Pharmacoepidemiology of Hormone Therapy: An Evolving Picture

ANNLOUISE R. ASSAF
Pfizer Inc., New London, CT, USA

LAURA E. BURLESON
Pfizer Inc., New London, CT, USA

JOANN E. MANSON
Brigham and Women’s Hospital, Harvard Medical School, MA, USA

When the FDA first approved menopausal hormone replacement therapy (Emmenin) in 1933, the formulation was not easy to mass-produce (Rothenberg, 2005). In 1942, Premarin, a more easily manufactured estrogen, was introduced to the US mass market (CDER, 1997). At that time, Premarin was known to contain two main estrogens, estrone and equilin, and some additional estrogens in smaller amounts. It was approved primarily for the short-term treatment of postmenopausal symptoms – for relief of vasomotor symptoms associated with menopause. According to the FDA website,

The drug’s approval in 1942 predated the current requirements for such comprehensive analysis of products under review for marketing approval. At that time, Premarin’s approval was based on acceptable chemistry, manufacturing, and controls information and a showing, from reports of clinical investigations, that the drug was safe for its intended use in the treatment of menopausal symptoms and related conditions (CDER, 2005).

Over the years, however, many US women began to use the product off label for prevention of certain diseases such as coronary heart disease (CHD) or osteoporosis and for much longer periods of times than first envisioned. By the mid-1960s, about 12% of all postmenopausal women were taking estrogen (Harvard, 2003). Eventually, in 1986, the FDA
approved the drug for prevention of osteoporosis. By the time WHI was initiated in 1993, progestins had been added to protect the endometrium and most postmenopausal women in the US had been prescribed hormone therapy (HT). In addition, most, if not all of these women, had used HT in a way not originally approved by the FDA. A study by Hersh, Stefanick and Stafford (2004) estimated that from 1999 to June 2002, approximately 15 million postmenopausal women per year were taking HT. This figure was up substantially from earlier years primarily due to an increase in oral oestrogen/progestin combinations. It is of interest to note that more than one-third of patients who were prescribed HT were older than 60 years and thus several years post menopause.

There is a long history of observational studies of HT, starting shortly after the drug’s approval to treat menopausal symptoms in 1942. Epidemiological studies, animal studies and studies of surrogate biomarkers of CHD, as well as clinical studies such as angiography and bone mineral density studies, all pointed in the direction that HT was beneficial for the prevention of CHD and osteoporosis. Some of the larger, and more influential, observational HT studies, such as the Million Women Study (MWS) and Nurses’ Health Study (NHS), are listed and described in Table F.1 (Banks et al., 2004; Beral, 2003; Beral, Bull and Reeves, 2005; Grodstein, Manson and Stampfer, 2001). In general, these observational studies, as well as several clinical trials, showed the efficacy of oestrogen, and later a combination of oestrogen and progesterone, for treating menopausal symptoms (NIH, 2005; Rothenberg, 2005). Most observational studies also indicated that HT conferred a protective effect on CHD and was associated with a decreased risk of fractures, higher bone mineral density (BMD) and a lower colorectal cancer risk than women of the same age and menopausal status who were not receiving therapy (Banks et al., 2004; Beral, 2003; Grodstein et al., 1999; Komulainen et al., 1999; Lufkin et al., 1992; Speroff et al., 1996).

In addition to demonstrating benefits of HT, early studies demonstrated several risks. The main risk was an increased risk of venous thromboembolic events (VTEs), or blood clots, in women on HT, a finding that has been borne out over time (BCDSP, 1974; Grady et al., 2000; Hulley, 2002; Nachtigall et al., 1979; PEPI, 1995; Petitti et al., 1979; WHI, 2002, 2004). Results of observational research studies were less conclusive about HT’s risks or benefits related to CHD, breast cancer and uterine cancer, though the results trended towards showing a benefit of HT on CHD risk and an increased risk of breast and uterine cancer (Beral, 2003; Grodstein, Manson and Stampfer, 2001; Lerner, 1986; Speroff et al., 1996; Wilson et al., 1985). The addition of progestins to the HT formulations in the mid-1980s was thought to be protective of the uterus, and some studies, though not all, demonstrated a decrease in the risk of uterine cancer in those women with a uterus who were taking combination oestrogen progestin therapy (Beral, Bull and Reeves, 2005) (Figure F.1)

