Effect of Risk Communication on Hormone Replacement Therapy Preferences

F. Reed Johnson
Senior Fellow and Principal Economist

Semra Özdemir
Health Economist

Elizabeth B. Andrews
Vice President of Pharmacoepidemiology and Risk Management

RTI Health Solutions
Research Triangle Institute
PO Box 12194
Research Triangle Park, NC 27709
frjohnson@rti.org

10/27/05

This study was funded by Wyeth Pharmaceuticals and by an unrestricted research grant from Research Triangle Institute.

RTI Health Solutions is presenting the following research study as a relevant case study because it highlights three important issues relating to the effectiveness of FDA's Center for Drug Evaluation and Research current risk communications strategy.

1) Patients and physicians can understand quantitative risk information and make informed choices. Their perspectives on willingness to trade risks for benefits of treatment should be considered in policy decisions on therapeutic guidelines, risk management programs, and in developing tools designed to communicate risks and benefits.
2) The perception of risks among patients varies with the methods of presenting the data. Relative risks focus attention on relatively large percentage changes in what may be relatively small base rates. Conversely, absolute risks focus attention on relatively small changes in the base rates themselves. It is not clear which of these approaches yields a more valid representation of patient preferences, but the mass media tends to report only relative risks. We recommend providing patients and other decision makers with full information on risks, including base rates, absolute changes, and relative changes.
3) Information on patient and physician understanding of risk levels, and willingness to trade risks for benefits can be measured in a scientifically robust manner and can be used to inform policy decisions.

The results of this study have been submitted for publication to a major medical journal.

SUMMARY

This study of women's tradeoff preferences for vasomotor symptom control versus serious adverse-event risks obtained two results of particular relevance for risk-communication interventions.

- Women in our sample were very concerned about adverse-event risks, but were willing to trade increases in risks for vasomotor symptom control if the perceived improvements are large enough.
- Women who were shown risks measured on an absolute scale had greater tolerance for adverse-event risks than women who were shown risks measured on a relative scale.

HRT physician advisories and public-information programs reflect a professional medical judgment that the risks of long-term HRT use exceed the benefits. Nevertheless, women in our sample who were well informed about possible risks, as well as personally experienced with the discomfort of vasomotor symptoms, appeared willing to accept explicit tradeoffs between therapeutic risks and benefits within the ranges observed in the Women's Health Initiative and other studies.

In addition, describing risks in different, but technically equivalent ways has differential impacts on women's willingness to trade risks for benefits. Relative risks focus attention on relatively large percentage changes in what actually are small base rates. Unfortunately, the mass media tends to report only relative risks. Providing absolute-risk information could make HRT more attractive to many women.

BACKGROUND

Menopause is defined as the permanent cessation of menses resulting from reduced ovarian hormone secretion. Vasomotor symptoms which are the most frequent symptoms of menopause are hot flashes and associated night sweats and sleep disturbances. The prevalence of these symptoms range from 14% to 51% for premenopause, around 50% for peri-menopause, and range from 30% to 80% for postmenopause (Nelson et al. 2005). Hot flashes are characterized by a feeling of pressure in the head that progresses into a flush; redness, warmth, and sweating on the face, neck, shoulders, and upper chest, sometimes followed by a slight chill; fast or pounding heartbeat. Night sweats are hot flashes that occur at night, and may lead to difficulty sleeping (NIH 2002). The severity and duration of hot flashes vary among women. Some women have hot flashes for a very short time during menopause. Other women may have hot flashes—at least to some degree—for life (Gold et al. 2004) Generally, hot flashes are less severe as time passes.

Hormone replacement therapies (HRT) are effective in reducing the incidence and duration of vasomotor symptoms. However, HRTs may have adverse effects. In 2002 results from the Women's Health Initiative (WHI) indicated significantly higher risks of several serious adverse events in women using estrogen plus progestin relative to placebo and the trial was terminated. Table 1 compares the observed number of cases, the absolute risks, and the relative risks of HRT. Observed incidence of hip fracture were lower for HRT relative to placebo (NIH 2002; Rossouw et al. 2002; Burkman et al. 2001).

The National Institutes of Health (NIH) issued a physician advisory and the Food and Drug Administration (FDA) initiated a public information program to publicize these results to the general population. The information was widely reported in the media. Most of the reports cited the relative risks rather than number of cases or absolute risks. Relative risk levels lead women to misinterpret the adverse effects of HRT (Levens and Williams 2004). A subsequent trial involving estrogen alone also was terminated because of elevated incidence of cognitive deficits in the treatment arm.

