

AD _____

GRANT NUMBER: DAMD17-94-J-4072

TITLE: Special Sabbatical for Training in Health Decision
Sciences with Application to Breast Cancer Treatment Evaluation

PRINCIPAL INVESTIGATOR: Richard D. Gelber, Ph.D.

CONTRACTING ORGANIZATION: Dana-Farber Cancer Institute
Boston, MA 02115

REPORT DATE: Feb 96

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

DTIC QUALITY INSPECTED 1

19960405 045

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE Feb 96	3. REPORT TYPE AND DATES COVERED Final (1 Feb 95 - 31 Jan 96)		
4. TITLE AND SUBTITLE Special Sabbatical for Training in Health Decision Sciences with Application to Breast Cancer Treatment Evaluation		5. FUNDING NUMBERS DAMD17-94-J-4072		
6. AUTHOR(S) Richard D. Gelber, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Dana-Farber Cancer Institute Boston, MA 02115		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command, Fort Detrick, Frederick, MD 21702-5012		10. SPONSORING / MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words) A special sabbatical year enabled Dr. Richard D. Gelber, Ph.D., an internationally respected clinical trial biostatistician, to receive training in contemporary methods and applications of health decision sciences. Dr. Gelber has been the Director of the Statistical and Data Management Center for the International Breast Cancer Study Group (IBCSG: formerly Ludwig Group) since 1978. With Dr. Aron Goldhirsch, IBCSG Scientific Director, he developed the Q-TWiST method (Quality-adjusted Time Without Symptoms of disease and Toxicity of treatment) for integrating aspects of quality of life into an evaluation of adjuvant therapies for operable breast cancer. Dr. Gelber also participates in the Worldwide Overview of randomized clinical trials evaluating adjuvant therapies for breast cancer. In addition, he currently coordinates biostatistics for the Division of Pediatrics at Dana-Farber Cancer Institute, directs the Pediatric Section of the Statistics and Data Analysis Center for the NIH AIDS Clinical Trials Group, and teaches a course in clinical trials at the Harvard School of Public Health. This award freed Dr. Gelber from his non-breast cancer related research responsibilities for one year, and enabled him to study methods in health decision sciences under the mentorship of Professor Milton Weinstein. This significantly enhanced Dr. Gelber's ability to influence treatment choices for patients with breast cancer by refining interpretation of clinical trial results. The special sabbatical coincided with the third update of the Worldwide Overview in which Dr. Gelber participated.				
14. SUBJECT TERMS health decision sciences, breast cancer adjuvant treatments, cost-effectiveness analysis, cost-benefit analysis, quality-adjusted survival analysis, breast cancer			15. NUMBER OF PAGES 73	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet optical scanning requirements.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

NASA - Leave blank.

NTIS - Leave blank.

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

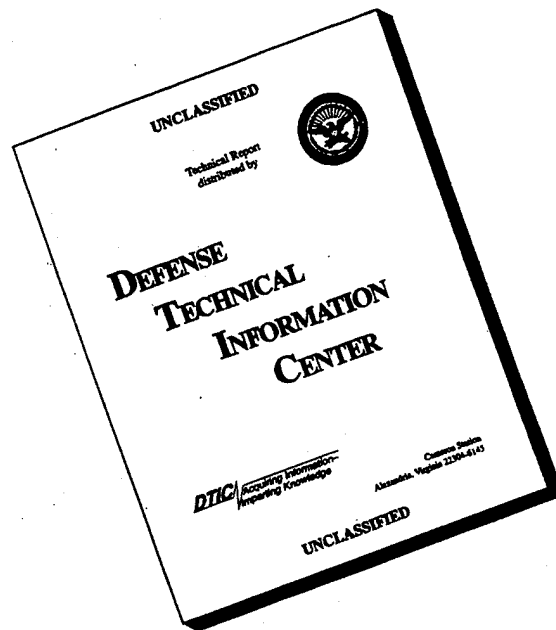
Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

N/A Where copyrighted material is quoted, permission has been obtained to use such material.

N/A Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

N/A Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

N/A For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories. ...

Richard D. Sellen 2/22/96
PI - Signature Date

FINAL PROGRESS REPORT

Special Sabbatical for Training in Health Decision Sciences with Application to Breast Cancer Treatment Evaluation

	PAGE
TABLE OF CONTENTS	4
INTRODUCTION - Background, Relevance, and Methods	5
BODY - Progress Report on Specific Aims	8
CONCLUSIONS	14
REFERENCES	15
APPENDIX (items copied on both sides of the page)	16
- Paper Presented at the EORTC Symposium on Methodological and Statistical Issues of Quality of Life and Economic Evaluation in Cancer Clinical Trials. - Grant prepared to support the development of breast cancer patient preferences for use in Q-TWiST analyses. - Paper published in The Cancer Journal of Scientific American on the Q-TWiST meta-analysis evaluating chemotherapy for premenopausal breast cancer patients. - Paper in press in The Lancet on the Q-TWiST meta-analysis evaluating chemotherapy plus tamoxifen versus tamoxifen alone for postmenopausal breast cancer patients.	

FINAL PROGRESS REPORT

Special Sabbatical for Training in Health Decision Sciences with Application to Breast Cancer Treatment Evaluation

INTRODUCTION

A crucial question both for the individual breast cancer patient and for society as a whole is identifying "optimal" treatment strategies for the medical care of each woman. Randomized clinical trials are now well-established as the gold standard methodology for comparative assessments of treatments for breast cancer. As Director of the Statistical Center for the International Breast Cancer Study Group (IBCSG: formerly Ludwig Group), Dr. Gelber has contributed substantially to the design, conduct, analysis and interpretation of randomized clinical trials for breast cancer adjuvant therapies.

The usual statistical approaches to the evaluation of such studies (e.g. assessment of disease-free survival and overall survival), however, do not adequately reflect the impact of therapies on the health-related quality of life of the patients. Dr. Gelber, in collaboration with Dr. Aron Goldhirsch, developed the Q-TWiST method (Quality-adjusted Time Without Symptoms of disease and Toxicity of treatment) to perform breast cancer adjuvant therapy treatment comparisons based on quality-adjusted survival analysis. The method has been applied in several clinical trials to illustrate the trade-off between quality-of-life decrements related to toxic effects of treatment and quality-of-life benefits associated with prolonging the time to breast cancer recurrence. Although the method is technically similar to the health decision science QALY approach (quality-adjusted life years), Q-TWiST was developed within the context of randomized clinical trials. Thus, Q-TWiST begins to bridge the gap between clinical trials research (used by medical oncologists) and clinical decision research (used by decision analysts).

Decision analysis is a systematic approach to decision-making under conditions of uncertainty. By learning more about contemporary methods and applications of health decision sciences, such as decision tree design, sensitivity analysis, Markov processes, cost-effectiveness analysis, utility assessment, and multi-attribute utility theory, Dr. Gelber took the next step to build on the Q-TWiST method and clinical trial applications. Dr. Gelber's new training enabled him - when designing clinical trials - to consider factors that influence subsequent clinical decisions. This special sabbatical was, therefore, directly relevant to improving methods used to evaluate treatments for women with breast cancer. The additional training enhanced Dr. Gelber's ability to contribute new ideas for the interpretation of results from the Worldwide Overview of adjuvant therapies for breast cancer. The Third Update of the Overview took place in September, 1995. The application of health decision sciences methods for performing supplemental analyses on this enormous database will provide important insights concerning costs and benefits of treatments for women with breast cancer.

The Candidate - Dr. Richard D. Gelber

Since acquiring his Ph.D. in 1975, Dr. Richard D. Gelber has been involved in a variety of activities which have secured his international reputation as a top biostatistical collaborator in clinical trial research. His career reflects his commitment to the application of appropriate statistical methods for the design and analysis of cancer clinical trials through successful collaboration with medical researchers. His leadership in international breast cancer trials, his innovative research in quality-of-life methods, his outstanding contributions to the pediatric oncology program at the Dana-Farber Cancer Institute (DFCI), and his service to the Department of Biostatistics, Harvard School of Public Health (HSPH) illustrate the quality of his work.

Dr. Gelber's main research interests involve the practice of biostatistics for the design, conduct, analysis and report of clinical trials in breast cancer, pediatric oncology and pediatric AIDS. A significant contribution has been the development, with Dr. Aron Goldhirsch, of the TWiST and Q-TWiST methods for comparing treatments incorporating aspects of quality of life. TWiST, which stands for Time Without Symptoms of disease and Toxic effects of treatments, was originally proposed to compare adjuvant chemotherapy versus observation for postmenopausal women with operable breast cancer. The overall survival time for patients in clinical trials is partitioned so that the positive effects of treatments which prolong the time to disease progression can be balanced against the increased subjective toxic effects of such therapies. This work received the 1987 Robert Wenner Prize awarded by the Swiss Cancer League, and the 1987 Farmitalia Carlo Erba Prize awarded by the German Cancer Society's Working Group for Medical Oncology.

Q-TWiST

Quality-adjusted TWiST (Q-TWiST) represents a generalization of the method which allows a spectrum of utility coefficient weights based on quality-of-life considerations to influence the interpretation of treatment comparisons in clinical trials. (See the first paper in the Appendix for a description of the method.) Recent work supported by a research grant from the American Cancer Society (awarded 1990-1994 with renewal pending) has extended the application of the method to a variety of clinical trial settings, and has developed further alternative methods for describing treatment benefit. These have focused on describing treatment effects in terms of the amount of quality time gained by patients on various treatments, and extrapolating available data to project potential future gains for treated patients.

Dr. Gelber's experiences with the development and application of the Q-TWiST method provided the motivation to pursue the proposed special sabbatical. To date the method has been used only with data routinely available in clinical trials to describe quality-of-life tradeoffs in treatment comparisons. Approaches to incorporate information obtained directly from the patients have not yet been used. Modern techniques of health decision sciences and patient input for clinical decision-making are required in order to apply the Q-TWiST concepts more effectively for evaluation of treatments for breast cancer patients.

Oncologists have come to rely on Kaplan-Meier curves for disease-free survival (DFS) and overall survival (OS) to evaluate efficacy of treatments in clinical trials. The physician can use these results to describe the "percent of patients who benefit" or "the relative reduction in the risk or odds of an event" such as relapse or death. However, both patients and physicians are also interested in how the various treatments may influence duration and quality of life.

Cancer patient care decision-making frequently involves evaluating the trade-off between the toxicity of therapy and the treatment-induced benefit of delayed recurrence of the disease. Q-TWiST provides the opportunity for the psychosocial measures of quality of life to influence the utility weights applied to the periods defined by TOX, TWiST and REL. In most studies, quality-of-life assessments are made independently of treatment efficacy assessments. The result is often that the treatment is shown to be beneficial in terms of disease-free survival or overall survival while the quality of life during treatment is lower in the more toxic therapy group. The Q-TWiST method bridges the gap between these two assessments and allows one to statistically compare treatments studied in clinical trials *directly* in terms of both quantity and quality of life. The method "speaks to oncologists" who are most comfortable with disease-free survival and overall survival curves, while also integrating quality-of-life issues into treatment comparisons.

SPECIFIC AIMS

1. To obtain training in health decision sciences methodologies by auditing courses, attending conferences and seminars, and reading key papers under the mentorship of Professor Milton C. Weinstein, Henry J. Kaiser Professor of Health Policy and Management, Harvard School of Public Health.
2. To participate in a research project involving application of health decision science methods to a problem relating to breast cancer treatment evaluation - specifically the development of a multi-attribute utility scale for quality of life of breast cancer patients.
3. To apply health decision science techniques for the analysis of some aspects of adjuvant therapy based on the data from the Third Update of the Worldwide Overview of Randomized Clinical Trials of Adjuvant Therapies for Breast Cancer.

BODY - PROGRESS REPORT ON THE SPECIFIC AIMS

The funding provided by this Special Sabbatical Award was sufficient to cover 45 percent of Dr. Gelber's salary for the one year period of February 1, 1995 through January 31, 1996. Other salary support was provided from sources directly related to ongoing research on studies evaluating treatments for breast cancer.

Specific Aim 1. To obtain training in health decision sciences methodologies by auditing courses, attending conferences and seminars, and reading key papers under the mentorship of Professor Milton C. Weinstein, Henry J. Kaiser Professor of Health Policy and Management, Harvard School of Public Health.

Professor Milton C. Weinstein served as the "mentor" for training in health decision sciences. He advised Dr. Gelber regarding courses to audit, conferences and seminars to attend, and literature to read in the field. Dr. Weinstein heads the Program in Health Decision Sciences at the Harvard School of Public Health. The material below, taken from the Course Catalogue, describes the program and the three courses which Dr. Gelber attended during the sabbatical year:

The objective of the Program in Health Decision Sciences is to offer integrated educational training in decision sciences within the context of health problems.

Decision sciences may be defined as the body of knowledge concerning quantitative techniques in decision making at the individual and collective level. It includes decision analysis, cost-benefit and cost-effectiveness analysis, and behavioral decision theory, as well as part of operations research, applied welfare economics, statistical inference, and computer science. Specific course offerings provide an operational definition of the field at the School of Public Health. However, while methods and theory will be taught rigorously, the primary emphasis of the program is on their application to decisions in medicine, health care policy, and the physical and social environments.

HPB 280b. Decision Analysis for Health and Medical Practices (Department of Health Policy and Management and the Department of Biostatistics)

Lectures. Two 2-hour sessions each week. 2.5 credits. Dr. Paltiel. Concerns the methods and applications of decision analysis, cost-effectiveness analysis, and cost-benefit analysis in the evaluation of medical technologies and health programs. Stresses applications and limitations. Examples used to illustrate techniques include: treatment decision for acute abdominal pain, coronary artery bypass surgery, cost effectiveness of pharmaceuticals, evaluation of immunization programs, and priority setting for AIDS prevention. Course emphasizes applications to medical technology assessment and health care resource allocation.

HPB 281c. Clinical Decision Analysis (Department of Health Policy and Management and the Department of Biostatistics)

Lectures, seminars. Two 2-hour sessions each week. 2.5 credits. Dr. Weinstein. An intermediate-level course on methods and applications of decision analysis

and other modeling techniques to clinical problems. Topics include: Markov models, life expectancy and survival modeling, dynamic models, ROC analysis and diagnostic technology assessment, quality of life valuation, multiattribute utility, and behavioral decision theory. Class sessions include presentation of methods and critiques of published analysis of medical decision problems. Students will learn computer software for decision modeling.

HPB 282d. Cost-Effectiveness and Cost-Benefit Analysis for Health Program Evaluation (Department of Health Policy and Management and the Department of Biostatistics)

Lectures, seminars. Two 2-hour sessions each week. 2.5 credits. Dr. Graham. Topics include: methods and applications of cost-effectiveness and cost-benefit analysis for health program evaluation, medical technology assessment, and environmental risk analysis; theoretical foundations; "shadow" pricing; economic valuation of life saving; choice of discount rates; cost accounting applied to economic evaluation in institutional settings; methods for assessing costs of environmental controls; economic evaluation of biomedical research; health status indexes; ethical issues; and modern critiques.

In addition to participating formally in the three courses described above, Dr. Gelber also attended regularly scheduled Risk Analysis and Decision Sciences seminars, sponsored by the Harvard Center for Risk Analysis, held every two weeks during the academic year.

In March, 1995, Dr. Jane Weeks, Assistant Professor in Medicine, Dana-Farber Cancer Institute, and an expert in the field of cost-effectiveness analysis, established the Center for Outcomes and Policy Research (COPR) at DFCI. As a result of the special sabbatical award, Dr. Gelber was able to become a member of COPR faculty. Other faculty members include Drs. Ezekiel Emanuel, Isaac Kohane, David Paltiel, Susan Parsons, and Milton Weinstein. Dr. Gelber attended the weekly seminars sponsored by COPR. He also presented his research at three of the sessions of this group during the sabbatical year. Topics included obtaining patient-centered assessments of preferences for adjuvant therapies of breast cancer and applications of the Q-TWiST method to data from the Worldwide Overview.

Dr. Gelber was invited to be a speaker at the December 1-2 symposium held in Brussels entitled "Methodological and Statistical Issues of Quality of Life and Economic Evaluation in Cancer Clinical Trials". The symposium was sponsored by the European Organization for Research on the Treatment of Cancer (EORTC). Dr. Gelber presented a paper entitled "Integration of length-of-life outcomes and quality-of-life data in cancer clinical trials" which is now in press and acknowledges the support of this special sabbatical award.¹ This work incorporated some of the principles of health decision sciences applied to the evaluation of clinical trial data.

Dr. Gelber is also on the planning committee for another symposium on Statistical Issues in Quality of Life Assessment in Clinical Trials. This symposium is scheduled for July, 1996, in St. Gallen, Switzerland. In addition to developing themes relating to patient-centered decision making for breast cancer therapies, Dr. Gelber is participating in a workshop on methodologies for evaluating longitudinal assessments of quality of life.

Specific Aim 2. To participate in a research project involving application of health decision science methods to a problem relating to breast cancer treatment evaluation - specifically the development of a multi-attribute utility scale for quality of life of breast cancer patients.

The goal of this aim was for Dr. Gelber to participate in a research project relating to breast cancer patient preferences and clinical decision making. The project was designed by Dr. Jane Weeks, Assistant Professor in Medicine, Dana-Farber Cancer Institute and was entitled: "Validation and Implementation of a Novel Approach to Quality of Life and Preference Assessment in Breast Cancer Patients." It was submitted for funding as a research project for support by the United States Army Medical Research and Development Command as part of the Broad Agency Announcement for Breast Cancer Research. Dr. Gelber planned to participate actively in this area of patient preference assessment research as it represents an important new initiative for enhancing the usefulness of the Q-TWiST methodology and improving the relevance of clinical trial reports for patient care decision-making.

Unfortunately, Dr. Weeks' project was not funded. Thus, we were not able to accomplish during the sabbatical year the development of a multi-attribute utility scale for quality-of-life assessment suitable for clinical trial applications. Nevertheless, we continued to pursue this important area of research by preparing another application for funding from the Susan B. Komen Foundation. This application was prepared in collaboration with Drs. Weeks and Weinstein. The summary of the grant application is included in the Appendix. We proposed an interview evaluation of postmenopausal women with breast cancer who had undergone chemotherapy. Our intention was to obtain information on their quality of life as well as on their preferences for the various treatment options available. Attitudes toward recurrence of the disease were also to be ascertained. The information to be obtained from this research would provide patient input for the appropriate interpretation of the analyses conducted on the Overview data (see Specific Aim 3). Unfortunately, we were just informed that the Komen Foundation was not able to support this patient-centered research project. We are continuing, however, to seek grant support that will incorporate patient preferences in evaluations of treatment decision-making.

An additional step was taken during the year to investigate the role of the Q-TWiST method for obtaining information on the comparative effectiveness of two treatment options. Dr. Gelber collaborated with Drs. Jane Weeks and Milton Weinstein to develop a strategy designed to compare a variety of ways to calculate cost-effectiveness ratios, the measures which are receiving increasing acceptance by health decision policy-makers. A grant to support this further research was prepared. The Q-TWiST method represents a potentially valuable user-friendly way to obtain the measure of effectiveness which goes into the denominator of a cost-effectiveness analysis (CEA). By defining clinical health states and estimating patient preferences for these health states, Q-TWiST avoids the need to collect frequent longitudinal assessments of patient preferences over time. In addition, a variety of ways to determine patient preferences (direct assessment, holistic approach, and multi-attribute approach) will be assessed to identify user-friendly approximations to the current standards. Finally, cost data will be collected to determine how different approaches can influence the CEA. The application has been submitted to NIH for review.

During this sabbatical year, Dr. Gelber contributed significantly to understanding the impact of the timing and duration of adjuvant chemotherapy for patients with breast cancer. Two large randomized clinical trials conducted by the International Breast Cancer Study Group (IBCSG: formerly Ludwig Group) on over 2700 patients reached five years of follow-up. Since 1978, this cooperative group has conducted large, randomized clinical trials evaluating adjuvant therapies for breast cancer. Over 8000 patients were enrolled in seven randomized trials that have completed recruitment, and patient entry is continuing in seven other newly activated trials. Participating centers are located in ten countries, and the operations office is in Bern, Switzerland.

For 1475 premenopausal patients with node-positive disease, a recently completed trial showed that three cycles of adjuvant CMF therapy were less effective at preventing disease recurrence than six or nine cycles of CMF.² The trial also showed that reintroduction of chemotherapy after a treatment-free interval was associated with a modest improvement in disease-free survival. For 1212 postmenopausal patients with node-positive disease, another recently completed trial showed that three cycles of adjuvant CMF added to tamoxifen was significantly more effective at preventing disease recurrence than tamoxifen alone.³ This trial also showed that late introduction of CMF after the initiation of tamoxifen therapy was detrimental, especially for patients with estrogen-receptor-negative tumors. This finding provides clinical evidence supporting the *in vitro* experimental data showing an antagonism between tamoxifen and specific cytotoxic agents; it also suggests that the common practice to start tamoxifen prior to a chemotherapy consult for postmenopausal patients could be dangerous.

Approximately eighty percent of the patients included in the two clinical trials summarized above had self-reported quality-of-life (QL) assessments obtained. Patients completed linear-analogue self-assessment scales for physical well-being, mood, appetite and effort to cope with the disease. Assessments were obtained at randomization and then at 3-month intervals throughout treatment and follow-up. We found that baseline QL scores were worse in patients with higher numbers of involved axillary nodes and those with estrogen-receptor negative tumors. During adjuvant therapy, all treatment groups showed substantial improvement in QL scores. Patterns of QL scores closely reflected presence, duration and timing of cytotoxic treatment. Longer initial cytotoxic therapy delayed improvement in QL. Later cytotoxic therapy had transient adverse effects. Anticipation of future therapy also affected QL scores. Overall, chemotherapy had a measurable adverse effect on QL, but this was transient and minor compared to patients' adaptation following diagnosis and surgery. The absence of a lasting loss of QL should encourage patients and doctors to choose appropriate therapy with less concern for initial toxicity. This research using patient self-assessments is in press in *The Lancet*⁴ and represents an important contribution of patient-derived information to be used for treatment selection.