Though these observational studies were helpful for evaluating some of the risks and benefits of HT, they had known biases that could impact the observed benefits and risks of HT (Grodstein, Clarkson and Manson, 2003). In general, women who are prescribed oestrogen are leaner, less likely to smoke, less likely to eat a high fat or high salt diet, more physically active, more highly educated, and more likely to visit physicians regularly and take part in recommended screenings than women who are not prescribed hormones (WHI, 1998). These factors in and of themselves tend to decrease the risk of both CHD and some cancers. Also, though not well documented, women taking HT and the physicians who prescribe it may be less likely to attribute ischaemic symptoms to CHD if they believe a priori that HT reduces CHD risk (Col and Pauker, 2003). This makes it difficult to draw conclusions about the effects of HT separate from the effects of these socio-behavioural factors.

In addition to the previously mentioned biases that make interpreting the results of observational trials of hormone replacement difficult, additional interpretability issues arise due to major differences between the studies. The age ranges of individuals in the various observational studies mentioned previously and their racial and socioeconomic makeup differed greatly from study to study. Major differences in these factors make comparing results (both within and across different observational studies) difficult, especially since these factors have been shown to be related to CHD development. In addition to these population differences, observational studies of HT differed in how they defined CHD and other
Grady studies demonstrated several risks. The main risk was et al., 1999; Komulainen therapy (Banks age and menopausal status who were not receiving lower colorectal cancer risk than women of the same fractures, higher bone mineral density (BMD) and a on CHD and was associated with a decreased risk of oestrogen, and later a combination of oestrogen and well as several clinical trials, showed the efficacy of (Banks Study (NHS), are listed and described in Table 46.1 of CHD and osteoporosis. Some of the larger, and the direction that HT was beneficial for the prevention of these socio-behavioural factors. HT and the physicians who prescribe it may be less difficult, especially since these factors have been interpreted due to an increase in oral oestrogen/progestin more influential, observational HT studies, such as of these women, had used HT in a way not orig- finally approved by the FDA. A study by Hersh, postmenopausal women in the US had been prescribed to CHD, breast cancer and uterine cancer, though the less conclusive about HT's risks or benefits related to these population differences, observational studies in the various observational studies mentioned previ- between the studies. The age ranges of individuals interpretability issues arise due to major differences in these factors make comparing results (both differed greatly from study to study. Major differ- in these socio-behavioural factors.

In addition to demonstrating benefits of HT, early of HT differed in how they defined CHD and other to these population differences, observational studies differences in these factors make comparing results (both esp. since these factors have been interpreted due to an increase in oral oestrogen/progestin more influential, observational HT studies, such as of these women, had used HT in a way not orig- finally approved by the FDA. A study by Hersh, postmenopausal women in the US had been prescribed to CHD, breast cancer and uterine cancer, though the less conclusive about HT's risks or benefits related to these population differences, observational studies in the various observational studies mentioned previ- between the studies. The age ranges of individuals interpretability issues arise due to major differences in these factors make comparing results (both differed greatly from study to study. Major differ- in these socio-behavioural factors.

In addition to the previously mentioned biases In Table F.1, major HT study description and results