Table 1: Hazard Ratio and Absolute Risk Seen in the HT Substudy of the WHI – Centrally Adjudicated Results

Event	Hazard Ratio HT vs placebo at 5.6 Years (95% CI*)	HT n = 8506	Placebo n = 8102	Relative Risk per 10,000 person-years
		Absolute Risk per 10,000 person-years		
CHD events[a,1]	1.24 (1.00-1.54)	39	33	0.18
Non-fatal MI	1.28 (1.00-1.63)	31	25	0.24
CHD death	1.10 (0.70-1.75)	8	8	0.00
Invasive breast cancer[b,2]	1.24 (1.01-1.54)	41	33	0.24
Hip Fracture[3]	0.67 (0.47-0.96)	11	16	-0.31

[a] CHD events included acute MI, silent MI, and coronary death.
[b] Includes metastatic and non-metastatic breast cancer with the exception of in situ breast cancer.
* Nominal confidence intervals unadjusted for multiple looks and multiple comparisons.
[1] Manson JE, Hsia J, Johnson KC, et al. 2003.
[2] Chlebowski RT, Hendrix SL, Langer RD, et al. 2003.
[3] Cauley JA, Robbins J, Chen Z, et al. 2003.

The WHI researchers concluded that the overall health risks are higher than the benefits. However, this study did not account for control of vasomotor symptoms, and the main reason women take HRT is to relieve vasomotor symptoms related to menopause. Some women may be willing to accept some additional risks in return for the benefits of vasomotor symptom control. In this study we quantify women's stated preferences for HRT using conjoint-analysis (CA) methods. Specifically, the objectives of the study are:

- To estimate women's willingness to trade risk for vasomotor symptom relief.
- To test whether the presentation of risk attributes in absolute-risk and relative-risk scales affect women's preferences.

Because CA yields quantitative estimates of trade-offs among treatment attributes, we can calculate equivalences between subjects perceived benefits and risks to identify the associated maximum acceptable risk. We calculate maximum acceptable risk for both the relative risk and

absolute risk subsamples and test whether stated preferences are different as a result of how risks are described.

SURVEY DESIGN AND METHODS

Regulators' treatment-approval decisions and physicians' and subjects' treatment decisions inevitably require explicit or implicit trade-offs between the potential efficacy of an intervention and the risk of side effects. RTI-HS developed, administered, and analyzed a CA survey to elicit women's preferences for a range of treatment attributes, including risks of serious adverse events, associated with medication and therapy for relief of vasomotor symptoms. We administered two versions of the survey, one that employed a description using absolute risks, and one that employed a description using relative risks.

CA is a technique specifically designed to provide useful and appropriate information about individuals' desires to make trade-offs between attributes of multi-attribute products. CA is based on the hedonic principal that products are composed of a set of various attributes and that the attractiveness of a product to an individual is a function of these attributes. People's relative preferences among product attributes and levels vary, and, thus, they are willing to accept trade-offs among them.[1]

CA methods recognize that products have value because of their characteristics or attributes. People have preferences for each attribute and are willing to accept tradeoffs among them. CA studies examine these tradeoffs to quantify the implicit weights people assign to various treatment attributes. These weights represent the impact of the different treatment attributes on subject's overall wellbeing and are likely to predict their satisfaction with treatments that include different mixes of attributes. Analysts have used CA to quantify preferences for a variety of market and nonmarket goods and services. These include medical interventions, pharmaceutical treatments, and environmental health risks (Gan et al. 2004, Bryan et al. 1998; Johnson et al. 2000; Johnson et al. 1998; Johnson and Desvousges 1997; Ryan and Hughes 1997; Viscusi, Magat, and Huber 1991; Wittink and Cattin 1989).

Understanding various dimensions of subject preferences requires a valid and reliable measurement method. Conjoint analysis (CA) data can be validated using a series of internal validity tests (Johnson et al. 2000; Ryan et al.1998). CA encourages subjects to explore their

preferences for various attribute combinations through a series of questions that require choosing among treatments with different features. This process of explicitly trading off attributes encourages subject introspection. Because each subject provides answers to multiple tradeoff questions, CA allows analysts to devise internal checks for attentiveness and consistency. We thus conducted several internal validity tests to check whether subjects in our sample were consistent in their stated preferences.

Implementing a valid and reliable CA study requires accurate treatment-feature definitions (attributes and levels),[2] attention to task format, careful pretesting, efficient experimental design, and appropriate statistical analysis. The remainder of this paper considers each of these elements of the study.

Survey Design

Attributes and Levels

Based on discussions with medical experts and evaluation of pretest results, we limited the tradeoff attributes to the most salient treatment features, including:

- severity and the frequency of daytime hot flashes;
- frequency of night sweats;
- duration of symptoms; and
- 10-year serious adverse-event risks, including risk of hip or back fractures, risk of heart attack and risk of breast cancer.

Table 2 lists the treatment attributes and levels used in the survey. We describe fracture risk as a reduction in the likelihood of a hip or back fracture, consistent with known osteoporosis benefits. The heart-attack and breast-cancer risks are described as decreases (based on previous evidence) and increases (based on more recent evidence). Finally, the same risk levels were described as relative risk and absolute risk in separate versions of the conjoint instrument.

[1] Attributes are generic product characteristics, such as color. Color may take on several possible levels, such as blue, yellow, and green.

[2] An attribute is a qualitative characteristic of the treatment, while a level is one of several values the attribute may have. Color and price are attributes. Blue and $25 are levels.

