

COMPARISON OF ADJUVANT THERAPIES USING QUALITY-OF-LIFE CONSIDERATIONS

Richard D. Gelber

Harvard Medical School and School of Public Health
Dana-Farber Cancer Institute, Boston

A. Goldhirsch

Ospedale San Giovanni, Bellinzona
Ospedale Civico, Lugano
Ospedale Beata Vergine, Mendrisio

Abstract

The benefit for patients with operable breast cancer treated with adjuvant systemic therapy is small, if reduction of early mortality within the context of randomized control trials is used for treatment comparison. One might consider that the 75%–85% of patients who die despite treatment are overtreated, as are patients who remain alive even without therapy *within a given time frame*. Larger treatment benefits in terms of avoided or delayed breast cancer relapse have been demonstrated even at early phases of follow-up in the vast majority of adjuvant trials. Exposure of all patients to adjuvant therapy at a time at which no symptoms of disease are present is detrimental in terms of quality of life.

Based on our assumption that the quality of life of the patient is typically altered both by subjective toxic effects of adjuvant treatment and by the appearance of relapse, we developed a method of comparing treatment effects in terms of time without symptoms of disease and toxicity of treatment (TWiST). Because the impact of treatment on relapse rates appears earlier than survival effects in all adjuvant therapy trials, and because the value of time without relapse in terms of the quality of life of the patients is as yet poorly defined, we have generalized our method of comparing treatment attitudes to include individual qualitative judgment values. The experience gained from integrating quality-of-life issues into clinical trials for breast cancer might also be applied to other diseases characterized by a chronic course, toxic treatments, and gains in periods of relative or absolute freedom from toxic effects and progressive disease.

INTRODUCTION

Most patients with operable breast cancer relapse and die of the disease. The annual relapse rate is relatively small and is related to many prognostic indicators which de-

We thank our colleagues Drs. R. J. Simes, P. Glasziou, and A. Coates for their significant contributions to developing the methodology.

pend upon biological factors. Even for patients with metastases in axillary nodes, the annual relapse rate is about 10% per year. The annual mortality rate for this population is about 5% per year (12). Improvements in outcome include a significant reduction of relapses seen in most of the individual trials (11) and a 15%–25% reduction in the odds of death within 5 years. An overview combining data from all available relevant randomized trials was required to demonstrate the presence of these treatment effects statistically (13). The considerable efforts being made presently to improve these results include, among others, investigation of the ideal timing for administration of adjuvant therapy, evaluation of intensive treatment regimens which require autologous bone marrow transplantation, use of recombinant colony-stimulating factors to reduce marrow toxicity, and study of hormonal agents used in association with chemotherapy. It is unlikely, however, that a definitive breakthrough will be achieved by these approaches, and it might be anticipated that advances in the treatment of operable breast cancer in terms of overall survival improvement will continue to be of small magnitude.

Such a prospect underscores the need for alternative evaluation methods which take into account the well-being of the patients. The following facts might be considered:

1. There is an increased tendency to treat patients with node-negative disease (1;7), whose risk for relapse is smaller than that of women with node-positive disease (for whom the value of systemic adjuvant therapy has been established during the past 2 decades).

2. Many physicians are more aware of issues of quality of life. An increasing number of health care professionals understand that the simplistic approach to cancer treatment, “better to live more decently even if for a shorter period,” represents a logical error: “shorter” in most cases is related to “worse” in advanced disease (4) and probably also in the adjuvant setting (6).

3. Adapting treatment results from clinical trials to the individual patient situation is very difficult and requires certain assumptions, a number of which are derived empirically. This, of course, is acceptable, since many aspects of modern medicine relate to empirical experience. It can very easily become a matter of controversy, however, especially if the treatment results in terms of disease control are of modest magnitude and borderline significance, while the side effects and costs of treatment are acute and obvious.

Additional endpoints that take into account factors related to the well-being of the patients represent an obviously important tool for decision-making when the treatment-choice situation is such as that currently applying for operable breast cancer.