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment(s)</th>
<th>Study type/follow-up</th>
<th>Study Population Characteristics</th>
<th>Primary Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronary Drug Project</td>
<td>E, placebo</td>
<td>Randomized controlled trial. Stopped after 2.5 years</td>
<td>Men with previous heart attacks</td>
<td>Slightly higher risk of death from all causes (19.9 in E group, 18.8 in placebo); equivalent risk of heart attack; and increased risk of VTE (1.5% in men taking E vs. 0.8% in men taking placebo)</td>
</tr>
<tr>
<td>Framingham Heart Study</td>
<td>E, none</td>
<td>Prospective cohort study. 8 years of follow up at time of first HT results</td>
<td>1234 postmenopausal women over age 50 from Framingham, MA who were free of CHD at baseline</td>
<td>Women who used E had increased risk of cardiovascular morbidity (adjusted RR = 1.76) and cerebrovascular events (adjusted RR = 2.27); no major benefits noted. Risk of CHD was shown to be higher in E users who smoked than in non-smoking E users</td>
</tr>
<tr>
<td>Nurses’ Health Study (NHS)</td>
<td>E, E + P</td>
<td>Prospective cohort study with 20 years of follow up</td>
<td>70,533 postmenopausal nurses on HT followed by periodic surveys. Substudy of 2489 women with previous MI</td>
<td>Current users of HT (E alone or E + P) were found to be at increased risk of breast cancer and decreased risk of CHD compared to those not taking HT. Women with previous MI who took HT for a long duration had lower rates of recurrent CHD events than women who never used HT (RR = 0.65). In this subset, CHD events increased slightly with short-term use of HT</td>
</tr>
<tr>
<td>Uppsala Sweden study</td>
<td>Medium potency or low-potency/short acting E with or without P</td>
<td>Prospective cohort study (survey + registry) started in 1987. 8 years follow-up</td>
<td>9236 women in Uppsala Sweden with a mean age of 61 who had received at least 1 prescription for HT between 1977–80</td>
<td>Decreased MI (RR = 0.75) and hip fracture risk (RR = 0.65) with medium-potency E or E + P compared to low potency E</td>
</tr>
<tr>
<td>Million Women Study</td>
<td>E, E + P in various formulations</td>
<td>Prospective cohort study</td>
<td>320,953 postmenopausal women (age 50–64) in the UK without history of hysterectomy or cancer</td>
<td>Risk of endometrial cancer varied by HT formulation. Compared to those who never used HT, users of continuous combined (continued)</td>
</tr>
<tr>
<td>Study</td>
<td>Treatment(s)</td>
<td>Study type/follow-up</td>
<td>Study Population Characteristics</td>
<td>Primary Results</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>----------------------</td>
<td>-----------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Postmenopausal Oestrogen/ Progestin Interventions Trial (PEPI).</td>
<td>CEE, CEE + MPA, CEE + MP, or placebo over various spans of a 28-day cycle</td>
<td>Randomized, controlled trial. 3-year follow-up</td>
<td>Recruited 1996–2001, mean follow-up 3.4 years</td>
<td>E &amp; P had a RR of 0.71; e alone RR of 1.45; and users of cyclic combined had a RR of 1.05. Users of any HT formulation had reduced risk of fracture and increased risk of breast cancer. Those on E alone were more likely to have endometrial hyperplasia compared to E + P and placebo groups. In women taking E alone, simple endometrial hyperplasia was noted in 27.7%, complex in 22.7%, and atypical in 11.8% compared to 0.8%, 0.8%, and 0% in placebo, respectively. All HT groups had lowered LDL and Lp (a) and increased triglyceride levels compared to placebo (PEPI, 1995, 1996)</td>
</tr>
<tr>
<td>Heart and oestrogen/progestin replacement study (HERS).</td>
<td>E, E + P, placebo</td>
<td>Randomized, controlled trial. Mean follow-up 4.1 years</td>
<td>2763 postmenopausal women age &lt; 80 years with a history of CAD</td>
<td>No significant differences between treated and placebo groups on death, MI, or secondary cardiovascular outcomes (RR = 0.99). VTE (RR = 2.89) and gall bladder disease (RR = 1.38) were increased in the HT groups</td>
</tr>
<tr>
<td>Heart and oestrogen/progestin replacement study II (HERS II).</td>
<td>E, E + P, placebo</td>
<td>Open-label follow-on study to HERS</td>
<td>2321 postmenopausal women with history of CAD who participated in HERS</td>
<td>No cardiovascular benefit of HT observed after following the HERS I cohort for additional time</td>
</tr>
<tr>
<td>Women’s Health Initiative (WHI)</td>
<td>E, E + P, placebo</td>
<td>Randomized Controlled Trial</td>
<td>For the E + P component, the population is 16,608 postmenopausal US women aged 50–79 with an intact uterus at baseline. For the E alone, it is 10,739 postmenopausal US women, age 50–79 with prior hysterectomy.</td>
<td>Neither E alone nor E + P reduced CHD risk (RR for E and E &amp; P vs. placebo = 0.91 and 1.29 respectively). Risks of E + P (including breast cancer, VTE and stroke) shown to outweigh benefits</td>
</tr>
</tbody>
</table>
endpoints. For example, some studies, such as the Framingham Heart Health Study, that examined the relationship between HT and CHD included angina as part of the constellation of conditions that comprise CHD while others, such as the Nurses Health Study, did not. This makes comparison of the results of these studies difficult at best.