Table 2: Treatment Attributes and Levels

Treatment Feature	Levels
Severity of Daytime Hot Flashes	▪ No daytime hot flashes ▪ Mild: a fleeting warm sensation with no sweating that does not disrupt normal daily activity ▪ Moderate: a warm sensation with sweating that does not disrupt normal daily activity ▪ Severe: a hot sensation with sweating that can disrupt normal daily activity
Frequency of Daytime Hot Flashes	▪ None (0 times) during the daytime ▪ 1–2 times during the daytime ▪ 3–6 times during the daytime ▪ More than 6 times during the daytime
Frequency of Night Sweats	▪ None (0 times) per night ▪ 1–3 times per night ▪ 4 or more times per night
Duration of Hot Flashes and Night Sweats	▪ 1 year ▪ 2 years ▪ 4 years ▪ 7 or more years
Risk of Hip or Back Fracture within 10 years	▪ 15/1,000 (1.5%) or 50% reduction ▪ 30/1,000 (3%) or no change
Risk of Heart Attack within 10 years	▪ 38/1,000 (3.8%) or 25% reduction ▪ 50/1,000 (5%) or no change ▪ 65/1,000 (6.5%) or 30% increase
Risk of Breast Cancer within 10 years	▪ 23/1,000 2.3%) or 25% reduction ▪ 30/1,000 (3%) or no change ▪ 39/1,000 (3.9%) or 30% increase

Experimental Design

The experimental design consists both of combinations of attributes and levels that describe a set of hypothetical treatment profiles and pairings of profiles that provide comparison sets for the tradeoff tasks. Most current CA applications use a D-optimal design to reduce the number of paired comparisons to the smallest number necessary for efficient estimation of preference weights (Dey, 1985; Huber and Zwerina, 1996; Kuhfeld, Tobias, and Garratt, 1994). Such efficient designs can be produced using an iterative computer algorithm (Zwerina, Huber, and Kuhfeld, 1996). We employed our own implementation of a D-optimal algorithm to search for a near-optimal experimental design of 3 blocks of 9 treatment pairs. Each subject was randomly assigned one of the design blocks. The sequence of tradeoffs also was randomly determined and the first two questions were repeated at the end. This procedure allows us to use the first two questions as a "warm-up", delete them from the estimation, and use the repeated questions for a consistency test.

Before the conjoint tasks we provided subjects with basic information about menopause, vasomotor symptoms, and possible treatment risks. We then asked about their own experience with vasomotor symptoms and perceived fracture, heart-attack, and cancer risks. The introduction to the tradeoff questions asked subjects to assume they were experiencing severe and frequent vasomotor symptoms. In each tradeoff question they were asked rate their preference for two hypothetical treatment alternatives. Figures 1 and 2 present example tradeoff tasks for the absolute and relative versions, respectively.

Figure 1: Tradeoff Task for Absolute-Risk Version

Considering the different results and risks associated with Treatments A and B, which would you prefer if these were the only options available?

	Results of Treatment A	Results of Treatment B
Intensity of daytime hot flashes	Mild	Severe
Frequency of daytime hot flashes	1 – 2 times a day	More than 6 times a day
Frequency of night sweats	None	1 – 3 times a night
Duration of hot flashes and night sweats	7 years	1 year
Risk of hip or back fractures within 10 years	15/1,000 (1.5%)	15/1,000 (1.5%)
Risk of heart attack within 10 years	38/1,000 (3.8%)	65/1,000 (6.5%)
Risk of breast cancer within 10 years	30/1,000 (3%)	23/1,000 (2.3%)

Check the box that best describes your opinion	☐	☐	☐	☐	☐
	A is much better	**A is somewhat better**	**A and B are the same**	**B is somewhat better**	**B is much better**

Figure 2: Tradeoff Task for Relative-Risk Version

Considering the different results and risks associated with Treatments A and B, which would you prefer if these were the only options available?

	Results of Treatment A	Results of Treatment B
Intensity of daytime hot flashes	Mild	Severe
Frequency of daytime hot flashes	1 – 2 times a day	More than 6 times a day
Frequency of night sweats	None	1 – 3 times a night
Duration of hot flashes and night sweats	7 years	1 year
Risk of hip or back fractures within 10 years	↓ 50% **decrease** in risk	↓ 50% **decrease** in risk
Risk of heart attack within 10 years	↓ 25% **decrease** in risk	↑ 30% **increase** in risk
Risk of breast cancer within 10 years	No change in risk	↓ 25% **decrease** in risk

Check the box that best describes your opinion	☐	☐	☐	☐	☐
	A is much better	**A is somewhat better**	**A and B are the same**	**B is somewhat better**	**B is much better**

Survey Administration and Subject Sample

After pretesting the instrument, the web-enabled survey instrument was administered to 523 US women between the ages of 46 and 60 drawn from the Harris Interactive Consumer Panel. Relative to the general population, Table 3 indicates the sample has a lower representation of racial and ethnic minorities, with 89% of the sample being white or Caucasian. The median years of education is 13 years. Sample incomes are below the national mean and median of $59,000 and $43,000, probably because the sample consists of older-age women.

