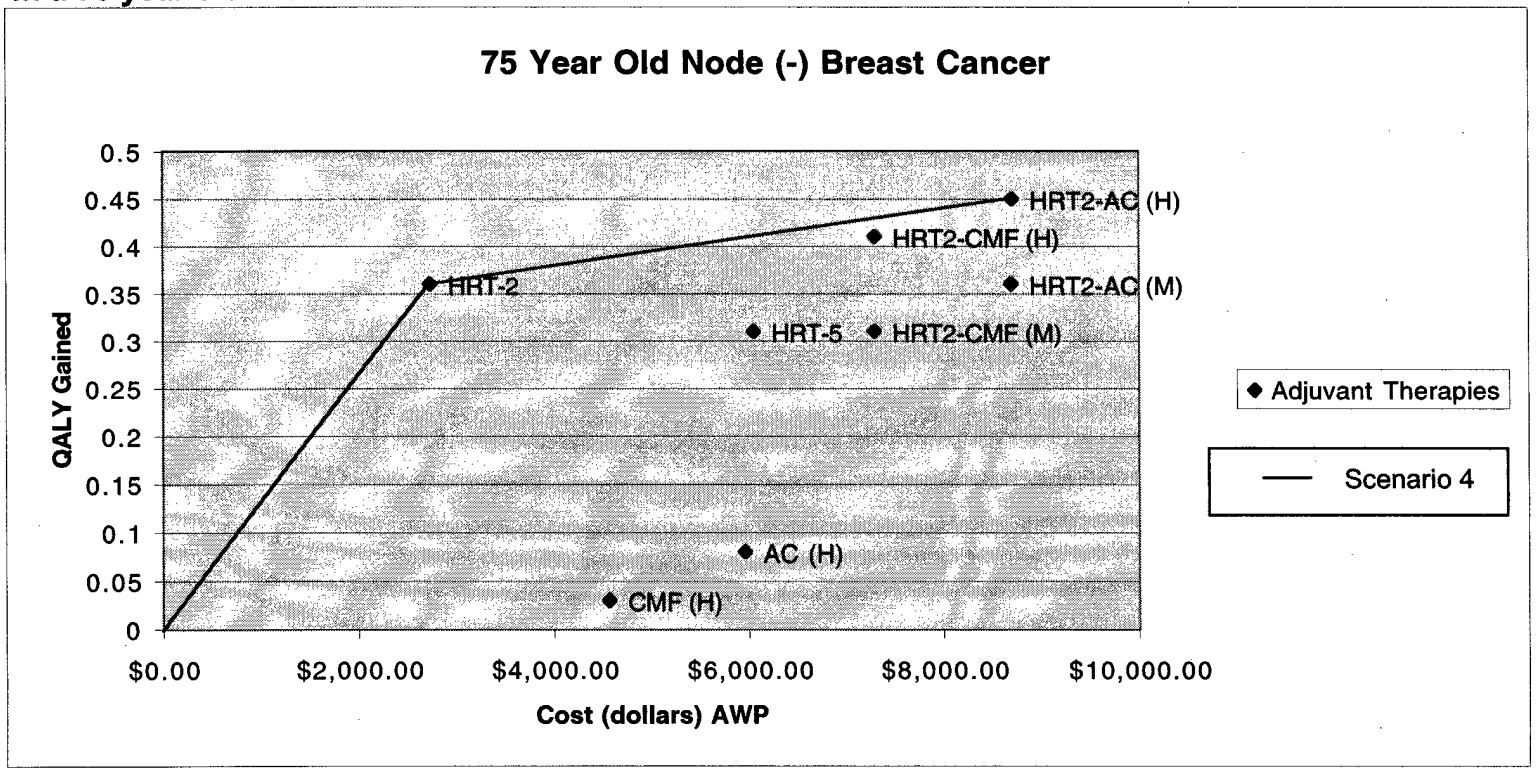
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 3b: Cost and Benefits of Diiferent Adjuvant Therapies
In a 75 year old with breast cancer



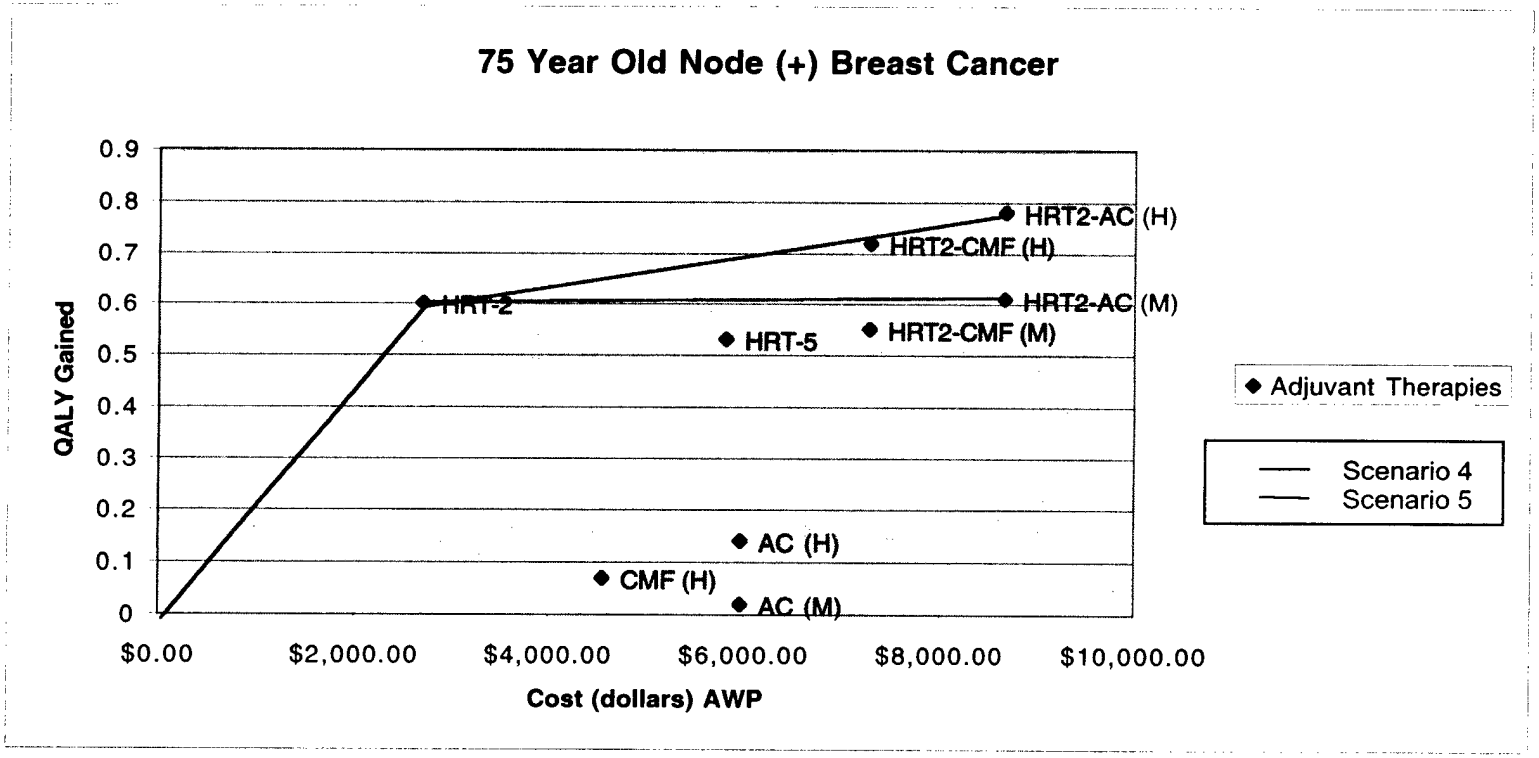
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 4a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 75 year old with breast cancer