In part because of these difficulties in interpreting data from observational studies (given the potential biases mentioned previously), a number of randomized controlled clinical trials (RCTs) of HT were initiated in an attempt to definitively answer questions about the risks and benefits of HT (Therapeutic Letter, 1999). These studies were conducted first in highly selected populations with trials such as the Coronary Drug Project (CDP) and the Heart and Estrogen/Progestin Replacement Study (HERS), and then in the more general population, with studies such as the Postmenopausal Estrogen/Progestin Interventions (PEPI) and the much larger, and more comprehensive, Women’s Health Initiative (WHI). The first RCT relating to this question – the HERS trial – was published in 1998 (Hulley et al., 1998; Petitti, 1998). Several additional small, short-term RCTs were also conducted. One of these trials, The Coronary Drug Project, was actually conducted in men (The Coronary Drug Project Research Group, 1973), though the others were conducted in women (Hall et al., 1998; Komulainen et al., 1999; Lufkin et al., 1992; Sporoff et al., 1996). These trials sought more conclusive evidence of the relationship between HT and CHD, uterine cancer, or other endpoints and all seem to corroborate evidence from the earliest published RCTs on HT. Table F.1 summarizes the study populations and results from some of these randomized clinical trials as well as several of the larger, more influential observational studies.

![Figure F.1. Hormone therapy timeline of major events: 1942–present.](image-url)
It was clear from the findings from these small, short-term trials that a larger, longer-term randomized clinical trial with a greater diversity of post-menopausal women and ‘hard’ disease endpoints needed to be undertaken to definitively answer these questions. And so, the timing was right for the Women’s Health Initiative, which was made possible because of the efforts of the NIH’s first female director, cardiologist Bernadine Healy, the Women’s Health Caucus, and other groups who successfully lobbied for a line item in congress’ budget to try to erase a 25-year gender gap in our knowledge of diseases which affect women in their later years. In addition, the FDA would not approve a statement in the HT label claiming a benefit for heart disease until a definitive clinical trial was conducted. This was the state of the field when the WHI was first proposed.

Flash ahead to 2002, when the WHI oestrogen plus progestin trial was stopped early because of an increased risk of breast cancer (HR = 1.26; CI = 1.00–1.59), CHD (HR = 1.29; CI = 1.02–1.63), Stroke (HR = 1.41; CI = 1.07–1.85), and PE (HR = 2.13; CI = 1.39–3.25) (WHI Writing Group, 2002). Immediately following the publication of these trial results in July of that year, oral HT prescriptions began a steady decline, while a slight increase was seen in the use of vaginal formulations. In 2003, the FDA, in conjunction with some members of Congress, launched a national campaign to provide information and increase awareness about the recent findings on menopausal use of HT (FDA, 2003). In April 2004, the results of the WHI oestrogen-alone trial were published (WHI Writing Group, 2004). The investigators noted an increased risk of fatal and non-fatal strokes (HR = 1.39; CI = 1.10–1.77) and venous thrombosis (HR = 1.47; CI = 1.04–2.08); no significant difference in risk of CHD HR = 0.91; CI = 0.75–1.12), colorectal cancer (HR = 1.08; CI = 0.75–1.55), total cancer (RR = 0.93; CI = 0.81–1.07), or all cause (HRR = 1.04; CI = 0.91–1.12).