DEVELOPMENT OF COMPARISON METHODS

TWiST Methodology

In 1986 (6), we described the TWiST methodology, which was applied to a data set, including 463 postmenopausal patients with N+ breast cancer who were randomized into Ludwig Trial III to receive either a chemoendocrine therapy (cyclophosphamide, methotrexate, 5-fluorouracil, prednisone and tamoxifen; CMFp+T) for 1 year *or* endocrine therapy alone (p+T) for one year *or* no adjuvant therapy (Observation). The *time without symptoms of disease and toxic effects of treatment* (TWiST) was considered as the endpoint. Its construction was based upon definitions which are sum-

Table 1. Definition of Quality-of-Life-Oriented Endpoints

<i>Event</i>	<i>Abbreviation</i>	<i>Time assigned to</i>
Any reversible subjective toxic effect ^a of any grade (excluding amenorrhea) reported during a cycle or month of treatment	TOX	Applied to the entire month
Alopecia and weight gain	TOX	Include an additional 3 months after last report to allow for recovery
Isolated mastectomy scar or contralateral breast cancer alone	LR ^b	Assign 3 months to allow for recovery from local treatment
Any relapse ^c (other than LR) or second primary tumour (not breast)	REL	Applies to entire remaining survival period
Leukopenia, thrombocytopenia, and grade 1 or 2 anemia (asymptomatic)		No penalty
None of the above	TWiST	Applies to any remaining survival time

^a Subjective toxic effects include those noted and graded by the investigators: nausea, vomiting, anorexia, diarrhea, mucositis, infections, epigastric pain, neurotoxicity, headache, euphoria, depression, allergic skin disorders, alopecia, cystitis, muscle weakness, hypercalcemia, hot flashes, thrombosis, thromboembolism, edema, lymphedema, weight gain, eye disorders, joint pain, symptomatic anemia, hemorrhage, nonmenstrual vaginal bleeding.

^b For convenience, this 3-month period has been incorporated within REL in the analysis.

^c Skeletal recurrences were considered only at the time of definitive confirmation.

marized in Table 1. TWiST was calculated by subtracting from overall survival all time periods for which some subjective toxicity (TOX) was noted, and all time following symptomatic systemic relapse (REL) that involved treatment intervention. In addition, an isolated mastectomy scar recurrence or contralateral breast cancer alone (LR) was considered separately. This pragmatic approach has been the basis for the development of the methodology, but it could hardly be considered an accurate way to evaluate the actual impact of treatment felt by the patients. Life with subjective toxicity and with some symptoms of relapse are obviously far from optimal conditions of life, but most women with breast cancer will, nevertheless, consider these periods to have some value greater than death (3).

Q-TWiST Methodology

In order to allocate some value greater than zero to time periods during which the patient is alive with or without the occurrence of a quality of life related event, the following considerations were made (using the same criteria as for the TWiST method).

The time period with subjective toxic effects (TOX) is a period of time for which the perception of well-being of the patients has a large spectrum of values. In fact, many patients are encouraged to be very active (i.e., choose toxic adjuvant therapy if this is an option) and accept with philosophy the subjective side effects associated

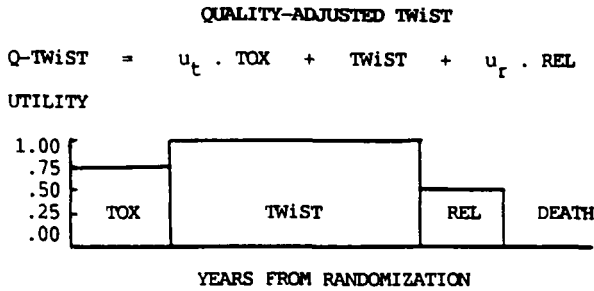


Figure 1. Calculation of quality-adjusted survival relative to TWiST (Q-TWiST): Weighted summation of time periods TOX, TWiST, and REL. Utility coefficients of $u_t = 0.75$ for TOX and $u_r = 0.50$ for REL are illustrated.

with adjuvant treatment. This attitude relates to the insurance mentality: invest now (in toxicity) to slightly reduce a heavy risk later. On the other hand, many patients are anxious to return to normal life as soon as possible after facing the initial diagnosis of breast cancer. This is typically obtained with various degrees of denial, a mechanism that is easily disturbed by adjuvant treatment.

The time following systemic relapse (REL) is shaded by the fact itself of confronting the disease. The need for palliation is dependent upon the manifestations of the relapse and is common. Furthermore, the almost universally fatal outcome reduces the spectrum of perceptions as compared with the range that is applicable to TOX.

The possibility to incorporate these concepts within the framework of the TWiST methodology is provided by introducing the notion of utility coefficients.

