Background: Hormone therapy (HT) provides the most effective relief of menopausal symptoms. This therapy is associated with a decreased risk of osteoporosis and colorectal cancer but increased risks of cardiovascular disease (CVD), venous thrombosis, and breast cancer. Our objective was to identify which women should benefit from short-term HT by exploring the trade-off between symptom relief and risks of inducing disease.

Methods: A Markov model simulates the effect of short-term (2 years) estrogen and progestin HT on life expectancy and quality-adjusted life expectancy (QALE) among 50-year-old menopausal women with intact uteri, using findings from the Women’s Health Initiative. Quality-of-life (QOL) utility scores were derived from the literature. We assumed HT-affected QOL only during perimenopause, when it reduced symptoms by 80%.

Results: Among asymptomatic women, short-term HT was associated with net losses in life expectancy and QALE of 1 to 3 months, depending on CVD risk. Women with mild or severe menopausal symptoms gained 3 to 4 months or 7 to 8 months of QALE, respectively. Among women at low risk for CVD, HT extended QALE if menopausal symptoms lowered QOL by as little as 4%. Among women at elevated CVD risk, HT extended QALE only if symptoms lowered QOL by at least 12%.

Conclusions: Hormone therapy is associated with losses in survival but gains in QALE for women with menopausal symptoms. Women expected to benefit from short-term HT can be identified by the severity of their menopausal symptoms and CVD risk.

Arch Intern Med. 2004;164:1634-1640

Decisions concerning menopausal hormone therapy (HT) remain difficult. Although the Women’s Health Initiative (WHI) trial has elucidated the impact of HT on many clinical outcomes, the complexity of balancing HT’s many risks and benefits has increased because more clinical end points must be considered. Although HT is the most effective treatment for menopausal symptoms and decreases risks of osteoporosis and colorectal cancer, HT increases risks of coronary heart disease (CHD), stroke, pulmonary embolism (PE), and breast cancer.

CME course available at www.archinternmed.com

For most women, the net risks of HT appear to outweigh its benefits, although the average difference is modest. For 10000 women in the WHI, 1 year of HT is expected to lead to 7 additional CHD events, 8 more strokes, 8 more PEs, and 8 more breast cancers, and 6 fewer colorectal cancers and 5 fewer hip fractures. A global index based on these end points suggested net harm for HT use. Although the absolute risk to healthy participants was low, the benefit-risk profile was considered inappropriate for primary prevention of chronic disease. However, the balance of benefits and risks will vary according to baseline risk and may vary when HT’s impact on menopausal symptoms and quality of life (QOL) is taken into account.

Because the treatment-associated risks for CHD, stroke, PE, and breast cancer increase with duration of HT, short-term HT (<5 years) is being considered for menopausal symptoms (eg, by The American College of Obstetricians and Gynecologists and the North American Menopause Society). However, since HT increased the risks of CHD, stroke, and PE within the first 1 to 2 years, even short-term use is not without potential danger. Considerations for QOL take on paramount importance in decisions about HT. Symptoms usually persist for more than...
1 year but abate within 2 to 3 years, substantially lowering QOL, with substantial improvement after HT.

This study, by exploring the trade-off between short-term symptomatic relief and risks of chronic disease, sought to determine which women might benefit from HT.

### THE DECISION MODEL

We developed a Web-based Markov simulation model based on published modeling techniques. The model simulates the impact of short-term HT (vs no HT) on estrogen deficiency symptoms as well as on life expectancy, quality-adjusted life expectancy (QALE), and clinical end points affected by HT. We modeled the impact of combination estrogen and progestin HT among a healthy cohort of 50-year-old white menopausal women with intact uteri. In the base-case scenario, we assumed persistence of perimenopause for 2 years and HT use for 2 years. In sensitivity analyses, we considered other durations of symptoms and treatment. Baseline results assume a 0 discount rate.

### MARKOV MODEL

The model simulates lifetime incidence of breast, colorectal, ovarian, and endometrial cancer, CHD, stroke, hip fracture, and PE and considers women without evidence of these diseases at baseline. With each year, cohort members can develop any one or a combination of these diseases or die of other causes, as in the general population.

### DISEASE INCIDENCE

We linked individual risk factors to future disease incidence for breast cancer, CHD, and stroke. Breast cancer risk was based on the model by Gail et al., adjusting baseline incidence rates according to the more representative Surveillance, Epidemiology, and End Results (SEER) data. Stroke and CHD risks were calculated from equations from the Framingham Study.

No composite risk models link risk factors to future risk levels for endometrial, colorectal, and ovarian cancer, hip fracture, and PE. These risks were based on age, sex, and race-adjusted incidence rates.