The effect on breast cancer was uncertain (HR = 0.77; CI = 0.59–1.01) and there was an increased benefit on bone fractures (HR = 0.70; CI = 0.63–0.79).

In March of 2005, an NIH State-of-the-Science Conference on the Management of Menopause-Related Symptoms took place to discuss and form consensus on menopause-related symptoms and preventive and treatment modalities. The report discusses HT as ‘menopause hormonal therapy’, reflecting the belief that menopause is a natural state of being for women of a certain age and not a disease state (NIH, 2005). After the release of results from the WHI, a black box statement pertaining to cardiovascular risks was added to the label for oestrogen. This statement read ‘The Women’s Health Initiative (WHI) reported increased risks of myocardial infarction, stroke, invasive breast cancer, pulmonary emboli, and deep vein thrombosis in postmenopausal women during 5 years of treatment with conjugated equine estrogens (0.625 mg) combined with medroxyprogesterone acetate (2.5 mg) relative to placebo.’ The warning goes on to state that the FDA assumes these findings will hold for all HT formulations containing oestrogen and suggests that HT drugs should be used in the lowest doses necessary for the shortest duration possible (FDA, 2003).

**CONTROVERSY OVER WHI RESULTS**

There is some controversy in the medical community about whether or not the results of the WHI were ‘valid’ or ‘relevant’, particularly in the context of divergent findings from earlier observational work and smaller clinical trials. Like most trials, the WHI may have had some minor limitations, which have been examined in an attempt to ascertain whether such problems might have biased the results sufficiently to change the direction of the effect of HT on CHD and on the global risk/benefit assessment. One of these limitations was the relatively high dropout and crossover rates that the study had. If dropouts and cross-over are differential, they would have had potential to bias the results of the study, particularly if dropouts or crossover occurred differentially based on women’s health status, which is plausible. However, these issues should not have biased the ‘intention-to-treat’ analyses.

Another concern voiced in the literature was the age of the WHI participants in the hormone trials compared with the average age of menopause. In most observational studies, HT is started at, or close to, the time of menopause. In WHI, as in most other
clinical trials, the therapy was often initiated more than a decade, on average, after menopause, with an average age of women in the E & P trial of 63.2 years and of 63.6 years in the E-alone trial (WHI, 1998, 2002, 2004). If the effects, both positive and negative, of HT vary depending on the age of the women (or the duration of time since menopause), as has been suggested, conducting the study in older women may have biased it away from seeing favourable effects and towards seeing increased risks associated with HT in high-risk women. Also, baseline absolute rates of disease are much lower in recently menopausal women than in older women, so the absolute number of excess events will also differ by age.

One issue that spurred much debate was the use of what the WHI investigators termed a ‘global index’ – a composite measure consisting of the earliest occurrence of CHD, invasive breast cancer, stroke, pulmonary embolism, endometrial cancer, colorectal cancer, hip fracture and death due to other causes – which was used to quantify the overall risk associated with the drug (WHI Writing Group, 2002). Issues raised regarding the global index include the criticism that it was not validated (Goldman, 2004); the selection of events to be included in the global index; and the index’s apparent lack of weighting of treatment types (despite very different health, quality of life, and economic impacts of each type of event).

Whatever the concerns, however, the findings (which seemed to contradict long-standing dogma), caused the medical and lay communities to express dismay and disbelief. As Elias Zerhouni, director of NIH, said about the HT controversy when WHI results were first published in 2002 – ‘Often in science the reaction to a new finding is directly proportional to the strength of the dogma it overturns. People are still in denial of the theory of relativity, too’ (Rothenberg, 2005; Spake, 2002). The way in which some of the WHI data was presented in the media may also have increased the concern expressed by some. There was an emphasis on Relative Risk (RR) versus rate difference (RD), which may have served to magnify the risks observed and led to misinterpretations of the results. For example, while the relative hazard of CHD was elevated by 29% for women taking E & P, in absolute terms this equates on only an additional 7 cases of CHD per 10,000 women per year (WHI, 2002). Regardless of these concerns, however, the fact remains that WHI is still the largest randomized clinical trial of HT to date, with the best ascertainment of outcomes. As such, the results of this landmark study should not be dismissed. As noted by the European Menopause and Andropause Society (EMAS) and by S. Barton, ‘it is not an easy task to opt between results of observational and clinical trials. High-quality observational studies may extend evidence over a wider population and are likely to be dominant in the identification of harms’, but the ‘best RCT still trumps the best observational study’ (Barton, 2000; Neves-e-Castro et al., 2002).