Quality-adjusted survival analysis is performed by considering utility coefficients u_t and u_r , taking values between 0 and 1, to represent the value, relative to TWiST, of toxic periods (TOX) and time after relapse (REL), respectively. Quality-adjusted survival relative to TWiST (Q-TWiST) for a patient who experiences an overall survival composed of TOX, TWiST, and REL is calculated as

$$Q\text{-TWiST} = u_t \times \text{TOX} + \text{TWiST} + u_r \times \text{REL}.$$

Figure 1 illustrates the way in which u_t and u_r influence the addition of time periods to the "baseline" TWiST for arbitrarily chosen values of 0.75 for u_t and 0.50 for u_r .

Population estimates of Q-TWiST achieved for a specific treatment regimen are obtained by partitioning overall survival up to a specified time from randomization into three regions (see Figures 2A–C). The regions are distinguished by curves for duration of toxic effects, time to systemic relapse, and overall survival time. The average amounts of time spent in TOX, TWiST, and REL for each treatment group up to a specified cut-off time from randomization are estimated by the areas between the curves, as shown. These averages are then combined as a weighted sum according to the above equation to obtain an average Q-TWiST for each treatment.

The average times accumulated in TOX, TWiST and REL during 7 years, which are graphically presented in Figure 2, may also be described numerically as shown in Table 2. The calculation of Q-TWiST is illustrated by a weighted combination of these

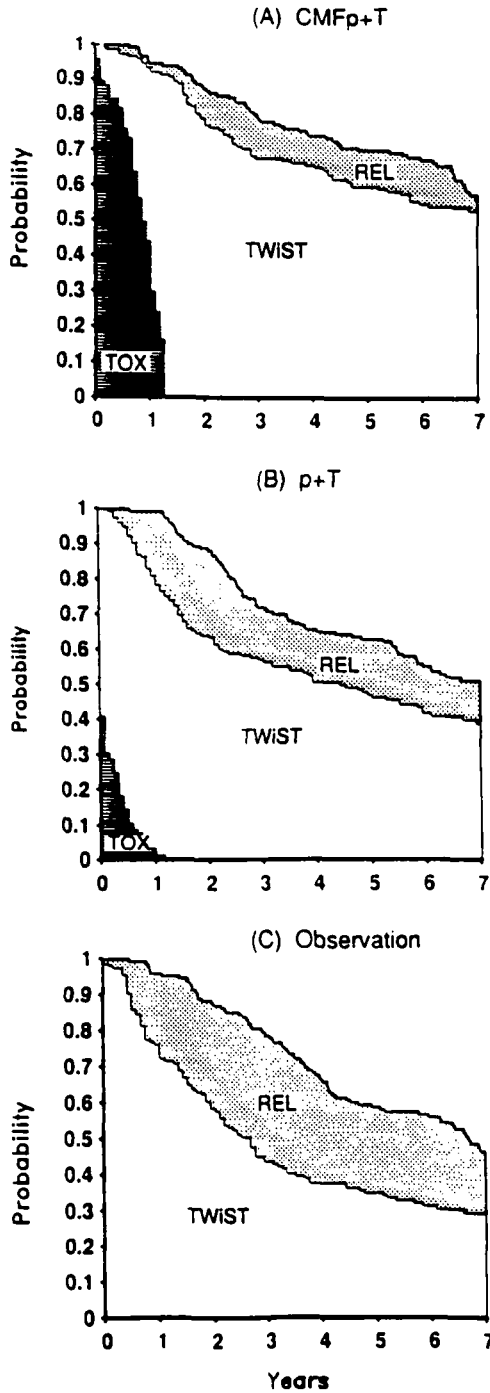


Figure 2. Time accumulated up to 7 years from randomization in Ludwig Trial III example. Time is partitioned for each treatment group by curves for time with toxicity, time to systemic relapse, and overall survival. Areas between these curves yield population average values for TOX, TWiST, and REL. Panel A = CMFp+T, panel B = p+T, and panel C = observation.

Table 2. Components of Q-TWiST for Treatment Evaluated in Ludwig Trial III.

	CMFp + T	p + T	Observation
TOX	9.6 (0.4)	2.0 (0.3)	0.0 (0.0)
TWiST	50.3 (2.5)	47.1 (2.7)	41.5 (2.6)
REL	7.1 (1.1)	12.9 (1.4)	20.9 (1.8)
Q-TWiST	58.7 (2.2)	54.6 (2.3)	51.9 (2.2)

($u_t = u_r = 0.5$)

Note: Average months of TOX, TWiST and REL accumulated within 7 years of randomization, with Q-TWiST calculated for arbitrary utility coefficients ($u_t = u_r = 0.5$). (Standard errors are shown in parentheses.)