SPECIFIC AIM 3. To apply health decision science techniques for the analysis of some aspects of adjuvant therapy based on the data from the Third Update of the Worldwide Overview of Randomized Clinical Trials of Adjuvant Therapies for Breast Cancer.

The special sabbatical year coincided with the 1995-1996 update of the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) Overview of therapies for breast

cancer. Dr. Gelber actively participated in the meeting of the Group, held in Oxford, England, in September, 1995. Proposals were made to access the data to perform analyses based on health decision science methodologies learned during the special sabbatical training. Specifically, a Q-TWiST analysis of the data based on patient-derived preferences is envisioned. The Overview data will also be used to form the basis for a cost-effectiveness analysis on the role of adjuvant therapies for breast cancer within the context of cost-containment in health care.

Dr. Gelber's interest in methods for meta-analysis has been stimulated by active participation since 1984 in the Early Breast Cancer Clinical Trials Group (coordinated by the Clinical Trials Service Unit, Oxford University, Richard Peto - Director). This group conducted two major Overviews of randomized trials evaluating adjuvant therapies, most recently in 1992. Dr. Gelber served on the planning committee for the Overview and is a member of the Hormone Receptor Working Group chaired by C. Kent Osborne. He has published several commentaries on the overview methodology and interpretation of results.

During the last decade, the meta-analysis, or overview, has become an increasingly popular method for making treatment comparisons. This method involves combining results of several randomized trials with common treatment arms in order to increase the statistical power to detect treatment effects. The second update of the breast cancer overview was published by the EBCTCG in the Lancet in January, 1992. This work used evidence derived from over 75,000 women with operable breast cancer enrolled in randomized trials between 1957 and 1985 to demonstrate that polychemotherapy, tamoxifen, and ovarian ablation were modalities that improved recurrence-free survival and overall survival for these patients. Perhaps the three most surprising features of this remarkable international collaboration were: (1) the overwhelming effect on recurrence-free and overall survival of ovarian ablation for younger women, (2) the demonstration of a continuously increasing survival benefit for treated patients more than 5 years after adjuvant systemic therapy, and (3) the evidence that chemoendocrine treatment yields an additional benefit compared with either modality alone.

The statistical methods used in the Overview provide estimates of the relative reduction in the risk of an event for the treated group compared with the control group. While these ratio measures are appropriate for investigating whether a treatment modality has any effect at all, absolute measures are crucial for assessing the magnitude of patient benefit. Health decision sciences methods focus on the latter types of measures as they are most relevant for recommending treatments for a patient population.

The data from the Third Update of the Overview are being evaluated now by the Oxford secretariat for the primary evaluations using previously used methods. The publications should be completed by the end of 1996, after which time data access will be sought. In the meantime, data from the 1992 Overview were obtained and used to perform two Q-TWiST analyses to incorporate aspects of quality of life into meta-analysis. These analyses demonstrate the potential use of Q-TWiST for patient-oriented treatment comparisons and provide the pilot study framework for conducting cost-effectiveness analyses based on the Overview data.

In collaboration with the EBCTCG and with the permission of all collaborating trialists, a Q-TWiST meta-analysis of adjuvant chemotherapy for premenopausal breast cancer

patients was performed. The paper acknowledged the support of this Special Sabbatical Award and was published in *The Cancer Journal of Scientific American*.⁵ Data were used from 1,229 node-positive breast cancer patients 49 years of age and younger who were randomized in all eight clinical trials comparing chemotherapy (CMF) versus no adjuvant systemic therapy.

Q-TWiST analyses were performed on the individual trials to obtain DFS and OS differences between treatments. These individual trial results were combined with a multiple regression model using the median follow-up time for each trial as the independent variable, and the corresponding OS and DFS as the dependent variables. For this analysis, the health state TOX was defined as six months, since the current standard CMF regimen for this disease is six months. The resulting threshold utility plots, at five- and ten-year timepoints, were drawn. The 10-year plot showed that, for all values of the utility coefficients (patient preferences), the treated group was preferred to the control group; furthermore, for most utility coefficients, the difference was statistically significantly ($p < 0.05$). (see manuscript in the appendix)

The second Q-TWiST meta-analysis, performed during the sabbatical year, investigated the role of combination chemotherapy plus tamoxifen compared with tamoxifen alone for the treatment of postmenopausal breast cancer patients. The paper reporting the results of the analysis acknowledges the support of the Special Sabbatical Award and is in press at *The Lancet*.⁶ In collaboration with the EBCTCG and with the permission of all collaborating trialists, a Q-TWiST analysis was performed based on data from 3920 patients 50 years of age or older with node-positive breast cancer randomized in nine clinical trials comparing combination chemotherapy plus tamoxifen versus tamoxifen alone.

The methods used for this analysis were the same as those for the premenopausal evaluation described above. In contrast to the positive results obtained for the evaluation of chemotherapy in premenopausal women, the results for postmenopausal women were less impressive. Within seven years of follow-up the modest benefit of increased relapse-free and overall survival for patients who received chemotherapy just balanced the burdens in terms of acute toxic side effects. Chemotherapy-treated patients gained an average of 5.4 months of relapse-free survival and 2 months of overall survival (neither statistically significant), but had to receive cytotoxic treatment for 2 to 24 months to achieve these gains. There were no values of utility weights for time spent undergoing chemotherapy and time after relapse for which the chemotherapy plus tamoxifen program produced significantly more Q-TWiST than tamoxifen alone. The nine trials included in this meta-analysis were mostly conducted during the early 80s and many included long duration chemotherapy and short duration tamoxifen. The use of shorter duration chemotherapy combined with longer duration tamoxifen, which will be reflected much more in the future Overview Update analyses, may improve the benefit to burden ratio for chemotherapy. Nevertheless, the Q-TWiST analysis performed on the available Overview data indicated that patient preference for or against adjuvant chemotherapy should be a factor influencing treatment selection for postmenopausal breast cancer patients.

The special sabbatical provided the opportunity to develop methods for a more precise and patient-oriented analysis of the updated results from the 1995 Overview. While individual utility assessments are not available from every patient included in the

Overview, some data are available from some of the clinical trials. For example, quality-of-life data has been obtained from all patients enrolled in IBCSG trials since 1986. The multi-attribute utility theory taught in the Harvard School of Public Health courses, and the preference weight formulas being developed by Drs. Weeks and Weinstein can be used to translate these data into utility estimates for treatment toxicity and disease progression endpoints. With these patient-derived estimates of preference, the Q-TWiST meta-analyses to be conducted on the third Overview update will have even more direct relevance for patient care.

CONCLUSIONS

In summary, the overall objectives of the Special Sabbatical Award were accomplished. Dr. Gelber received much needed additional training in health decision sciences research (Specific Aim 1). He was also able to perform Q-TWiST analyses based on data from the Worldwide Overview to answer patient-oriented questions concerning the benefits of adjuvant chemotherapy for premenopausal breast cancer patients, and adjuvant chemoendocrine therapy for postmenopausal patients (Specific Aim 3). Unfortunately, funding needed to fully realize data collection from patients to develop a user-friendly multi-attribute utility function was not forthcoming (Specific Aim 2). However, plans for such a study were developed and applications to secure the needed funding were prepared. Preliminary data on the impact of chemotherapy on patients' self-report of their quality of life were analyzed in the International Breast Cancer Study Group trials. This will provide further justification to develop methods designed to incorporate patient preferences into treatment selection.

Q-TWiST is a method designed to include aspects of quality of life into an evaluation of treatment comparisons. The results of Q-TWiST analyses of randomized clinical trials can provide unbiased estimates of the relative effectiveness of two competing treatment approaches. These estimates are the values for treatment effectiveness which go into the denominator of cost-effectiveness ratios. The method highlights how the choice of patient preference weights can influence treatment selection. The next step in this approach is to design studies which elicit these patient preferences. In this way, the points of view of breast cancer patients can be incorporated into the treatment decision-making process.

REFERENCES

- 1 Gelber RD, Gelber S, Cole BF, Goldhirsch A. Integration of length-of-life outcomes and quality-of-life data in cancer clinical trials. Proceedings of the EORTC Symposium on Methodological and Statistical Issues of Quality of Life and Economic Evaluation in Cancer Clinical Trials, December 1-2, 1995, Brussels (in press).
- 2 International Breast Cancer Study Group. Duration and reintroduction of adjuvant chemotherapy for node-positive premenopausal breast cancer patients. *J Clin Oncol* 1996; (in press).
- 3 International Breast Cancer Study Group. Effectiveness of adjuvant chemotherapy in combination with tamoxifen for node-positive postmenopausal breast cancer patients. (submitted).
- 4 International Breast Cancer Study Group. Counting the costs of adjuvant therapy: impact of treatment on quality of life in node-positive patients with operable breast cancer. *Lancet* 1996;(in press).
- 5 Gelber Rd, Cole BF, Goldhirsch A, Bonadonna G, Howell A, McCardle CS, Mouridsen HT, Rubens RD, Welvaart K. Adjuvant chemotherapy for premenopausal breast cancer: a meta-analysis using quality-adjusted survival. *Cancer J Sci Am* 1995;1:114-121.
- 6 Gelber RD, Cole BF, Goldhirsch A, Rose C, Fisher B, Osborne CK, Boccardo F, Gray R, Gordon NH, Bengtsson N-O, Sevelde P. Adjuvant chemotherapy plus tamoxifen compared with tamoxifen alone for postmenopausal breast cancer: a meta-analysis using quality-adjusted survival. *The Lancet* 1996; (in press).

APPENDIX (items copied on both sides of the page)

- Paper Presented at the EORTC Symposium on Methodological and Statistical Issues of Quality of Life and Economic Evaluation in Cancer Clinical Trials.
- Grant prepared to support the development of breast cancer patient preferences for use in Q-TWiST analyses.
- Paper published in The Cancer Journal of Scientific American on the Q-TWiST meta-analysis evaluating chemotherapy for premenopausal breast cancer patients.
- Paper in press in The Lancet on the Q-TWiST meta-analysis evaluating chemotherapy plus tamoxifen versus tamoxifen alone for postmenopausal breast cancer patients.

Integration of length-of-life outcomes and quality-of-life data in cancer clinical trials

Richard D. Gelber¹, Shari Gelber², Bernard F. Cole³, Aron Goldhirsch⁴

(prepared for the EORTC symposium on methodological and statistical issues of quality of life and economic evaluation in cancer clinical trials, December 1-2, 1995, Brussels)

- 1 Harvard Medical School, Harvard School of Public Health, and Dana-Farber Cancer Institute, Boston, MA.
- 2 Frontier Science and Technology Research Foundation, Brookline, MA.
- 3 Department of Community Health and Division of Applied Mathematics, Brown University, Providence, RI.
- 4 International Breast Cancer Study Group, Ospedale Civico, Lugano, Switzerland.

Corresponding Author: Richard D Gelber, Division of Biostatistics, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, U.S.A.

Partial support for this work was provided by Grant PBR-53 from the American Cancer Society, Grant CA-06516 from the National Cancer Institute, and Grant DAMD17-94-J-4072 from the U.S. Department of the Army.

I. Introduction

The evaluation of treatments in cancer clinical trials utilizes a variety of end points, including overall survival, disease-free survival, tumor response, treatment toxicity and quality-of-life assessment. If all measured end points favor the same therapeutic regimen, then the clinical trial results are clear and patient care is positively influenced. However, it is not uncommon for end points to provide conflicting information as to the most beneficial treatment for the patient. Frequently the most efficacious cancer therapy is also the most toxic, thus making it necessary to consider the tradeoffs between an overall survival, or disease-free survival, advantage and a quality-of-life decrement.

There are two general approaches to quality-of-life (QOL) assessments: psychometric or health status assessment and the decision theory or economic approach. The first evaluates QOL using instruments which assess a patient in multiple domains (e.g., physical, psychological, social) and reports a profile of QOL on these areas, sometimes with a summary, global health score. The second approach evaluates the patient's overall health-related QOL using a single number to express the net benefit of treatment. This approach is not limited by the need to determine which domains are relevant to measure for a particular trial nor whether all the domains should be given equal importance. It is this second approach which is most commonly used to integrate length of life and quality of life.

This paper is intended to be a brief review of some key aspects relating to integrating quality and quantity of life with the purpose of raising a variety of questions to stimulate discussion.

II. Quality-adjusted Survival

Quality-adjusted survival (QAS) analysis makes treatment comparisons in terms of quality and quantity of life by penalizing treatments which have negative quality-of-life effects and rewarding those which increase survival and have other positive quality-of-life effects.¹ As in an ordinary survival analysis, the focus of the method is on time, but, rather than look at a single end point such as overall survival, or disease-free survival, multiple outcomes corresponding to changes in quality of life are considered. Periods of time with the negative side effects of treatment are weighted according to the severity of the side effects, and periods of time with disease progression are weighted according to the severity of the symptoms. A weight of zero indicates the period of time is as bad as death and a weight of one indicates perfect health. Weights between zero and one indicate degrees between these extremes. These weights are called *utility scores* or *utility coefficients*. In order for the multiplication of duration and quality to be meaningful, the value associated with a period of time must be measured in units on an interval scale. A composite measure of quality and quantity of life (i.e., quality-adjusted survival) is obtained by summing the weighted periods of time. For example, if u_i represents the utility score and s_i represents the duration for time period i , and there are k time periods over the course of a lifetime, then the quality-adjusted survival calculated over all k time periods is given by

$$QAS = \sum_{i=1}^k u_i s_i. \quad (1)$$

This summary measure can be used for treatment comparisons and for the effectiveness component in cost-effectiveness analysis.

The above QAS model makes three main assumptions about individual preferences: (1) utility independence, (2) risk neutrality, and (3) context independence. *Utility independence* implies that the utility score for each period of time does not depend on the length of the period; that is, the utility scores can be specified independently of the duration of the survival period. *Risk neutrality* implies that utility is directly proportional to the length of life. This assumption can be relaxed by discounting future gains in the model; with discounting, an extra year of life in the distant future is made less valuable than an extra year of life in the immediate future. *Context independence* means that the utility assigned to a given period of time is independent of the previous (or future) quality of life experienced before (or after) the given time period.

Assessment of the utility scores has several limitations:² (1) An individual's preferences may vary depending on circumstances. For example, a person's preferences may change if he or she is experiencing economic stress. (2) Utility scores are sometimes elicited from individuals who have not experienced the condition under study. This is usually the case when multi-attribute techniques³ are used. (3) It is difficult for some individuals to respond to questions involving risks and gambles. For example, Fowler et al.⁴ found that for people who are reluctant to give up any life at all, any questions involving risk or trading life years may be poor measures of health status. (4) Bias results when patients provide the answers that they believe are preferred by the investigator.

There is much debate regarding whose preferences should be used: patients, health care providers, relatives or the general public. Patients often adapt to their illness and their utility scores improve even though their condition remains stable. Health care providers have knowledge of the entire disease process but rarely have personal experience. Relatives are sometimes used as surrogates for patients and utility values from the general public are frequently used for health policy decisions.

QAS analyses have been performed in two different ways: the QALY method and the Q-TWiST method. In a QALY (Quality-Adjusted Life Year) analysis, specified intervals are defined prospectively based on the calendar time points at which utility assessments are to be obtained. In a Q-TWiST (Quality-adjusted Time Without Symptoms and Toxicity) analysis, the course of the patients' treatment and follow-up is partitioned into health states that are clinically relevant for the therapies under consideration, and a utility score is assigned to each clinical health state. In the next two sections, we describe these methodologies in more detail.

III. QALY Analysis

QALYs were first proposed in the medical decision analysis field.⁵ Using serial utility scores obtained from each patient, the total QALYs can be computed for an individual patient. The total QALYs is defined as the sum of the products of each assessed utility score multiplied by the duration of the time period being measured. The time period for each assessment is generally taken to begin and end at the mid-points between the scheduled sequential assessments. For example, a particular trial lasting 18 months might obtain patient derived utility scores at months 3, 9 and 15, and the values would be

applied to the six month intervals 0 to 6, 6 to 12, and 12 to 18. Thus, if the utility assessment at month 3 were 0.82, the assessment at month 9 were 0.90, and the assessment at month 15 were 0.84, the patient would have experienced 1.28 QALYs during the 1.50 year period of study⁶ (see Figure 1). Utility scores of zero are used for patients who have died. The QALYs for all patients in a given treatment arm are then averaged to provide a QALY estimate for that therapy. If complete utility assessments are available, treatment comparisons can be performed using t-tests or Wilcoxon tests as appropriate.

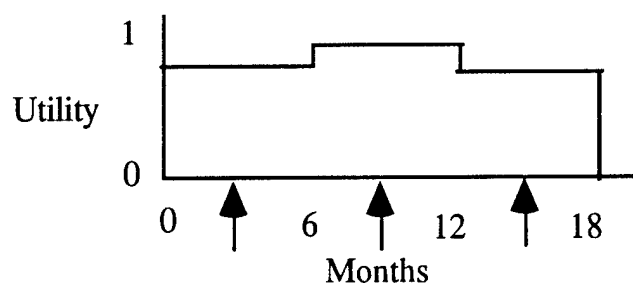


Figure 1. QALYs equal the area under the curve. The height of the curve is based on the utility assessments obtained at 3, 9, and 12 months as indicated by the arrows.

QALY analyses have not frequently been performed in clinical trials due to several limitations. First, the method requires serial utility assessments. The most accepted methods for obtaining these assessments are the standard gamble and time tradeoff,⁷ both of which require lengthy individual interviews that are often not feasible. The more commonly used cancer quality-of-life measures (EORTC QLQ-C30,⁸ FACT,⁹ FLIC¹⁰) are not currently transformable into utility scores. Utilities may be obtained, however, by using certain health status questionnaires which map health states into utility scores. The McMaster Health Utilities Indexes (HUI),³ Quality of Well-Being Scale (QWB)¹¹, York Questionnaire,^{12,13} Spitzer's Quality of Life Index¹⁴ and EuroQol¹⁵ have been linked to utility scores and several clinical trials have used these instruments to elicit utilities.

A second limitation to QALY analyses is the problem which arises from missing data in longitudinal assessments. If a patient has not furnished a utility assessment at a particular time point, then there will be incomplete data for that individual. Patients who are too sick to provide a utility would thus be eliminated from the analysis, yielding a summary QALY score reflecting the healthier patients. One possibility is to impute a value for missed assessments, but this involves making assumptions as to why the assessment was missed and what the utility score might have been.

A third limitation for calculating QALYs in clinical trials relates to the handling of censored observations. Not all patients in the trial will be followed long enough to have completed all utility assessments. In this case, it might be tempting to use Kaplan-Meier methodology¹⁶ directly on the accumulated QALY score for each patient up to his or her follow-up time. Such an estimate is biased, however, because the distribution of QALY scores and the censoring distribution (related to administrative follow-up) are not statistically independent; the analysis is subject to "informative" censoring. Patients

with low QOL weights will accumulate QALYs more slowly than patients with better QOL and, thus, will be censored earlier in the Kaplan-Meier analysis.¹⁶ Therefore, even if serial utility assessments are available within a clinical trial, methods must be developed to appropriately account for missing and censored observations.

IV. Q-TWiST

The Q-TWiST method was developed to evaluate adjuvant therapies for breast cancer^{17,18} and has been applied to treatment comparisons for rectal cancer,¹⁹ melanoma²⁰ and HIV infection.^{21,22} Instead of forming QAS estimates from each individual's history, the average amounts of time patients spend in QOL-oriented clinical health states are estimated, and these are recombined using a weighted average to form the Q-TWiST estimates. This procedure can be used with censored data without the difficulties associated with informative censoring described above. The three steps involved in applying the Q-TWiST method are reviewed briefly below.²³

- 1) The first step is to define progressive quality-of-life oriented clinical health states that highlight differences between the treatments being compared. For example, in the case of adjuvant chemotherapy for resectable breast cancer, the time with toxicity (TOX) is represented by the period in which the patient is exposed to subjective side effects of therapy; the time without symptoms of disease or toxicity of treatment (TWiST) is a state of relatively good quality of life; and time spent living with overt metastatic disease or time in relapse (REL) represents all time after the diagnosis of systemic spread of the disease. The definitions of TOX and REL reflect the fact that these periods of time have a negative impact on the overall quality of life of the patient. Furthermore, their definition is designed to emphasize the contrasting properties of the different treatments under study. Each Q-TWiST application requires a careful definition and data collection to identify individual patient transitions between the clinical health states.
- 2) The second step is to consider each treatment separately and to partition the overall survival time into the clinical health state durations using data for the transition times between the states defined in Step 1. For example, areas between the Kaplan-Meier estimates¹⁶ for overall survival (OS), disease-free survival (DFS), and time to end of toxicity (TOX) represent the average amounts of time spent in the respective clinical health states: TOX, TWiST=DFS-TOX, and REL=OS-DFS. These are calculated up to a specific point in time determined by the follow-up limits of the study cohort. The resulting estimates are called *restricted means*¹⁶ because they represent the mean survival time within the follow-up interval. The survival curves for the outcomes can be plotted on the same graph to illustrate the partitioning according to treatment group (Figure 2). This is known as a *partitioned survival analysis*.¹
- 3) The third step is to compare the treatment regimens in terms of quality-adjusted survival. This composite measure is obtained by summing the average clinical health state durations calculated in Step 2 multiplied by utility coefficients. The bootstrap method may be used to produce confidence intervals for the treatment effect (mean difference between treatment groups) on quality-adjusted survival.¹

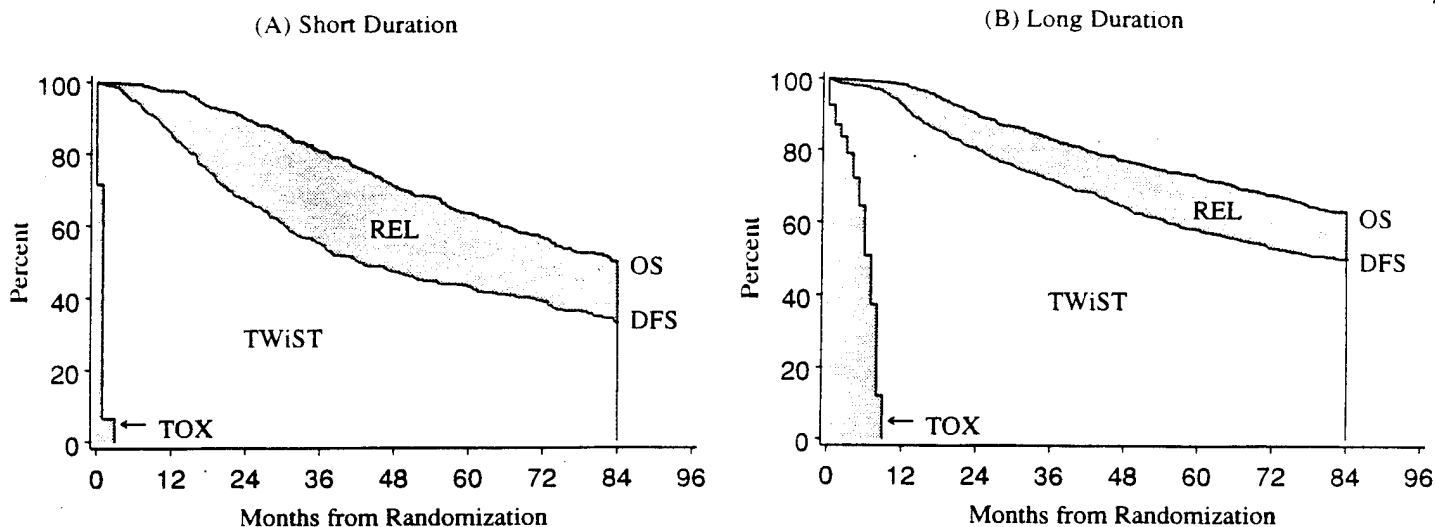


Figure 2. Partitioned survival for the long duration treatment (A) and for the short duration treatment (B) for 1,229 patients in International Breast Cancer Study Group (IBCSG) Trial V at 7 years of median follow-up.^{18,23} In each graph, the area under the overall survival curve (OS) is partitioned by the survival curves for disease-free survival (DFS) and time with treatment toxicity (TOX). The areas between the survival curves give the average months spent in TOX, TWiST and REL as indicated.