Table 3. Socio-Economic Characteristics

Subject Characteristic	Distribution
Race	89% White 11% Other
Employment	41% full employed 15% part-time employed
Age	Mean = 52 Median = 52
Years of Education	Mean = 14 years Median = 13 years
Annual Income	Mean = $55,000 Median = $42,500

Table 4 summarizes background information on subjects' own experience with menopausal symptoms and perceived risks. 26% of the subjects have not experienced menopause or they are not sure. Of the remainder, the majority experienced moderately severe symptoms and hot flashes 1 to 2 times a day and night sweats 1 to 3 times a night. Most women in the sample reported menopausal symptoms for 3 years or less. About the half of women think their personal risks of fractures, heart attack, and breast cancer are the same as population risks.

Table 4. Subjects' Current Treatment

Treatment Attribute	Distribution
Experience with menopause	16% Have never experienced 27% Experienced in the past 47% Currently experience 10% Not sure
Intensity of daytime hot flashes	20% Mild 60% Moderate 20% Severe
Frequency of daytime hot flashes	56% 1-2 times a day 31% 3-6 times a day 13% More than 6 times a day
Frequency of night sweats	83% 1-3 times a night 17% 4 or more times a night
Duration of hot flashes and/or night sweats (past experience)	40% Less than 1 year 40% 1-3 years 6% 3-5 years 8% 5-7 years 6% More than 7 years
Duration of hot flashes and/or night sweats (current experience)	18% Less than 1 year 38% 1-3 years 14% 3-5 years 12% 5-7 years 18% More than 7 years
Perceived risk of getting hip or back fractures compared to other women at same age	3% Much higher 12% Somewhat higher 48% About the same 23% Somewhat lower 14% Much lower
Perceived risk of getting heart attack compared to other women at same age	5% Much higher 28% Somewhat higher 48% About the same 14% Somewhat lower 5% Much lower
Perceived risk of getting breast cancer compared to other women at same age	5% Much higher 12% Somewhat higher 56% About the same 17% Somewhat lower 10% Much lower

Model Estimation

We employ an ordered-probit procedure that is appropriate for the naturally ordered rating variable. This approach does not require that the distances between rating categories be equal. The estimates also account for our having multiple ratings for each subject.

Maximum Acceptable Risk

If there is a nonzero risk of a serious adverse event, subjects' expected level of wellbeing or expected conjoint utility from treatment is

$$EU = P_{AE} \cdot U_{AE} + U_B \qquad (1)$$

Where EU is expected conjoint utility, P_{AE} is the probability of the adverse event, U_{AE} is the conjoint utility of the adverse event, and U_B is the conjoint utility of the treatment benefit. Maximum acceptable risk simply is the probability P^*_{AE} that makes EU equal to the pretreatment conjoint utility U_o.

$$U_O = EU^* = P^*_{AE} \cdot U_{AE} + U_B \qquad (2)$$

Any adverse-event probability less than P^*_{AE} makes EU greater than the health outcome without treatment, so perceived treatment benefits are greater than the perceived adverse-event risks. Substituting into Equation (1) and solving,

$$P^*_{AE} = \frac{U_B - U_o}{-U_{AE}} \qquad (3)$$

We use Equation (3) to derive maximum acceptable risk from the estimated conjoint satisfaction weights. Appendix B contains additional details on the conceptual derivation of maximum acceptable risk.

RESULTS

Internal Validity

Because CA subjects complete a series of tradeoff tasks, we are able to devise various internal checks for attentiveness and consistency. For example, stability of preferences requires that if subjects prefer treatment A to treatment B at one point in the sequence of choice questions, then they should prefer A to B at any subsequent point. About 1/3 of the sample failed at least one of several consistency tests. However, we found no significant biases in preference estimates when we included subjects who failed a test. Thus the results reported here include all subjects who completed the survey.

Satisfaction Weights

The rescaled ordered probit estimates and 95% confidence intervals for severity and frequency of daytime hot flushes and night sweats and for risk of fractures, heart infarct, and breast cancer, are presented graphically in Figures 3 and 4, respectively. These figures depict satisfaction weights for both the absolute and relative risk versions of the survey. The length of the line segments between points in Figures 3 and 4 indicate improvements in satisfaction. For example, the line segments between Severe and Moderate symptom severity indicate the corresponding improvement in satisfaction score is twice as large for subjects who evaluated absolute risks as for subjects who evaluated relative risks ((62.4-38.5)/(57.8-45.6) = 1.98).

Figure 3 indicates that estimated satisfaction weights are consistent with the natural ordering of the categories, so no symptoms or milder symptoms and fewer symptoms of each type generally have significantly higher satisfaction weights than less attractive outcomes. Similarly, Figure 4 indicates that estimated satisfaction weights penalize higher risk levels. These results indicate that subjects understood the natural ordering of attribute levels and could discriminate effectively among them. The only exception was that the current level of absolute breast-cancer risk was slightly more preferred than the 25% reduction in risk. However, the difference was statistically insignificant and we combined these 2 levels into a single level for subsequent analysis.

