

Women's and Physicians' Utilities for Health Outcomes in Estrogen Replacement Therapy

MARGARET M. HOLMES, PhD, DAVID R. ROVNER, MD,
MARILYN L. ROTHERT, RN, PhD, ARTHUR S. ELSTEIN, PhD,
GERALD B. HOLZMAN, MD, RUTH B. HOPPE, MD,
WILLIAM P. METHENY, PhD, MICHAEL M. RAVITCH, PhD

Estrogen replacement therapy (ERT) prevents fractures and relieves vasomotor symptoms, but it increases the risk of endometrial cancer. Previous studies and national prescribing patterns show that physicians are conservative in their approach to this therapy. The authors interviewed physicians and perimenopausal women to assess their utilities for the various health outcomes of estrogen replacement therapy. On all outcomes, physicians rated illness episodes followed by recovery as being closer to perfect health than did perimenopausal women. Physicians, in judging which outcomes were most important to women, estimated relief of symptoms above fracture prevention, whereas women rated fracture prevention above symptom relief. These results emphasize the need to assess patients' utilities directly, particularly when utilities for the outcome of a particular therapy may influence the choice of a therapeutic regimen. Key words: estrogen replacement; utilities for health outcomes; decision making. J GEN INTERN MED 1987;2:178-182.

ESTROGEN REPLACEMENT THERAPY (ERT) for menopausal women is controversial. Although it is associated with an increase in the incidence of endometrial cancer, ERT helps prevent fractures, which cause over ten times the annual mortality of endometrial carcinoma.^{*1} Epidemiologic data suggest that estrogen therapy decreases mortality from all causes.² Based on this evidence, expert opinion has shifted toward recommending ERT for postmenopausal women. However, clinicians remain reluctant to prescribe estrogen.^{3,4}

In a previous study, physicians' reluctance to prescribe ERT was not related to incorrect knowledge of probabilities, but might have been due to their estimates of utilities for fracture, vasomotor symptoms, and cancer. Because in that study physicians were asked to estimate utilities for the average perimenopausal patient, we investigated the difference between physician estimates of perimenopausal women's values and women's directly assessed values related to particular outcomes of ERT.

Received from the Department of Medicine, Michigan State University, East Lansing, Michigan (MMH, DRR, RBH); the College of Nursing, Michigan State University (MLR); the Center for Educational Development, University of Illinois at Chicago, Chicago, Illinois (ASE); the Department of Obstetrics and Gynecology, Medical College of Georgia, Atlanta, Georgia (GBH, WPM); and Northwestern University Medical School, Evanston, Illinois (MMR).

Supported in part by a grant from the National Fund for Medical Education, #47/81.

Address correspondence and reprint requests to Dr. Holmes: Department of Medicine, B220 Life Sciences, Michigan State University, East Lansing, MI 48824.

METHODS

This study built on a previous study of physicians' decision making in regard to estrogen replacement therapy for menopausal women. The methods of decision analysis and utility assessment used in that study are described elsewhere.⁴ In the present study we used the utility assessment portion to compare estimates of women's utilities provided by physicians with women's utilities assessed directly.