In the breast cancer example, we may suppose that the utility coefficient for TWiST is one because it is a period of relatively perfect health. On the other hand, TOX and REL represent periods of possibly poorer quality of life. Therefore, the quality-adjusted outcome can be formed by

$$Q\text{-TWiST} = u_{\text{TOX}} \times \text{TOX} + \text{TWiST} + u_{\text{REL}} \times \text{REL}, \quad (2)$$

where u_{TOX} and u_{REL} are the utility coefficients for TOX and REL, respectively. Figure 3 illustrates the weighted time periods according to assumed utility coefficients of 1.0 for TWiST, 0.75 for TOX and 0.50 for REL. This example represents a scenario in which, on average, one month spent in TOX would be exchanged for 3 weeks in TWiST and one month spent in REL would be exchanged for 2 weeks in TWiST.

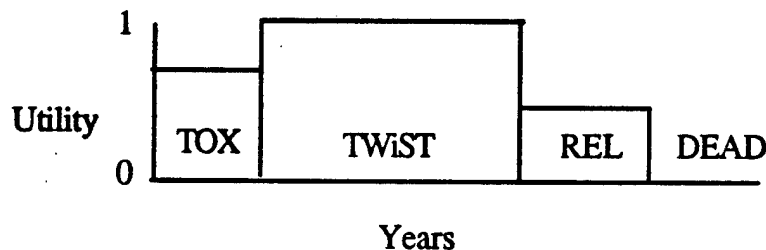
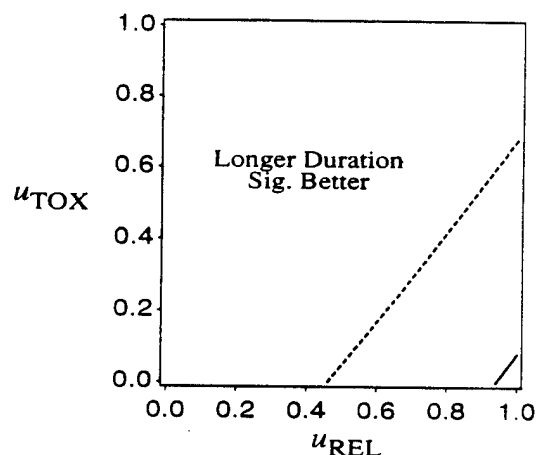


Figure 3. Components of Quality-adjusted Time Without Symptoms and Toxicity (Q-TWiST) for the example of adjuvant chemotherapy for breast cancer. The division of overall survival into TOX (subjective toxic effects), TWiST, and REL (relapse), and the weighting of these time periods using utility coefficients u_{TOX} and u_{REL} are illustrated.

Q-TWiST is considered to be the best possible clinical health state for the particular disease being studied. In most analyses it is assigned the value of one for all treatment arms. In some disease settings, however, different therapies may be less successful in returning patients to states of relatively good health and in this case Q-TWiST may be assigned a value less than one.

When patient-derived utility values are not available, the treatment comparison results are best presented as a threshold utility analysis, a sensitivity analysis of the treatment comparison ranging over all possible values of the utility coefficients. A threshold line is computed by equating the Q-TWiST formula for the two treatments (i.e., setting Q-TWiST treatment difference equal to zero) and solving for u_{TOX} in terms of u_{REL} . Figure 4 shows the threshold utility plot for the Q-TWiST treatment comparison based on the partitioned survival analyses presented in Figure 2.

Figure 4. Threshold utility analysis for 1,229 patients in IBCSG Trial V. The vertical axis shows the value of u_{TOX} , and the horizontal axis shows the value of u_{REL} . Both u_{TOX} and u_{REL} range between 0 and 1, where the value 1 indicates that the time is worth the same as Q-TWiST, while the value 0 indicates that the time is worth nothing. The solid line is the threshold (based on values of u_{TOX} and u_{REL}) for which the treatments have equal Q-TWiST. The dashed line shows the 95% confidence band for the threshold. The region denoted by "Longer Duration Sig. Better" indicates the values of utility coefficients for which average Q-TWiST at 7 years after randomization was statistically significantly greater for the long duration chemotherapy treatment compared with the short duration chemotherapy treatment.^{18,23}



Several recent Q-TWiST extensions have been developed. The gain function describes Q-TWiST treatment differences according to increasing follow-up time and enables extrapolation of results using parametric models.^{24,25} Prognostic factors can be incorporated in Q-TWiST analyses using Cox proportional hazards models.²⁶ Q-TWiST meta-analyses can also be performed using data from several randomized clinical trials.^{27,28}

In addition to the general assumptions for QAS, Q-TWiST also assumes that the defined clinical health states are progressive. The Q-TWiST model described in equation (2) also assumes that health state scores are independent from treatment assignment. If critical for making treatment comparisons, these assumptions may be satisfied by adding more clinical health states to the model. This, however, increases the complexity of the analysis and makes it more difficult to focus on fundamental tradeoffs. If a large number

of clinical health states is required to reasonably describe the major aspects of the disease and treatment course, or if these health states are not necessarily progressive, semi-Markov models can be used to estimate the health state durations or average sojourn times.²⁹ These estimates can then be multiplied by utility scores to obtain a QAS evaluation.

Q-TWiST has most frequently been criticized for using arbitrary utility weights rather than incorporating patient-derived utilities.^{30,31} There are now two ongoing clinical trials which are collecting utility scores to be used in a Q-TWiST analyses. One is a comparison of surgical techniques for the removal of colon tumors and employs Q-utilities,³² a multi-attribute utility theory translation from Spitzer's Quality-of-Life Index. The second is an evaluation of therapeutic regimens for childhood acute lymphoblastic leukemia and uses the McMaster HUI2/3.³

Q-TWiST has several advantages as a QAS methodology. Q-TWiST can be applied to previously conducted clinical trials to provide information on QOL issues even without patient-derived QOL measures. Its graphical presentations are understandable to clinicians who are familiar with (KM) survival curves. The method works with censored data. The threshold utility plots allow patients to identify which area of the threshold plot might best reflect their individual situation and thereby see which might be the treatment preferred with regard to QOL.

Kaplan Meier

V. Conclusions

The evaluation of treatment effectiveness in terms of quality of life is becoming increasingly important in clinical trials. For chronic illnesses with no cure, new treatments will need to be evaluated not only for a survival effect, but also for possible palliative advantages. In this case as well, quality-adjusted survival is the way to incorporate quality-of-life considerations into the usual assessment of treatment comparisons based on length of life.

Some have argued that attempts to integrate length of life outcomes and quality-of-life data are misguided and fundamentally wrong.³³ It is claimed that QOL and length of life should be presented as multiple end points to allow patients and physicians to combine these outcomes in ways that are meaningful for them. We believe that integrating quality and quantity of life provides more information than either outcome alone, especially if the results are presented with an emphasis on sensitivity analysis.

If, within a particular treatment comparison, it is reasonable to define progressive clinical health states, then a Q-TWiST analysis can be performed with or without patient-derived utility assessments. If such assessments are available, these can be used to estimate plausible values for the Q-TWiST utility coefficients. If frequent and complete serial assessments are available, QALY analysis would be preferred.

VI. Some Questions for Discussion

How can QAS best be used for individual patient treatment decisions? What graphical presentations would be useful?

For cost-effectiveness analysis, QALY generally refers to expected value over the entire life time. Can QALY analyses in clinical trials have meaning if restricted to specified follow-up intervals?

How can QALYs be estimated from clinical trial data subject to censoring and missing observations?

In a QALY analysis is it best to average the utilities across patients or across assessment periods?

What are the pros and cons of conducting quality-adjusted survival analyses based on clinical health state durations (i.e., Q-TWiST analysis)?

How can patient derived utilities best be included in a Q-TWiST analysis?

VII. References

- 1 Glasziou PP, Simes RJ, and Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990;9:1259-1276.
- 2 Weinstein MC, Feinberg HV, Elstein AS, et al. *Clinical Decision Analysis*. Philadelphia, PA: W.B. Saunders, 1980.
- 3 Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems: Health Utilities Index. *PharmacoEconomics* 1995;7:490-502.
- 4 Fowler FJ, Cleary PD, Massagli MP, et al. The role of reluctance to give up life in the measurement of the values of health states. *Med Decis Making* 1995;15:195-200.
- 5 Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977;296:716-721.
- 6 Feeny D, Barr RD, Furlong W, et al. Quality of life of the treatment process in pediatric oncology: an approach to measurement. In: D. Osoba (ed.) *Effect of Cancer on Quality of Life*. Boston, MA: CRC Press, 1991.
- 7 Torrance G. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;5:1-30.
- 8 Aaronson NK, Ahmedzai SA, Bergman B, et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Nat Cancer Inst* 1993;85:365-376.
- 9 Cella DF, Tulsky DS, Gray G, et al. The functional assessment of cancer therapy scale: development and validation of the general measure. *J Clin Oncol* 1993;11:570-579.
- 10 Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients: the functional living index - cancer: development and validation. *J Clin Oncol* 1984;2:472-483.
- 11 Kaplan RM, Anderson JP. An integrated approach to quality of life assessment: the general health policy model. In: B. Spilker (ed.) *Quality of Life Assessments in Clinical Trials*, New York: Raven Press, 1990.
- 12 Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? *Int J Epi* 1978;347-358.
- 13 Gudex C, Kind P. The QALY toolkit. Discussion Paper 38, Centre for Health Economics, University of York, 1988.
- 14 Spitzer WO, Dobson AJ, Hall L, et al. Measuring the quality of life of cancer patients. *J Chron Dis* 1981;34:585-597.
- 15 EuroQol Group. EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;16:199-208.

- 16 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Amer Stat Ass* 1958;54:457-481.
- 17 Goldhirsch A, Gelber RD, Simes RJ, et al. for the Ludwig Breast Cancer Study Group. Costs and benefits of adjuvant therapy in breast cancer: a quality adjusted survival analysis. *J Clin Oncol* 1989;7:36-44.
- 18 Gelber RD, Goldhirsch A, Cavalli F for the International Breast Cancer Study Group. Quality-of-life-adjusted evaluation of a randomized trial comparing adjuvant therapies for operable breast cancer. *Ann Intern Med* 1991;114:621-628.
- 19 Gelber RD, Goldhirsch A, Cole BF, et al. Costs and Benefits of adjuvant radiochemotherapy for resectable rectal cancer: a quality-of-life evaluation using a Q-TWiST analysis. 1995 (submitted)
- 20 Cole BF, Gelber RD, Kirkwood JM, et al. A quality-of-life adjusted survival analysis of Interferon Alfa-2b adjuvant treatment for high-risk resected cutaneous melanoma: an ECOG Study (E1684). 1995 (submitted)
- 21 Gelber RD, Lenderking WR, Cotton DJ, et al. for the AIDS Clinical Trials Group. Quality-of-life evaluation in a clinical trial of zidovudine therapy in patients with mildly symptomatic HIV infection. *Ann Intern Med* 1992;116:961-966.
- 22 Lenderking WR, Gelber RD, Cotton DJ, et al. Evaluation of the quality-of-life assessment in asymptomatic human immunodeficiency virus infection. *N Engl J Med* 1994;330:738-743.
- 23 Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *Amer Stat* 1995;49:161-169.
- 24 Gelber RD, Goldhirsch A, Cole BF. Parametric extrapolation of survival estimates to quality-of-life evaluation of treatments. *Control Clin Trials* 1993;14:485-489.
- 25 Cole BF, Gelber RD, Anderson KM for the International Breast Cancer Study Group. Parametric approaches to quality-adjusted survival analysis. *Biometrics* 1994;50:621-631.
- 26 Cole BF, Gelber RD, Goldhirsch A for the International Breast Cancer Study Group. Cox regression models for quality-adjusted survival analysis. *Stat Med* 1993;12:975-987.
- 27 Cole BF, Gelber RD, Goldhirsch A for the International Breast Cancer Study Group. A quality-adjusted survival meta-analysis of adjuvant chemotherapy for premenopausal breast cancer. *Stat Med* 1995;14:1771-1784.
- 28 Gelber RD, Cole BF, Goldhirsch A, et al. Adjuvant chemotherapy for premenopausal breast cancer: a meta-analysis using quality-adjusted survival. *Cancer J Sci Amer* 1995;1:113-121.
- 29 Lagakos SW, Sommer CJ, Zelen M. Semi-Markov models for partially censored data. *Biometrika* 1978;65:311-317.
- 30 Kaplan RM. Quality of life assessment for cost/utility studies in cancer. *Cancer Treat Rev* 1993;19(Suppl A):85-89.

- 31 Selby P. Measuring the quality of life of patients with cancer. In: SR Walker, RM Rosser (eds.) *Quality of Life Assessment: Key Issues for the 1990's*. Boston, MA: Kluwer Academic Publishers, 1993.
- 32 Weeks J, O'Leary J, Fairclough D, et al. The 'Q-tility Index': a new tool for assessing health-related quality of life and utilities in clinical trials and clinical practice. *Proceedings of ASCO* 1994;13:436.
- 33 Cox DR, Fitzpatrick R, Fletcher AE, et al. Quality of life assessment: can we keep it simple? *J Royal Stat Society, Part A*, 1992;155:353-393.

Weeks, J.

Last name, First Initial of Principal Investigator

Application category: ☐ Basic Research ☒ Clinical Research

**SUSAN G. KOMEN BREAST CANCER FOUNDATION
REQUEST FOR FUNDING
FOR BREAST CANCER RESEARCH PROJECT**

Principal Investigator:

JANE C. WEEKS, M.D.

Director, Center for Outcomes and Policy Research

Institution:

DANA-FARBER CANCER INSTITUTE

Address:

44 Binney Street, Bldg 454 Brkl, Suite 23

Boston, Mass 02115

Phone:

(617) 632-2509

FAX:

(617) 632-2270

Total amount requested:

119,926

Signature/title of approving
institutional personnel:

Mary E. McElroy

Mary E. McElroy, Director, Grants and Contracts

Title of Project: "Incorporating Patient Preferences in the Evaluation of
Adjuvant Therapies"

Please provide a short summary paragraph suitable for release to the general public if this application is chosen for funding.

A quality-adjusted survival analysis, Q-TWiST, will evaluate the role of chemotherapy for postmenopausal breast cancer patients based on the 1995 worldwide meta-analysis of clinical trials. Using several measures of patient derived preferences, benefits in terms of improved survival will be balanced against the costs of toxic side effects. This will facilitate patients' and their physicians' interpretation of the meta-analysis results for treatment selection.

(Photocopies of this form are acceptable.)

Grant #R10195

Incorporating Patients' Preferences in the Evaluation of Adjuvant Therapies

1.0 INTRODUCTION

The 1990 update of the Early Breast Cancer Trialists' Cooperative Group (EBCTCG) overview reported important results from a meta-analysis of all randomized clinical trials evaluating adjuvant therapies for breast cancer patients¹. The results of this analysis contributed to the adoption of tamoxifen as the standard adjuvant therapy for postmenopausal women. The addition of chemotherapy to tamoxifen did not significantly improve survival for those patients in the 1990 overview. Thus an advantage for using adjuvant chemotherapy for postmenopausal breast cancer patients was not clearly demonstrated. The next update of the overview is now underway, and the results are scheduled for presentation to the clinical trialists in September, 1995. These results are again expected to have a profound influence on breast cancer therapeutic recommendations. The proper interpretation of the 1995 overview update will require evaluation of quantity and quality of life.

The primary purpose of the proposed study is to incorporate patient preferences in a quality-adjusted survival analysis, Q-TWiST (Quality-adjusted Time Without Symptoms of disease and Toxicity of treatment), using the 1995 overview database and focusing on the issue of chemotherapy treatment for postmenopausal patients. Specifically, we propose to address the tradeoff between increased toxicity associated with chemotherapy and possible benefits in terms of disease-free and overall survival in an effort to provide a guide for treatment decision making.

2.0 SPECIFIC AIMS

- 1) To perform a Q-TWiST threshold utility analysis on the 1995 overview database designed to estimate the quality-adjusted survival provided by adjuvant chemotherapy for postmenopausal women with breast cancer.
- 2) To perform a comprehensive survey comparing five methods for eliciting patient-derived preferences and utilities during the course of patients' treatment and disease.
- 3) To incorporate patient-derived preferences in Q-TWiST analyses in order to assist women and their physicians with interpretation of the overview results based on considerations of quantity and quality of life.

3.0 BACKGROUND AND SIGNIFICANCE

The 1990 overview, based on 3932 patients fifty years of age and older, reported that chemotherapy added to tamoxifen did not significantly improve overall survival, although there was some evidence suggesting an improvement in disease-free survival¹. The absolute benefit in overall survival was estimated to be less than four percent at ten years for node-negative patients and less than six percent at ten years for node-positive patients. Therefore the use of adjuvant chemotherapy for postmenopausal breast cancer patients has remained controversial. Based on the positive results of individual studies published since the 1990 report, it is quite likely that the 1995 overview update will show a statistically significant, but clinically modest improvement in overall survival with the addition of chemotherapy to tamoxifen. This outcome would heighten the debate concerning whether statistical significance in this setting implies clinical significance in terms of the usefulness of chemotherapy to improve net patient benefit. The use of chemotherapy requires consideration of the tradeoffs between the potential for modest gains in overall survival against the decrease in quality of life (QOL). Quality-adjusted survival analyses of chemotherapy data, such as Q-TWiST, are designed to highlight these tradeoffs.

A quality-adjusted survival analysis by definition requires an evaluation of patients' QOL, but there is disagreement as to what type of QOL assessment is most appropriate. Although some investigators believe that QOL scales based on descriptive health state measures, such as the Breast Cancer Chemotherapy Questionnaire, are valid for use as weights in calculating quality-adjusted survival, the dominant view is that the health-state weights should be based on preferences². These preferences, called utilities, are measured by comparison with an external metric, such as time or risk. Utilities are conventionally measured on a scale from 0 to 1, with 0 representing death and 1, perfect health. Utilities are appealing measures of QOL for several reasons. They are global measures reflecting the respondent's overall health. Also, because of the way they are defined, the product of a utility and time is a meaningful number. This allows the calculation of quality-adjusted life expectancy.

There are two basic approaches to utility assessment: direct, holistic utility assessment and indirect weights. In the holistic approach, patients are asked a time-tradeoff, standard gamble, or rating scale question to determine how they value their current QOL. A time-tradeoff question asks respondents to identify the minimum length of life in perfect health that they would be willing to accept in exchange for a longer length of life in their current state of health. The ratio of these two life expectancies at which they are indifferent is the utility for the respondents' current state of health. A standard gamble utility is the probability of death in a gamble between excellent health and death, such that respondents are indifferent between the gamble and their current state of health. There are practical problems in using these measures, such as the difficulty of administering these complex questions to diverse respondents in an inexpensive manner. The rating scale item, which asks patients to value their current state of health between 0 and 100, is easily administered in the clinical trial setting, but has questionable theoretical standing as a utility measure. There is also a theoretical debate about the ideal timing of assessment. Patients who are currently experiencing the health state are most knowledgeable about QOL in that state, but may have accommodated to the disease and may have biases regarding the treatment choice. Respondents who have not yet experienced the state are better able to weigh one hypothetical outcome against another, as a decision maker must, but may not fully understand the disease experience.

Indirect weights derived from health-state classification systems offer an appealing alternative. One example of a generic system for which preference weights have been obtained is the Health Utilities Index (HUI), developed at McMaster University². Preference weights in the HUI are based on multi-attribute utility theory where the domains of the classification system are regarded as attributes in a utility function³. Health status domains included in the Mark II/III version of the HUI are vision, hearing, speech, emotion, pain/discomfort, ambulation/mobility, dexterity, cognition, self-care and fertility. We will also evaluate a cancer-specific classification system, the Q-tility Index, developed by Drs. Jane Weeks and Milton Weinstein⁴, and based on the five-item Quality of Life Index designed by Dr. Walter Spitzer⁵. As in the HUI, preference weights for the Q-tility Index are based on multi-attribute utility theory.