The contribution of different attributes to overall well-being is indicated by the difference in the satisfaction scale between the best and worst attribute levels. For example, the attribute with the smallest overall effect on satisfaction is the frequency of night sweats. The distance

Figure 3. Estimated Satisfaction Weights for Severity and Frequency of Vasomotor Symptoms

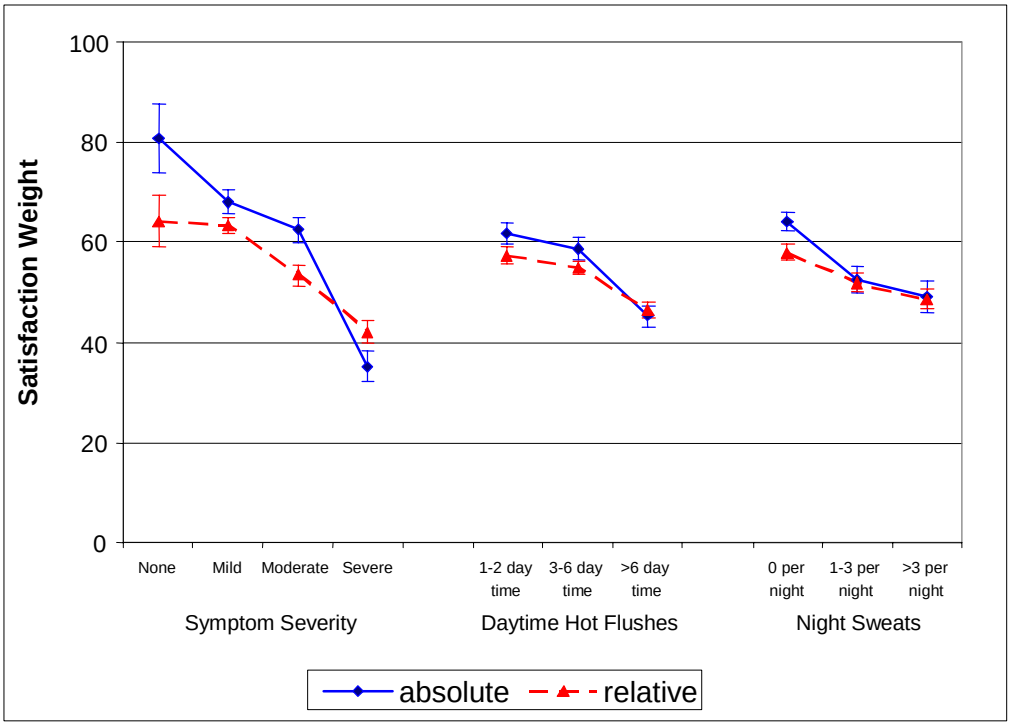
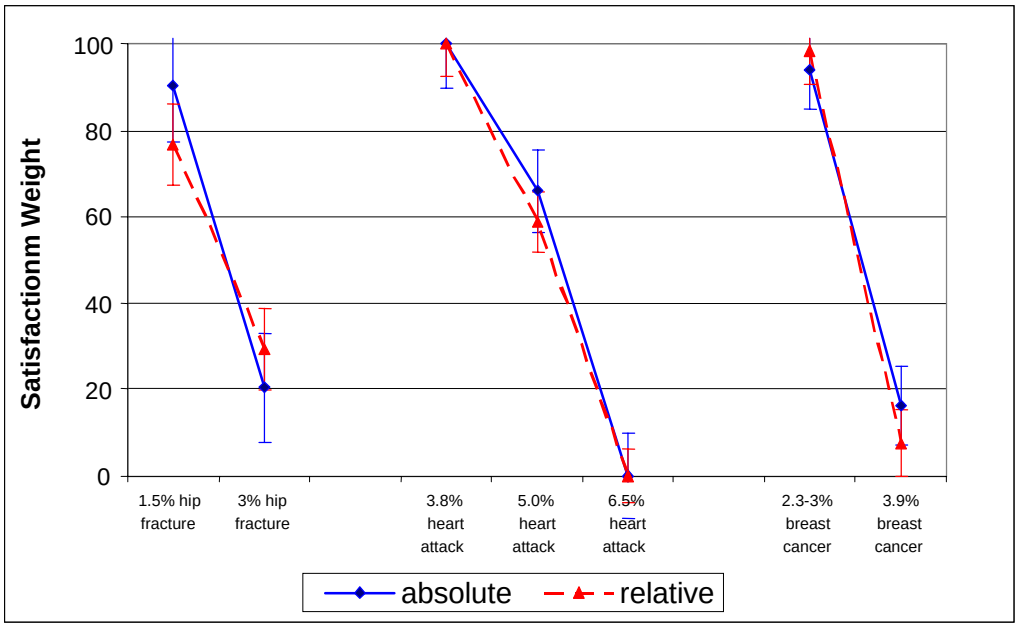


Figure 4. Estimated Satisfaction Weights for Risk Factors



between best and worst levels for this attribute is about 15 in the absolute risk version and 9 in the relative risk version. The attribute with the greatest effect on satisfaction is the risk of heart infarct, with a distance between best and worst levels of about 100, more than 6 times larger than the range for frequency of night sweats.