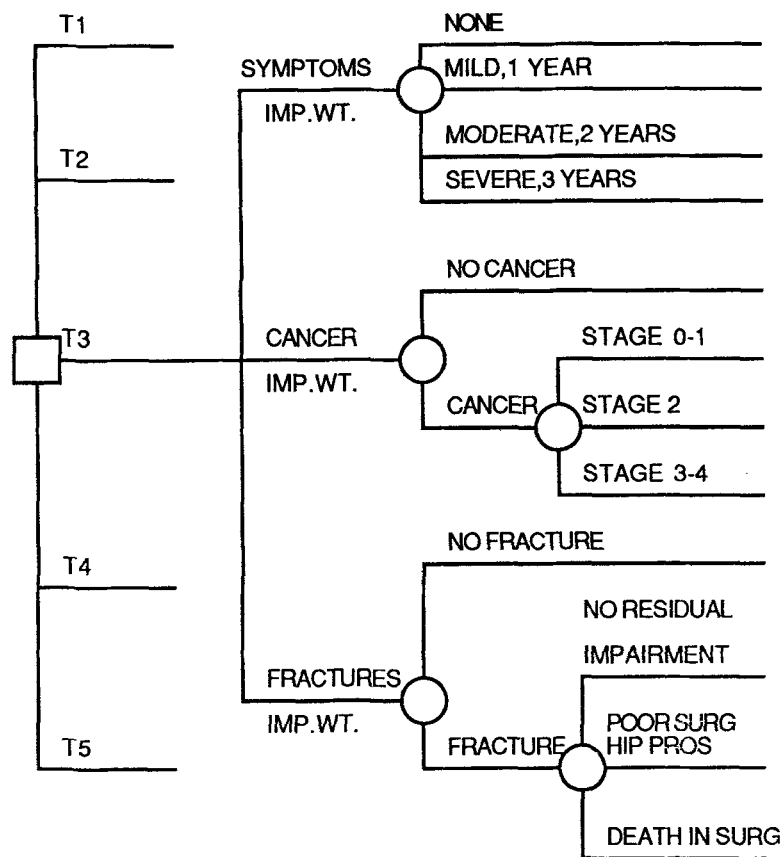
Subjects

A stratified random sample was drawn from employee rolls of Michigan State University (MSU). Recruitment was done through the Provost's Office to protect access to employee rosters. Two strata, faculty and custodial union members, were sampled to allow comparison by social class. Population percentages of 10% faculty and 20% union were drawn to produce 25 in each category. One subject was dropped from the sample because her primary classification was student, reducing N to 49. One subject did not complete the importance weight procedure, making N 48 for importance weight calculations only. Selection criteria were: age 40 to 50 years; female; uterus intact; menstruating. Each participant was paid \$15.00 for her time. Mean ages were 44.8 years for union volunteers and 43.1 years for faculty volunteers.

The group of physicians is described elsewhere.⁸ There were 50 community physicians in private practice from five mid-Michigan cities, 25 family physicians and 25 gynecologists, with 21 men and four women in each group. The list from which the physician sample was drawn was compiled from the Directory of Medical Specialists for Michigan and the list of adjunct faculty of Michigan State University. Since the utility assessment procedure was an interview, the sample was limited to those physicians practicing within a 100-mile radius of Michi-

*Annual mortality from endometrial carcinoma = 3,000.⁵ Annual incidence of hip fractures = 210,000.⁶ Case fatality rate = 18%.⁷ Annual hip fracture mortality is therefore approximately 38,000. The ratio of hip fracture mortality to mortality from endometrial carcinoma is 13:1. This estimate is a conservative measure of the relationship of the two conditions. If morbidity and mortality from other fractures were included, the ratio would increase. The relationship of estrogen use to cardiovascular disease and breast cancer is not well established.⁷

FIGURE 1. Decision tree for estrogen replacement therapy. T₁ = no treatment; T₂ = estrogen only for up to three years; T₃ = estrogen and progestin for up to three years; T₄ = estrogen only for five years or more; T₅ = estrogen and progestin for five years or more.



gan State University. An initial letter, followed by a telephone call, yielded response rates of 80% for obstetricians and 60% for family physicians. Mean age of the gynecologists was 44.9 years and that of the family physicians, 46.5 years. Consent was obtained from the subjects after the nature of the procedures was fully explained.

Decision Analysis

We used decision-analytic techniques to investigate women's and physicians' preferences for trade-offs among the three major outcomes affected by ERT: reduced risk of fractures, relief of symptoms (hot flashes and flushes), and increased risk of endometrial cancer. A standard gamble technique⁹ was used to assess the women's and physicians' utilities for each outcome. Edward's Simplified Multi-Attribute Rating Technique¹⁰ was used to elicit importance weights.

In an interview with a staff person trained to do utility assessment, each participant was asked to evaluate four possible scenarios describing levels of disability from perfect health for each of the three outcomes: hip fracture, vasomotor symptoms, and endometrial cancer. Figure 1 outlines the three outcomes and the four levels of severity for each outcome. Table 1 presents a detailed example of the

four levels of severity for one of the three outcomes (hip fracture).

The scenarios presented to women used lay language for various medical terms. Women were instructed to "rate the desirability of the outcomes you are going to see. We want to get an idea of what outcomes you think you would prefer." Physicians were asked to "assess how most 50-year-old women in your community would rate the desirability of the outcomes you are going to see. We want to get an idea of what outcomes you think patients would prefer." Otherwise, the scenarios presented to women and physicians were the same.

A standard gamble technique^{8, 9} estimated women's and physicians' utilities for each outcome at each level of severity. In the standard gamble, for each outcome respondents rank-order the four levels of severity, with the least preferred assigned a utility of 0 and the most preferred assigned a utility of 1.0. For each of the two intermediate severity levels, X, the respondent must choose between accepting that outcome at that severity or taking a gamble. The gamble gives a chance of P_x for the most preferred level of severity and a chance of 1 - P_x for the least preferred. The probability P is then varied until the respondent is indifferent or can no longer choose between the outcome at severity X and the gamble. This probability is numerically equivalent to the respondent's utility for that outcome at severity X.