### RISK STRATIFICATION

We stratified analyses by factors with the greatest impact on outcomes: cardiovascular disease (CVD) (both CHD and stroke) and severity of menopausal symptoms. Low CVD risk was defined as the absence of any risk factors: systolic blood pressure of 120 mm Hg, total serum cholesterol level of 180 mg/dL (4.7 mmol/L), and high-density lipoprotein cholesterol level of 65 mg/dL (1.7 mmol/L). High risk was defined as the presence of 3 risk factors: hypertension (systolic blood pressure of 150 mm Hg), hypercholesterolemia (total cholesterol level of 260 mg/dL [6.7 mmol/L] and high-density lipoprotein level of 35 mg/dL [0.9 mmol/L]), and smoking. Average women were defined as having population averages for each CVD and breast cancer risk factor for women in the Framingham Study and the model by Gail et al.

### IMPACT OF HT

Effects of HT on chronic disease end points were based on data from the WHI (Table 1), using time-dependent risk functions reflecting 5.2 years of follow-up. In sensitivity analyses, we applied estimates from 2 meta-analyses. Hormone therapy appeared to decrease the risk of breast cancer during the first 2 years, likely reflecting a masking by HT of its early detection (among HT users, detected breast cancers were larger and more likely to be node positive). We therefore assumed a linear risk function, interpolating from the final relative risk. In sensitivity analyses, we assumed 4% higher breast cancer mortality rates among those who developed breast cancer while receiving HT. We assumed that HT had no impact on endometrial or ovarian cancer. Because several observational trials reported that HT increased the risk of ovarian cancer, in sensitivity analyses we applied the relative risk of 1.51.

### MORTALITY ESTIMATES

For each disease, we modeled short-term (first year) and long-term (after the first year) disease-specific mortality. We calculated disease-specific mortality rates according to age and time since diagnosis from SEER survival data (Table 2), using cancer-associated deaths, and drew hip fracture mortality from a large population database. The mortality rates among patients surviving the first year after hip fracture approach those of age-matched individuals without hip fracture. We based CHD mortality on data from the Minnesota Heart Study and stroke mortality on data from the Framingham Study. We assumed 9% higher CVD mortality rates for black women compared with white women after adjusting for differences in income. All-cause mortality rates were based on recent age- and race-specific life tables for women, subtracting mortality from disease explicitly modeled to avoid double counting.

### QUALITY OF LIFE

Quality adjustment of life expectancy considers not only length of life but the QOL of the extended period and is expressed in

<table>
<thead>
<tr>
<th>Outcome</th>
<th>WHI, HR (95% CI)</th>
<th>Nelson et al, Meta-analysis, RR (95% CI)</th>
<th>Beral et al, Meta-analysis, RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>1.29 (1.02-1.63)</td>
<td>0.97 (0.82-1.16)</td>
<td>1.11 (0.96-1.3)</td>
</tr>
<tr>
<td>Pulmonary embolus</td>
<td>2.13 (1.39-3.25)</td>
<td>2.14 (1.64-2.81)</td>
<td>2.16 (1.47-3.18)</td>
</tr>
<tr>
<td>Stroke</td>
<td>1.41 (1.07-1.85)</td>
<td>1.12 (1.01-1.23)</td>
<td>1.27 (1.06-1.51)</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>1.24 (1.01-1.54)</td>
<td>1.21 (1.40)</td>
<td>1.27 (1.03-1.56)</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>0.83 (0.47-1.47)</td>
<td>0.8 (0.6-1.2)</td>
<td>0.76 (0.45-1.31)</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>0.63 (0.45-0.92)</td>
<td>0.64 (0.59-0.74)</td>
<td>0.72 (0.52-0.98)</td>
</tr>
<tr>
<td>Hip fracture</td>
<td>0.66 (0.45-0.98)</td>
<td>0.64 (0.32-1.04)</td>
<td>0.64 (0.32-1.04)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; CHD, coronary heart disease; HR, hazard ratio; HT, hormone therapy; RR, relative risk; WHI, Women’s Health Initiative.

†Represents the range of RR; no CI was reported.

‡Represents the range of RR; no CI was reported.

*Thromboembolism.
quality-adjusted life years. We derived QOL estimates from utility scores for the chronic diseases affected by HT through a literature search, using utilities reported for affected persons in the base-case scenario and estimates from unaffected persons for sensitivity analyses (Table 3). When QOL estimates were available only according to disease stages, we based weighted averages on the prevalence of each stage. Scores from a time tradeoff (TTO) assessment of patient subjects were preferred to other techniques. Utility scores from visual analog scales (VASs) were converted to TTO. 