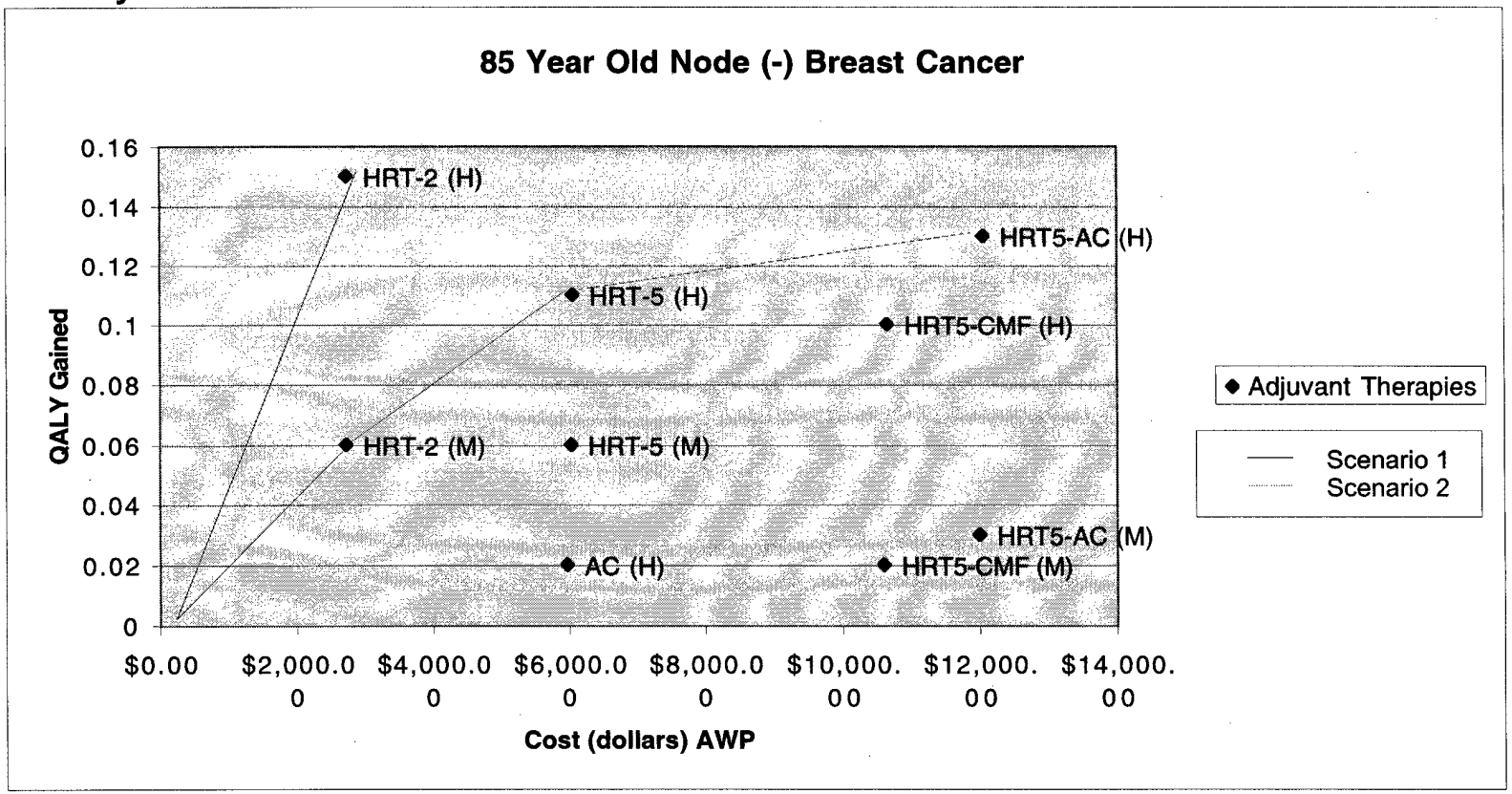
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 4b: Cost and Benefits of Diiferent Adjuvant Therapies In a 75 year old with breast cancer



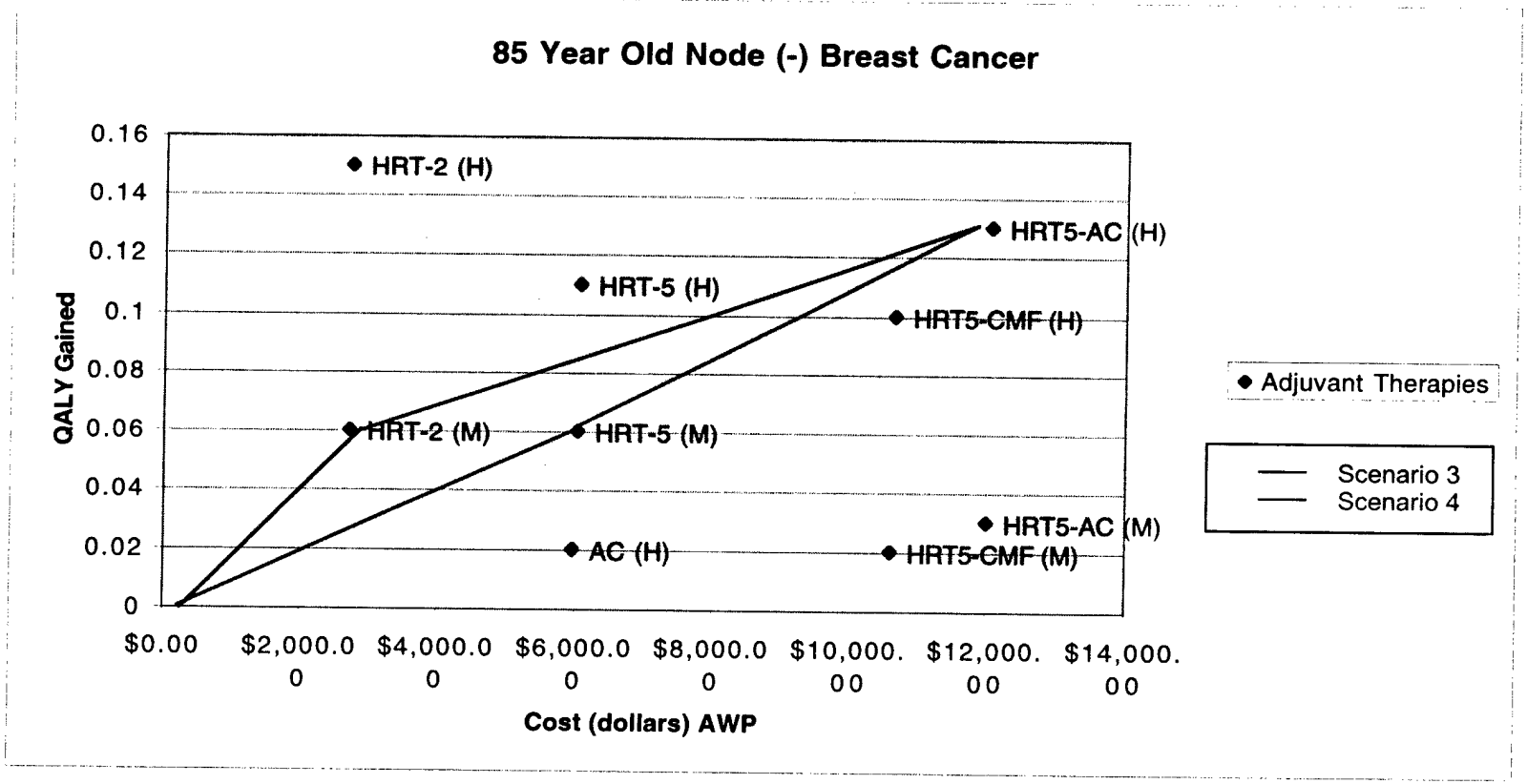
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 5a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 85 year old with breast cancer