TABLE 1
Fracture Scenarios

	Patient Scenario	Physician Scenario
No fracture	You have 25 years of good health followed by a sudden heart attack and death during sleep at age 75.	25 years of good health followed by acute MI and death during sleep at age 75.
Fracture; no residual impairment	You have 15 years of good health (to age 65). You have a broken hip at age 65. The break is operated on; you are hospitalized one week and recover at home for four weeks. You have little pain or crippling from the fracture over the next few years. You die in your sleep of a sudden heart attack at age 75.	15 years of good health (age 65). Fracture of right femoral neck at age 65. Repaired surgically. Hospitalized one week. Four weeks of convalescence. Minimal discomfort or lack of mobility for the next ten years. Sudden death due to acute MI during sleep at age 75.
Fracture, poor surgical result, hip prosthesis	You have 15 years of good health. You have a broken hip at age 65. The hip doesn't heal well and you have to use a walker for six months. Over the next 2½ years, you have pain which doesn't go away taking aspirin. You have trouble walking up and down stairs. You have more pain when standing and thus are not able to do much shopping. At age 65 you have an operation to get an artificial hip. After three months recovering at home, you are able to do all the things you could do before the hip broke. You walk with a cane which slows you down during the winter. Slight pain in your other hip and lower back is better when you rest and take aspirin. You are in otherwise good health from age 69 to 75. You die in your sleep at age 75.	15 years of good health. Fracture of right femoral neck at age 65. Poor surgical result necessitating use of walker for six months. Over next 2½ years, patient has moderate pain poorly controlled by aspirin, Tylenol or Darvon. Has trouble going up and down stairs. Notices pain when standing. Thus is not able to do much shopping. At age 68 patient undergoes surgery for implantation of a prosthetic hip joint. After three months' recuperation, patient is able to resume normal activities. She walks with a cane, which restricts her activities during winter. Occasional minor pain in her other hip joint and in her lower back is relieved by a brief rest and aspirin. Patient is in otherwise good health from age 69 to 75. She dies in her sleep of acute MI at age 75.
Fracture, death in surgery	You have 15 years of good health (age 65). You have a broken hip at age 65. You die during surgery at age 65.	15 years of good health (age 65). Fracture of right femoral neck at age 65. Patient dies during surgery from an embolus at age 65.

Because utilities for fracture reduction, alleviation of symptoms, and increased cancer risk are obtained on separate scales and are not directly comparable, we used Edward's simplified multiattribute rating technique to put the three outcome categories on a single scale. The multiattribute technique requires respondents to rank order in terms of importance the different outcomes (cancer risk, fracture reduction, and symptom alleviation). The least important is assigned a value of 10 and the question asked "How many times more important is the next higher ranked outcome?" This is repeated for the

third outcome. The relative importance weights were then scaled to a 0 to 1 interval.

Analysis

We used multiple analysis of variance (MANOVA) to compare physicians' and women's utilities. Separate analyses were performed for fracture reduction, vasomotor symptoms alleviation, and cancer risk; the null hypothesis was that physicians and potential patients did not differ. In each analysis the two intermediate utility values were the independent variables and the two subject groups were the

TABLE 2
Perimenopausal Women's and Physician's Utilities for Outcome States

	Perimenopausal Women's Utility Mean \pm SD	Physicians' Utility Mean \pm SD
Fracture		
Fracture, surgical repair	0.77 \pm 0.22*	0.85 \pm 0.17
Fracture, hip prosthesis	0.35 \pm 0.41*	0.59 \pm 0.32
Fracture, premature death	0.29 \pm 0.35*	0.05 \pm 0.17
Symptoms		
1 year/mild	0.75 \pm 0.28	0.82 \pm 0.19
2 years/moderate	0.61 \pm 0.28	0.60 \pm 0.23
Cancer		
Stage I, H & BSO	0.78 \pm 0.27*	0.92 \pm 0.14
Stage II, H & BSO recurrence	0.69 \pm 0.27	0.75 \pm 0.24

*Significantly different, $p < 0.05$, univariate physician vs. perimenopausal women.

†Significantly different, $p < 0.001$, univariate physician vs. perimenopausal women.

dependent variables. In each case univariate comparisons were evaluated only when the multivariate tests of significance (Pillais, Hotellings, Wilks) were less than 0.05. Chi square was used to determine whether differences in their rankings of outcome categories were statistically significant.