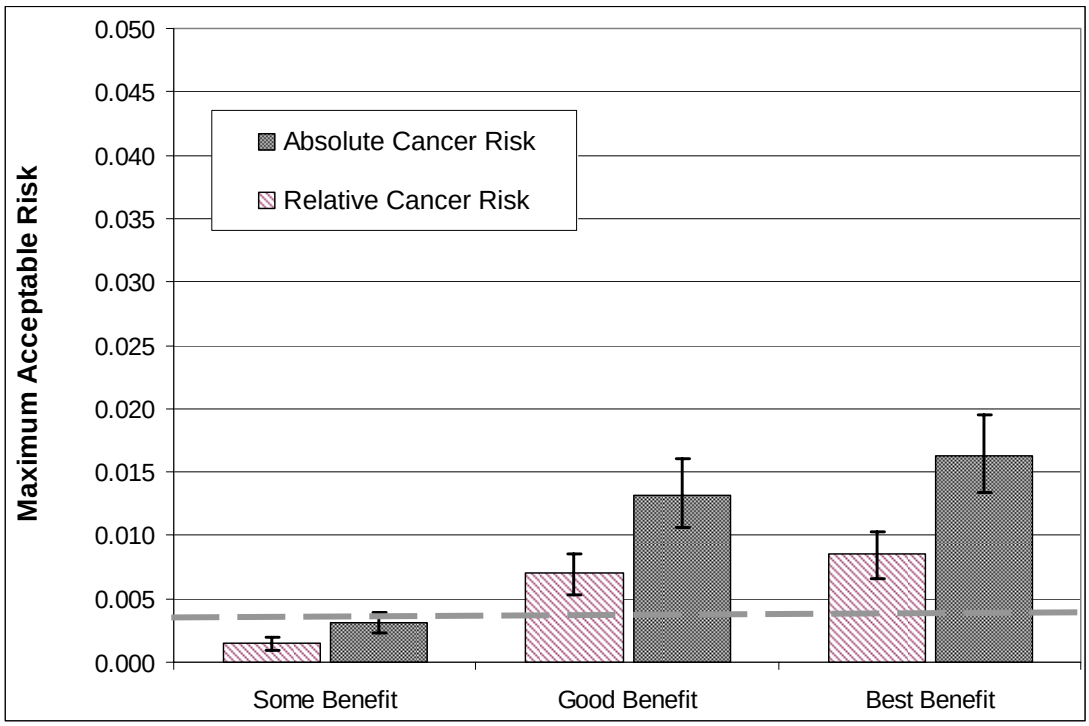
Preferences are generally similar between the absolute and relative risk versions, except for the important interaction between symptom severity and duration. In this case, the subjects who evaluated absolute risks perceived symptom relief as having a significantly greater impact on satisfaction relative to risk attributes than subjects who evaluated relative risks. Specifically, the difference between the "none" and "severe" weights is nearly twice as large for the absolute risk version as for the relative risk version (42 versus 22, $p<0.001$). This difference affects the degree to which women are willing to trade off symptom relief for potential adverse events. Women presented with absolute risks tended to be relatively less concerned about adverse-event risks compared to symptom relief. This is not surprising since risks expressed as percentage differences seem large compared to risks expressed as absolute decimals when base-rate risk is small.

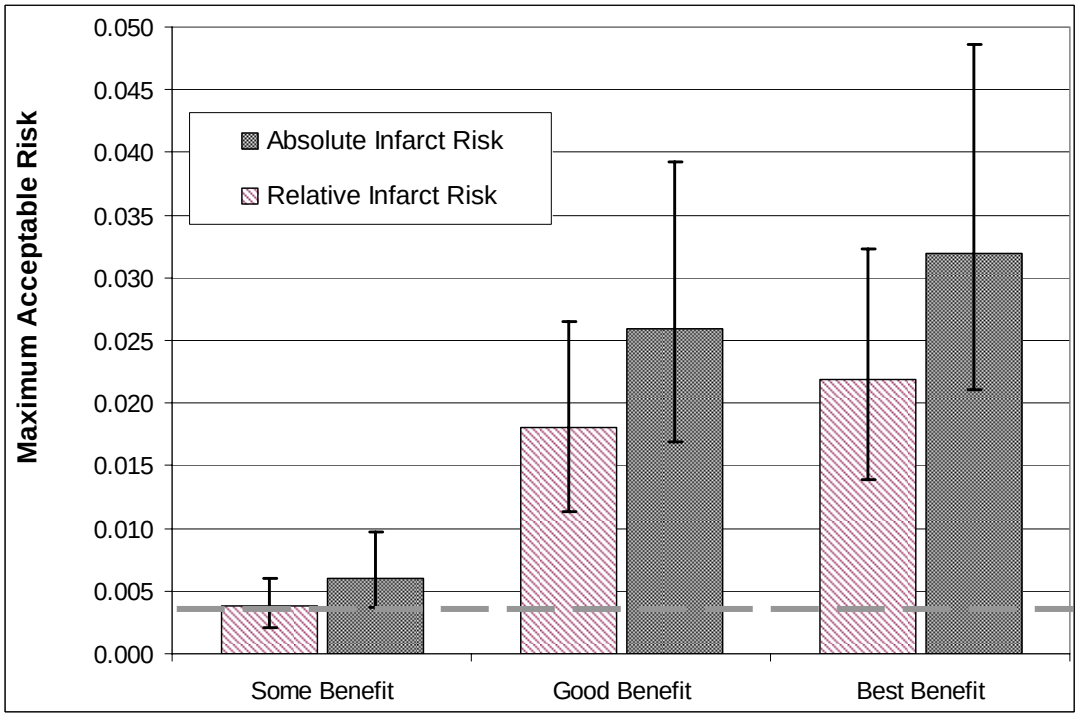
We also examined whether preferences differed between younger and older women. There were no differences in satisfaction weights between younger and older women for vasomotor symptoms or duration, but older women were more concerned about fracture and breast cancer risks than younger women ($p=0.004$).