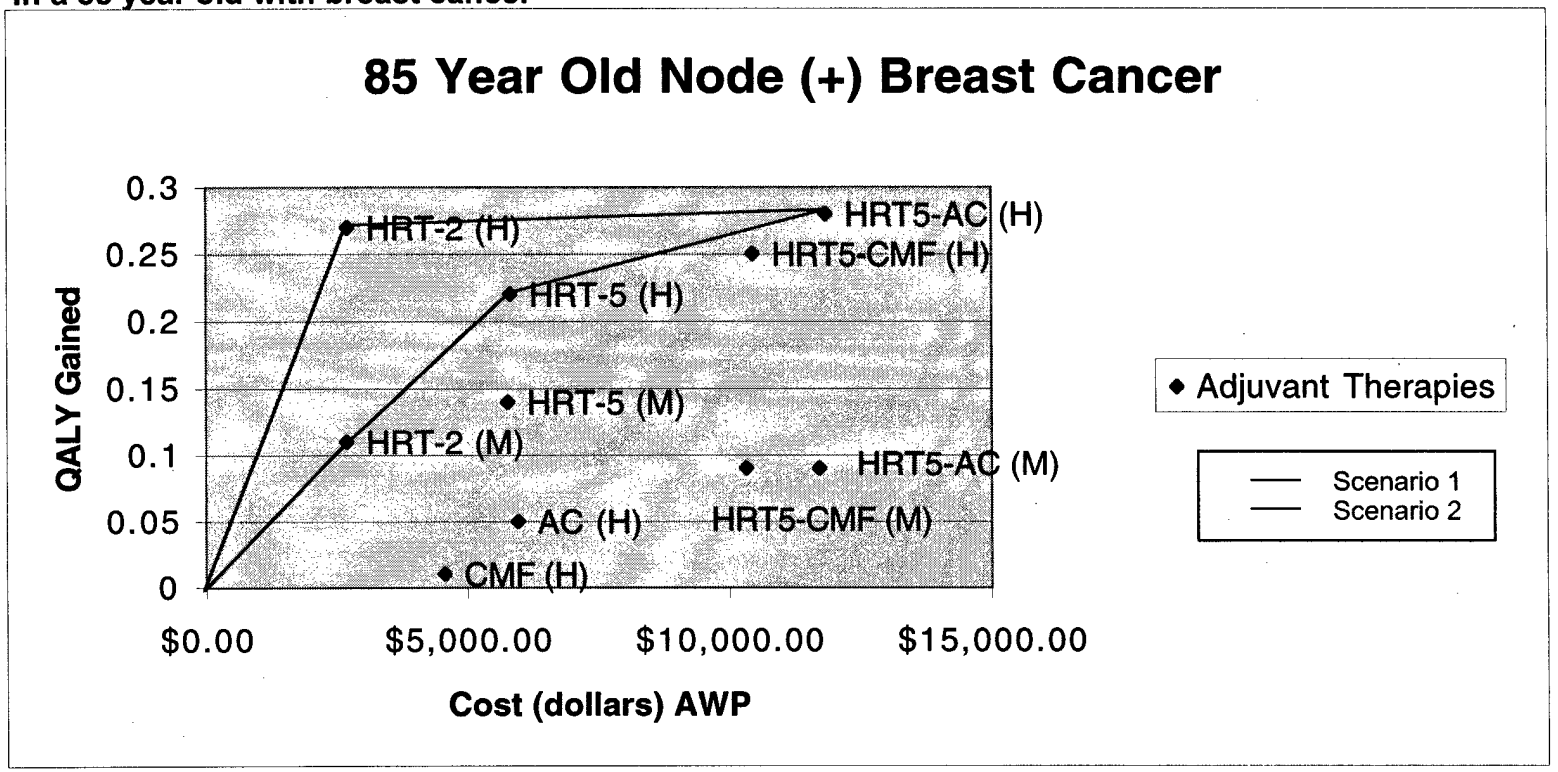
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 5b: Cost and Benefits of Diiferent Adjuvant Therapies In a 85 year old with breast cancer



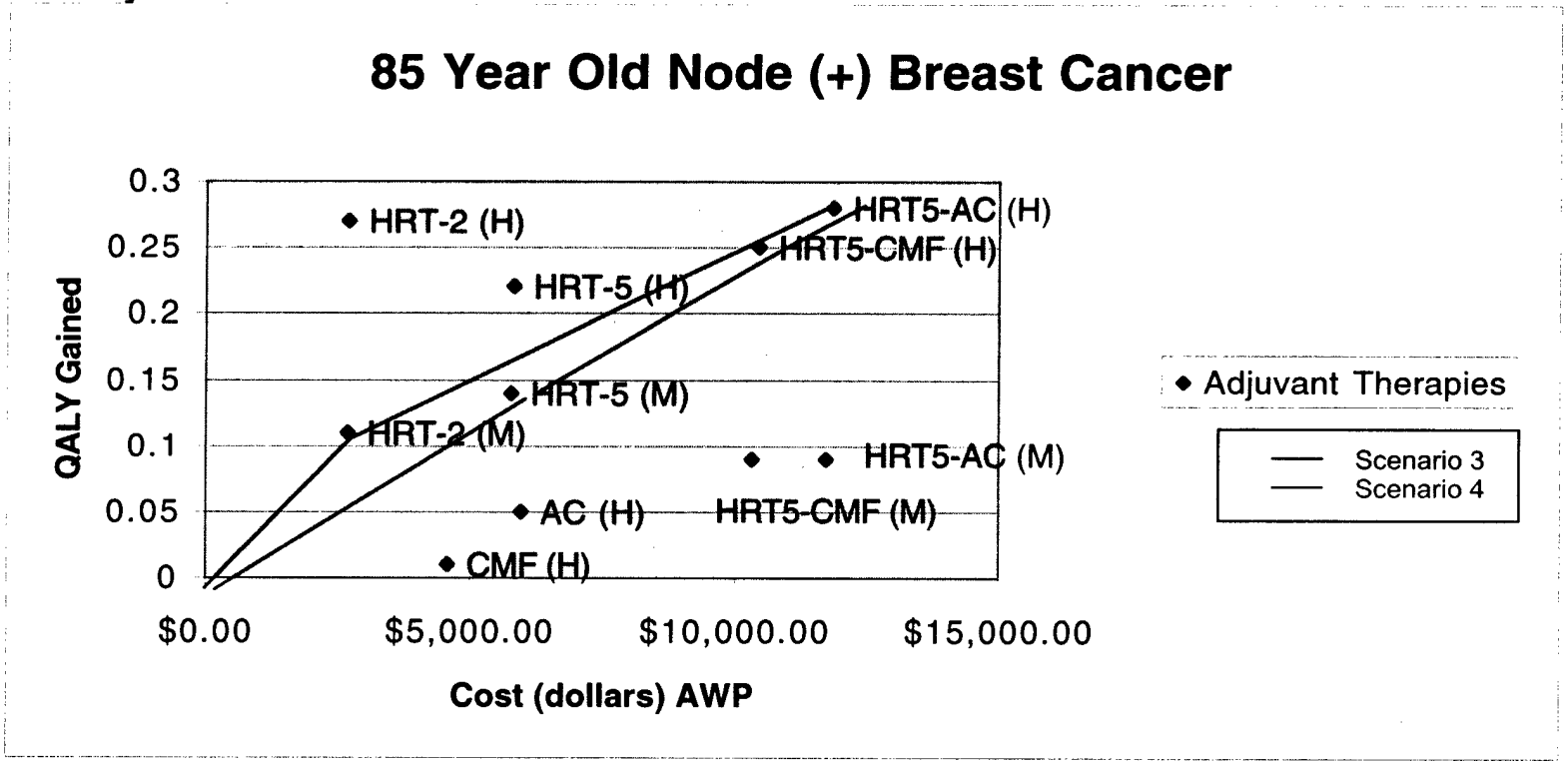
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 5c: Cost and Benefits of Diiferent Adjuvant Therapies
In a 85 year old with breast cancer