Regardless of the outcome of the 1995 overview update, the issue of toxicity associated with chemotherapy will remain a major consideration for patients and physicians when making treatment decisions for postmenopausal breast cancer. This will be especially true if the overall survival advantage for chemotherapy is modest. In this case, patients and physicians will still be left with the major question: "Is the survival benefit worth the side effects of chemotherapy?" This question can only be answered by incorporating QOL considerations into the treatment evaluation. By making use of the overview data, along with patient-derived preference data, we will be able to estimate the percent of patients who will benefit from chemotherapy in terms of quality-adjusted survival time and the percent of patients who will not. We believe this information will be critically important for making treatment decisions.

4.0 PRELIMINARY STUDIES

A. Q-TWiST. Drs. Richard Gelber, Bernard Cole, and colleagues pioneered the Q-TWiST methodology, which is designed to highlight tradeoffs from clinical trials incorporating QOL considerations for treatment decision making⁶. It has been used extensively in the evaluation of adjuvant therapies for operable breast cancer. In this setting, the benefits associated with a delay in the time to recurrence of symptomatic breast cancer have to be balanced against the side effects of treatment. Q-TWiST is currently being applied to evaluate treatments for rectal cancer, colon cancer, melanoma, myeloma, childhood acute lymphoblastic leukemia, HIV infection, and multiple sclerosis. Q-TWiST has been modified for use in meta-analyses and has been applied to evaluate data from the 1990 breast cancer overview⁷. We will use this innovative methodology in the current proposal.

In the adjuvant breast cancer setting three QOL-oriented clinical health states have been defined to highlight the differences between the regimens being compared. These are time spent with treatment toxicity (TOX), time without symptoms of the disease and without toxicity of treatment (TWiST), and time following

the diagnosis of systemic spread of the disease or relapse (REL). TWiST represents a period of relatively uncompromised QOL reflecting the best QOL available for the study patients and is assigned a utility weight of one. Utility coefficients assigned to the other two clinical health states reflect the average values of time spent in each state relative to TWiST. The Q-TWiST outcome is calculated as the weighted sum of the clinical health state durations and the utility coefficients according to the following equation: $Q\text{-TWiST} = u_{\text{TOX}} \times \text{TOX} + \text{TWiST} + u_{\text{REL}} \times \text{REL}$. To illustrate the method, we describe a Q-TWiST analysis of chemotherapy versus no adjuvant treatment for 1229 premenopausal patients based on data from the 1990 overview. The clinical health state durations, TOX, TWiST, and REL are estimated by partitioning the overall survival time of patients assigned to the treatment options being compared. Figure 1 shows this partitioning of overall survival for the patients in the chemotherapy group⁷. Areas between the curves provide estimates of TOX, TWiST, and REL which are substituted into the equation separately for each treatment group. The effectiveness of chemotherapy versus no chemotherapy is estimated by computing differences in Q-TWiST for any pair of values of the utility coefficients. The influence of patient preferences on treatment choice can be examined by a

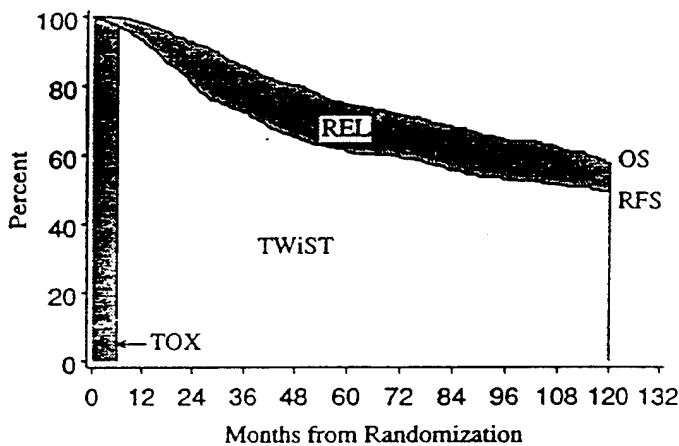


Figure 1: Partitioned Survival for Chemotherapy Group

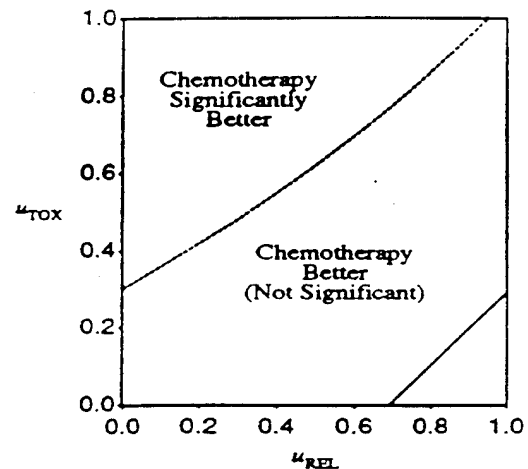


Figure 2: Threshold Utility Plot

sensitivity analysis, called a threshold utility analysis (see Figure 2), which displays treatment comparisons for varying values of the utility coefficients⁸. The solid line represents the set of utility coefficients for which the treatment groups have equal Q-TWiST after 60 months of median follow-up. The area above the solid line represents utility coefficients for which the chemotherapy group has more Q-TWiST, while the area below the solid line has more Q-TWiST for the no chemotherapy group. As illustrated in Figure 2, adjuvant chemotherapy for premenopausal patients is better than no chemotherapy for a wide range of utility coefficients. To date, Q-TWiST analyses have been performed using threshold utility plots to define the range of utility coefficients for which one treatment would be preferred to another according to data on efficacy obtained from clinical trials. The next step, to be provided by this research proposal, is to obtain and incorporate patient-derived utility coefficients for the Q-TWiST analysis.

B. Utility Assessment. Drs. Weeks and Weinstein have been developing a cancer-specific multi-attribute utility system, the "Q-tility Index"⁴. This system employs the Quality of Life Index (QLI)⁵ as the health state classification instrument. The QLI is a brief (5 item), self-administered, cancer-specific QOL questionnaire that has been shown to provide valid and reliable assessments of QOL in all stages of illness. It assesses QOL in five domains with three levels of functioning in each domain: work, self-care, general health, social support and outlook. Therefore, it classifies respondents into 243 unique health states. The Q-tility Index represents a mapping formula which assigns a utility to each of those 243 states reflecting the preference weights supplied by a reference population. In contrast to the HUI which used a general population as the reference population.

relatives (and close friends) of cancer patients were selected to provide preference weights for the Q-tility Index. Using relatives and close friends of cancer patients is advantageous because these individuals are more knowledgeable than members of the general population about cancer treatment and recovery, yet are still in a position to provide weights necessary to generate comparative ratings of a variety of hypothetical health states as required in the multi-attribute approach. The ability of this model to predict elicited utilities for composite states among respondents represents a significant improvement over the prior work in this field³.

A recent study⁹ compared the Q-tility score to a holistic utility derived from a time-tradeoff question. An in-person, interviewer-administered time-tradeoff question using visual aids resulted in a relatively high rate of respondent confusion despite ideal conditions. This study confirmed anecdotal data suggesting that these measures are not practical for use in clinical trial or population survey settings. The current proposal will survey postmenopausal breast cancer patients to ascertain patient-derived utilities for use in the proposed Q-TWiST analysis and to establish recommended procedures for eliciting patient preferences for future clinical situations.

5.0 RESEARCH DESIGN AND METHODS

Efficacy Data: The proposed study will use data from the EBCTCG 1995 overview of all randomized clinical trials on disease-free survival and overall survival of postmenopausal breast cancer patients (50 years of age and older) who were enrolled in trials comparing chemotherapy plus tamoxifen versus tamoxifen alone. In addition, individual patient covariates will be available to stratify the analysis according to age, estrogen receptor status, and nodal status. Dr. Gelber has been an active participant in the breast cancer overview collaboration since 1984. He has obtained permission from the EBCTCG to conduct Q-TWiST analyses.

Survey Patient Population: A cross-sectional, convenience sample of postmenopausal breast cancer patients at the Dana-Farber Cancer Institute (DFCI) who are at various points during the course of their disease and treatment will be surveyed to elicit utilities. Potentially eligible postmenopausal breast cancer patients will be identified by study personnel using the DFCI's Clinical Research Information System, a comprehensive database of longitudinal clinical information maintained on all breast cancer patients. The study protocol will be submitted for approval by the DFCI Human Protection Committee. A letter describing the study will be sent to eligible patients prior to a scheduled clinic visit. Patients who do not wish to participate will be asked to return an enclosed self-addressed postcard. All other patients will be approached for consent at their clinic visit.

Survey Data Collection: Patient data will be collected by interview after informed consent has been obtained. The QOL information collected in patient interviews will consist of the following: a standard gamble utility question, a time-tradeoff utility question, a 0-100 rating scale question, the Mark II/III version of the HUI, the 5-item Quality of Life Index, and sociodemographic information. Interviews will be conducted by a trained interviewer on the study staff, in private, in a room adjacent to the clinic. In addition, data on disease status at the time of the interview and treatment history will be obtained from the DFCI Clinical Research Information System.

Analytic Plan for Aim 1: Quality-adjusted survival will be computed using the following formula: $Q\text{-TWiST} = u_{\text{TOX}} \times \text{TOX} + \text{TWiST} + u_{\text{REL}} \times \text{REL}$, where the utility weight for TWiST is set to one and the coefficients, u_{TOX} and u_{REL} , hold places for the values of the respective clinical health states relative to TWiST. The second term of each product is the length of time spent in each clinical health state. Mean durations for the clinical health states will be computed according to treatment group based on data from the overview. Information on duration of toxicity, time to disease progression and time to death will be used. A threshold utility analysis will be performed to illustrate the preferred treatment for all pairs of utility coefficients. Patient preference data will not be used in this analysis. Subgroup analyses based on age, nodal status and estrogen receptor status will be performed to explore the influence of these factors on treatment outcome.

Analytic Plan for Aim 2: The goal of this portion of the analysis is to estimate the differences among the five preference measurements: the standard gamble, time tradeoff, 0 to 100 rating scale, Q-tility Index, and the HUI. A pilot study (n=124) suggested that the standard deviation of the paired differences between the time-tradeoff question and rating scale item (both rescaled to a range of 0 to 1) is 0.26. Adjusting for multiple comparisons using Scheffe's method (10 pairwise comparisons among the 5 estimates) the 95% confidence intervals for mean difference between any two measures will have a half-width equal to $(2.6 \times 0.26) / \sqrt{n}$. A total of 183 surveys will be required to provide 95% confidence intervals with half-widths less than 0.05. Based on the census of newly diagnosed and follow-up cases at DFCI, over 300 patients will be eligible for study during the eight month recruitment period, and at least 200 are anticipated to participate.

Analytic Plan for Aim 3: Each patient in the survey will be classified into one of the three Q-TWiST clinical health states (TOX, TWiST, and REL) using information on disease and treatment status. To estimate utility coefficients for these states, we will average the utility estimates separately from each of the five patient-derived preference measures. These average utility values will be used to obtain five point estimates for Q-TWiST evaluations of chemotherapy versus no chemotherapy from DFCI patients according to the following formula: $Q-TWiST = u_{TOX} \times TOX + u_{TWiST} \times TWiST + u_{REL} \times REL$. The preference measures' influence on the Q-TWiST point estimate will be evaluated. The variability of patient preferences will also be used to evaluate the range of preferred treatment choices based on the Q-TWiST threshold plot obtained from Specific Aim 1. The threshold utility plot will be divided into a grid and the patient preference data will be used to define the proportion of patients falling into each region of the grid. This will indicate the most prevalent treatment selections based on DFCI patient preferences.

Timetable for Investigation

The study will be completed in one year. During the first two months of the project, we will obtain data from the meta-analysis of the 1995 EBCTCG overview and design the survey instrument and protocol. All primary data collection of patient preference estimates will take place during months 3 to 10 of the project. Data analysis will take place over the final two months of the project. During the last month of the project, we will prepare the final report as well as manuscripts for publication.

6.0 LITERATURE

1. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 relapses and 24,000 deaths among 75,000 women. *Lancet* 1992;339:1-15;71-85.
2. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;5:1-30.
3. Torrance GW, Boyle MH, and Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Op Res* 1982; 30:1043-1069.
4. Weeks J, O'Leary J, Fairclough D, et al. "The Q-tility Index": A new tool for assessing health-related quality of life and utilities in clinical trials and clinical practice. *Proceedings of ASCO* 1994;13:436.
5. Spitzer WO, Dobson AJ, and Hall J, et al. Measuring the quality of life of cancer patients. *J Chron Dis* 1981;34:585-597.
6. Gelber RD, Goldhirsch A, Cavalli F, for the International Breast Cancer Study Group. Quality-of-life-adjusted evaluation of adjuvant therapies for operable breast cancer. *Ann Int Med* 1991;114:621-628.
7. Gelber RD, Cole BF, and Goldhirsch A. Adjuvant chemotherapy for premenopausal breast cancer: A meta-analysis using quality-adjusted survival. *Cancer J Sci Am* 1995;1:114-121.
8. Glasziou PP, Simes RJ, and Gelber RD. Quality adjusted survival analysis. *Stat in Med* 1990;9:1259-1276.
9. O'Leary JF, Fairclough DL, Jankowski MK, and Weeks JC. Comparison of time-tradeoff utilities and rating scale values in cancer patients and their relatives: Evidence for a possible plateau relationship. *Med Dec Making* 1995;15:132-137.

Adjuvant Chemotherapy for Premenopausal Breast Cancer: A Meta-Analysis Using Quality-Adjusted Survival

Richard D. Gelber, PhD,^a Boston, Massachusetts; Bernard F. Cole, PhD,^b Providence, Rhode Island; Aron Goldhirsch, MD,^c Lugano, Switzerland; Gianni Bonadonna, MD,^d Milan, Italy; Anthony Howell, MD,^e Manchester, U.K.; Colin S. McArdle, MD,^f Glasgow, Scotland; Henning T. Mouridsen, MD,^g Copenhagen, Denmark; Robert D. Rubens, MD,^h London, U.K.; Kees Welvaart, MD,ⁱ Leiden, The Netherlands

PURPOSE

Adjuvant chemotherapy for early breast cancer has been shown to offer an improvement in recurrence-free and overall survival, especially for younger women, but the acute toxic effects of this treatment discourage some physicians from prescribing it. The purpose of this analysis was to determine whether the benefit of 6 months of adjuvant CMF (cyclophosphamide, methotrexate, fluorouracil) treatment outweighs its costs in terms of toxic effects.

METHODS

A meta-analysis of quality-adjusted survival was performed based on data from 1229 patients, aged 49 years or younger, randomized in eight trials comparing CMF versus no adjuvant systemic therapy. The eight trials were included in the worldwide overview conducted by the Early Breast Cancer Trialists' Collaborative Group. The Q-TWiST method was used in a meta-analysis that provided treatment comparisons incorporating differences in quality of life associated with the amount of time patients spend with subjective toxic effects, after relapse, and without symptoms of relapse.

RESULTS

Within 6 years of follow-up evaluation for patients with node-positive disease, the benefit in terms of

increased relapse-free and overall survival balanced the costs in terms of acute toxic side effects. This was true even for the extreme case in which a zero value was assigned to all 6 months during which patients might receive adjuvant CMF chemotherapy. Within 10 years of follow-up evaluation, treated patients gained an average of 1.5 years of relapse-free survival time, almost 1 year of overall survival time, and 1 year of time without symptoms and toxicity.

CONCLUSIONS

Adjuvant chemotherapy for younger women with node-positive breast cancer provided substantial amounts of quality-adjusted survival time, even after accounting for costs associated with toxic effects of the treatment. The Q-TWiST method represents a valuable tool for comparing treatments because it incorporates patients' perceptions of their quality of life for therapeutic decision-making. (*Cancer J Sci Am* 1995; 1:114-121)

Keywords: Breast cancer, adjuvant chemotherapy, quality-adjusted survival analysis, quality of life, Q-TWiST, meta-analysis

Adjuvant chemotherapy has been shown to increase survival and relapse-free survival for patients with resectable breast cancer, especially for women 49 years of age or less.¹ For this age group the effects of ovarian abla-

tion and of tamoxifen are similar to that of polychemotherapy when indirect comparisons are made using the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) overview data. In fact, estimated typical reductions in the annual odds of relapse or prior death ($\pm$ SD) were 37% ($\pm$ 5) for polychemotherapy versus no adjuvant systemic therapy, 30% ($\pm$ 9) for ovarian ablation versus no adjuvant systemic therapy, and 27% ($\pm$ 7) for tamoxifen versus no adjuvant systemic therapy. The corresponding estimated typical reductions in the annual odds of death from any cause ($\pm$ SD) were 27% ($\pm$ 6), 28% ($\pm$ 9), and 17% ($\pm$ 10), respectively.

The many physicians who use indirect comparisons in order to select treatment for patients condemn-

From the Division of Biostatistics, Dana-Farber Cancer Institute, Boston, Massachusetts^a; Department of Community Health and Division of Applied Mathematics, Brown University, Providence, Rhode Island^b; Division of Oncology, Ospedale Civico, Lugano, Switzerland^c; Istituto Nazionale per lo Studio e la Cura dei Tumori, Milan^d; Italy; Department of Medical Oncology, Christie Hospital, Manchester, U.K.^e; University Department of Surgery, The Royal Infirmary, Glasgow, Scotland^f; Rigshospitalet, Copenhagen, Denmark^g; ICRF Clinical Oncology Unit, Guy's Hospital, London, U.K.^h; Department of Surgery, Academic Hospital, Leiden, The Netherlandsⁱ.

Received on Jan. 18, 1995; accepted for publication March 3, 1995.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

Reprint requests: Richard D. Gelber, PhD, Division of Biostatistics, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

Copyright © 1995 Scientific American Inc.

For related commentary, see page 101

chemotherapy because of quality-of-life considerations relating to its early adverse side effects. For these physicians, endocrine therapies also reduce relapse and death and therefore represent preferred therapeutic options.

No direct comparison data exist to solve the controversy concerning which treatment is "the best" from among the three modalities available for younger women. Although some clinical trials are being conducted currently to provide such data, this is not likely to be a worthwhile endeavor; instead, efforts should be made to improve upon all results.

This continuing controversy, even for the subgroup of patients for whom chemotherapy has been proven to be the most effective, motivated us to develop a tool for a cost and benefit analysis of relevance to the patient. Our evaluation of chemotherapy effectiveness in premenopausal patients was based on a meta-analysis of quality-adjusted survival.

We analyzed data from all no adjuvant systemic treatment-controlled trials of CMF (cyclophosphamide, methotrexate, fluorouracil) polychemotherapy included in the EBCTCG overview update. CMF was the most frequently assessed specific chemotherapy regimen. All trials included started before 1985. Treatment effects were estimated exclusively for patients aged 49 years or younger. We provided an additional type of analysis that incorporated aspects of quality of life into the treatment comparison.

MATERIALS AND METHODS

All data were obtained from the EBCTCG secretariat with permission of the individual trialists. The EBCTCG records were used to identify all randomized clinical trials that were included in the EBCTCG overview update¹ and that compared CMF versus no adjuvant systemic treatment in women aged 49 years or younger. From the identified trials, we obtained patient-level data for all patients meeting the eligibility criteria of the EBCTCG overview. The data from each patient included the following variables: treatment, time to relapse, and time to death or censoring. The data collection and quality control checks are described in the publication of the overview results.

The Quality-adjusted Time Without Symptoms and Toxicity (Q-TWiST) method²⁻⁴ was used to analyze each trial individually. The individual Q-TWiST results were then combined using a meta-analysis methodology⁵ in order to produce overall summary results. The Q-TWiST methodology used in this paper to analyze the individual trials has been previously described²⁻⁴ and used to evaluate treatments for breast cancer.⁶ A brief overview of the three steps involved is given here.

Step 1

A series of clinical end-points were defined, which corresponded to transitions between states of health

Table 1. Q-TWiST Meta-Analysis Trials

Trial Name	Treatment Comparison	Patient Population
INT Milan 7205 ¹⁶⁻¹⁹	12 CMF vs. nil	N+
Manchester 120 ²¹	12 CMF vs. nil	N+
DBCG 77B22-24	12 CMF + RT vs. RT	N+ or T>50mm
Glasgow ^{25,28}	12 CMF + RT vs. RT	N+
EORTC 09771 ²⁷	24 CMF vs. nil	N+
Guy's/Manch. II ²⁸⁻³⁰	12 CMF vs. nil	N+
INT Milan 8004 ^{31,32}	12 CMF vs. nil	N-; ER-
UK/Asia Collab. ³³	24 CMF vs. nil	N+

Abbreviations: N, nodal status; T, tumor; ER, estrogen receptor; RT, radiation therapy; CMF, combined chemotherapy.

during the course of follow-up evaluation for individual patients. TOX was defined as the time from randomization until the end of treatment toxicity, any relapse, or death, whichever occurred first; RFS was defined as the time to any relapse or death, whichever occurred first; and OS was defined as the time to death from any cause. The clinical health states corresponding to these end-points were: (1) time spent with subjective toxic effects of treatment (TOX), (2) time without symptoms of disease relapse or toxicity of treatment (TWiST = RFS - TOX), and (3) time following disease relapse (REL = OS - RFS).

Step 2

The clinical trial data were used to estimate the mean duration of each health state according to treatment group. For the health state TOX, individual patient level data on the duration of toxicity were not available. As a surrogate, we used 6 months as the mean TOX duration for the chemotherapy arm of each trial.