DISCUSSION

The evolving attitude towards HT, from a treatment many believed unethical not to prescribe to post-menopausal women to one associated with negligible benefit and moderate risk, is a good example of how scientific thinking about a drug may change over time. The HT debate illustrates the various scientific methods used to understand drug safety and how each can either increase understanding or produce confusion. While it is difficult to make decisions on HT use now because of uncertainty, the evolution of knowledge about HT provides a good example of current thinking about the tradeoffs of using data from observational studies or clinical trials to understand the population impact of drug treatment. Some of these principles are discussed below using examples from the study of HT.

STRENGTHS OF CLINICAL TRIALS COMPARED TO OBSERVATIONAL EPIDEMIOLOGY STUDIES

Despite their strengths, which make them valuable tools in attempting to determine relationships between drugs and outcomes, observational epidemiologic studies have a number of limitations. These limitations are a large part of the reason that clinical trials remain the gold standard for evaluating drug-outcome relationships.

A major concern with observational studies is that they are not randomized and treatment may well vary according to characteristics linked to the disease process. This could occur through selective
prescribing or through issues related to medical care access. Confounding by demographics, socioeconomic status, and variables related to health, as described previously, can be a substantial problem in observational studies.

A second problem with observational epidemiology studies, especially the larger ones, is that they tend to have less robust ways of evaluating endpoints than clinical trials, relying on self-report or simple clinical reports, rather than a methodical AE reporting system. Short-term effects of treatment may be particularly difficult to ‘capture’ in observational settings. There may also be inconsistent endpoint definitions between studies, making comparability difficult. This problem may have contributed to the divergent estimates of CHD in women taking HT in the NHS and Framingham studies, which defined CHD differently.

In addition to problems ascertaining outcomes, observational studies also frequently have poor or limited characterization of health status or lifestyle factors, particularly those that change over time. For example, over-the-counter medications that can affect outcomes are not always ascertained and variables such as physical activity levels and diet are often not measured as often or with as much precision as would be desirable. Because of potential differences in these factors between those prescribed and those not prescribed drugs, residual confounding is likely to be present in observational studies. If some of the relevant confounding variables are measured at baseline (and, if possible, throughout the study), adjustment for some confounding is certainly possible, but residual confounding remains likely.

Another potential limitation of observational studies is assessment of drug exposure, in terms of both dose and duration. The duration of use of a drug is sometimes poorly defined in observational studies. Often treatment use is measured at the beginning of the study or only at irregular periods throughout the study, and constant use is assumed, whether this is valid or not. Observational studies often have poor information on the dose or particular formulation of a drug that is being used. Strategies, including having patients bring all of their medicines with them to intake visits, have been developed to help with this problem, but many studies, particularly those that are survey-based, have limitations related to exposure assessment. Clinical trials, on the other hand, frequently employ systems such as pill counts or blood level monitoring that allow researchers to monitor actual dose received on an ongoing basis.

Finally, in observational studies, particularly cross-sectional, case-control and prevalence studies, it is often not possible to establish a temporal relationship between drug and disease, which is crucial to establishing cause-and-effect relationships, rather than simple associations. This limitation can impede interpretation, such as when a drug improves survival with a condition or increases its latency period, both benefits, rather than causing the condition itself.