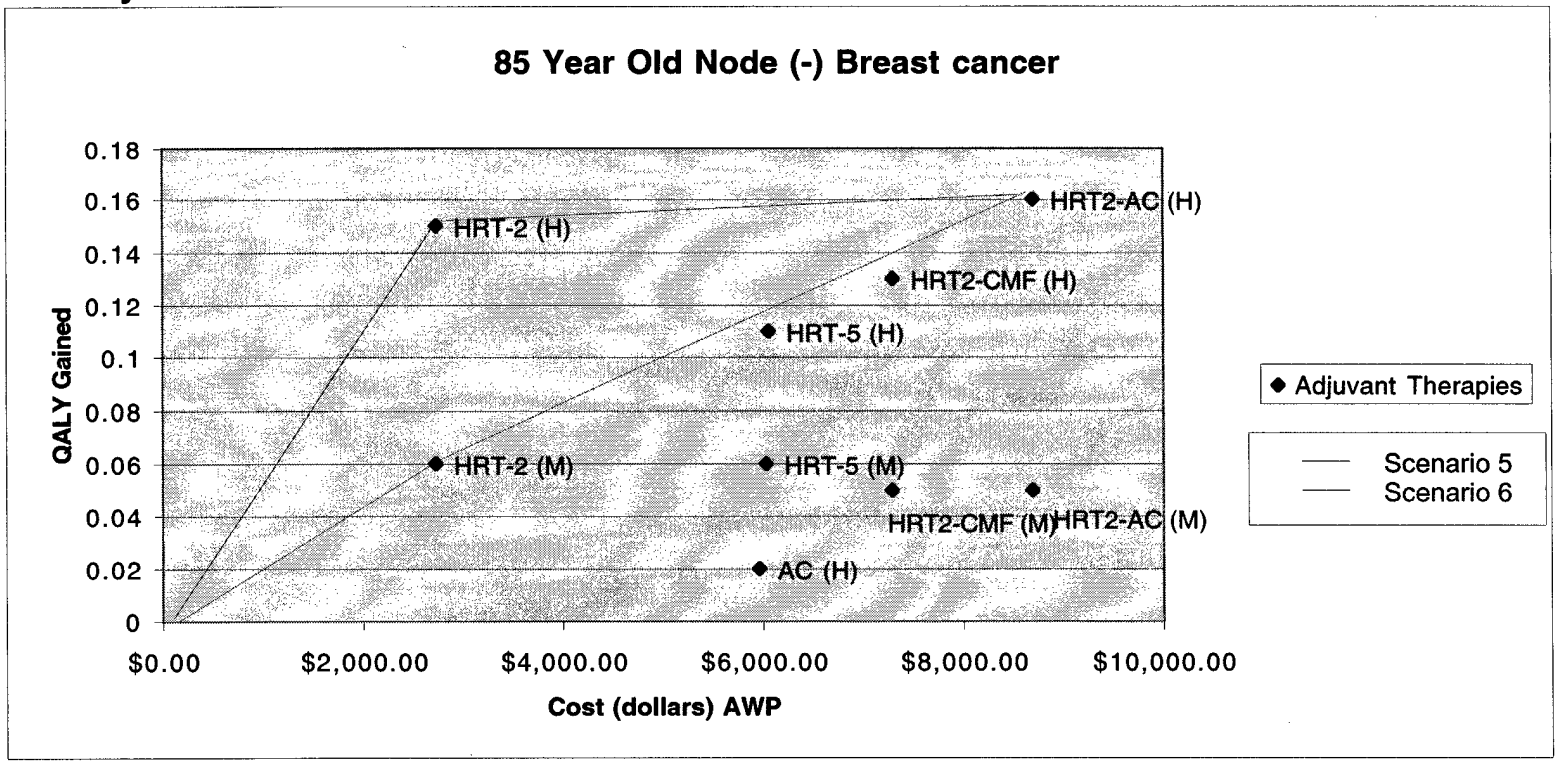
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 5d: Cost and Benefits of Diiferent Adjuvant Therapies In a 85 year old with breast cancer



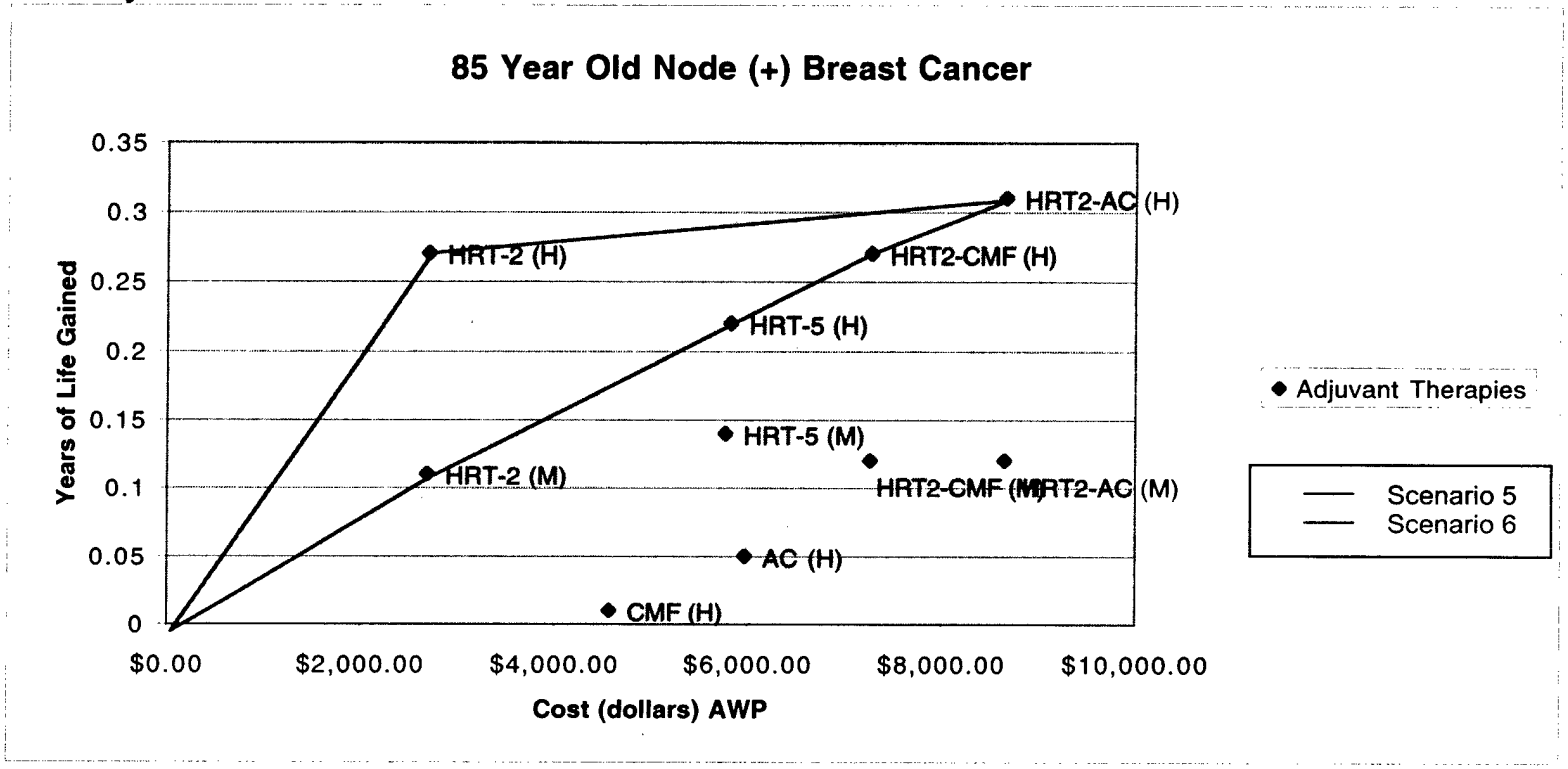
Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 6a: Cost and Benefits of Diiferent Adjuvant Therapies
In a 85 year old with breast cancer



Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

Figure 6b: Cost and Benefits of Diiferent Adjuvant Therapies In a 85 year old with breast cancer



Decision-makers should fund along the marked line until the money runs out or up to a designated cut-off point

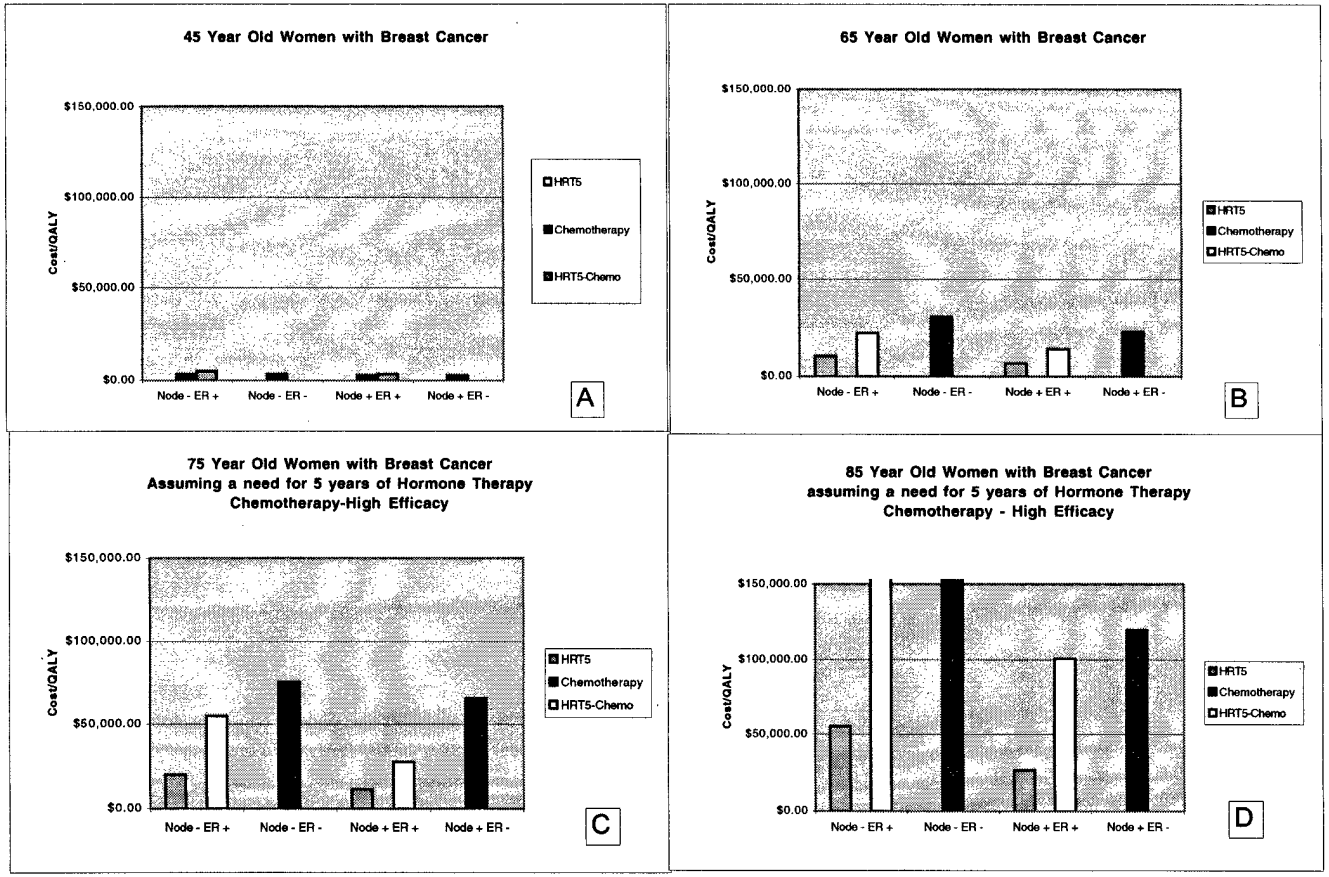


Figure 7: (Panel A) - In a 45 y/o patient all modalities of adjuvant therapy are cost-effective, but chemotherapy dominates hormone therapy alone in those patients who are ER+. Combination therapy is incrementally cost-effective. (Panel B) - In a 65 y/p patient hormone therapy dominates chemotherapy and comination therapy is incrementally cost-effective in patients who are ER+. In ER - patients, chemotherapy is cost-effective as adjuvant therapy. (Panel C and D) - If one assumes the same efficacy with chemotherapy as a 65 year old, the effect of age (and associated comorbidity and functional decline) can be evaluated. (Panel C) - In a 75 year old, hormone therapy is still very cost effective in ER+ patients, and combination therapy is still incrementally cost-effective. Chemotherapy alone is also cost-effective in ER - patients. (Panel D) - In an 85 year old, even though hormone therapy is still cost-effective, combination therapy is not in Node - patients. Furthermore, chemotherapy is not cost-effective in node - ER - patients and marginal in Node + patients.
Bars indicate incremental cost-effectiveness of the dominate therapies. Inferio, or dominated adjuvant strategies are note shown (bars empty).