Table 2. Sample Sizes, Relapses and Deaths For Trials in the Q-TWiST Meta-Analysis

Trial Name	No. of Patients		No. of Relapses/Deaths		No. of Deaths	
	Chemo	Control	Chemo	Control	Chemo	Control
INT Milan 7205	95	75	57	55	46	48
Manchester I	21	24	15	15	12	12
DBCG 77B	149	130	69	79	60	69
Glasgow	45	35	23	24	22	23
EORTC 09771	98	95	50	52	34	38
Guy's/Manch. II	86	93	37	71	26	54
INT Milan 8004	28	29	4	12	3	9
UK/Asia Collab.	129	97	46	47	29	22
Total	651	578	301	355	232	275

Abbreviation: Chemo, chemotherapy.

Table 3. Individual Q-TWiST Results

Trial Name	Median Follow-up	Treatment Differences ^a					
		TOX	TwIST	REL	OS	RFS	Q-TWiST ^b
INT Milan 7204	174	6	23	-11	18	29	20 (2 - 38)
Manchester I	132	6	-4	-3	-1	-2	-3 (-28 - 22)
DBCG 77B	112	6	8	-9	5	14	7 (-3 - 17)
Glasgow	109	6	13	-2	17	19	15 (-3 - 33)
EORTC 09771	99	6	2	-2	6	8	4 (-4 - 12)
Guy's/Manch. II	86	6	22	-11	17	28	20 (12 - 28)
INT Milan 8004	68	6	8	-7	7	14	7 (-3 - 17)
UK/Asia Collab.	49	6	-1	-4	2	5	0 (-4 - 4)

^aMean differences in months (chemotherapy group minus control group) for the components of Q-TWiST restricted to follow-up times for the individual trials.

^bQ-TWiST differences are given for the arbitrary utility values of $u_{TOX} = u_{REL} = 0.5$, with 95% C.I. in parentheses.

Abbreviations: TOX, time spent with adverse side effects due to chemotherapy; TwIST, time without symptoms of relapse and toxicity of treatment; REL, time following disease relapse; OS, overall survival; RFS, relapse-free survival; Q-TWiST, quality-adjusted TwiST.

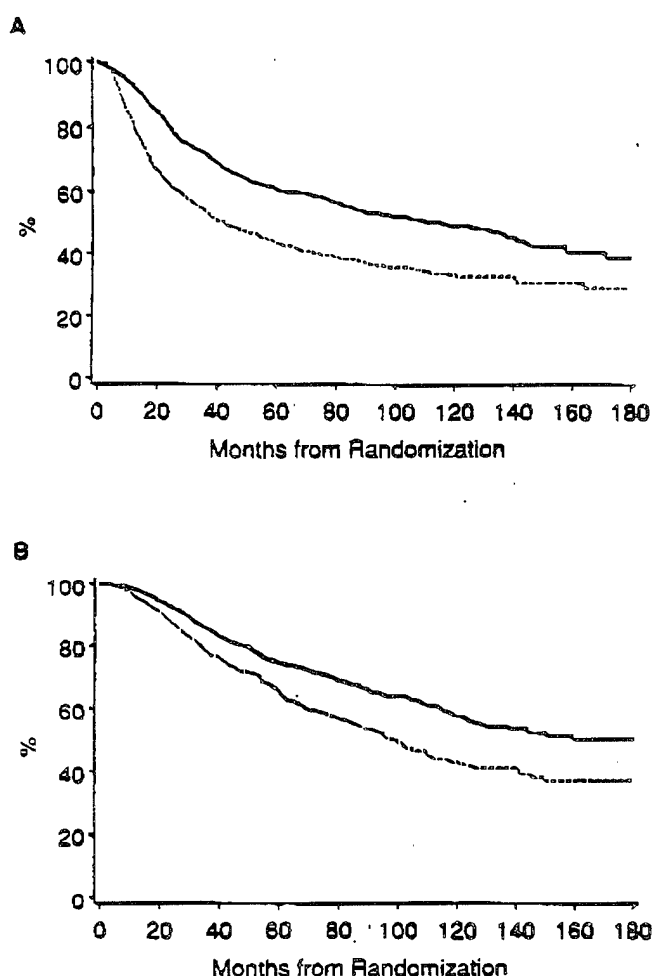


Figure 1 (A) Relapse-free survival and (B) overall survival for the entire, 1229-patient sample according to treatment group (dashed line, control group; solid line, chemotherapy group).

Although the trials tested 12 or 24 months of CMF, 6 months is today's standard duration, and has been shown to be equivalent to longer durations.^{7,8} For the control arm, no toxicity was assumed (i.e., TOX = 0). For the end-points RFS and OS, the Kaplan-Meier product limit method⁹ was used to estimate the respective transitional survival curves according to treatment group. The areas between these curves, restricted to the median follow-up duration of the trial, represented the estimated mean duration of each of the clinical health states (for example, see Fig. 2).

Step 3

Utility coefficients were introduced to reflect each health state's quality-of-life value relative to time in TwiST. The utility coefficient scales range from 0 to 1, where 0 represents a state as bad as death, 1 represents a state as good as TwiST, and values between 0 and 1 represent degrees between these extremes. Mean quality-adjusted survival (Q-TWiST) for each treatment arm was then calculated from the mean clinical health state durations as follows:

$$Q-TWiST = u_{TOX} \times TOX + TwiST + u_{REL} \times REL$$

where TOX, TwiST, and REL represent the mean health state durations, and u_{TOX} and u_{REL} denote the utility coefficients for the states TOX and REL, respectively. Q-TWiST treatment comparisons were made by subtracting the estimated mean for the control group from the estimated mean for the CMF-treated group. Variance estimates for the health state durations and Q-TWiST were also obtained. The utility coefficients used in this analysis were not estimated from patient-derived data, nor were they assumed to be any

particular set of values. Instead, the final results are displayed as a sensitivity analysis over the full range of utility coefficient values between 0 and 1. This is called a threshold utility analysis.³

Once these three steps were performed for each of the individual trials, the results were combined in a meta-analysis to provide overall estimates of quality-adjusted survival. The technical details of the meta-analysis methodology are reported elsewhere⁵; however, we give a brief description here. The individual trial results were combined using regression models. The dependent variables were mean RFS and mean OS, and the independent variable was median follow-up duration. This allowed the meta-analysis to accommodate the varying median follow-up intervals among the trials. The regression parameters were estimated by generalized least squares¹⁰ in order to weight the trials according to their respective variance estimates. Separate regression models were estimated for each treatment group.

The estimated regression models were then used to predict the mean clinical health state durations for follow-up durations of 60 and 120 months. Covariation among these estimates was also derived from the regression model. The regression analyses were performed separately for each treatment group. The results were substituted into the above equation for Q-TWiST, and the Q-TWiST treatment difference was determined for all possible values of the unknown utility coefficients in a threshold utility analysis.

The threshold utility analysis consisted of determining all pairs of utility coefficient values for which the two treatments had equal Q-TWiST (i.e., the treatment difference was zero). This provided a threshold line which separated the range of possible utility

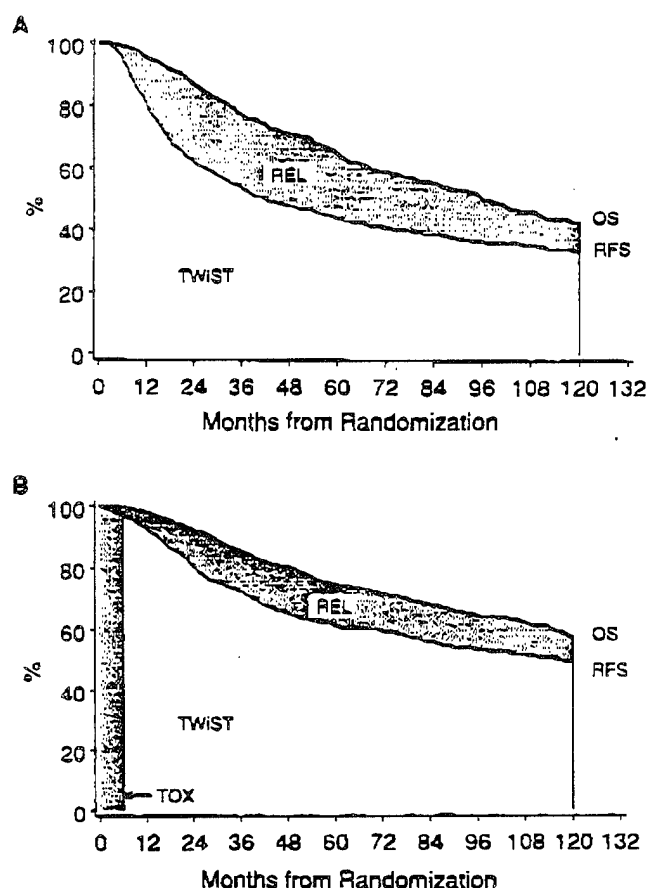


Figure 2 Partitioned survival curves for the entire, 1229-patient sample according to treatment group: (A) control and (B) chemotherapy. Areas between the curves correspond to the mean health-state durations within the follow-up interval. TOX corresponds to time spent with adverse side effects due to chemotherapy; TWiST denotes time spent without toxicity and symptoms of relapse; REL denotes the time following disease relapse.

Table 4. Overall Means in Months for the Components of Q-TWiST^a

Follow-up Duration	Group	TOX	TWiST	REL	OS	RFS	Q-TWiST $\mu_{TOX} = \mu_{REL} = 0.5$
60	Chemotherapy	6	41.7	5.2	52.9	47.7	47.3
	Control	0	37.8	10.8	48.6	37.8	43.2
	Difference	6	3.9 (2.9)	-5.6 ^b (1.9)	4.3 (2.3)	9.9 ^b (2.9)	4.1 (2.4)
120	Chemotherapy	6	75.4	12.2	93.5	81.4	84.5
	Control	0	63.0	20.5	83.6	63.0	73.3
	Difference	6	12.4 ^b (4.8)	-8.3 ^b (2.5)	9.9 ^b (4.0)	18.4 ^b (4.8)	11.2 ^b (4.2)

^aMeta-analysis results for follow-up durations of 60 and 120 months; standard errors for the treatment differences appear in parentheses.

^bTreatment difference is significantly different from zero ($P < 0.05$).

Abbreviations: TOX, time spent with adverse side effects due to chemotherapy; TWiST, time without symptoms of relapse and toxicity of treatment; REL, time following disease relapse; OS, overall survival; RFS, relapse-free survival; Q-TWiST, quality-adjusted TWiST.

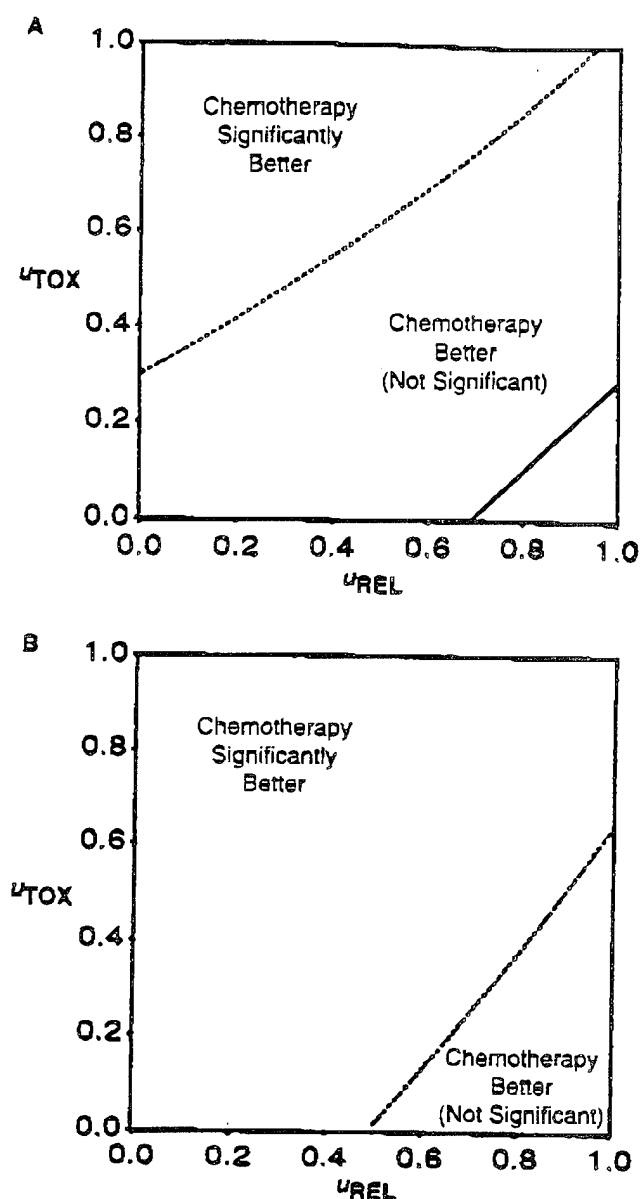


Figure 3 Threshold utility plots for (A) 60 months and (B) 120 months. The solid line in Fig. A indicates pairs of utility coefficient values for which the two treatments have equal Q-TWiST. The dashed lines are upper 95% confidence limits for the threshold line.

coefficient pairs into two regions. These regions corresponded to pairs of utility values for which the chemotherapy group provided more Q-TWiST than the control group and vice versa. Ninety-five percent confidence limits for the threshold line were also computed (for example, see Fig. 3).

RESULTS

A total of eight clinical trials that compared CMF versus no adjuvant systemic therapy in women under 50 years of age were found in the EBCTCG database (1229 patients in total). Table 1 describes the treatment comparison and patient population for each trial,

and Table 2 gives the number of patients randomized to each treatment group as well as the number of relapses and deaths on each treatment arm. Relapse-free survival and overall survival are shown in Figure 1 for the entire 1229 patient sample.

Table 3 shows the results of the individual Q-TWiST treatment comparisons performed on each clinical trial. The treatment effect (i.e., the mean for the chemotherapy group minus the mean for the control group) is shown for each of the components of Q-TWiST: RFS, OS, TWiST and REL, as well as for Q-TWiST with the arbitrary utility values of $u_{TOX} = u_{REL} = 0.5$. These values were selected to illustrate the range of Q-TWiST results across the trials and do not represent patient-derived estimates. Using these values, the Q-TWiST treatment effects for chemotherapy compared with no adjuvant therapy ranged from -3 months to +20 months. These results describe the mean amount of time gained for the chemotherapy group, up to a maximum determined by the median follow-up duration, for each of the Q-TWiST components. The median follow-up duration for each clinical trial is also shown in Table 3. These follow-up intervals ranged from 49 months to 174 months.

Table 4 presents the estimated means for the components of Q-TWiST corresponding to median follow-up durations of 60 months and 120 months based on the meta-analysis. The treatment effect (chemotherapy group minus control group) is also given for each component as well as Q-TWiST with the arbitrary utility coefficient values of $u_{TOX} = u_{REL} = 0.5$. The data in Table 4 demonstrate that within 10 years, almost 1 year of overall survival and 1.5 years of relapse-free survival can be gained by treated patients compared to controls.

Based on these results, the chemotherapy treatment provides more time in each of the clinical health states TOX and TWiST, but less time in REL. This is due to the larger effect that chemotherapy has on RFS than on OS. The treatment tends to delay the time to disease relapse more than it delays the time to death. Therefore, patients undergoing chemotherapy spend on average less time in REL as compared to patients receiving no adjuvant systemic therapy (Fig. 2).

Figure 3 displays threshold utility analyses based on the meta-analysis for follow-up durations of 60 months and 120 months. In the 60-month plot, the solid line indicates pairs of utility coefficient values for which the two treatments have equal Q-TWiST. For pairs that lie below this threshold line, the chemotherapy treatment provided less Q-TWiST than the control. Conversely, for pairs above the threshold line, the chemotherapy treatment provided more Q-TWiST than the control. The dashed line is an upper 95% confidence limit for the threshold line. For pairs of utility coefficient values that lie above the dashed line, the chemotherapy treatment provided significantly more Q-TWiST than the control. In the 120-month

analysis, the chemotherapy treatment provided more Q-TWIST than the control for all pairs of utility coefficient values (i.e., no threshold line appears on the plot). Both analyses indicate that the chemotherapy treatment provides significantly more quality-adjusted survival time for a wide range of utility coefficient pairs. For example, patients who place a fairly high value on time with treatment toxicity ($u_{\text{TOX}} = 0.8$) and a lower value on time with disease relapse ($u_{\text{REL}} = 0.6$) can expect a benefit from chemotherapy in terms of quality-adjusted survival. The larger region above the 95% confidence limit in the 120-month analysis as compared to the 60-month analysis is due to the longer median follow-up time chosen and the longer-term benefits of chemotherapy. As the median follow-up time increases, we observe the longer-term benefits of chemotherapy, and these longer-term benefits provide additional compensation for the initial 6 months of toxicity.

This effect is also illustrated by the Q-TWIST gain function shown in Figure 4. This diagram illustrates the predicted months gained for the chemotherapy group according to follow-up duration based on the meta-analysis. The dark line within the shaded region corresponds to Q-TWIST with $u_{\text{TOX}} = u_{\text{REL}} = 0.5$. As follow-up time increases, the gains for the chemotherapy group associated with delayed relapse and improved survival become apparent. The shaded region in the graph represents the range of possible gains as the utility coefficients range from zero to one. Initially, the gain in Q-TWIST for the chemotherapy group is negative because of the toxicity experienced early in follow-up. For short follow-up durations (of 20 months or less), the benefits of chemotherapy have not had sufficient time to compensate for the diminished quality of life associated with initial toxicity.

DISCUSSION

There is no doubt that adjuvant chemotherapy for younger patients has a major effect on the disease and is useful for patient management. But since the controversy raised by many concerning "which adjuvant therapy is the best" will probably continue into the future, it will be important to provide a tool for comparison of such modalities based on values that reflect patient quality-of-life considerations.

We developed a method to investigate, in terms of early cost and delayed benefit, a treatment program that is known to yield a large overall survival and relapse-free survival improvement for a target population. To illustrate the method, we have chosen data from all trials for premenopausal breast cancer patients studying treatment effects of a CMF program and comparing it to a control group of no adjuvant systemic therapy. Each randomized trial provides estimates for mean relapse-free survival time (RFS) and mean overall survival time (OS) for treated and control groups. These estimates are used to obtain

the quality-of-life meta-analysis based on fitting multiple multivariate regression models for RFS and OS as functions of follow-up time. Partitioning the OS into health states of TOX, TWIST, and REL, and weighting these time periods in proportion to their consequences on quality of life, facilitates treatment comparisons that highlight the value judgments that should influence treatment choice.

The findings of this comparison demonstrate that within 10 years, almost 1 year of overall survival and 1.5 years of relapse-free survival can be gained by treated patients compared with controls (Table 4). These benefits are sufficiently large in terms of average number of months gained for the treated patients to justify the use of treatment in many cases even when subtracting time with early subjective toxicity experienced by the entire population.

The results for the current analysis were primarily obtained for patients with node-positive disease. In fact, the 10-year estimates for relapse-free and overall survival for the studied controls were 33% and 43%, respectively. Although the relative treatment effects of CMF are likely to be similar for patients with node-positive and node-negative disease (a 46% reduction in the odds of relapse and a 39% reduction in the odds of death), the magnitude of benefit from CMF in terms of the average number of quality-adjusted months gained is likely to be much smaller for patients with a lower risk of relapse. For example, the current analysis indicated that within 10 years of follow-up time, a treatment that reduced the odds of relapse by 46%

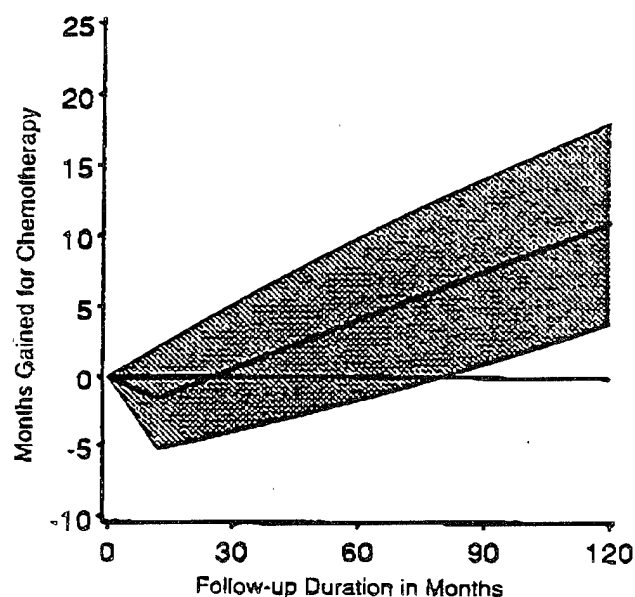


Figure 4 Q-TWIST gain function. The months of Q-TWIST gained for the chemotherapy group is plotted according to the follow-up duration. The shaded region indicates the range of months gained as the utility coefficients u_{TOX} and u_{REL} vary between zero and one. The solid line within the shaded region corresponds to Q-TWIST with $u_{\text{TOX}} = u_{\text{REL}} = 0.5$.

provided an average of 18 months more relapse-free survival time for a patient cohort with a 33% 10-year relapse-free survival. If the same treatment effect were achieved for a patient cohort with a 70% 10-year relapse-free survival, the average number of relapse-free time gained would be approximately 6 months. In this case, the average benefit might be too small to balance the initial burden of subjective toxic effects of treatment defined by the patient's judgment concerning the cost (or disutility) of 6 months of adjuvant CMF therapy.