Table 2. Mortality Rates (Annual Probabilities) 

<table>
<thead>
<tr>
<th>Disease</th>
<th>Mortality Rate</th>
<th>Long-term (After Year 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Black</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.20</td>
<td>0.27</td>
</tr>
<tr>
<td>Black</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.07</td>
<td>0.33</td>
</tr>
<tr>
<td>Black</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.23</td>
<td>0.24</td>
</tr>
<tr>
<td>Black</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Hip fracture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>Black</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pulmonary embolus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Black</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>Black</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.15</td>
<td>0.16</td>
</tr>
<tr>
<td>Black</td>
<td>0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

When HT’s impact on QOL is not considered, an average 50-year-old woman using HT for 2 years will experience a small net loss in life expectancy (12 days). Losses are larger for women at high CVD risk (37.5 days) and smaller for those at low risk (6 days).

Incorporating QOL adjustments for perimenopause and for conditions affected by HT substantially alters these findings. Women with mild menopausal symptoms gain 4.3 months of QALE from HT if at low CVD risk and 3.3 months if at high risk (Figure 1). Those with severe symptoms gain 8.3 or 6.9 months, respectively. The higher the CVD risks, the lower the expected QALE gains. In comparison, asymptomatic women who use HT experience net losses in QALE of 25 days to 2.9 months, depending on CVD risk.

These results are based on estimates of the utility associated with menopausal symptoms from 2 studies that may not be representative of utilities assigned by individual women to their symptoms. In threshold analyses, we explored the minimum severity of symptoms warranting HT, using any gain in QALE to define a benefit. Women at low CVD risk would benefit if symptom distress lowered their QOL during perimenopause to 0.96 or less (compared with perfect health, assigned a utility of 1.0). Among women at high CVD risk, menopausal QOL would need to be 0.88 or less on a TTO scale (Figure 2). The greater the severity of symptoms, the greater the gains in QALE from HT, although gains are smaller for women at higher CVD risk.

The impact of 2 years of HT on various chronic disease outcomes after 25 years of follow-up was relatively small, even among women at high CVD or breast cancer risk; cumulative risk of CHD increased by 0.008.

The longer HT is used, the larger the expected gains in QALE if HT relieves symptoms during the entire treatment period (although incremental gains diminish each year). For example, among women at high CVD risk with severe symptoms that would have persisted the entire treatment period, increasing HT from 2 to 5 years increases the gains in QALE from 6.9 months to 16.9 months. If symptoms would have been severe during the first 2 years but mild for the next 3 years, the gain in QALE from 5 years of HT is 11.1 months.

Sensitivity Analysis

Results were relatively insensitive to whether the utilities for the chronic conditions in the model represented those of persons affected or unaffected by the condition itself (QALE differences of <1 day) and insensitive to
changing short- and long-term mortality rates for both CHD and stroke by 20% (gains in QALE of \( \geq 3 \) days).

Substituting the WHI's estimates of HT’s impact on clinical outcomes with those from the Beral et al\(^{17} \) or Nelson et al\(^{18} \) meta-analyses changed the threshold utility of menopausal symptoms from 0.88 to 0.93 or 0.97, respectively, for women at high CVD risk.

Applying a 4% higher breast cancer mortality rate to HT users resulted in differences of less than 0.1 day. The model was relatively insensitive to changes in baseline risk for PE, breast cancer, hip fracture, and colorectal cancer and to whether HT increased risk of ovarian cancer by 50%.

**IMPACT OF HT ON BLACK WOMEN**

We explored the impact of HT on black women, accounting for racial differences in incidence and mortality rates. We assumed that the relationship between risk factors and disease incidence mirrors that occurring in white women. The gains in QALE from HT were consistently smaller than those for white women, but differences were inconsequential (\(< 1 \) day).

**IMPACT OF DISCOUNTING**

The higher the discounting of future years, the smaller the amount of symptom distress associated with gains from HT (eg, among women at high CVD risk, increasing the discount rate from 0 to 0.05 to 0.2 changed the TTO threshold for symptom distress from 0.88 to 0.92 to 0.95).

**COMMENT**

Whether short-term HT is beneficial or harmful depends primarily on a woman’s treatment goals, the severity of her estrogen-responsive symptoms, and her CVD risk. If the goal is to maximize longevity, HT is not advisable, since it is associated with small losses in life expectancy. However, if the goal is to maximize QALE, HT can be beneficial, especially among women at low CVD risk, among whom HT is associated with gains in QALE even when menopausal symptoms are mild. Women at high CVD risk can benefit from HT if their symptom burden is high enough to justify the risks of treatment (cor-
The decision whether to use HT amounts to a toss-up among women with low CVD risk and mild symptoms, with personal preferences driving the decision. However, less than 15% of US women are at low CVD risk. Short-term HT carries greater risks among most women with elevated CVD risk. Because the impact of short-term HT on breast cancer is uncertain, caution is warranted, especially among women at high risk for breast cancer.19

Incorporating women’s values into the HT decision may improve its quality. Women more concerned about the risks of cancer may be more likely to forego HT compared with women who place greater value on the benefits of symptom relief.61 The importance of eliciting patient values and concerns increases when we consider that HT may mask the diagnosis of breast cancer during the first few years.18 The benefits of HT outweigh its risks only when menopausal symptom distress is sufficient. The present analyses help women balance these risks and benefits.