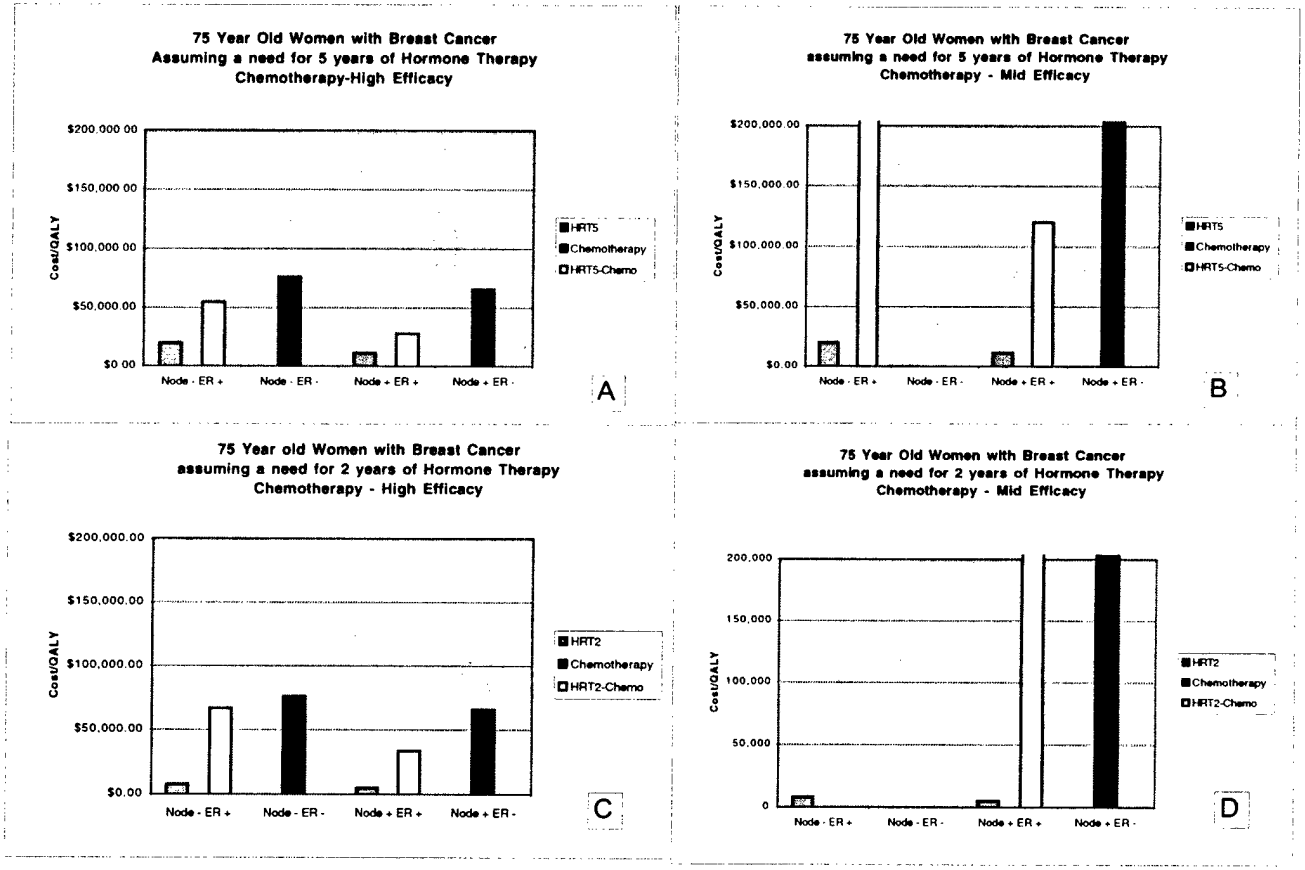


Figure 8: Models using different assumptions in 75 year old women are shown here. The based model is shown in Panel A. (Panel B) - If one assumes only mid-efficacy of chemotherapy, chemotherapy is not cost-effective in ER - patients. Combined therapy may be incrementally cost-effective. In Node + ER + patients (Panel C) If one assumes 2 years hormone therapy provides maximum benefit and chemotherapy has high efficacy, then hormone therapy is very cost-effective and combined therapy is incrementally cost-effective in ER + women. In ER - patients chemotherapy is still cost-effective. (Panel D) - In a scenario where 2 years of hormone therapy is sufficient and the benefits from chemotherapy are reduced, only hormone therapy in ER+ women is cost-effective. The use of chemotherapy or combined therapy is not cost-effective.

Bars indicate incremental cost-effectiveness of the dominate therapies. Inferio, or dominated adjuvant strategies are note shown (bars empty).

Healthcare Cost-Effectiveness Analysis for Older Patients: Using Cataract Surgery and Breast Cancer Treatment Data

Chapter 6

Current Issues and Future Work

Arash Naeim, M.D.
Geriatric-Oncology Fellow
RAND-UCLA Doctoral Fellow in Public Policy
Division of Hematology-Oncology
Department of Medicine
University of California, Los Angeles
Los Angeles, California

Introduction

Patients, physicians, nurses, insurers, and policymakers would probably all agree that providing high quality cost-effective medical care to older patients is essential. Differences in opinion may exist though on how to define "high quality" and the standards and methods used to assess cost-effectiveness. This dissertation has highlighted several aspects of performing cost-effectiveness analysis using: (a) a randomized control trial of cataract surgery and (b) modeling using data derived from meta-analyses of trials with mainly patients under 70 years of age. Each analyses demonstrated inherent limitations.

Although randomized control trials are considered the "gold standard" for evidence-based data, there may be problems with adequate accrual (statistical power), selection bias, and precise measurement of outcomes using generic instruments (such as the Health Utilities Index Mark3 in the randomized cataract surgery trial). Furthermore, randomized clinical trials are expensive and slow. On the other hand, any modeling exercise, even though relatively inexpensive, relies on a set of assumptions, the validity of which can always be challenged.

Several issues need to be resolved in order to determine optimal treatment choices and pathways for older patients. These issues include the following: (1) improving clinical trial recruitment of a representative population of older patients; (2) developing outcome measurement instruments reliable and valid in an older population; (3) defining approaches to communicate uncertainty in treatment outcome to patients; (4) soliciting treatment preferences from older patients, (5) creating a decision aid that incorporates the best available clinical evidence with patient preferences in order to individualize treatment, and (6) collecting functional status*, comorbidity, patient preference, treatment intensity, and outcome data (disease progression, mortality, and quality of life**) prospectively to further define the complex interactions that shape the healthcare treatment of older patients. Future work on these issues may lay the foundation required to overcome many of the limitations described in this dissertation.

---

* Functional Status refers to the ability of an individual to perform required daily tasks at home and at work. It reflects the level of physical strength, mobility, and energy level of an individual.

** Quality of Life is a subjective measure of one's overall enjoyment of life. It includes some of the components of functioning, similar to Functional Status, but also includes other physical (pain, nausea,...), social, and emotional components. Therefore, even though an individual may objectively have a poor functional status, they may rate their overall quality of life high, or vice versa.

Performing Clinical Trials on Older Patients

Clinical trial recruitment played a central issue in both the cataract surgery trial and the breast cancer modeling analysis. In the randomized trial of cataract surgery, the recruitment, enrollment, and follow-up of older patients was very labor intensive. These obstacles limited the number of patients in the trial. This resulted in adequate power[ϕ] to measure significant changes in visual functioning but not enough power to measure changes in overall utility (although the sensitivity of the instrument to small changes plays a role as well). The lack of recruitment and enrollment of older patients in clinical trials was also a feature of most breast cancer trials [1] looking at adjuvant treatment. However, the critical contributing factor in these trials were the protocol exclusion criteria in many trials limiting participation to: (a) those under 70, (b) with good performance status, and (c) minimal comorbidities.

Can the information in the cataract and breast cancer clinical trials be extended to older patients in the community? In the cataract surgery trial the enrollment of only cognitively and hearing intact patients was required in order to complete the outcome surveys. Even though the patients were all over 65, there was probably selection bias. The data from breast cancer trials is even more suspect since older patients represented such a small percentage of the participants, and those that were enrolled probably represented the healthiest subgroup.

Several possibilities exist to improve clinical trial data for older patients. A more stringent application of the FDA guidelines for the study of drugs likely to be used in the elderly would be a good start. Two approaches are available for improving the implementation of these guidelines. First, there could be a requirement placed that all phases of clinical trials should have samples that reflect the percentage of older patients with the disease entity being studied. Alternatively, there could be a requirement that Phase IV clinical studies be performed specifically to determine treatment response in older patients. This secondary approach though suffers from the lack of data on treatment safety and toxicity specific to older patients derived from earlier phases of clinical trials.

Including older patients, who on aggregate have more comorbidities and functional limitations, requires a reformulation of clinical trial protocols. If only the healthiest of older patients

---

[ϕ] Power is defined as the ability of a study to find true differences of between two groups or arms (ex control and treatment group). It is determined by the sample size and the "effect size", the standardized difference between the two groups at a specific confidence level (95%). Power=0.8, α=0.05.

are enrolled in clinical trials, then healthcare providers will still be in a quandary as to how to treat the average patient presenting in their clinics. On the other hand, including a frail subgroup would require dose and interval adjustments of many drugs in clinical trials. Perhaps companion protocols for clinical trials designed specifically to address these issues would be helpful. The International Conference on Harmonization of Technical Requirements (ICH) published a guideline for industry for Registration of Pharmaceuticals for Human Use on studies in support of older populations. This document highlighted the importance of including patients over the age of 75 and avoiding unnecessary exclusion of patients with concomitant illnesses [2]. Specifically highlighted was the need to recognize pharmocokinetic differences between younger and older patients, often related to impairment in renal or hepatic function or to drug-drug interactions [2].

Once older patients are appropriately represented in clinical trials rather than an excluded, focus can be shifted in developing better strategies for patient recruitment. Further research needs to focus on how to better integrate primary care physicians and geriatricians in the clinical trial process. Educating community physicians on the existence of trials, their criteria, and the pros and cons of a trial for an older patient is essential. Since so many clinical trials are performed, the task may initially seem overwhelming. However, creating user-friendly databases that physicians can access via the Internet or on a Compact Disk that focus specifically on trials for which older patients are eligible would be of great help. These databases may be a resource not only for healthcare providers but also older patients and their families seeking cutting-edge treatment for their illness.