CMFp + T = cyclophosphamide, methotrexate, 5-fluorouracil, prednisone + tamoxifen (chemoendocrine therapy) given for 12 months.

p + T = prednisone + tamoxifen (endocrine therapy alone) given for 12 months.

Observation = no adjuvant therapy

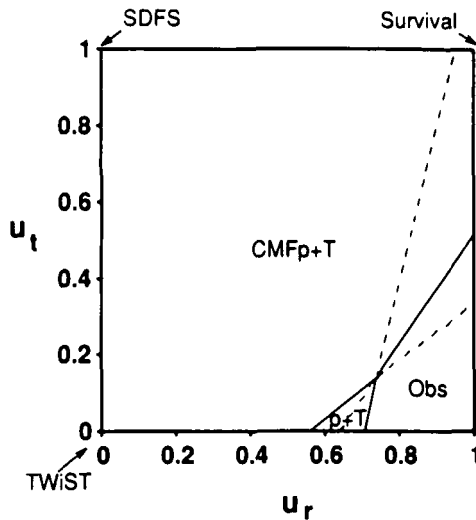


Figure 3. Threshold analysis for Ludwig Trial III comparing average Q-TWiST values accumulated up to 7 years from randomization. Lines are drawn for values of u_t and u_r which yield equal values of average Q-TWiST between treatment pairs.

components using $u_t = u_r = 0.50$. Based on these arbitrary values for the utility coefficients, the average Q-TWiST accumulated during 7 years (84 maximum possible months) for chemoendocrine therapy was 58.7 months—4.1 months longer than for endocrine therapy alone ($p = 0.20$) and 6.7 months longer than for the observation group ($p = 0.05$).

In order to avoid relying on arbitrary selection of specific values of u_t and u_r , treatment comparisons can be made using a *threshold analysis*. The methodology is described in greater detail in a recent publication (8). Threshold values of u_t and u_r between treatment pairs are those for which the treatments have equal average Q-TWiST values. Figure 3 shows the threshold analysis for Ludwig Trial III with cut-off at 7 years from randomization. The bold lines divide the figures into three regions where

each treatment is favored according to the values placed on time with toxicity (u_t) and time following systemic relapse (u_r) relative to TWiST. The dashed lines show the thresholds between pairs of treatments. For the specific values at the corners of Figure 3, treatment comparisons are for systemic-disease-free survival ($u_t = 1, u_r = 0$), TWiST ($u_t = 0, u_r = 0$), and overall survival ($u_t = 1, u_r = 1$). The threshold analysis for our specific example indicates that CMFp+T is the preferred treatment for a wide range of plausible values of u_t and u_r . It is important to note that these thresholds are based upon 7-year figures, and further follow-up might change the relative Q-TWiST values associated with the three treatment options.

In order to provide some data about the statistical uncertainty of the results, we display the threshold lines together with their confidence intervals. The comparisons of CMFp+T versus Observation (Figure 4A) and CMFp+T versus p+T (Figure 4B) are shown. For the first comparison, there is a statistically significant difference in favor of CMFp+T for most values of u_t and u_r . The confidence regions for CMFp+T versus p+T are wide, and the range of values of u_t and u_r , for which the chemoendocrine therapy is statistically significantly preferred, is small.

The chemoendocrine therapy involves early toxicity but later gains, which exceed those of the endocrine therapy alone. The evolution of the differences in average Q-TWiST (for $u_t = u_r = 0.5$) between treated and control groups is displayed in Figure 5A. As time from randomization increases, the early losses caused by toxic effects are counterbalanced by the gains associated with delayed relapse, so that an average overall gain for the population of treated patients begins to emerge at 3–4 years after operation. The evolution of results for other endpoints of TWiST (Figure 5B) and overall survival (Figure 5C) can be displayed in similar fashion. This practice of displaying the results in terms of different endpoints illustrates the clinical utility of the methods, since the dimension of time from diagnosis (or from primary treatment) is added for comparison with the conventional assessments of disease-free survival and overall survival.

It is important to note that the treatment comparisons displayed in Table 2 and Figures 2–5 are based upon the specific data available from Ludwig Trial III. The endocrine therapy was given for the duration of 1 year only, a suboptimal use of this treatment modality. Standard chemotherapy regimens of today include 6 months of cytotoxics and not the 1 year used in our example. Other forms of adjuvant therapy or prolonged duration of tamoxifen might improve both the length and quality of life compared with the 12-month CMFp+T regimen.