STRENGTHS OF OBSERVATIONAL EPIDEMIOLOGY STUDIES COMPARED TO CLINICAL TRIALS

As shown by some of the analyses from the WHI, confounding can be a problem even in clinical trials, especially when the blind is imperfect. This can lead to some problems of differential follow-up and ascertainment, but these are usually less prominent than those seen in observational epidemiologic studies. Critics cite this as a potential flaw of the WHI that may, in addition to other factors (such as the age structure of the population the trial was conducted in), have biased the results.

Though the WHI looked at hard outcomes, many clinical trials use surrogates as their primary outcomes of interest. This can cast doubt about whether the results of such trials are clinically meaningful. By contrast, observational epidemiologic studies can, and usually do, look at ‘real’ events (such as MI) rather than their surrogates (e.g., cholesterol). This is due in part to the fact that, unlike clinical trials, they can be retrospective (mitigating the need for costly follow-up) or long-term prospective follow-up of a large-scale cohort may be feasible.

The WHI, which looked at actual events in a large population of women over a long time period, had many of the advantages of a large-scale prospective follow-up usually associated with observational studies rather than with clinical trials, but duration of follow-up tended to be shorter than in many observational studies. However, some of the earlier, smaller clinical trials that form an important part of the evidence base about the risks and benefits of HT used surrogate endpoints and were limited to very
short-term follow-up. Thus, the duration of observational studies is frequently longer than is feasible for clinical trials, which allows evaluation of ‘hard’ outcomes instead of surrogates and of rare or time-delayed effects. Part of the controversy following the results of the early, short-term randomized controlled trials of HT stemmed from their short duration (especially compared to some of the observational studies) and, in some trials, use of surrogate markers rather than clinical disease states for some outcomes.

Another strength is that observational studies occur in ‘real life’, using drugs in the particular dose and schedule used by patients in the field. This may more appropriately represent usage patterns than more controlled studies. One criticism of the WHI is that only one combination of oestrogen and progestin was evaluated, though other doses and formulations exist. Arguably, the effects of a drug, in terms of both risks and benefits, may vary by the dose, duration of exposure and route of administration, as well as the demographics and health status of those treated. Thus, an observational study may be better equipped to evaluate several factors that are relevant to usage of the drug in ‘real life’.

Finally, because of their decreased demands in terms of cost, observational studies, particularly large simple studies, often allow for larger sample sizes and longer duration of follow-up than clinical trials. Although some consider observational studies more ‘cost-effective’, tradeoffs related to confounding and selection biases must be given careful consideration, as discussed above.

THE FUTURE

There is no doubt that future drugs being developed for postmenopausal prevention and treatment of disease will be undergoing intense scrutiny by the FDA, the medical community and consumers. Each of these groups is better informed now than in the mid-1990s, and as new replacements for oestrogen therapy (such as new oestrogen receptor modulators or SERMS) are explored, the lessons of HT will remain in the forefront, influencing the way future drugs are developed, approved, marketed and prescribed.

Several professional societies have weighed in on these issues (and continue to do so). The American College of Obstetricians and Gynecologists (ACOG), for example, formed a task force to examine the evidence from WHI and other studies and in 2004 issued the following statement: ‘The risks of HT exceed the benefits for the prevention of chronic diseases in postmenopausal women. Hormone therapy remains an effective therapy for treating women with vasomotor symptoms and vaginal atrophy.’ The ACOG task force went on to state that

The EMAS has also carefully weighed the risk/benefit ratio of HT and has revised their earlier recommendation statements for clinical practitioners regarding peri- and postmenopausal HT to reflect the changing state of research following WHI (EMAS 2005; Neves-e-Castro et al., 2002).

As noted by the ACOG HT Task Force (2004) and by others, ‘Virtually all medications carry risks as well as benefits, and as detailed in the preceding chapters, HT is no exception. Balancing these beneficial and harmful effects is a challenging but important task for making informed decisions about the prescribing and use of HT.’ Despite the many questions answered by WHI and the even more questions raised by this study, it is clear that professional societies in both Europe and the US feel that there is no one solution for all postmenopausal women. It is also true that one piece of clarity in all of this controversy is that the WHI has paved the way for more open communication between the postmenopausal women and her health care provider. That is most certainly a good thing.

REFERENCES