For this analysis, the utility coefficients were chosen arbitrarily because no patient-derived information was available to indicate the "worth" of life with toxic effects of treatment or with relapse. Values of 0, 1, and 0.5 were used to illustrate treatment comparisons for different end-points and threshold utility analyses were done to determine how the coefficient values influenced the results. Thresholds were values of u_{TOX} and u_{REL} where the groups had equal Q-TWIST. In future studies, standard methods^{11,12} can be used to obtain utility values that specifically reflect patients' perceptions of their quality of life. Such a study has been performed using a time trade-off method to obtain patient preferences from 104 women who had received adjuvant CMF treatment.¹³⁻¹⁵ Patients were presented with hypothetical scenarios of the general form: "Suppose that without treatment you would live 5 years. Based on your experience of chemotherapy, what period of survival would make 6 months of initial treatment worthwhile?" Seventy-seven percent of the patients would accept the chemotherapy for as little as 12 additional months of survival time (6 years total), and 89 percent would accept it for an additional 24 months of survival. Thus, the utility associated with TOX (u_{TOX}) might be greater than 0.83 (5 years/6 years) for 77% of the patients, and greater than 0.71 (5 years/7 years) for 89% of the patients. For a utility value of 0.75 for TOX and a value of 0.50 for REL, the current meta-analysis indicates an average Q-TWIST gain of 14 months for treated patients within 10 years of follow-up time.

Perhaps the most important feature of the Q-TWIST analysis is the accumulation of gains for the treated patients over time (Fig. 3). The toxic costs of adjuvant CMF are felt early after the diagnosis of breast cancer. Even if late adverse events occur, these are not frequent enough to detract from the demonstrated benefit in terms of delayed relapse and improved survival. These benefits continue to accrue to treated patients as long as the relapse-free survival and overall survival curves for treated and control groups remain separated. Thus, early costs of adjuvant CMF have been justified by increasing long-term benefits. The Q-TWIST method is useful for illustrating the cost and benefits in comparisons between two modalities that substantially differ in terms of exposure to early toxic effects. For example, it might be useful for comparing programs of chemotherapy versus endocrine therapy versus their

combinations incorporating patients' perceptions. By specifically including a subtraction of amounts of time to account for toxic costs, the Q-TWIST method incorporates into the treatment comparison aspects of quality of life that are hidden in the usual relapse-free survival and overall survival analyses.

■ ACKNOWLEDGMENT

We thank the patients, nurses, physicians, and data managers who participated in the randomized clinical trials whose results are reported. We also gratefully acknowledge the efforts of Richard Gray, Jon Godwin, and the other members of the Early Breast Cancer Trialists' Collaborative Group Secretariat, Clinical Trials Service Unit (R. Pero, Director), Radcliffe Infirmary, Oxford University, for their kind assistance and encouragement of this project.

■ REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 relapses and 24,000 deaths among 75,000 women. *Lancet* 1992;339:1-15;71-85.
2. Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *J Natl Cancer Inst* 1989;7:36-44.
3. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990;9:1259-1276.
4. Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWIST method. *Amer Stat*, 1995; in press.
5. Cole BF, Gelber RD, Goldhirsch A. A quality-adjusted survival meta-analysis of adjuvant chemotherapy for premenopausal breast cancer. *Stat Med* 1995; in press.
6. Gelber RD, Goldhirsch A, Cavalli F. Quality-of-life adjusted evaluation of adjuvant therapies for operable breast cancer. *Ann Intern Med* 1991;114:621-628.
7. Tancini G, Bonadonna G, Valagussa P et al. Adjuvant CMF in breast cancer: comparative 5-year results of 12 versus 6 cycles. *J Clin Oncol* 1983;1:2-10.
8. Bonadonna G. Evolving concepts in the systemic adjuvant treatment of breast cancer. *Cancer Res* 1992;52:2127-2137.
9. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Statist Assoc* 1958;53:457-481.
10. Weisberg S. *Applied Linear Regression*. New York: John Wiley, 1985: 73-96.
11. Weinstein MC, Fineberg H. *Clinical Decision Analysis*. Philadelphia: WB Saunders, 1980: 253.
12. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;40:593-603.
13. Simes RJ, Cocker K, Glasziou P et al. Costs and benefits of adjuvant (adj) chemotherapy for breast cancer: an assessment of patient preferences [abstract]. *ASCO Proceedings* 1989;8:52.
14. Coates AS, Simes RJ. Patient assessment of adjuvant treatment in operable breast cancer. In: Williams CJ, ed. *Introducing new treatments for cancer: Practical, ethical and legal problems*. New York: John Wiley, 1992:447-458.
15. Gelber RD, Goldhirsch A, Hurny C, Bernhard J, Simes RJ for the International Breast Cancer Study Group. Quality of life in clinical trials of adjuvant therapies. *NCI Monographs* 1992;11:127-135.

16. Bonadonna G, Brusamolino E, Valagussa P et al. Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* 1976;294:405-410.
17. Bonadonna G, Rossi A, Valagussa P, Banfi A, Veronesi U. The CMF program for operable breast cancer with positive axillary nodes. *Cancer* 1977;39:2904-2915.
18. Rossi A, Bonadonna G, Valagussa P, Veronesi U. Multimodal treatment in operable breast cancer: five-year results of the CMF programme. *Br Med J* 1981;282:1427-1431.
19. Bonadonna G, Valagussa P, Rossi A et al. Ten-year experience with CMF-based adjuvant chemotherapy in resectable breast cancer. *Br Cancer Res Treat* 1985;5:95-115.
20. Howat JMT, Hughes R, Durning P et al. A controlled clinical trial of adjuvant chemotherapy in operable cancer of the breast. In: Salmon SE, Jones SE, eds. *Adjuvant Therapy of Cancer III*. New York: Grune and Stratton, 1981: 371-376.
21. Rubens RD, Knight RK, Fendiman IS, et al. Controlled trial of adjuvant chemotherapy with melphalan for breast cancer. *Lancet* 1983;ii:839-843.
22. Brincker H, Mouridsen HT, Anderson KW. Adjuvant chemotherapy with cyclophosphamide or CMF in premenopausal women with Stage II breast cancer. *Br Cancer Res Treat* 1983;3:91-95.
23. Mouridsen HT, Rose C, Brincker H et al. Adjuvant system therapy in high-risk breast cancer: the Danish Breast Cancer Cooperative Group's trials of cyclophosphamide or CMF in premenopausal and tamoxifen in postmenopausal patients. In: Senn H-J, ed. *Adjuvant Chemotherapy of Breast Cancer*. Berlin: Springer-Verlag, 1984:117-128.
24. Anderson M, Storm FH, Mouridsen HT. Adjuvant chemotherapy and subsequent malignant disease. *Lancet* 1991;338:885-886.
25. Smith DC, Crawford D, Dykes EH, Calman KC, Russell AR, McArdle CS. Adjuvant radiotherapy and chemotherapy in breast cancer. In: Jones SE, Salmon SE, eds. *Adjuvant Therapy of Cancer IV*. Orlando, FL: Grune and Stratton, 1984:283-289.
26. McArdle CS, Crawford D, Dykes EH et al. Adjuvant radiotherapy and chemotherapy in breast cancer. *Br J Surg* 1986;73:264-266.
27. Repelaer van Driel OJ, Welvaart K, Zwaveling A, van de Velde CJH. Influence of low-dose adjuvant CMF chemotherapy on survival and locoregional relapses in operable breast cancer. *Eur Surg Res* 1987;19:[suppl]:93.
28. Howell A, George WD, Crowther D, et al. Controlled trial of adjuvant chemotherapy with cyclophosphamide, methotrexate, and fluorouracil for breast cancer. *Lancet* 1984;ii:307-312.
29. Howell A, Rubens RD, Bush H et al. A controlled trial of adjuvant chemotherapy with melphalan versus cyclophosphamide, methotrexate, and fluorouracil for breast cancer. In: Senn H-J, ed. *Adjuvant Chemotherapy of Breast Cancer*. Berlin: Springer-Verlag, 1984:74-89.
30. Richards MA, O'Reilly SM, Howell A et al. Adjuvant cyclophosphamide, methotrexate, and fluorouracil in patients with axillary node-positive breast cancer: an update of the Guy's/Manchester trial. *J Clin Oncol* 1990;8:2032-2039.
31. Bonadonna G, Valagussa P, Tancini G et al. Current status of Milan adjuvant chemotherapy trials for node-positive and node-negative breast cancer. *NCI Monographs* 1986;1:45-49.
32. Zambetti M, Valagussa P, Brambilla C, Di Fronzo G, Silvestrini R, Bonadonna G. Adjuvant CMF in node-negative and estrogen receptor negative breast cancer: updated results. *ASCO Proceedings* 1989;8:21.
33. Senanayake F. Adjuvant hormonal chemotherapy in early breast cancer: early results from a controlled trial. *Lancet* 1984; ii:1148-1149.

- of cisplatin, fluorouracil and leucovorin for advanced non-small cell lung cancer: a randomized phase II study of the Cancer and Leukemia Group B. *Ann Oncol* 1992;3:727-732.
27. Beahrs OH, Henson DE, Hutter RVP, Kennedy BJ, eds. *Handbook for Staging of Cancer*. 4th Ed. Philadelphia: J.B. Lippincott, 1993:129-136.
28. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976;16:31-41.
29. Hoffman P, Drinkard L, Masters G et al. Ifosfamide plus Navelbine: a phase I study in advanced non-small cell lung cancer. *Proc Am Soc Clin Oncol* 1995;13:358.
30. Crawford J, O'Rourke MA, Hohnaker J, Burman S. A phase I/II trial of dose escalation of Navelbine in combination with carboplatin in patients with advanced non-small cell lung cancer [abstract]. *Proc Am Soc Clin Oncol* 1994;13:351.
31. Brooks BJ, Gralla RJ, McGaw HJ et al. Cisplatin plus vinorelbine combination chemotherapy for advanced non-small cell lung cancer testing the efficacy of a regimen designed to reduce toxicity and increase dose-intensity [abstract]. *Proc Am Soc Clin Oncol* 1994;13:348.
32. Vokes EE, Rosenberg RK, Jahanzeb M. Multicenter phase II study of oral vinorelbine for stage IV non-small cell lung cancer. *J Clin Oncol* 1995;13:637-644.
33. Adams DJ. Synergy of Navelbine-taxol combination treatment in two human breast cancer cell lines. *Proc Am Acad Cancer Res* 1994; 35:327.
34. Masters G, Drinkard L, Hoffman P et al. A phase I study of concomitant chemoradiotherapy with cisplatin and Navelbine for advanced malignancies of the chest. *Proc Am Soc Clin Oncol* 1995;13:378.

ERRATUM

✱ The authors wish to add the acknowledgments that follow to the article by Gelber et al: "Adjuvant chemotherapy in premenopausal breast cancer: a meta-analysis using quality-adjusted survival (*Cancer J Sci*

Am 1995;1:114-121)."

This research was partially supported by American Cancer Society grant PBR-53, and Department of the Army grant DAMD17-94-J-4072.

IN FUTURE ISSUES

Antiproliferative effect in vitro and antitumor effect in vivo of brefeldin A, Edward A. Sausville, Kimberly L.K. Duncan, Adrian Senderowicz, et al.

In vivo murine tumor gene delivery and therapy by systemic recombinant vaccinia virus encoding interleukin 1 β , Gary R. Peplinski, Kangla Tsung, Matthew J. Casey, et al.

Human papillomavirus DNA in tumor-free regional lymph nodes: A potential prognostic marker in cervical carcinoma, Hans Ikenberg, Irena Wiegner, Jacobus Pfisterer, et al

Endometrial pathology at dilatation and curettage in breast cancer patients: Comparison of tamoxifen users and nonusers, L.E. Gibson, R.R. Barakat, E.S. Venkatraman, W.J. Hoskins

The impact of the sequence of radiation and chemotherapy on local control after breast-conserving surgery, Beryl McCormick, Larry Norton, Ty Jyun Yao, Joachim Yahalom, Jeanne Ann Petrek

Adjuvant Chemotherapy Plus Tamoxifen Compared with Tamoxifen Alone
for Postmenopausal Breast Cancer: A Meta-Analysis Using Quality-
Adjusted Survival

Richard D. Gelber, Ph.D., Bernard F. Cole, Ph.D., Aron Goldhirsch, M.D.,
Carsten Rose, M.D., Bernard Fisher, M.D., C. Kent Osborne, M.D.,
Francesco Boccardo, M.D., Richard Gray, M.Sc., Nahida H. Gordon, Ph.D.,
Nils-Olof Bengtsson, M.D., and Paul Sevela, M.D.

Division of Biostatistics and Center for Outcomes and Policy Research, Dana-Farber
Cancer Institute, Boston (R.D.G.); Department of Community Health and Division of
Applied Mathematics, Brown University, Providence (B.F.C.); Division of Oncology,
Ospedale Civico, Lugano (A.G.); Department of Oncology R, Odense University
Hospital, Odense (C.R.); Department of Surgery, University of Pittsburgh (B.F.);
Department of Medicine, The University of Texas Health Science Center, San Antonio
(C.K.O.); Istituto Nazionale per la Ricerca Cancro, Genova (F.B.); Clinical Trial Service
Unit, Oxford (R.G.); University Hospitals of Cleveland, Cleveland (N.H.G.); Department
of Oncology, Regionsjukhuset, Umea (N-O.B.); University Hospital, Vienna (P.S.)

Correspondence to Dr. Richard D. Gelber, Division of Biostatistics, Dana-Farber Cancer
Institute, 44 Binney Street, Boston, MA 02115, USA telephone (617) 632-3603; FAX
(617) 632-2444

Keywords: Breast cancer, adjuvant chemotherapy, quality-adjusted survival analysis, quality of life, Q-TWiST, meta-analysis, overview.

Acknowledgment: We thank the patients, nurses, physicians, and data managers who participated in the randomized clinical trials whose results are reported. We also gratefully acknowledge the efforts of Mike Clarke, Jon Godwin and the other members of the Early Breast Cancer Trialists' Collaborative Group Secretariat, ICRF Clinical Trials Service Unit (R. Peto, R. Collins, co-directors), Radcliffe Infirmary, Oxford University, for their kind assistance and encouragement of this project. This work was partially supported by grants from the American Cancer Society (PBR-53), from the United States Department of the Army (DAMD17-94-J-4072) and from the National Cancer Institute (CA-06516).

Summary

Purpose: Adjuvant tamoxifen for early breast cancer has been shown to provide an improvement in relapse-free and overall survival, especially for older women. The purpose of this analysis was to determine whether the benefit of adding chemotherapy to tamoxifen outweighed its costs in terms of toxic effects for postmenopausal patients.

Methods: A meta-analysis of quality-adjusted survival was performed based on data from 3920 patients 50 years of age or older with node-positive breast cancer randomized in nine trials comparing combination chemotherapy plus tamoxifen versus tamoxifen alone. The nine trials were included in the worldwide overview conducted by the Early Breast Cancer Trialists' Collaborative Group. The Q-TWiST method, Quality-adjusted Time Without Symptoms of disease and Toxicity of treatment, was used to provide treatment comparisons incorporating differences in quality of life associated with times patients spend with subjective toxic effects, after relapse, and without symptoms of relapse or toxicity.

Results: Within seven years of follow-up the modest benefit of increased relapse-free and overall survival for patients who received chemotherapy just balanced the costs in terms of acute toxic side effects. Chemotherapy-treated patients gained an average of 5.4 months of relapse-free survival and 2 months of overall survival (neither statistically significant), but had to receive cytotoxic treatment for 2 to 24 months to achieve these gains. There were no values of utility weights for time spent undergoing chemotherapy and time after relapse for which the chemotherapy plus tamoxifen program produced significantly more Q-TWiST than tamoxifen alone.

Conclusions: Within seven years of follow-up, adjuvant chemoendocrine therapy did not provide more quality-adjusted survival time compared with tamoxifen alone for women 50 years of age or older with node-positive breast cancer. A better selection and administration of chemotherapy regimen, different scheduling of chemotherapy and tamoxifen, and appropriate use of patient and tumor characteristics might increase the therapeutic advantage for the combination. (Word count = 320)

Introduction

Adjuvant tamoxifen has been shown to increase relapse-free survival and overall survival for patients with resectable breast cancer, especially for women 50 years of age or more.¹ For this age group the effect of combination chemotherapy appeared to be less than that of tamoxifen when indirect comparisons were made using the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) overview data. In fact, the estimated typical reductions in the annual odds of relapse or prior death ($\pm$ s.d.) were 22 percent ($\pm$ 4) for combination chemotherapy versus no adjuvant systemic therapy, and 30 percent ($\pm$ 2) for tamoxifen versus no adjuvant systemic therapy. The corresponding estimated typical reductions in the annual odds of death from any cause ($\pm$ s.d.) were 14 percent ($\pm$ 5), and 19 percent ($\pm$ 3), respectively. [Note that a reduction in the annual odds of 15 percent translates into approximately 2 to 6 deaths avoided per 100 patients treated depending on baseline risk of death.^{1]}

The many physicians who use indirect comparisons in order to select treatment for patients condemn chemotherapy due to quality-of-life considerations relating to its early adverse side effects. For these physicians endocrine therapies also reduce relapse and death and therefore represent preferred therapeutic options. In fact, tamoxifen is now prescribed as adjuvant therapy for virtually all postmenopausal women with operable breast cancer.

Several studies have been conducted to test the comparison between combination chemotherapy and tamoxifen, to determine which one is "the best" for postmenopausal women.^{2,3} Some of these trials indicated that chemotherapy produced superior results for subgroups of patients with estrogen receptor negative tumors, while tamoxifen was superior for patients with estrogen receptor positive tumors. Although the competition between modalities has some interest, a more worthwhile question is whether the addition of chemotherapy to tamoxifen could improve outcome compared with tamoxifen alone. Based on data from 3920 women 50 years of age or older, the EBCTCG overview showed a reduction in annual odds of recurrence or death of

26%(±5) and a reduction in annual odds of death of 10%(±7) from the addition of combination chemotherapy to tamoxifen. Thus, chemotherapy provided a statistically significant advantage in relapse-free survival, but the effect on overall survival was not statistically significant. The question of whether the observed benefit for chemotherapy outweighs the costs in terms of toxic effects requires an answer before the regimen should be used for all postmenopausal patients who require adjuvant systemic therapy.

Our evaluation of chemotherapy effectiveness in postmenopausal patients was based on a meta-analysis of quality-adjusted survival. We analyzed data from all trials evaluating combination chemotherapy plus tamoxifen versus tamoxifen alone included in the EBCTCG overview update. Treatment effects were estimated exclusively for patients 50 years of age or older. We provided an additional type of analysis that incorporated aspects of quality of life into the treatment comparison.

Subjects and methods

All data were obtained from the Early Breast Cancer Trialists' Cooperative Group (EBCTCG) secretariat in collaboration with the individual trialists. The EBCTCG records were used to identify all randomized clinical trials that were included in the EBCTCG overview update,¹ and which compared combination chemotherapy plus tamoxifen versus tamoxifen in women 50 years old or older. From the identified trials, we obtained patient-level data for all patients meeting the eligibility criteria of the EBCTCG overview. The data from each patient included the following variables: treatment, time to relapse, and time to death or censoring. The data collection and quality control checks are described in the publication of the overview results.

The Quality-adjusted Time Without Symptoms and Toxicity (Q-TWiST) method^{4,5,6} was used to analyze each trial individually. The individual Q-TWiST results were then combined using a meta-analysis methodology⁷ in order to produce overall summary results. The Q-TWiST methodology used in this paper to analyze the

individual trials has been previously described⁴⁻⁶ and used to evaluate treatments for breast cancer.^{8,9} A brief overview of the three steps involved is given below.

Step 1: A series of clinical end points were defined which corresponded to transitions between states of health during the course of follow-up for individual patients. TOX was defined as the time from randomization until the end of treatment toxicity, any relapse or death, whichever occurred first; RFS was defined as the time to any relapse or death, whichever occurred first; and OS was defined as the time to death from any cause. The clinical health states corresponding to these end points were: (i) time spent with subjective toxic effects of treatment (TOX); (ii) time without symptoms of disease relapse or toxicity of treatment (TWiST = RFS - TOX) and (iii) time following disease relapse (REL = OS - RFS).

Step 2: The clinical trial data were used to estimate the mean duration of each health state according to treatment group. For the health state TOX, individual patient level data on the duration of toxicity were not available. As a surrogate, we used 6 months as the mean TOX duration for the chemoendocrine therapy arm of each trial. Although some trials tested longer durations of chemotherapy, 6 months is today's standard duration, and it has been shown to be equivalent to longer durations.¹⁰ For the duration of tamoxifen use in either arm, no toxicity was assumed (i.e., TOX = 0) because whatever tamoxifen-related toxicity occurred was assumed to be present in both treatment groups. For the end points RFS and OS, the Kaplan-Meier product limit method¹¹ was used to estimate the respective transitional survival curves according to treatment group. The areas between these curves, restricted to the median follow-up duration of the trial, represented the estimated mean duration of each of the clinical health states (see Figure 2).

Step 3: Utility coefficients were introduced to reflect each health state's quality-of-life value relative to time in TWiST. The utility coefficient scales range from 0 to 1, where 0 represents a state as bad as death, 1 represents a state as good as TWiST, and values between 0 and 1 represent degrees between these extremes. Mean quality-

adjusted survival (Q-TWiST) for each treatment arm was then calculated from the mean clinical health state durations as follows:

$$Q\text{-TWiST} = u_{\text{TOX}} \times \text{TOX} + \text{TWiST} + u_{\text{REL}} \times \text{REL}$$

where TOX, TWiST and REL represent the mean health state durations, and u_{TOX} and u_{REL} denote the utility coefficients for the states TOX and REL, respectively. Note that for $u_{\text{TOX}}=u_{\text{REL}}=1$, Q-TWiST equals mean OS, and for $u_{\text{TOX}}=1$, $u_{\text{REL}}=0$, Q-TWiST equals mean RFS.