Maximum Acceptable Risk

Figure 5 presents MAR values for particular improvements in health status. As expected, women in our sample were willing to accept higher levels of risk in return for greater improvements in symptom relief. However, those women who evaluated relative risks had lower MAR values than women who evaluated absolute risks. For example, an improvement in health status from "severe symptoms" to "full symptom relief" (best benefit) was associated with a mean MAR for breast cancer of 0.009 when risk was given in relative terms and 0.016 when risk was given in absolute terms ($p=0.002$). The differences in mean MAR values between absolute and relative risk versions were statistically significant across benefit levels for breast cancer, but not for heart infarct.

Figure 5. Mean Maximum Acceptable Risks for Breast Cancer and Heart Infarct for Women Who Have Severe Vasomotor Symptoms Before Treatment





Note: The horizontal line in each graph corresponds to actual event risks as reported by WHI trial investigators

Subjects were more willing to accept increased heart infarct risk compared to increased breast cancer risk for all three levels of symptom improvement (none to maximum relief, none to some relief, and some relief to maximum relief). This pattern was observed in both the absolute risk version ($p = 0.028$, 0.029, and 0.049 for the three improvement levels, respectively) and the relative risk version ($p = 0.007$, 0.008, 0.026, respectively). Overall, the MAR estimates are 0.002 to 0.016 points lower for breast cancer risks than for heart infarct risks. Nevertheless, even for relative cancer risk, MAR exceeds the actual risk as reported by WHI investigators for the two larger benefit levels ($p = 0.001$ and 0.001, respectively). MAR values for the smallest benefit level were not significantly greater than the larger benefit levels in either the absolute or relative risk version.

DISCUSSION

Our study of women's tradeoff preferences for vasomotor symptom control versus serious adverse-event risks obtained several results of interest.

- Most women in our sample provided informative answers to the tradeoff tasks. Controlling for observed consistency errors resulted in no significant biases in our estimates.
- Women in our sample were very concerned about adverse-event risks. They were more concerned about breast-cancer risks than about heart-attack risks, but were willing to trade increases in risks for vasomotor symptom control if the perceived improvements are large enough.
- Women who traded off adverse-event risks measured on an absolute scale had larger maximum acceptable risks than women who traded off risks measured on a relative scale.

While many pharmaceuticals and medical devices have demonstrated clinical value in alleviating symptoms of disease, such benefits often are accompanied by risks of adverse events. Risk managers often must weigh the potential risks of medical interventions to a small number of patients against the potential benefits of the same interventions to a large number of patients. Risk assessments generally focus primarily on clinical outcomes. However, effective risk management also requires understanding the behavioral context in which adverse events occur. Problems with adherence to therapeutic guidelines can occur when there are systematic differences between physicians' and patients' perceptions, and regulators' explicit or implicit judgments regarding relative risks and benefits.

The HRT physician advisories and public-information programs reflect a professional medical judgment that the risks of long-term HRT use exceed the benefits. Physicians, pharmaceutical regulators, and pharmaceutical companies traditionally have been reluctant to accept patients' own assessment of acceptable risk-benefit tradeoffs. This reluctance presumably stems from doubts about whether patients are sufficiently knowledgeable to make such judgments and concerns about potential legal liabilities associated with severe adverse events. Nevertheless, women in our sample who were well informed about possible risks, as well as personally experienced with the discomfort of vasomotor symptoms, appeared willing to accept explicit tradeoffs between risks within the ranges observed in the WHI and clinically realistic health-state improvements.

There are several important implications of our results. First, describing risks in different, but technically equivalent ways has differential impacts on women's willingness to trade risks for benefits. Relative risks focus attention on relatively large percentage changes in what may be relatively small base rates. Conversely, absolute risks focus attention on relatively small changes in the base rates themselves. It is not clear which of these approaches yields a more valid representation of patient preferences, but the mass media tends to report only relative risks. We recommend providing patients and other decision makers with full information on risks, including base rates, absolute changes, and relative changes.

The second important implication of these results is that the apparent discrepancy between HRT therapeutic guidelines and the risk-benefit preferences of women in our sample raises questions about the validity of the guidelines themselves. Health scientists are trained to collect and analyze data to detect statistically significant effect differences in medical treatments. It is not clear this expertise extends to subjective judgments about relative values. In many other areas of public policy, the preferences of affected members of the public often weigh more heavily than they do in the determination of therapeutic guidelines. For example, federal agencies such as the Occupational Safety and Health Administration must invite and respond in writing to public comments on all new regulations governing workplace safety.