A 12% loss in TTO during a 2-year perimenopause corresponds to a woman’s willingness to give up 2.9 months (12% of 2 years) of perfect health to avoid 24 months of menopausal symptoms. More practical would be the use of a VAS for which a transformation function to TTO has been proposed60:

\[
\begin{align*}
(1) \quad & \text{TTO} = 1 - (1 - \text{VAS})^{1.61} \\
& \text{or} \\
(2) \quad & \text{VAS} = 1 - (1 - \text{TTO})^{0.62}.
\end{align*}
\]

A practical approach to assessing symptom severity with VAS is to have women rate current health status on a scale of 0 (dead) to 100 (perfect health). A rating of less than 100 due to menopausal symptoms is divided by 100 to yield menopausal QOL. If she reports other conditions, she should rate her present health and her health just before menopause. The QOL due to menopausal symptoms is the complement (1 minus) of the difference between these scores, divided by 100 (this retrospective assessment of health status involves a response shift bias). For those at low CVD risk and high CVD risk, HT is appropriate if the VAS score is less than 0.86 and less than 0.74, respectively.

The value women place on current vs subsequent years (the discount rate) may affect the HT decision. We conservatively assumed a discount rate of 0, biasing against HT. Women with a higher discount rate (who value current more than future years) benefit more than those with a lower discount rate from HT, but differences are small. Some of our findings are based on methodological weak studies in which perimenopause was incorrectly treated as a chronic condition.7,62 One study asked women to recall symptoms before HT.7 Such retrospective ratings in respondents with improved QOL may be biased downwards owing to response shift (a scale recalibration phenomenon).63 Additionally, the terms characterizing symptoms in these studies are subjective. Nonetheless, their results were in accordance with VAS utilities elicited, strengthening confidence in their utilities.

The model does not include dementia because data concerning HT’s adverse impact on dementia are limited.35 It also does not consider any anxiety persisting after HT discontinuation. Findings pertaining to nonwhite women are less robust because of limited data; reported differences may be attributable to socioeconomic differences.

The decision regarding HT also depends on its efficacy in alleviating menopausal symptoms. Not all symptoms during perimenopause are due to declining estrogen levels or are responsive to HT. Some studies report that HT has a deleterious effect on QOL.65 Additionally, the placebo effect in the treatment of menopausal symptoms is large.66

In clinical trials examining the impact of HT on menopausal symptoms, women responding to HT did so within 2 to 5 weeks,67 suggesting that a 1-month trial period would be sufficient to determine response to HT. If response is unsatisfactory, HT should be withdrawn and other approaches to symptom relief explored, including lifestyle changes and administration of selective serotonin reuptake inhibitors. Since HT is associated with gains in QALE only when used for symptom relief, we recommend an annual drug holiday to assess the continued presence of menopausal symptoms.

Individual risk levels for osteoporosis, breast cancer, and colorectal cancer had little effect on our analyses, reflecting their low prevalence among the cohort examined, the small relative risk of HT on these outcomes, and the short duration of HT. These analyses reinforce that prevention of colorectal cancer or hip fracture is not an appropriate indication for HT. Effective treatments with fewer serious adverse effects than HT are available for preventing and treating osteoporosis, including bisphosphonates and parathyroid hormone. The limited sensitivity of our findings to the utilities used for chronic disease end points affected by HT suggests that further attention to determining utilities for these outcomes may not be warranted.

Small gains or losses in QALE at the population level can be important at the individual level. The low prevalence of CVD, PE, and breast cancer among women in this age group dilutes the overall impact of HT on life expectancy or QALE. However, in individuals who develop these adverse events, the impact on personal QALE will be substantial.

The potential increase in QALE due to HT is related linearly to the value a woman assigns to perimenopausal QOL. While these gains can be substantial, they need to be balanced with HT’s unfavorable impact on overall survival. Informed decision making requires patients to understand that any gains in QALE reflect subjective gains in QOL accrued during the early years of perimenopause but at the risk of incurring complications such as CHD, stroke, PE, and breast cancer.

Accepted for publication April 9, 2004

This work was supported in part by grant RO1 HS01332901 from the Agency for Healthcare Research and Quality, Rockville, Md; the American Cancer Society Breast Cancer Prevention Forum, Atlanta, Ga; and Generalist Physician Faculty Scholars Award No. 033958 from the Robert Wood Johnson Foundation, Princeton, NJ.

Portions of this work were presented at the 25th Annual Meeting of the Society for Medical Decision Making.
REFERENCES