Adding Other Arbitrary Utility Coefficients to Q-TWiST

Some other ideas related to economical sciences may be borrowed for our attempt to integrate quality-of-life-oriented values to the analysis of breast cancer therapies. One year commencing from the present time might have intrinsically more value than 1 year commencing from some time (5 years, say) in the future. A discount rate per year of life without relapse may be considered to express the fact that the loss of present time because of immediate toxicity might be too high a price for gaining an equal amount of time in the future. Similarly, delaying a relapse that would have occurred late might not be as valuable as delaying an early recurrence. A delayed relapse also has a better prognosis (9) in terms of survival, and, therefore, some kind of discount might be justified.

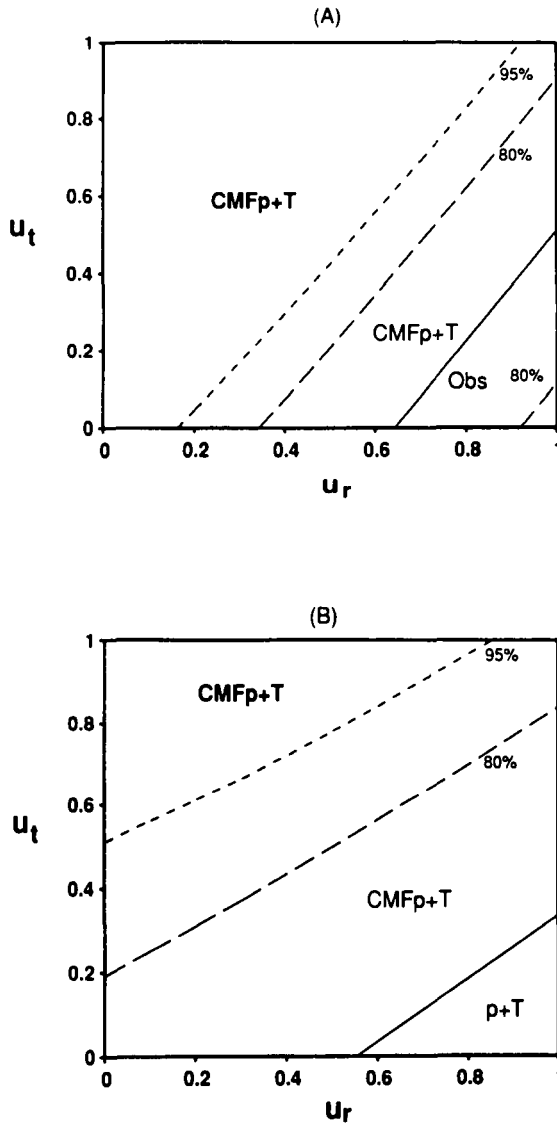


Figure 4. Pairwise treatment comparisons for (A) CMFp+T versus Observation and (B) CMFp+T versus p+T showing 80% and 95% confidence regions.

Criticism may obviously be raised against this discounting policy because of the fact that many patients might reason according to the “insurance mentality” – an investment today (in terms of toxic effects to all) for a chance of reduction of odds of relapse in the future should not be penalized regardless of when the relapse eventually occurs. It is important to evaluate the prospective perception of the patient regarding a delayed relapse, whether early or late. The effect of the discounting of time by 10% per year on our Ludwig Trial III example within 7 years is displayed in the threshold analysis presented in Figure 6. As expected, if later gains are assigned lower values, early toxic effects of adjuvant therapy are more difficult to justify.