Finally, a concerted effort must be made to facilitate an older patient's continued participation in a trial once they have agreed to participate. Clinical trials often require more follow-up visits, more paperwork, and more lab tests. Older patients who live alone and have difficulty with transportation are a great risk for either loss to follow-up or poor compliance to the clinical trial regiment. A multidisciplinary approach to patient care utilizing the additional services of a social worker, physician and occupational therapist, pharmacist, and nutritionist would greatly supplement the standard care provided older patients on clinical trials. Considering that one in five individuals will be over the age of 65 within the next several decades, it is important that researchers and policymakers view older patients not as an obstacle that needs to be overcome but more an untapped exciting opportunity to provide more generalizeable clinical trial data.

Outcome Measurement Instruments in Older Populations

There is a plethora of health-related quality of life (HRQOL) measures currently available for use. In the cataract surgery trial, the key outcome measure for the cost-effectiveness analysis was the Health Utilities Index, Mark 3, HUI3. The HUI3 is a generic instrument. One concern with selecting a generic instrument measuring overall health was its sensitivity in measuring changes related to just vision. Alternatively a condition-specific (vision) measure might have been more sensitive to visual changes. The decision to choose either a generic or condition-specific HRQOL is controversial [3-6]. In older patients, who have many competing illnesses, there is a strong argument for using generic HRQOL measures since the goal is to determine the benefit of an intervention in the context of overall health. Some researchers argue that the standard should be to include both generic and disease-specific instruments [7], while others argue there is no clear guideline on deciding how to use the results from two different instruments [1].

More importantly, it is note clear if HRQOL measured in a clinical trial population is representative for a non-clinical trials population. One interesting study focused on the HRQOL in two cohorts of patients with human immunodeficiency virus (HIV) disease: (a) multi-center AIDS Clinical Group Trials in which most subjects are white, privately insured, and high-income (n = 1,907); and b) a study of ethnically diverse, low-income patients recruited from public clinics (n = 205) [8]. HRQOL scores were significantly lower in the non-trial sample ($P < 0.001$) by about one standard deviation, even after direct adjustment for clinical and demographic characteristics [8], raising concerns about generalization of HRQOL results from clinical trials.

The psychometric properties of these instruments are usually estimated in patient and population settings where older individuals are under-represented. An argument can be made that these instruments should be assessed in an older population. Additionally, the use of aggregate quality of life data and preference weights from younger patients may be biased. For example, the modeling of adjuvant therapy of breast cancer used quality of life weights derived from clinical trials, which excluded patients over the age of 70. Aggregate values may not be the best approach to adjusting for quality of life even if, derived in an older population (see preference solicitation discussion below).

Communicating Risk and Benefit with Uncertainty

Communicating risk and benefit to patients when there are insufficient data or too many complicating variables is an understudied area. The communication of risk and benefit is essential with both cataract surgery and adjuvant therapy for breast cancer. Previous studies using the Cataract Surgery Index, CSI, have shown that the probability of benefiting from surgery can be predicted. The manner in which probabilities should be expressed to patients and how they are adjusted for other factors, such as competing illnesses or underlying functional status, is unclear. These issues are magnified when discussing adjuvant breast cancer for an older patient where there are several complex sequential steps. First, there needs to be extension of the benefits and side effects of treatment from clinical trials in younger patients to older patients. Next, a baseline life expectancy needs to be derived for the patient without treatment. This baseline is derived from aggregate life table data from a very heterogeneous older population based on chronological age. There is no accepted method of adapting life expectancy for individual levels of comorbidity and functional status to derive a true individualized physiologic life expectancy. This baseline then needs to be adjusted for additional mortality risk from the acute disease. Finally the uncertain benefit and baseline life expectancy need to be combined to determine the quality-adjusted life years gained with treatment.

Communicating the uncertainty behind this entire process to an older patient may be confusing. One might argue that uncertainty should not be communicated at all, just a physician's best recommendation. However, given that informed consent to treatment is a cornerstone in medicine, and adequate informed consent requires a thorough discussion of risk and benefits, the discussion of uncertainty seems a prerequisite.

Most of the sparse literature on communicating uncertainty comes from cancer screening and treatment. In cancer screening, testing for a gene or protein can help demonstrate that someone has a higher probability for the development of a future cancer. Communicating cancer risk information from this type of testing is germane to a number of health professions including physicians, geneticists, genetic counselors, psychologists, nurses, health educators and social workers [9]. Some recent work has focused on techniques in communicating risk and benefits, but none of these studies have specifically focused on an older population nor have they incorporated the communication of degrees of uncertainty [10-13]

Preference Solicitation in Older Patients

Some have argued strongly that preference measures may not be appropriate for older patients. Threats to validity of these instruments include: construct under-representation and construct-irrelevant variance [14]. Construct under-representation occurs when a stimulus presented to a judge fails to fully represent the depth and complexity of information required in actual judgments [14]. Construct-irrelevant variation occurs when factors irrelevant to preferences influence measurements of utilities. Among several factors that cause construct-irrelevant variation are cognitive abilities, calculation skills, emotions and prejudices, and the elicitation procedure [14]. Cognitive abilities and the ability to perform numerical calculations are often diminished in older patients. Furthermore, commonly used elicitation methods (visual-analog scales, time tradeoff, and standard gamble) capture different preference facets (desirableness of states, time preferences, and risk attitude) to different degrees [14].

Although there are reports of patient preferences for different types of cancer treatments [15], little work has been performed on the preferences of older cancer patients. Unfortunately, communication with older patients regarding treatment options and benefits is often time consuming for oncologists. As a result, patient preferences are often not incorporated adequately into the decision making process [16]. Previous research has shown that with methods adapted for their limitations, health preferences can be successfully elicited in patients over the age 80, but these preferences varied greatly depending on baseline health. Preference solicitation in the elderly must consider how patients feel about health in the absence of the disease being studied (i.e. how bothered patients are about their specific comorbidities) [17].

Decision Aids and Individualizing Treatment

Although eliciting preferences can be challenging, several studies have shown that incorporating individual preferences into decisions may help with physician and patient education, empower patients, and help establish treatment recommendations [18-20]. A study of 60 early-stage

breast cancer patients showed that patient participation in deciding cancer treatment empowered many patients and promoted responsibility for their own care [21]. Since the amount of acceptable risk or tolerance for uncertainty is likely to be heterogeneous in an older population of patients, models that allow incorporation of individual preferences may be more representative than those that use general population weights [22].

The use of preferences and decision analysis has been used successfully in the treatment of patients with T3 laryngeal lesions, prophylactic oopherectomy, and atrial fibrillation[23, 24] [25] [26]. Tools incorporating patient preferences have lead to better-informed patients, better care, and better health outcomes from the patient's point of view. In addition to helping tailor care to individuals, these decision tools will help link evidence-based medicine into clinical practices and may improve the quality of care [27].

Recently, computer programs have been developed that use decision models and automatically create evidence-based guidelines and recommendations and that can be individually tailored and updated. An example has been a software system called ALCHEMIST that utilizes decision models to create evidence-based guidelines. This tool has been used with success in studying the need for implantable cardioverter defibrillators (ICD) and BRCA breast cancer mutation testing in women. The study showed that such a web-based system could easily incorporate individual preferences, weighting for relevant health states, and create patient-specific recommendations that result in an increase in quality-adjusted life expectancy [28].

Database Development

The fact that older patients tend to have additional comorbidity and functional limitations creates additional challenges in determining if high quality of care is provided. Treatment selection, under-treatment and over-treatment, is related not only to important factors such as disease stage, comorbidity and functional status, but also less justified factors such as age, gender, and race. Breast cancer is a perfect example since older patients are often under-treated in terms of surgery, radiation, and hormone therapy, and sometimes over-treated when it comes to chemotherapy [29] [30, 31]. Most of the studies trying to separate the impact of these multiple factors in older patients have been retrospective.