Finally, effective risk management requires controls that ensure that physicians and patients adhere to strategies designed to reduce risks. Such strategies are likely to be less effective if patients' preferences are inconsistent with the strategy. If patients' risk preferences differ from those of risk managers, it may be necessary to impose physical restrictions on pharmaceutical access and use. Such restrictions will improve adherence, but also are likely to result in patient

dissatisfaction and potential conflicts between regulatory authorities and patient advocacy groups.

REFERENCES

Bryan S., Martin B., Robert S., and Alison G. (1998). Magnetic Resonance Imaging for the Investigation of Knee Injuries: An Investigation of Preferences. *Health Economics* 7: 595 – 603.

Burkman R.T., Collins J.A., Greene R.A. (2001). Current Perspectives on Benefits and Risks of Hormone Replacement Therapy. *American Journal of Obstetrics and Gynecology* 188 (2 Suppl): S13 – 23.

Cauley JA, Robbins J, Chen Z, et al. Effects of estrogen plus progestin on risk of fracture and bone mineral density: the Women's Health Initiative randomized trial. *JAMA* 2003; 290(13):1729-1738.

Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in health postmenopausal women. The Women's Health Initiative Randomized Trial. *JAMA*. 2003; 289(24):3243-3253.

Dey A. (1985). *Orthogonal Fractional Factorial Designs*. New York: Halstead Press.

Gan T. J., Lubarsky D. A., Flood E. M., et al. (2004). Patient Preferences for Acute Pain Treatment. *British Journal of Anesthesia* 92 (5): 681- 688.

Gold E.B., Block G., Crawford S., Lachance, L., FitzGerald G., Miracle H., and Sherman S. (2004) Lifestyle and Demographic Factors in Relation to Vasomotor Symptoms: Baseline results from the Study of Women's Health Across the Nation. *American Journal of Epidemiology* 159: 1189 – 1199.

Huber J. and Zwerina K. (1996). The Importance of Utility Balance in Efficient Choice Designs*. Journal of Marketing Research* 33: 307-317.

Johnson F.R. and Matthews K.E. (2001). Sources and Effects of Utility-Theoretic Inconsistencies in Stated-Preference Surveys. *American Journal of Agricultural Economics* 83(5):1328-1333.

Johnson F.R., Banzhaf M., Desvousges W. (2000). Willingness to Pay for Improved Respiratory and Cardiovascular Health: A Multiple-Format Stated-Preference Approach. *Health Economics* 9: 295–317.

Johnson F.R. and Desvousges W.H. (1997). Estimating Stated Preferences With Rated-Pair Data: Environmental, Health, and Employment Effects of Energy Programs. *Journal of Environmental Economics and Management* 34: 79–99.

Kuhfeld W., Tobias F., Garratt M. (1994). Efficient Experimental Design with Marketing Research Applications. *Journal of Marketing Research* 31: 545 - 557.

Levens E., Williams R. S. (2004). Current Opinions and Understanding of Menopausal Women about Hormone Replacement Therapy (HRT) – The University of Florida Experience. *American Journal of Obstetrics and Gynecology* 191: 641 – 647.

Manson JE, Hsia J, Johnson KC, et al. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med 2003; 349(6):523-534.

National Institutes of Health, National Heart, Lung, and Blood Institute. (2002). *Information for the Public: Facts About Postmenopausal Hormone Therapy*, NIH Publication No. 02-5200.

Nelson H.D., Haney E., Humphrey L., et al. (2005). Management of Menopause-Related Symptoms. Summary, Evidence Report/Technology Assessment No. 120. AHRQ Publication No. 05-E016-2. Rockville, MD: Agency for Healthcare Research and Quality.

Rossouw J.E., Anderson G.L., Prentice R.L., et al. (2002). Risks and Benefits of Estrogen plus Progestin Healthy Postmenopausal Women: Principal Results from the Women's Initiative Randomized Controlled Trial. Writing Group for the Women's Health Initiative. *Journal of American Medical Association* 288: 321 – 333.

Ryan M., and Hughes J. (1997). Using Conjoint Analysis to Assess Women's Preferences for Miscarriage Management. *Health Economics* 6: 261 – 273.

Strine T.W., Okoro C.A., Chapman D.P., Balluz L.S., Ford E.S., Ajani U.A., Mokdad A.H. (2005) Health-Related Quality of Life and Health Risk Behaviors Among Smokers. *American Journal of Preventive Medicine* 28(2): 182- 187.

Viscusi W., Magat W., Huber J. (1991). Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-offs for Chronic Bronchitis. *Journal of Environmental Economics and Management* 21: 32 – 51.

Wittink D.R., and Cattin P. (1989). Commercial Use of Conjoint Analysis: An Update. *Journal of Marketing* 53: 91–96.

Zwerina K., Huber J., Kuhfeld W. (1996). *A General Method for Constructing Efficient Choice Designs*. Fuqua School of Business, Duke University.