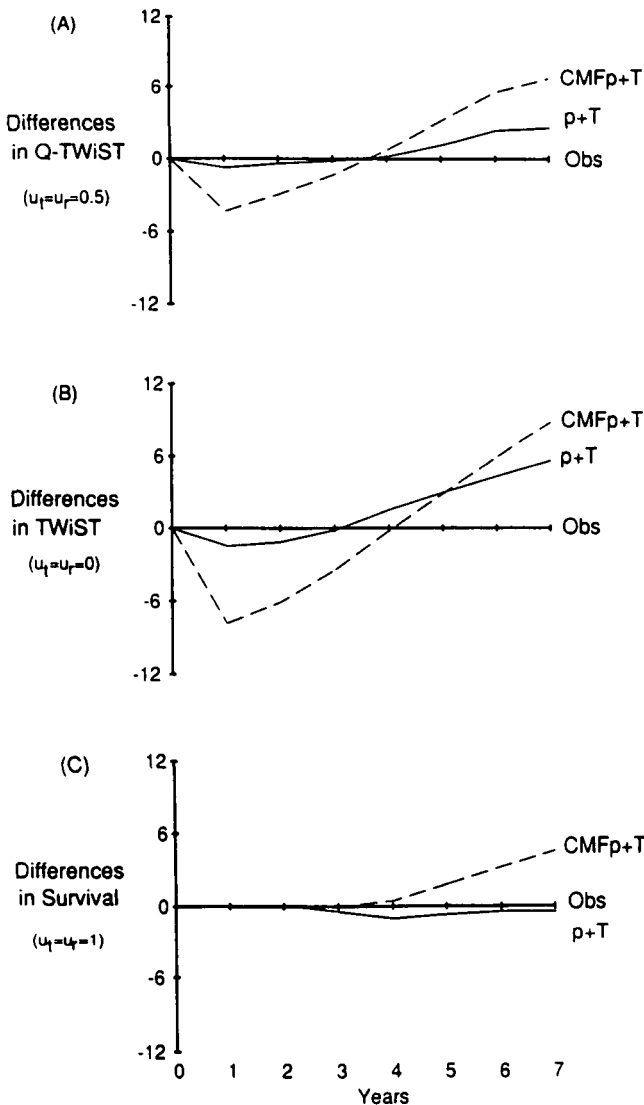


Figure 5. Treatment differences relative to observation as functions of time from randomization. Differences are in terms of average months of Q-TWiST accumulated since randomization for (A) Q-TWiST with arbitrary values $u_l = u_r = 0.5$; (B) TWiST (i.e., Q-TWiST with $u_l = u_r = 0$); (C) overall survival (i.e., Q-TWiST with $u_l = u_r = 1$).

But What is the Single Answer Provided by a Q-TWiST Analysis?

Four values must be specified in order to obtain a unique solution for average Q-TWiST treatment comparisons derived from a clinical trial data set: the two utility coefficients u_l and u_r , the time from diagnosis considered for the comparison, and the discount factor. Some might be uncomfortable with the prospect that a clinical trial not yield a definitive answer recommending a treatment for application to all patients. Yet, quality of life is a philosophical issue that in fact cannot have such a unique answer. The

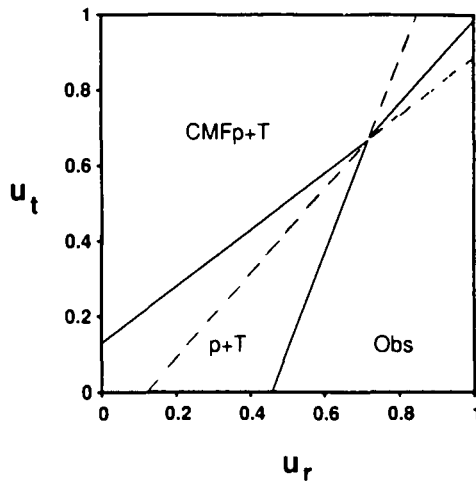


Figure 6. Threshold analysis for Ludwig Trial III within 7 years from randomization with the value of future time discounted at 10% per year.

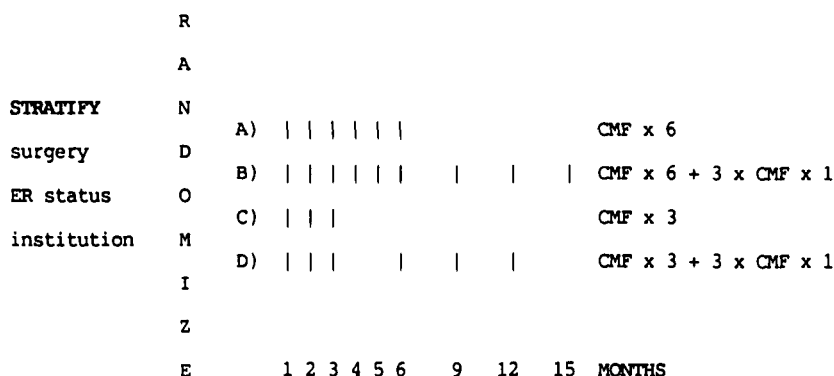
Q-TWiST methodology is developed with this spirit, recognizing that choices are required to be individualized and establishing a framework within which the parameters for treatment choice can be explicitly introduced.