Q-TWiST treatment comparisons were made by subtracting the estimated mean for the tamoxifen group from the estimated mean for the chemoendocrine therapy group. Variance estimates for the health state durations and Q-TWiST were also obtained. The utility coefficients used in this analysis were not estimated from patient-derived data, nor were they assumed to be any particular set of values. Instead, the final results are displayed as a sensitivity analysis over the full range of utility coefficient values between 0 and 1. This is called a threshold utility analysis (see Figure 3).⁵

Once these three steps were performed for each of the individual trials, the results were combined in a meta-analysis to provide overall estimates of quality-adjusted survival. The technical details of the meta-analysis methodology are summarized in the appendix and are reported in detail elsewhere.⁷ Briefly, the individual trial results were combined using regression models. The dependent variables were mean RFS and mean OS, and the independent variable was median follow-up duration. This allowed the meta-analysis to accommodate the varying median follow-up intervals among the trials. The regression parameters were estimated by generalized least squares¹² in order to weight the trials according to their respective variance estimates. Separate regression models were estimated for each treatment group. The estimated regression models were then used to predict the mean clinical health state durations for the follow-up duration of 84 months (7 years). Covariation among these estimates was also derived from the

regression model. The results were substituted into the above equation for Q-TWiST, and the Q-TWiST treatment difference was determined for all possible values of the unknown utility coefficients in a threshold utility analysis.

The threshold utility analysis consisted of determining all pairs of utility coefficient values for which the two treatments had equal Q-TWiST (i.e., the treatment difference was zero). This provided a threshold line which separated the range of possible utility coefficient pairs into two regions. These regions corresponded to pairs of utility values for which the chemoendocrine therapy group provided more Q-TWiST than the tamoxifen group and vice versa (see Figure 3). Ninety-five percent confidence limits for the threshold line were also computed.

Results

Chemoendocrine therapy with multiagent cytotoxics and tamoxifen was compared with tamoxifen alone in women 50 years of age or older within the framework of nine clinical trials included in the 1990 update of the EBCTCG database. Table 1 describes the treatment comparison and patient population for each trial^{2,13,14,15,16,17,18,19,20,21} and Table 2 gives the number of patients randomized to each treatment group as well as the number of relapses and deaths on each treatment arm. Relapse-free survival and overall survival are shown in Figure 1 for the entire population of 3920 randomized patients.

Table 3 shows the results of the individual Q-TWiST treatment comparisons performed on each clinical trial. The treatment effect (i.e., the mean for the chemoendocrine therapy group minus the mean for the tamoxifen group) is shown for each of the components of Q-TWiST: RFS, OS, TWiST and REL as well as for Q-TWiST with the arbitrary utility values of $u_{TOX} = u_{REL} = 0.5$. These utility values were selected to illustrate the range of Q-TWiST results across the trials and do not represent patient-derived estimates. Using these values, the Q-TWiST treatment effects for chemoendocrine therapy compared with tamoxifen ranged from -2 months to +7 months. These results describe the mean amount of time gained for the chemoendocrine

therapy group, up to a maximum determined by the median follow-up duration, for each of the Q-TWiST components. The median follow-up duration for each clinical trial is also shown in Table 3. These ranged from 24 months to 108 months of follow-up.

Table 4 presents the estimated means for the components of Q-TWiST corresponding to a median follow-up duration of 84 months based on the meta-analysis. The treatment effect (chemoendocrine therapy group minus tamoxifen group) is also given for each component as well as Q-TWiST with the arbitrary utility coefficient values of $u_{TOX} = u_{REL} = 0.5$.

Based on these results, the chemoendocrine therapy provides more time in the clinical health state TOX, but less time in REL. This is due to the larger effect that chemotherapy has on RFS than on OS. The amount of time spent in TWiST is essentially identical for the two treatment groups. This is based on the results restricted to 7 years of follow-up and is influenced by our selection of 6 months for the duration of TOX associated with the chemoendocrine therapy.

Subgroup analyses were also performed for two cohorts defined by estrogen receptor content of the primary tumor; 932 patients with less than 50 fmol/mg cytosol protein and 1898 patients with 50 fmol/mg cytosol protein or more. Within 84 months of follow-up, the estimated mean Q-TWiST gained for chemoendocrine therapy compared with tamoxifen alone (with $u_{TOX} = u_{REL} = 0.5$) and 95% confidence intervals were 3.4 months (-2.5 to 9.2) and -6.6 months (-12.9 to -0.2), respectively.

Figure 3 displays the threshold utility analyses based on the meta-analysis for follow-up duration of 84 months. The solid line indicates pairs of utility coefficient values for which the two treatments have equal Q-TWiST. For pairs which lie below this threshold line, the chemoendocrine therapy provided less Q-TWiST than tamoxifen. Conversely, for pairs above the threshold line, the chemoendocrine therapy provided more Q-TWiST than tamoxifen. In particular, chemoendocrine therapy provided more Q-TWiST for patients who value time following relapse lower than they value time spent with toxic effects of adjuvant chemotherapy. For example, patients who place a fairly

high value on time with treatment toxicity ($u_{TOX} = 0.8$) and a lower value on time with disease relapse ($u_{REL} = 0.6$) can expect a benefit from chemoendocrine therapy in terms of quality-adjusted survival. The threshold utility plot also allows threshold lines to be drawn to indicate the 95 percent confidence intervals for the utility coefficients that provide statistically significant differences in Q-TWiST. These are usually displayed as dashed lines on the figure.^{5,7,9} The fact that no such lines appear on Figure 3 indicates that the Q-TWiST differences are not statistically significant for any pair of utility coefficients.

The Q-TWiST gain function illustrates the predicted months gained for the chemoendocrine therapy group according to follow-up duration (Figure 4). The shaded region represents the range of Q-TWiST differences calculated for all pairs of possible utility coefficients, and the dark line within the shaded region corresponds to Q-TWiST with $u_{TOX} = u_{REL} = 0.5$. Initially, the gain in Q-TWiST for the chemoendocrine therapy group is negative because of the toxicity experienced early in follow-up. As follow-up time increases, the gains for the chemoendocrine therapy group associated with delayed relapse and improved survival become apparent. These gains are modest and not statistically significant.

Discussion

Chemoendocrine therapy consisting of cytotoxic agents combined with tamoxifen was always hypothesized to provide an additive effect of the two modalities. In the Early Breast Cancer Trialists' Cooperative Group (EBCTCG) overview update in 1990, prolonged cytotoxic chemotherapy added to tamoxifen was shown to provide statistically significant reduction in the annual odds of recurrence or death compared with tamoxifen alone for women 50 years of age or older. The benefit in terms of overall survival was not statistically significant. Because tamoxifen alone provides an overall survival advantage for these patients, this adjuvant treatment, which is associated with only minimal toxic side effects, has become the standard. If cytotoxic chemotherapy

also had virtually no toxic side effects, the benefits with respect to relapse-free survival available today might be sufficient to justify the routine addition of this modality to tamoxifen. Cytotoxic chemotherapy, however, is associated with unpleasant adverse side effects against which the benefits of using it in conjunction with tamoxifen must be weighed. It is, therefore, important to provide a tool for comparison of treatment regimens based on values reflecting patient quality-of-life considerations.

We developed a method, Q-TWiST, to investigate, in terms of early cost and delayed benefit, a treatment program which is known to yield a relapse-free survival improvement for a target population. We applied the method to data from all randomized clinical trials for postmenopausal breast cancer patients studying treatment effects of chemoendocrine therapies using different types of cytotoxic agents combined with tamoxifen and compared with tamoxifen alone for the same duration. Each trial provided estimates for mean relapse-free survival time (RFS) and mean overall survival time (OS) for treated and control groups. These estimates were used to obtain the quality-of-life meta-analysis based on fitting multiple multivariate regression models for RFS and OS as functions of follow-up time. Partitioning the OS into health states of TOX, TWiST and REL, and weighting these time periods in proportion to their consequences on quality of life facilitate treatment comparisons which highlight the value judgements that influence treatment choice.

The findings of this comparison demonstrate that within seven years, the improvement in relapse-free survival is counter-balanced by a period of cytotoxic treatment duration set at six months. While quality-adjusted survival analysis favored chemoendocrine therapy for patients who place a low value on time after relapse and are not averse to experiencing the toxic effects of chemotherapy, the gain in Q-TWiST was not statistically significant for any pair of utility coefficients.

For this analysis, the utility coefficients were chosen arbitrarily because no patient-derived information was available to indicate the "worth" of life with toxic effects of treatment or with relapse. Values of 0, 1, and 0.5 were used to illustrate

treatment comparisons for different endpoints and threshold utility analyses were done to determine how the coefficient values influenced the results. Thresholds were values of u_{TOX} and u_{REL} where the groups had equal Q-TWiST. In future studies, standard methods^{22,23} can be used to obtain utility values which specifically reflect patients' perceptions of their quality of life. Such a study has been performed using a time trade-off method to obtain patient preferences from 104 women who had received adjuvant CMF treatment.^{24,25,26} Patients were presented with hypothetical scenarios of the general form: "Suppose that without treatment you would live 5 years. Based on your experience of chemotherapy, what period of survival would make six months of initial treatment worthwhile?" Seventy-seven percent of the patients would accept the chemotherapy for as little as 12 additional months of survival time, and 89 percent would accept it for an additional 24 months of survival. Thus, the utility associated with TOX (u_{TOX}) might be greater than 0.83 (5 years/6 years) for 77 percent of the patients, and greater than 0.71 (5 years/7 years) for 89 percent of the patients. Based on these findings, for a utility value of 0.75 for TOX and a value of 0.50 for REL, the current meta-analysis indicates an average Q-TWiST gain of 2.2 months for chemoendocrine treated patients within seven years of follow-up (95% confidence interval, -4.5 to 8.9 months).

The Q-TWiST analysis takes into account the overall benefits and costs of the combined modality and the endocrine therapy alone as expressed by the partitioning (Figure 2). The administration of all cytotoxic regimens was associated with a relevant subjective toxicity which was represented in our calculations by the initial costs (TOX). The subjective and objective costs as well as the efficacy of the cytotoxic component of the chemoendocrine regimens are heterogeneous across trials and in some of the clinical trials the cytotoxic doses administered were suboptimal.^{27,28} Lower doses of chemotherapy still represent a cost in terms of subjective toxicity but are associated with significantly lower efficacy. Thus, the Q-TWiST analysis indicates that in order to be

efficient we must limit the period of TOX to the minimum while maintaining efficacy with full-dose chemotherapy.

The chemotherapy regimens were also different in terms of duration. In fact, cytotoxic regimens were given for durations which varied between 63 days and 24 months. We did not take into account such heterogeneity in duration of treatment because "standard" cytotoxic chemotherapies today indicate for postmenopausal patients a proper treatment duration to be of 3 to 6 monthly courses.^{10,29} Based on this analysis, the additional toxicity provided by the prolonged chemotherapy (CMFVP or the melphalan regimens) would do more harm than good. The longer durations of chemotherapy also increase the opportunity for a negative drug interaction between cytotoxic agents and tamoxifen. Such interaction has been postulated from data *in vivo*³⁰ and *in vitro*.³¹

In this analysis, all patients in the control groups received endocrine therapy with tamoxifen alone or combined with low-dose prednisone for durations of one year to five years. This treatment has been shown to be effective for postmenopausal patients, especially for those with elevated levels of estrogen receptor content in the primary tumor. Many of the trials included in this overview restricted enrollment to patients with estrogen receptor-positive tumors (Table 1). The magnitude of effectiveness achieved by the endocrine therapy alone is very high for these patients, leaving less opportunity for chemotherapy to further improve upon outcome.

Individual trials which use short duration, full-dose chemotherapy given prior to initiation of tamoxifen might identify an optimal chemoendocrine regimen. The use of the Q-TWiST method on overview data could quantify net patient benefit even before overall survival differences become apparent. Chemoendocrine therapy combinations remain investigational for postmenopausal women, especially for patients selected by disease and patient characteristics which predict a high level of response to endocrine therapies alone.

(Word count = 3378)

REFERENCES

- 1 Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 relapses and 24,000 deaths among 75,000 women. *Lancet* 1992;339:1-15;71-85.
- 2 Boccardo F, Rubagotti A, Bruzzi P, Cappellini M, Isola G, Nenci I, et al. for the Breast Cancer Adjuvant Chemo-Hormone Therapy Cooperative Group. Chemotherapy versus tamoxifen versus chemotherapy plus tamoxifen in node-positive, estrogen receptor-positive breast cancer patients: results of a multicentric Italian study. *J Clin Oncol* 1990;8:1310-1320.
- 3 Kaufmann M, Jonat W, Caffier H, Kreienberg R, Hilfrich J, Abel U, Maass H, Kubli F for the Gynecological Adjuvant Breast Group (GAGB). Adjuvant systemic risk adapted cytotoxic $\pm$ tamoxifen therapy in women with node positive breast cancer. In: Salmon SE, editor. *Adjuvant Therapy of Cancer V*. Orlando: Grune and Stratton, 1987: 337-346.
- 4 Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *J Clin Oncol* 1989;7:36-44.
- 5 Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990;9:1259-1276.
- 6 Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *Amer Stat* 1995;49:161-169.
- 7 Cole BF, Gelber RD, Goldhirsch A. A quality-adjusted survival meta-analysis of adjuvant chemotherapy for premenopausal breast cancer. *Stat Med* 1995;14:1771-1784.

- 8 Gelber RD, Goldhirsch A, Cavalli F. Quality-of-life adjusted evaluation of adjuvant therapies for operable breast cancer. *Ann Int Med* 1991;114:621-628.
- 9 Gelber RD, Cole BF, Goldhirsch A, Bonadonna G, Howell A, McArdle CS, Mouridsen HT, Rubens RD, Welvaart K. Adjuvant chemotherapy for premenopausal breast cancer: a meta-analysis using quality-adjusted survival. *Cancer J Sci Am* 1995;1:114-121.
- 10 Tancini G, Bonadonna G, Valagussa P, et al. Adjuvant CMF in breast cancer: comparative 5-year results of 12 versus 6 cycles. *J Clin Oncol* 1983;1:2-10.
- 11 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Statist Assoc* 1958;53:457-81.
- 12 Weisberg S. *Applied Linear Regression*. New York: John Wiley and Sons; 1985.
- 13 Ludwig Breast Cancer Study Group. Randomized trial of chemo-endocrine therapy, endocrine therapy, and mastectomy alone in postmenopausal patients with operable breast cancer and axillary node metastasis. *Lancet* 1984;i:1256-1260.
- 14 Castiglione-Gertsch M, Johnsen C, Goldhirsch A, Gelber RD, et al for the International Breast Cancer Study Group: The International (Ludwig) Breast Cancer Study Group Trials I-IV: 15 years follow-up. *Ann Oncol* 1994;5:717-724.
- 15 Larsson, L-G. Trials with adjuvant hormonal and cytotoxic treatment in operable breast cancer patients (Breast Cancer Group in Northern Sweden). EORTC - Breast Cancer Cooperative Group, 30th June - 3rd July, 1987, abstracts book, p C2D.7.
- 16 Pearson OH, Hubay CA, Gordon NH, Marshall JS, Crowe JP, Arafah BA, McGuire WL. Endocrine versus endocrine plus 5-drug chemotherapy in postmenopausal women with stage II, estrogen receptor positive disease. *Cancer* 1989;64:1819-1823.
- 17 Sevela P, Zielinski CC, Salzer H, Langer M, Kubista E, Spona J, Staffen A, Aiginger P. Combined adjuvant treatment with CMF-chemotherapy and tamoxifen in postmenopausal patients with stage II breast cancer: a prospective randomized

- study. In: Senn HJ, editor. Adjuvant Therapy of Primary Breast Cancer, St. Gallen. 1988:64 (abstract).
- 18 Rivkin SE, Green S, Metch V et al. Adjuvant CMFVP versus tamoxifen versus concurrent CMFVP and tamoxifen for postmenopausal, node-positive, and estrogen receptor-positive breast cancer patients: a Southwest Oncology Group study. *J Clin Oncol* 1994;12:2078-2085.
 - 19 Senanayake F. Adjuvant hormonal chemotherapy in early breast cancer: early results from a controlled trial (letter). *Lancet* 1984;ii:1148.
 - 20 Overgaard M, Christensen JJ, Johansen H, Nybo-Rasmussen A, Rose C, van der Kooy P, Panduro J, Laursen F, Kjær M, Sørensen NE, Gadeberg CC, Hjelm-Hansen M, Overgaard J, West Andersen K, Zedeler K. Evaluation of radiotherapy in high-risk breast cancer patients: report from the Danish Breast Cancer Cooperative Group (DBCG 82) Trial. *Int J Radiation Oncology Biol Phys* 1990;19:1121-1124.
 - 21 Fisher B., Redmond C., Legault-Poisson S., et al: Postoperative chemotherapy and tamoxifen compared with tamoxifen alone in the treatment of positive-node breast cancer patients aged 50 years and older with tumors responsive to tamoxifen: results from the National Surgical Adjuvant Breast and Bowel Project B-16. *J. Clin. Oncol* 1990; 8: 1005-1018.
 - 22 Weinstein MC, Fineberg H. *Clinical Decision Analysis*. Philadelphia:WB Saunders Company; 1980.
 - 23 Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;40:593-603.
 - 24 Simes RJ, Cocker K, Glasziou P, et al. Costs and benefits of adjuvant (adj) chemotherapy for breast cancer: an assessment of patient preferences (abstract No 201). *Proc ASCO* 1989;8:52.

- 25 Coates AS, Simes RJ. Patient assessment of adjuvant treatment in operable breast cancer. In: Williams CJ, editor. *Introducing New Treatments for Cancer: Practical, Ethical and Legal Problems*. New York: John Wiley and Sons; 1992.
- 26 Gelber RD, Goldhirsch A, Hürny C, Bernhard J, Simes RJ for the International Breast Cancer Study Group. Quality of life in clinical trials of adjuvant therapies. *JNCI Monogr* 1992;11:127-35.
- 27 Wood WC, Budman DR, Korzun AH, Cooper MR, Younger J, Hart RD, et al. Dose and dose intensity of adjuvant chemotherapy for stage II, node-positive breast carcinoma. *N Engl J Med* 1994;330:1253-1259.
- 28 Goldhirsch A, Gelber RD. Understanding adjuvant chemotherapy for breast cancer (editorial). *N Engl J Med* 1994;330:1308-1309.
- 29 Schumacher M, Bastert G, Bojar H, Huebner K, Olschewski M, Sauerbrei W, et al. for the German Breast Cancer Study Group. Randomized 2x2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. *J Clin Oncol* 1994;12:2086-93.
- 30 Fisher B, Redmond C, Brown A et al. Adjuvant chemotherapy with and without tamoxifen: five-year results from the National Surgical Adjuvant Breast and Bowel Project Trial. *J Clin Oncol* 1986;4:459-471.
- 31 Osborne C.K. Interactions of tamoxifen with cytotoxic chemotherapy for breast cancer. In Jordan VC, editor. *Long Term Tamoxifen For Breast Cancer*. The University of Wisconsin Press, 1994, pp181-198.

Table 1. Trials in the Q-TWiST Meta-Analysis Comparing Chemoendocrine Therapy versus Endocrine Therapy Alone for Postmenopausal Patients

<u>Trial Name</u>	<u>Treatment Comparison</u> ⁺	<u>Patient Population</u>
Ludwig III ^{4 13 14}	12 CMF + 12 pT vs. 12 pT	N+, all ER
N Sweden BCG ¹⁵	8 AC T ₂₄ vs. 24 T	N+, all ER
Case Western B ¹⁶	12 CMFVP T ₃₆ vs. 36 T	N+, ER+
Vienna Gyn. ¹⁷	6 CMF T ₁₂ vs. 12 T	N+, all ER
SWOG ¹⁸	12 CMFVP T ₁₂ vs. 12 T	N+, ER+
U.K. Asia ¹⁹	24 CMF T ₂₄ vs. 24 T	N+, all ER
GROCTA I Italy ²	6 -> 4 CMF -> Epidox T ₆₀ vs. 60 T	N+, ER+
Danish BCG ²⁰	8 CMF T ₁₂ vs. 12 T	N+, all ER
NSABP B16 ²¹	4 AC T ₆₀ vs. 60 T	N+, ER+ ^{**}
	17 PF* T ₆₀ vs. 60 T	
	17 PAF* T ₆₀ vs. 60 T	

⁺ Numbers refer to number of cycles of chemotherapy and number of months of tamoxifen, respectively.

^{*} From 1984 to 1985, chemotherapy included melphalan and 5-fluorouracil; from 1985 to 1988, chemotherapy included melphalan, 5-fluorouracil and adriamycin.

^{**} ER+ if < 60 years old; any ER if ≥ 60 years old.

Table 2. Sample Sizes, Relapses and Deaths for the Trials in the Q-TWiST Meta-Analysis.

<u>Trial</u>	<u>Number of Patients</u>		<u>Number of Relapses/Deaths</u>		<u>Number of Deaths</u>	
	CT + Tam	Tam	CT+ Tam	Tam	CT + Tam	Tam
Ludwig III	165	158	91	105	74	87
N Sweden BCG	35	37	19	24	16	18
Case Western B	44	47	14	22	11	19
Vienna Gyn.	30	29	16	11	12	10
SWOG 7827	301	292	94	97	78	67
UK/Asia Collab.	46	67	23	26	19	19
GROCTA I Italy	106	104	21	25	10	11
Danish BCG	602	631	193	251	142	158
NSABP B16	817	409	98	83	47	35
<u>Total</u>	<u>2,146</u>	<u>1,774</u>	<u>569</u>	<u>644</u>	<u>409</u>	<u>424</u>

CT + Tam = chemoendocrine therapy (cytotoxic chemotherapy + tamoxifen)

Table 3. Mean Differences in Months (chemoendocrine group minus tamoxifen group)
for the Components of Q-TWiST Restricted to the Median Follow-up Duration for the
Individual Trials.

	Median follow-up (months)	Treatment Differences in Months					Q-TWiST ($\mu_{TOX} = 0.5$ $\mu_{REL} = 0.5$)
		TOX	TwIST	REL	OS	RFS	
Ludwig III	108	6	8	-7	7	13	7
N Sweden BCG	84	6	4	0	9	10	7
Case Western B	72	6	2	-5	3	8	3
Vienna Gyn.	72	6	-10	1	-2	-4	-6
SWOG 7827	60	6	-5	-2	-1	1	-3
UK/Asia Collab.	60	6	-10	2	-2	-4	-6
GROCTA I Italy	48	6	-4	-1	1	2	-2
Danish BCG	36	6	-4	-2	0	2	-2
NSABP B16	24	6	-5	-1	0	1	-2

Table 4. Overall Means (in months, restricted to 7 years) Based on the Meta-Analysis.