RESULTS

There were statistically significant differences between physicians and perimenopausal women in the utilities assigned to several outcomes (Table 2). In five of seven instances, physicians rated intermediately severe outcomes showing some degree of disability as being closer to the most preferred outcome (a value of 1) than did perimenopausal women. The difference was statistically significant for the one cancer outcome (a grade 1, stage 1 adenocarcinoma of the endometrium, a hysterectomy and no further complications.) Significant differences were also found for both intermediate fracture outcomes.

The worst fracture outcome, premature perioperative death at age 65, showed a very different pattern. The pattern of physicians' giving higher values was reversed, although differences between physicians and women were still statistically significant. Women gave a mean value of 0.29 to this state, while physicians rated it 0.05. The utility of premature death was not 0. Five physicians (of 50) and 21 premenopausal women (of 48) rated surgical repair of a fracture of the right femoral neck and subsequent surgical implantation of a prosthetic hip joint, with minor residual pain, as worse than perioperative death for a fracture of the right femoral neck. This reversal made the standard deviation of the figure for fracture states large. Some would argue that no state of poor health could be worse than death.¹¹ However, in this analysis we made whatever state the subject called worst the bottom anchor. In general, the standard deviations for women were larger than those for physicians.

Both physicians and perimenopausal women considered cancer the most important outcome. However, women reported that fracture was more important than vasomotor symptoms, while physicians indicated that women would value symptom relief more than reduction of the risk of fracture (Table 3). The rank order of the three categories showed a marked difference in priorities (Table 4). While both groups ranked cancer as most important, 35 perimenopausal women ranked fracture reduction second, whereas only 23 physicians did. Seven women compared with 14 physicians ranked vasomotor symptoms second. All of these differences were statistically significant. Third ranks were not calculated because the third rank was dependent on the first and second.

TABLE 3

Rank Orders of Mean Importance Weights for Outcomes of Estrogen Replacement Therapy among Perimenopausal Women and Physicians

	Importance Weight		
	Rank	Mean	±SD
Perimenopausal women			
Cancer	1	0.63	±0.22
Fractures	2	0.26	±0.19
Symptoms	3	0.11	±0.13
Physicians			
Cancer	1	0.52	±0.20
Symptoms	2	0.30	±0.12
Fractures	3	0.18	±0.22

TABLE 4

Rank Orders of Three Outcomes of Estrogen Replacement Therapy by Women and Physicians

	Outcome		
	Cancer	Symptoms	Fracture
Number of subjects who rank category first			
Perimenopausal women	40	2	6
Physicians	34	16	0
$\chi^2 = 17.34, df = 2, p 0.001$			
Number of subjects who rank category second			
Perimenopausal women	6	7	35
Physicians	13	14	23
$\chi^2 = 7.36, df = 2, p 0.05$			

DISCUSSION

Several discrepancies occur between women's and physicians' importance weights and utility ratings for outcomes associated with ERT. The utility ratings showed that intermediate-level-severity outcomes were rated more highly by physicians than by women. For physicians, fracture that is surgically repaired and cancer that does not recur were close to the most preferred outcome. For women, these states represented substantial departures from the preferred outcome. More women than physicians ranked fracture reduction as the most important outcome. Analysis of mean importance weights also showed women giving greater importance to fracture reduction than to alleviation of vasomotor symptoms, while physicians reversed these two.

Why might these discrepancies have occurred? With regard to the physicians' assumption that relief of hot flashes is more important to women than reduction of fractures, it may be that the women arriving in physicians' offices complain about hot flashes. It may be far less likely for a woman to complain that she may develop fractures, although the popular and medical press indicates a growing concern with

this issue. We also sampled a presumably healthy population of women who work at a variety of jobs, including cook, custodian, college teacher, and radiologist. These women may not represent the case-loads of gynecologists and family practitioners or the population of non-working women.

Women may have different preferences for health outcomes with ERT than physicians estimate they do. There is a growing awareness that it is possible to measure and consider patients' preferences in medical decision making.¹² Prior studies have shown that patients' attitudes toward risks of medical interventions vary and do not always support measures that extend life but diminish its quality.¹³ Decision analysis is a particularly useful technique for considering patient values in clinical decisions because it focuses the patient's attention on her values for the health outcomes (utilities) rather than on the overall analysis of the clinical problem.¹⁴

In clinical decisions where the ethical demands of the problem require a physician to substitute the patient's values for his or her own, differences between physician and patient preferences for outcomes become critical. We hypothesized that physicians, as proxy decision makers for menopausal women in making the prescription decision, might well base their decisions on an overestimate of concern for cancer. However, in this study, both groups ranked cancer as the most important. In general, women rated reduction of fractures as more important than alleviation of vasomotor symptoms, and within the fracture category, women rated disability from fracture as less preferred than premature death.