Specifically, consider the choice of u_t and u_r , the utility coefficients associated with time in toxicity and time after relapse. Using the threshold analyses displayed in Figure 3, one can determine the values of u_t and u_r for which the chemoendocrine therapy is preferred to either of the other two treatment options. If, based on subjective feelings, toxic effects are not considered to be too debilitating compared with the impact on quality of life of a systemic relapse, then u_t is higher than u_r and the data in this example favor chemoendocrine therapy. Note that the subjective feelings of relative worth might be derived from the physician (motivated by a prejudice to treat or not to treat), and/or by the patient (who clearly expresses a preference relating to her personal life status at the time).

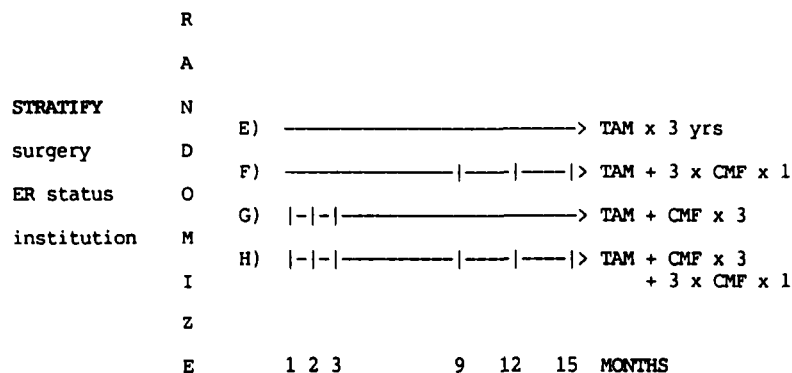
In fact, given that a particular treatment choice is made (e.g., use chemoendocrine therapy for one year rather than no adjuvant treatment), the threshold analysis in Figure 3 defines the implicit relative values of u_t and u_r that apply. In this way the prejudices of physicians who choose a particular treatment strategy are described in terms of how they must weigh the relative value of the time periods TOX, TWiST, and REL.

The other two variables (time from diagnosis and discount rate) are easier to handle. The time from diagnosis is constrained by the maturity of the data set. In order to consider results that are less subject to statistical fluctuations, we generally extrapolate no further than the median follow-up of the study cohort. The choice of discount factor is related to the patient's age, family situation at the time of diagnosis, and the relative value of the next year as compared with a year that is several years into the future. A time trade-off methodology might be useful for determining the discount rate. The exact specification of the discount rate is, however, not crucial for the application of the methodology because even with a discount of 10% per year (about 40%

INTERNATIONAL BREAST CANCER TRIAL VI
PRE- AND PERIMENOPAUSAL PATIENTS WITH
OPERABLE NODE-POSITIVE BREAST CANCER



INTERNATIONAL BREAST CANCER TRIAL VII
POSTMENOPAUSAL PATIENTS WITH
OPERABLE NODE-POSITIVE BREAST CANCER



Oral CMF
 C - cyclophosphamide 100 mg/m² orally days 1-14 |
 M - methotrexate 40 mg/m² i.v. days 1 and 8 | -q 28 days
 F - 5-fluorouracil 600 mg/m² i.v. days 1 and 8 |
 TAM - tamoxifen 20 mg daily for 3+ years

Figure 7. Designs for International Breast Cancer Study Group (IBCSG) Trials VI and VII: Adjuvant therapy for women with operable node-positive breast cancer.

over a cumulative 5-year period), conclusions are not much altered from a zero discount analysis.

The importance of the analysis is to recognize its subjective nature and explicitly define the parameters that impact upon treatment choice. Integrating individual patient perceptions into the analysis of treatment choice is the next logical step to increase the clinical application of the methodologic framework.

Integrating Perceptions and Opinions of Individual Patients into the Analysis of Treatment Results

The development of the TWiST and Q-TWiST methodologies was based upon the perception and biases of physicians as to the values of life periods. The direct possibility to integrate the patients' own value judgments into the analysis was not possible because of the lack of prospectively collected information. The idea, however, has matured rather early during the developmental process, and some instruments are being used in recent trials of systemic adjuvant therapy. The objective is to assess patient coping, well-being, and subjective burden of some toxicities to obtain relative values for TOX, TWiST, and REL directly from the population of study patients. The trials in which these instruments are being tested are outlined in Figure 7 (10). Some initial cycles of chemotherapy may be administered with reintroduction of three single cycles (days 1–14) delivered at 3-month intervals later in the patient follow-up (e.g., months 9, 12, and 15).