Variable	Treatment group		Difference	95% CI
	CT + Tam	Tam		
TOX	6.0	0.0	6.0	-
TWiST	54.7	55.3	-0.6	-8.3 to 7.1
REL	8.2	11.6	-3.4	-7.9 to 1.1
OS	68.9	66.9	2.0	-3.2 to 7.2
RFS	60.7	55.3	5.4	-2.3 to 13.1
Q-TWiST	61.8	61.1	0.7	-6.0 to 7.4

($u_{TOX} = 0.5$,

$u_{REL} = 0.5$)

FIGURE LEGENDS

Figure 1. (A) Relapse-free survival and (B) overall survival for the entire 3920-patient sample according to treatment group (solid line, chemoendocrine therapy group; dashed line, tamoxifen alone group). Seven-year relapse-free survival percents were 52% and 45%, respectively, and seven-year overall survival percents were 57% and 55%, respectively.

Figure 2. Partitioned survival curves for the entire, 3920-patient sample according to treatment group: (A) tamoxifen alone and (B) chemoendocrine therapy. Areas between the curves correspond to the mean clinical health-state durations within the follow-up interval. TOX corresponds to time spent with adverse side effects due to chemotherapy; TWiST denotes time spent without symptoms of relapse and toxicity of treatment; REL denotes the time following disease relapse.

Figure 3. Threshold utility plot for 84 months of follow-up. The solid line indicates pairs of utility coefficient values for which the two treatments have equal Q-TWiST. Chemoendocrine therapy provides more Q-TWiST than tamoxifen alone for pairs of utility coefficients above the line, and less Q-TWiST for coefficients below the line. (N.S. denotes not significant.)

Figure 4. Q-TWiST gain function. The months of Q-TWiST gained for the chemoendocrine therapy group compared with tamoxifen alone is plotted according to the follow-up duration. The shaded region indicates the range of months gained as the utility coefficients (u_{TOX} and u_{REL}) vary between zero and one. The solid line within the shaded region corresponds to Q-TWiST with $u_{TOX} = u_{REL} = 0.5$.

Figure 1A

(A) Relapse-Free Survival

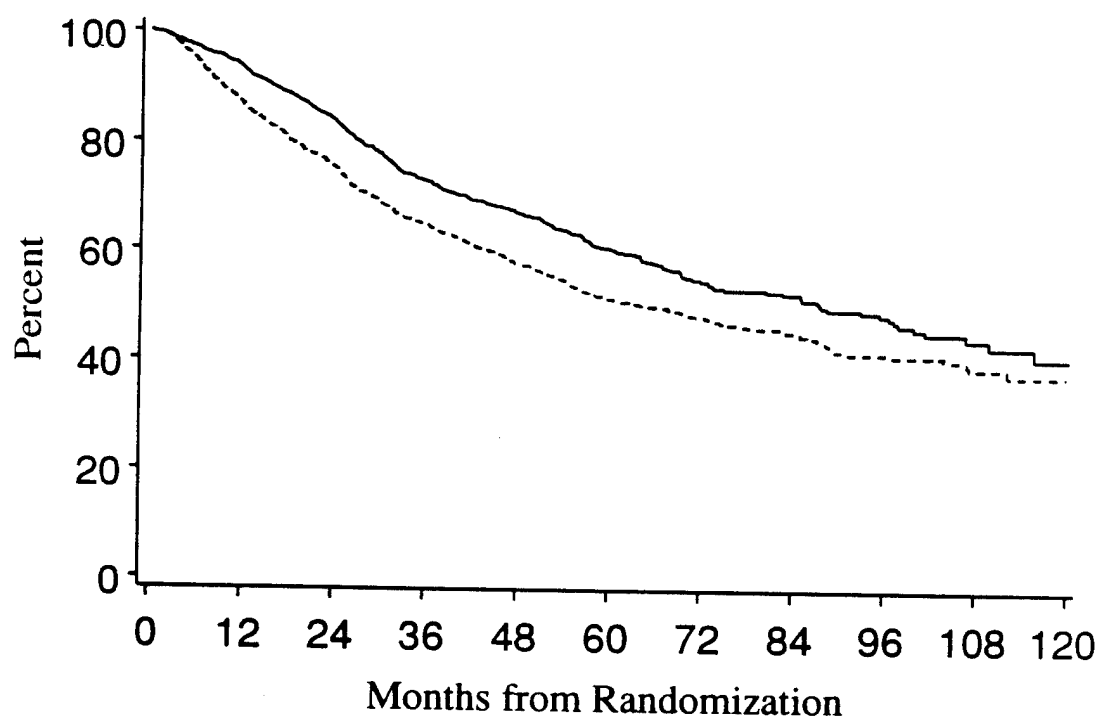


Figure 1B

(B) Overall Survival

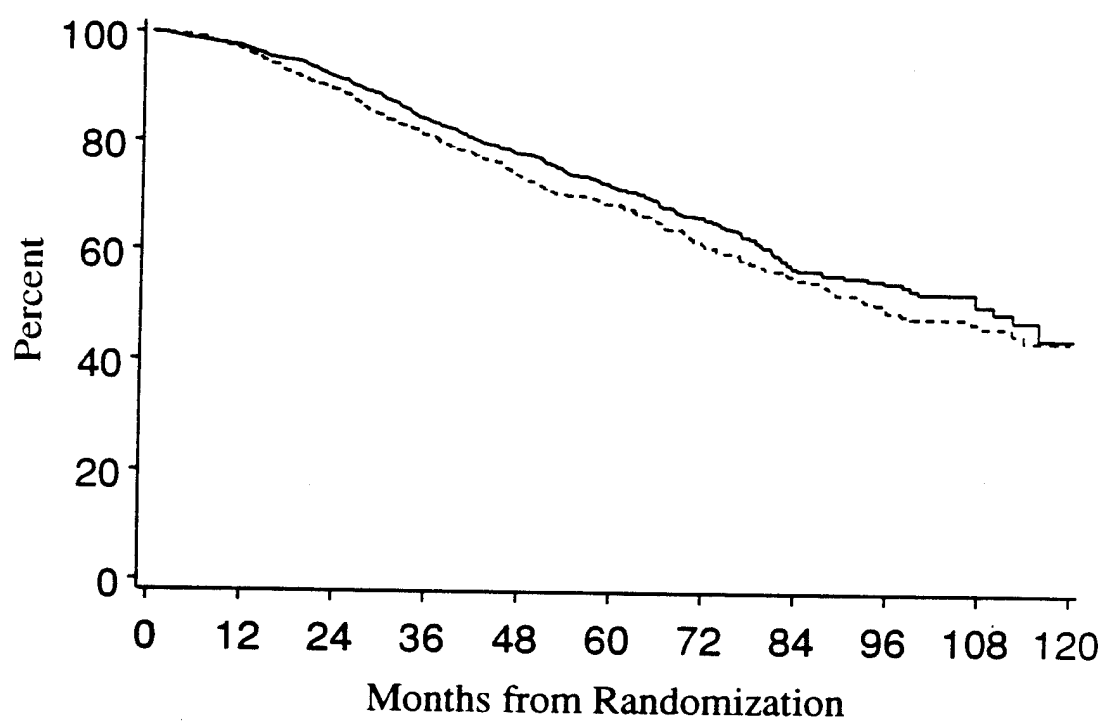


Figure 2A

(A) Tamoxifen Alone

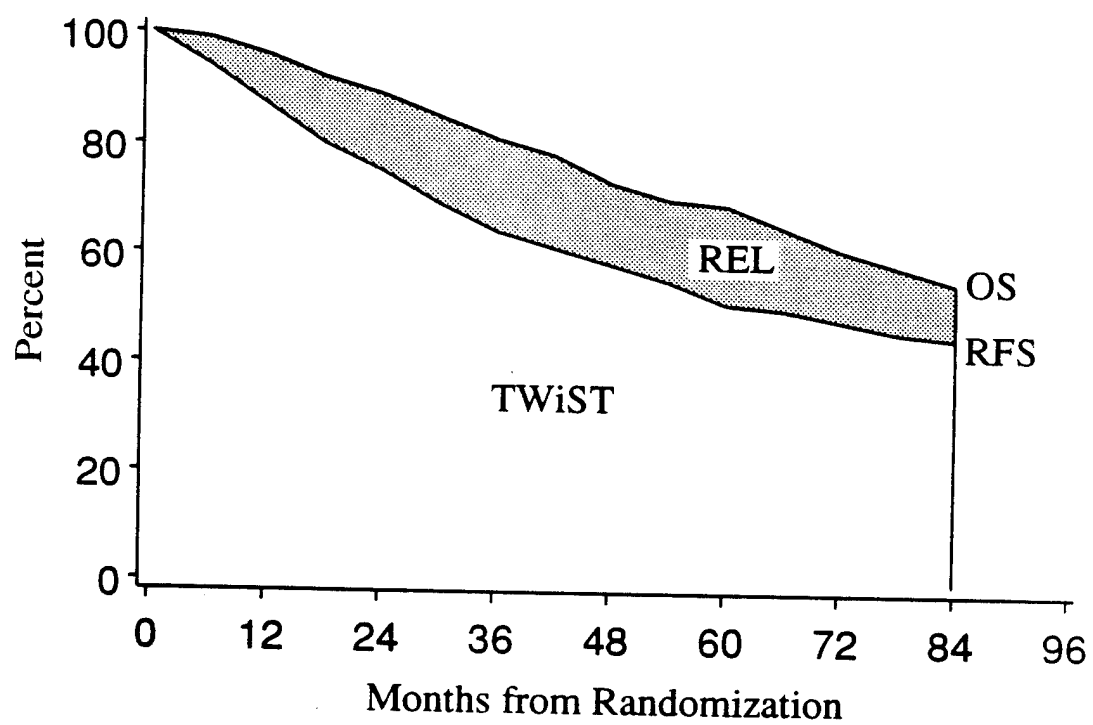


Figure 2B

(B) Chemoendocrine Therapy

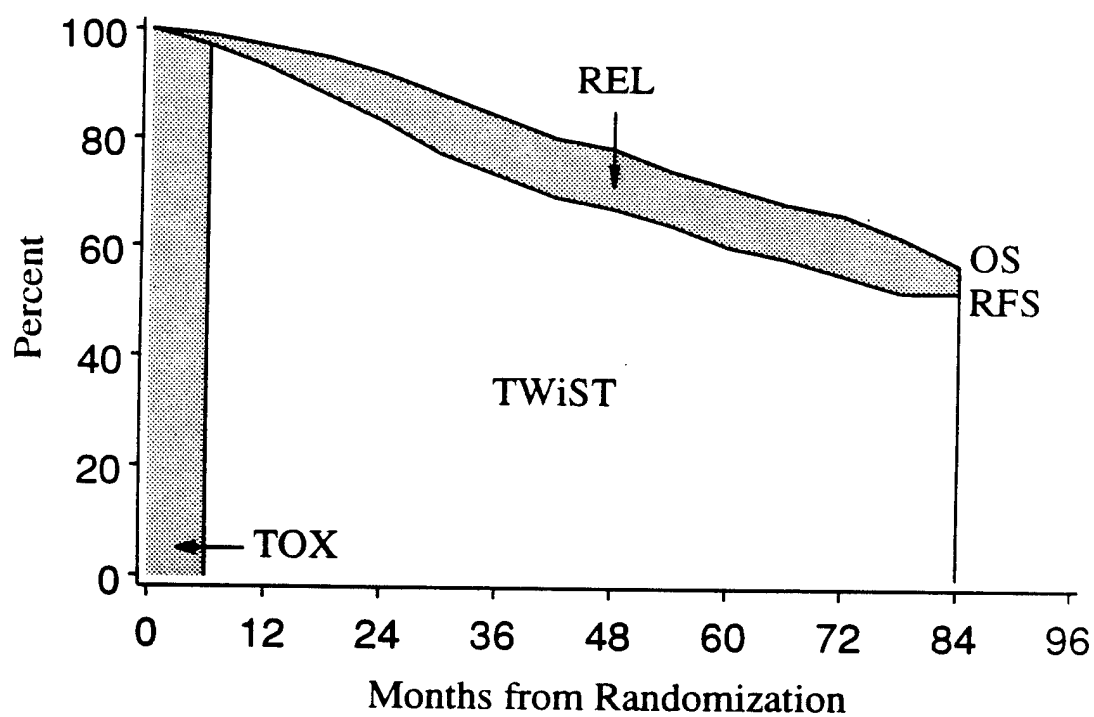


Figure 3

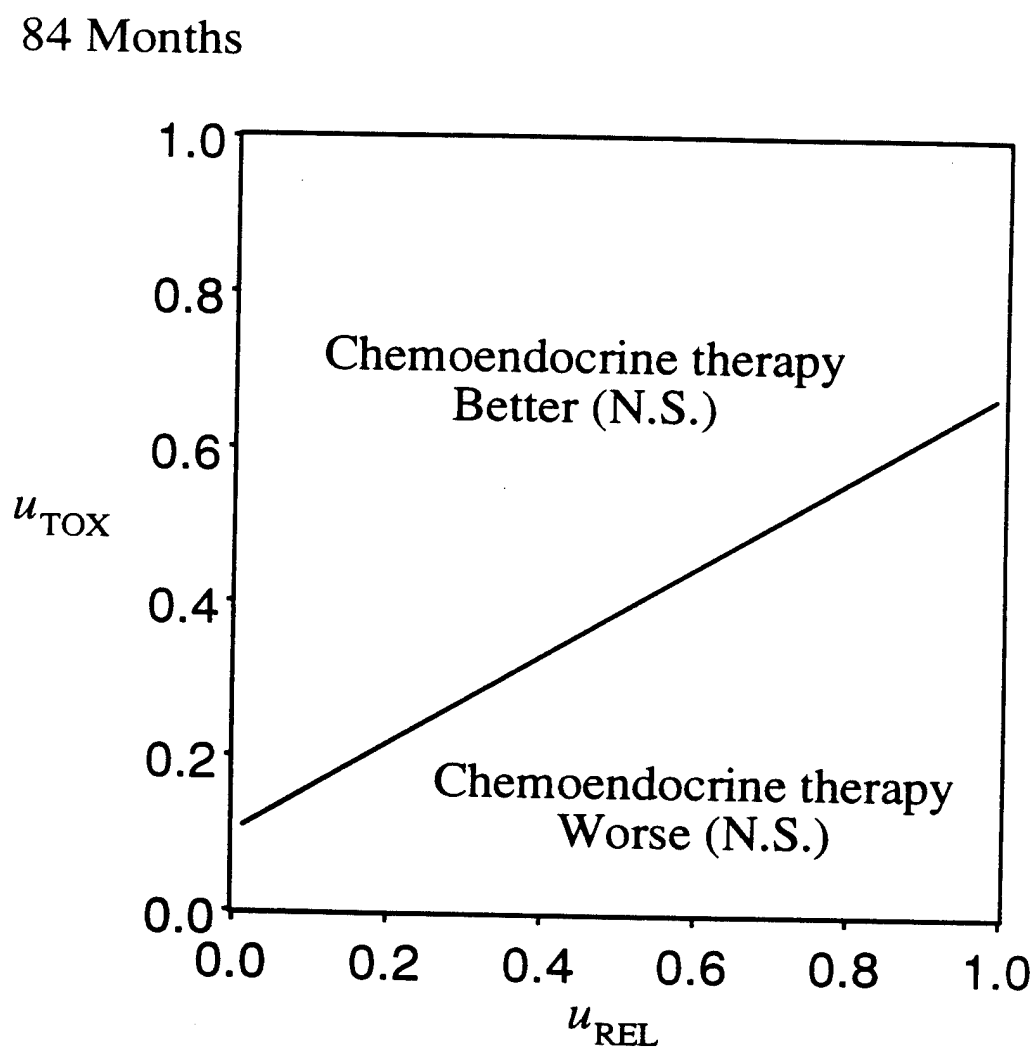
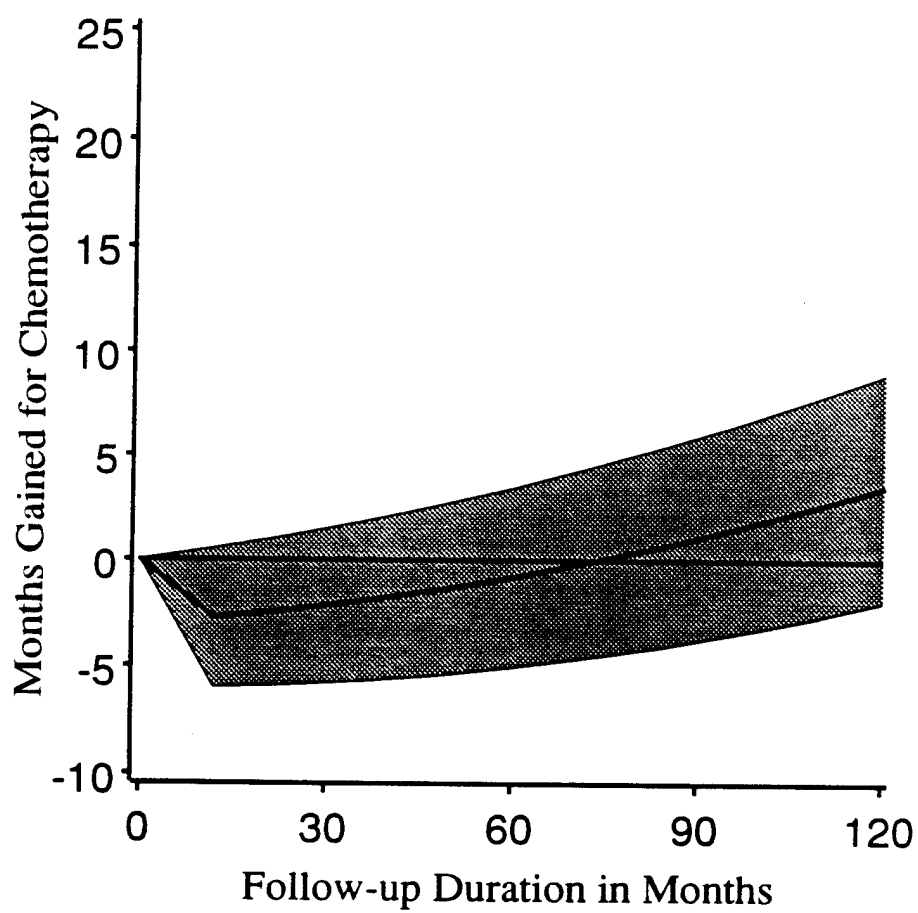


Figure 4



APPENDIX

In this appendix we describe briefly the computational methods used in the Q-TWiST meta-analysis (see Reference 7 for more details). Individual Q-TWiST analyses were first carried out on each clinical trial. These results were combined using the following regression models:

$$\text{RFS} = \beta_1 M + \beta_2 M^2$$

$$\text{OS} = \theta_1 M + \theta_2 M^2,$$

where RFS and OS denote respectively the mean relapse-free survival time and the mean overall survival time, and M denotes the median follow-up duration (in months) for the clinical trial. The unknown regression parameters are denoted by β_1, β_2 and θ_1, θ_2 .

To perform the meta-analysis, we used the RFS and OS results (restricted to the median follow-up duration) for each clinical trial as data for estimating the regression models. The regression parameters were estimated by generalized least squares according to treatment group. This consisted of minimizing the total sum of the weighted squared differences between the observed data and the means predicted by the regression models, where the weights were derived from the inverse of the variation observed within each of the clinical trials. Using these techniques, the following regression models were estimated:

CT + TAM

TAM

$$\text{RFS} = 1.0332M - 0.00370M^2$$

$$\text{OS} = 1.0496M - 0.00273M^2$$

$$\text{RFS} = 0.9929M - 0.00398M^2$$

$$\text{OS} = 1.0509M - 0.00303M^2$$

We then substituted $M = 84$ months into these equations to obtain predicted mean RFS and OS based on the meta-analysis for each treatment group. For example, substituting $M = 84$ months in the first equation for TAM gives $\text{RFS} = 55.3$, corresponding to the RFS estimate given in Table 4 for the TAM group. The remaining estimates in Table 4 were obtained using the equations:

$$\text{TWiST} = \text{RFS} - \text{TOX}$$

$$REL = OS - RFS$$

$$Q-TWiST = u_{TOX}TOX + TWiST + u_{REL}REL$$

Treatment differences were obtained by subtracting the results for the TAM group from the results for the CT+TAM group. In particular, the mean treatment differences corresponding to TOX, TWiST and REL are 6.0, -0.6 and -3.4, respectively. Therefore, the amount of Q-TWiST gained for the CT+TAM group within 84 months based on the meta-analysis is

$$6.0u_{TOX} - 0.6 - 3.4u_{REL}.$$

The threshold line appearing in Figure 3 was obtained by setting this equation equal to zero and solving for the unknown utility coefficients. This gives the following equation for the threshold line:

$$u_{TOX} = \frac{0.6 + 3.4u_{REL}}{6.0}.$$

This equation describes the pairs of values of u_{TOX} and u_{REL} for which the two treatments provide the same amount of Q-TWiST.

The Q-TWiST gain function shown in Figure 4 was obtained by plotting the results as M ranged from 0 to 120 months. The shaded region in the plot was computed by determining the range of Q-TWiST treatment differences as the utility coefficients varied between 0 and 1 for each value of M . For example, when $M = 84$ months, the maximum Q-TWiST difference occurs when $u_{TOX} = 1.0$ and $u_{REL} = 0.0$, and the minimum occurs when $u_{TOX} = 0.0$ and $u_{REL} = 1.0$. The region between these extremes is shaded in the figure as M ranges from 0 to 120 months. The solid line within the shaded region corresponds to $u_{TOX} = u_{REL} = 0.5$. In this way, the Q-TWiST gain function shows the range of Q-TWiST treatment differences over a range of median follow-up durations.