Physicians, on the other hand, when estimating perimenopausal women's preferences for health outcomes, rated various degrees of disability as closer to preferred health than did the women themselves. One might argue that the physicians have a better sense of how women will cope when they actually encounter these health outcomes. They have, after all, seen women cope successfully with states very much like those in our scenarios. Almost half of the perimenopausal women ranked premature death as better than serious prolonged disability from fractures, while only 10% of the physicians did so. We conclude from these data that the physicians and the perimenopausal women in our sample did not make the same assumptions about the values of the outcomes affected by ERT. Women's values for intermediate (nonhealthy) states were lower than physicians' estimates.

Estrogen replacement therapy continues to be controversial because of the complex tradeoff between gains that can be achieved in symptom relief and reduced risk of osteoporosis compared with the increased risk of endometrial cancer. Decision analytic models are illuminating in considering such

problems because they allow the explicit quantification of expected mortality from all relevant factors, as well as patient preferences for the likely outcomes. Formal models incorporating expected mortality and quality of life have demonstrated a moderate preference for ERT, in large part due to the high frequency of fractures in postmenopausal women compared with a much lower rate of endometrial cancer.⁴

Our analysis suggests that when considering ERT, perimenopausal women rate outcome fractures as more important than physicians estimate that they do. Perimenopausal women more frequently reported preferring premature death in surgery to a longer life with disability from a hip repair than physicians estimated they would. These data suggest that the attitudes of perimenopausal women toward fracture reduction, risk of endometrial cancer, and alleviation of vasomotor symptoms support, rather than discourage, the use of estrogens to prevent osteoporosis. Further, they suggest that physicians should proceed with caution in decision making where they assume, rather than directly elicit, patient values. To incorporate our findings into routine clinical decision making, however, requires an informed patient. The issue of how to present information related to possible outcomes to patients is of continuing concern in medical decision making.¹⁵

REFERENCES

1. Osteoporosis. National Institutes of Health Consensus Conference Statements. JAMA volume 5, number 3
2. Bush T, Cowan L, Barrett-Connor E, et al. Estrogen use and all-cause mortality. JAMA 1983;249:903-6
3. Kennedy DL, Baum C, Forbes MB. Noncontraceptive estrogens and progestins: use patterns over time. Obstet Gynecol 1985;65:441-6
4. Elstein AS, Holzman GB, Ravitch MM, et al. Comparison of physicians' decisions regarding estrogen replacement therapy for menopausal women and decisions derived from a decision analytic model. Am J Med 1985;80:246-8
5. Cancer Statistics, 1986. CA 1986;36:9-25
6. Cummings SR, Kelsey JL, Nevitt MC, O'Dowd KJ. Epidemiology of osteoporosis and osteoporotic fractures. Epidemiol Rev 1985;7:178-207
7. Weinstein M, Schiff T. Cost-effectiveness of hormone replacement therapy in the menopause. Obstet Gynecol Survey 1983;38:445-55
8. Holzman G, Ravitch M, Metheny W, Rotherth M, Holmes M, Hoppe R. Physicians' judgments about estrogen replacement therapy for menopausal women. Obstet Gynecol 1984;63:303-11
9. Raffia H. Decision analysis: introductory lectures on choices under uncertainty. Reading, MA: Addison-Wesley, 1968
10. Edwards W. How to use multi-attribute utility measurement for social decisions. IEEE Transactions on Systems, Man and Cybernetics 1977;(7):326-40
11. Rovner DR, Holmes MM, Rotherth ML, Elstein AS. States worse than death: premenopausal women and physicians preferences for death and disability. Med Decis Making 1985;4:541
12. Leighton-Reid J, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes comparison of assessment methods. Med Decis Making 1984;4:315-29
13. Eraker SA, Sox HC. Assessment of patient's preferences for therapeutic outcomes. Med Decis Making 1981;1:29-39
14. Weinstein MC, Feinberg HV, Elstein AS, et al. Clinical decision analysis. Philadelphia: W. B. Saunders, 1980
15. McNeil BJ, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. N Engl J Med 1982;306:1259-63