Prospective population-based study looking at the factors that determine treatment of older patients is needed. For example in cancer, it would be very valuable to look at treatment selection and chemotherapy dosing in older patients. Older patients are often not offered chemotherapy, and when given, doses are often reduced. To determine whether dose-reductions are justified or these reflect ageism, one would need to collect information on a wide range of confounders including functional status, comorbidities, physiologic (kidney and liver) function, and patient preferences. Furthermore, it would be important to determine if the treatment variation that occurs among older patients has an effect on outcomes (quality of life, disease progression, hospitalizations, and mortality). This information would be essential in confirming or refuting many of the assumptions used in the modeling of benefits from adjuvant chemotherapy in older breast cancer patients.

These types of studies can be helpful in the development of interventions to improve the care of older patients. Ideally a database with these important variables would be national project or at least representing a large collaborative group of institutions. Unfortunately, a project of this magnitude is costly and may have to begin at local institutions where research on the care of older patients is a priority.

Conclusion

Health services research, particularly cost-effectiveness analysis, on older patients will play a critical role in the next several decades. As the population of older individuals grows, policymakers will need to make difficult resource decisions in order to provide for this community. More clinical trials need to be designed to incorporate older patients and reflect the characteristics of general community. The lack of information on treatment benefit for older patients leads to uncertainty, which makes communication of treatment benefits to older patients difficult. Uncertainty in treatment benefits and costs can be mapped using decision analysis modeling. However, all modeling exercises utilize a set of assumptions that need to be evaluated and verified by future clinical trials. Nevertheless, this process may identify key questions for which future clinical trials can be designed.

In order to provide high quality medical care in the future several issues specific to older patients need to be resolved. These issues include: (1) the structure and selection criteria of clinical trials, (2) appropriate outcome measures for older patients, (3) improving communication regarding

benefits and side effects of treatment, (4) validated methodology to solicit treatment preferences from older patients, (5) individualizing treatment plans by taking into account both patient preference and the heterogeneity of comorbid disease burden and functional limitations, and finally (6) creating a prospective database to follow the many contributing factors that determine the type and care provided to older patients. These areas represent the future of outcomes and health services research in an older population.

References

1. EBCTCG, *Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group.* Lancet, 1998. 352(9132): p. 930-42.
2. ICH, *Guideline for Industry. Studies in Support of Special Populations: Geriatrics.* Federal Register, 1994. 59(102): p. 39398-39400.
3. Dowie, J., *Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions.* Health Econ, 2002. 11(1): p. 1-8.
4. Feeny, D., *Commentary on Jack Dowie, "Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions".* Health Econ, 2002. 11(1): p. 13-6; discussion 21-2.
5. Guyatt, G., *Commentary on Jack Dowie, "Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions".* Health Econ, 2002. 11(1): p. 9-12; discussion 21-2.
6. Brazier, J. and R. Fitzpatrick, *Measures of health-related quality of life in an imperfect world: a comment on Dowie.* Health Econ, 2002. 11(1): p. 17-9; discussion 21-2.
7. Patrick, D.L. and R.A. Deyo, *Generic and disease-specific measures in assessing health status and quality of life.* Med Care, 1989. 27(3 Suppl): p. S217-32.
8. Cunningham, W.E., et al., *Comparison of health-related quality of life in clinical trial and nonclinical trial human immunodeficiency virus-infected cohorts.* Med Care, 1995. 33(4 Suppl): p. AS15-25.
9. Bottorff, J.L., et al., *Communicating cancer risk information: the challenges of uncertainty.* Patient Educ Couns, 1998. 33(1): p. 67-81.
10. Levine, M. and T. Whelan, *Decision-making process--communicating risk/benefits: is there an ideal technique?* J Natl Cancer Inst Monogr, 2001(30): p. 143-5.
11. Fortin, J.M., et al., *Identifying patient preferences for communicating risk estimates: A descriptive pilot study.* BMC Med Inform Decis Mak, 2001. 1(1): p. 2.
12. Lipkus, I.M., et al., *Testing different formats for communicating colorectal cancer risk.* J Health Commun, 1999. 4(4): p. 311-24.
13. Edwards, I.R. and B. Hugman, *The challenge of effectively communicating risk-benefit information.* Drug Saf, 1997. 17(4): p. 216-27.
14. Lenert, L. and R.M. Kaplan, *Validity and interpretation of preference-based measures of health-related quality of life.* Med Care, 2000. 38(9 Suppl): p. II138-50.
15. Jansen, S., et al., *Patients' utilities for cancer treatments: a study of the chained procedure for the standard gamble and time tradeoff.* Med Decis Making, 1998. 18: p. 391-399.
16. Ganz, P., *Interaction between the physician and the older patient: the oncologist's perspective.* Cancer, 1997. 80: p. 1323-25.
17. Tsevat, J., et al., *Health values of hospitalized patients 80 years or older. HELP Investigators. Hospitalized Elderly Longitudinal Project.* Jama, 1998. 279(5): p. 371-5.
18. Sherbourne, C.D., et al., *Relationship between age and patients' current health state preferences.* Gerontologist, 1999. 39(3): p. 271-8.
19. Harris, R.A. and R.F. Nease, Jr., *The importance of patient preferences for comorbidities in cost-effectiveness analyses.* J Health Econ, 1997. 16(1): p. 113-9.

20. Feldman-Stewart, D., M.D. Brundage, and L. Van Manen, *A decision aid for men with early stage prostate cancer: theoretical basis and a test by surrogate patients.* Health Expect, 2001. 4(4): p. 221-34.
21. Goel, V., et al., *Randomized trial of a patient decision aid for choice of surgical treatment for breast cancer.* Med Decis Making, 2001. 21(1): p. 1-6.
22. Cher, D.J., J. Miyamoto, and L.A. Lenert, *Incorporating risk attitude into Markov-process decision models: importance for individual decision making.* Med Decis Making, 1997. 17(3): p. 340-50.
23. van der Donk, J., et al., *Patient participation in clinical decision-making for treatment of T3 laryngeal cancer: a comparison of state and process utilities.* J Clin Oncol, 1995. 13(9): p. 2369-78.
24. Protheroe, J., et al., *The impact of patients' preferences on the treatment of atrial fibrillation: observational study of patient based decision analysis.* Bmj, 2000. 320(7246): p. 1380-4.
25. Protheroe, J., et al., *Effects of patients' preferences on the treatment of atrial fibrillation: observational study of patient-based decision analysis.* West J Med, 2001. 174(5): p. 311-5.
26. Pell, I., et al., *Development and preliminary evaluation of a clinical guidance programme for the decision about prophylactic oophorectomy in women undergoing a hysterectomy.* Qual Health Care, 2002. 11(1): p. 32-8; discussion 38-9.
27. Djulbegovic, B., I. Hozo, and G.H. Lyman, *Linking evidence-based medicine therapeutic summary measures to clinical decision analysis.* MedGenMed, 2000. 13.
28. Sanders, G.D., R.F. Nease, Jr., and D.K. Owens, *Publishing web-based guidelines using interactive decision models.* J Eval Clin Pract, 2001. 7(2): p. 175-89.
29. Mandelblatt, J.S., et al., *Breast cancer screening for elderly women with and without comorbid conditions. A decision analysis model.* Ann Intern Med, 1992. 116(9): p. 722-30.
30. Mandelblatt, J.S., et al., *Patterns of breast carcinoma treatment in older women: patient preference and clinical and physical influences.* Cancer, 2000. 89(3): p. 561-73.
31. Mandelblatt, J.S., et al., *Constructs of burden of illness in older patients with breast cancer: a comparison of measurement methods.* Health Serv Res, 2001. 36(6 Pt 1): p. 1085-107.