The patients are questioned every 3 months concerning their well-being and the extent of disturbance caused by treatment. These are compared to a baseline evaluation obtained before the start of any adjuvant systemic treatment. Patients are requested to give the same information one month after the first relapse. Linear Analogue Self Assessments (LASA) (2;4) provide patient perceptions of adjustment/coping, physical well-being, mood, and appetite. A self-administered adjective checklist is also used to evaluate general well-being at each assessment. These data will help to define values of u_i and u_r that apply to a population of breast cancer patients.

CONCLUSION

The Q-TWiST methodology, which in this case has been applied to a breast cancer trial of adjuvant therapies, can be extended to other fields which contain aspects of value judgment for balancing symptoms of disease against toxic effects of treatment. A major development in this methodology might be provided by refining the instruments that help the patient to express his or her own perceptions of well-being.

REFERENCES

1. Anonymous. Clinical alert stimulates changes in practice. *Journal of the National Cancer Institute*, 1988, 80, 1185–87.
2. Baum, M., Priestman, T., West, R. R., et al. A comparison of subjective responses in a trial comparing endocrine with cytotoxic treatment in advanced carcinoma of the breast. In H. T. Mouridsen & T. Palshof (Eds.), *Breast cancer—Experimental and clinical aspects*. Elmsford, NY: Pergamon Press, 1980, 223–26.
3. Coates, A., GebSKI, V., Bishop, J. F., et al., for the Australian-New Zealand Breast Cancer Trials Group, Clinical Oncological Society of Australia. Improving the quality of life during chemotherapy for advanced breast cancer: A comparison of intermittent and continuous treatment strategies. *New England Journal of Medicine*, 1987, 317, 1490–95.
4. Coates, A., Fischer, Dillenbeck, C., McNeal, D. R., et al. On the receiving end—II. Linear analogue self-assessment (LASA) in evaluation of the quality of life of cancer patients receiving therapy. *European Journal of Cancer and Clinical Oncology*, 1983, 19, 1633–37.
5. Gelber, R. D., Goldhirsch, A., Castiglione, M., Price, K., Isley, M., & Coates, A., for the Ludwig Breast Cancer Study Group. Time without symptoms and toxicity (TWiST): A quality-of-life-oriented endpoint to evaluate adjuvant therapy. In S. E. Salmon (Ed.), *Adjuvant therapy of cancer*, vol. 5. Philadelphia: Grune & Stratton, 1987, 455–65.

6. Gelber, R. D., & Goldhirsch, A., for the Ludwig Breast Cancer Study Group. A new end-point for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. *Journal of Clinical Oncology*, 1986, 4, 1772-79.
7. Goldhirsch, A., & Gelber, R. D. Adjuvant systemic therapy for breast cancer patients without lymph node metastases (N-). In F. Cavalli (Ed.), *Endocrine therapy of breast cancer*, vol. 3. U. Veronesi (Series ed.). Berlin: Springer-Verlag and ESO [European School of Oncology Monographs], 1989, 1, 37-43.
8. Goldhirsch, A., Gelber, R. D., Simes, R. J., Glasziou, P., & Coates, A. Costs and benefits of adjuvant therapy in breast cancer: A quality-adjusted survival analysis. *Journal of Clinical Oncology*, 1989, 7, 36-44.
9. Goldhirsch, A., Gelber, R. D., & Castiglione, M. Relapse of breast cancer after adjuvant treatment in premenopausal and perimenopausal women: Patterns and prognoses. *Journal of Clinical Oncology*, 1988, 6, 89-97.
10. Goldhirsch, A., & Gelber, R. D., for the Ludwig Breast Cancer Study Group. Adjuvant therapy for breast cancer: The Ludwig Breast Cancer Trials, 1987. In S. E. Salmon (Ed.), *Adjuvant therapy of cancer*, vol. 5. Philadelphia: Grune & Stratton, 1987, 297-309.
11. Henderson, I. C. Adjuvant systemic therapy for early breast cancer. *Current Problems in Cancer*, 1987, 2, 125-207.
12. Nemoto, T., Vana, J., Bedwani, R. N., et al. Management and survival of female breast cancer. Results of a national survey by the American College of Surgeons. *Cancer*, 1980, 45, 2917-24.
13. Early Breast Cancer Trialists' Collaborative Group. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer: An overview of 61 randomized trials among 28,896 women. *New England Journal of Medicine*, 1988, 319, 1681-92.