Estrogen Replacement Therapy Moderates the Sleep Disruption Associated with Nocturnal Blood Sampling

Karen E. Moe PhD,1 Lawrence H. Larsen PhD,1 Michael V. Vitello PhD,1 and Patricia N. Prinz PhD1,2

1Department of Psychiatry and Behavioral Sciences; 2Department of Biobehavioral Nursing and Health Systems, University of Washington, Seattle, Washington, 98195

Study Objectives: To determine whether chronic oral estrogen replacement therapy (ERT) (1) improves the sleep of older, non-symptomatic postmenopausal women; and (2) reduces the sleep disruption associated with a stressor (frequent remote nocturnal blood-sampling through an intravenous catheter).

Design: Descriptive, cross-sectional, secondary analysis of a larger study.

Setting: The General Clinical Research Center at the University of Washington Medical Center.

Participants: Women aged 57—80 (mean age = 70) at least 5 years past menopause were recruited from the community. Hot flashes and significant sleep difficulties were exclusion criteria. The ERT group (n=37) consisted of women on chronic oral ERT for ≥2 years. The NERT group (n=56) consisted of women not using estrogen (NERT) for ≥2 years.

Interventions: N/A

Measurements and Results: Following an adaptation night, polysomnographic measures were collected for 2 consecutive nights. A blood sample was collected every 20 minutes for the last 24 hours (including Night 2), through an intravenous catheter. The only group difference in sleep on the baseline (non-catheter) night was that NERT women had a shorter sleep latency. Sleep on the catheter night was characterized by increased wakefulness, longer sleep latency, and decreased REM sleep for both groups relative to the baseline. However, the impact of nocturnal blood sampling was much greater for NERT than for ERT women: they experienced significantly greater percent changes in more sleep-wake variables, particularly slow-wave sleep (SWS).

Conclusions: In this cross-sectional study, the use of chronic oral ERT was associated with little effect on the sleep of older postmenopausal women not experiencing hot flashes, except in the presence of a challenge to sleep. ERT ameliorated the disruptive effect of nocturnal blood sampling on both objectively assessed and subjectively assessed sleep.

Key words: Estrogen; polysomnography; catheterization; aging

INTRODUCTION

The perimenopausal years are associated with a sharp increase in the frequency of sleep complaints by women.1,2 Although the exact prevalence of self-reported sleep problems varies widely across studies,3-5 perimenopausal women typically cite insomnia, disturbed sleep, and fatigue as some of their most frequent and pressing health concerns. The initiation of these sleep difficulties is usually attributed to the presence of hot flashes and related perimenopausal symptoms.2 This hypothesis is supported by laboratory sleep studies,6,7 though there is at least one contradictory report.5 Perimenopausal women experiencing hot flashes have more arousals during sleep and more sleep stage changes compared with women of equivalent age and hormonal status who are not experiencing hot flashes.7

Estrogen replacement therapy (ERT) has been widely prescribed for the alleviation of hot flashes and consequent sleep disruption. Numerous studies have shown that ERT reduces hot flashes and improves self-reported sleep quality (e.g., 9). Only a few of these studies examined ERT effects on objective sleep measures (e.g., polysomnography). These studies found that ERT could reduce wakefulness6,10 and movement arousals,11 and increase REM sleep.10,12 Collectively these studies suggest that estrogen may be of use in alleviating the sleep complaints of older women, especially those who are experiencing hot flashes.13,14

Sleep complaints do not subside after the menopausal transition years. Many older postmenopausal women continue to report significant sleep disturbance, including nighttime awakenings, long sleep onset latencies, and overall poor sleep quality.1 Little is known about the impact of ERT on the sleep of older, postmenopausal women who are not experiencing hot flashes. Previous studies using polysomnography to assess ERT effects were typically based on women experiencing hot flashes. One purpose of this study was to examine the effect of long-term oral ERT on objective and subjective sleep measures in postmenopausal women who were not experiencing hot flashes and who were well past the hormonal fluctuations that occur during the three to four years on either side of menopause.15,16

The other purpose of this study was to determine whether long-term oral ERT might ameliorate the impact of stress or environmental disruption on the sleep of older women. Sleep quality can be negatively affected by a variety of stressors or environmental disruptions: noise,17 change in ambient temperature,18 sleeping in a novel environment,19 and periodic blood sampling through an intravenous catheter.20 However, the sleep impact of these circumstances varies significantly among individuals. This
variation is the result of interplay between psychosocial and biological factors.

Two such factors are age and gender. Sleep in older individuals is more vulnerable to disruption. Auditory awakening thresholds decline with age, and older subjects tend to report more awakenings in response to noise. A novel sleeping environment is associated with increased sleep latency, decreased sleep time, and decreased REM and slow-wave sleep in older compared with younger individuals. Gender also accounts for significant variation in sleep responses to disruption. Most studies have found that environmental factors disrupt sleep quality more in women than in men. This is particularly true of older women. Such increased susceptibility to exogenous disruption of sleep may partially explain why older women have more sleep complaints than older men.

Estrogen may be another factor that influences susceptibility to sleep disruption, through its effects on stress responsivity. Low estrogen states, such as in the postmenopausal years, are associated with enhanced cardiovascular and hormonal responses to stress (e.g., postmenopausal women have increased heart rate and blood pressure responses to stress compared to age-matched premenopausal women). Conversely, ERT seems to lower stress responsivity in older women. Estrogen replacement lowered cardiovascular, neuroendocrine, and immune responses to stress in women. ERT also significantly reduced cardiovascular and hormonal responses to stress in a well-designed study of postmenopausal women.

The data reported here are from a secondary analysis of a larger study that provided an opportunity to test this possible ameliorating effect of ERT on sleep disruption. The larger study protocol included a 24-hour blood sampling period, when a small blood sample was drawn every 20 minutes through an intravenous catheter. We have previously shown that there is wide individual variation in how sleep of older individuals is affected by this protocol, ranging from no apparent impact on sleep to significant disruption. Overall, sleep quality in that previous study was more disrupted in older women than older men. The present study tested the hypothesis that the sleep of women on ERT would be less affected by periodic nighttime blood sampling than the sleep of women not on ERT.

### METHODS

#### Subjects and Recruitment

The results presented here are from a secondary analysis of data collected for a large cross-sectional study on hormones, sleep and cognition in normal aging. Subjects were recruited from western Washington State via public service announcements in local news media as well as letters to senior centers, retirement communities, and groups of retirees. Recruitment announcements emphasized the requirement for healthy non-smokers with no memory impairment, obesity, or frank sleep disorders. Individuals who responded were screened via a three-step protocol to ensure that all subjects met the study inclusion/exclusion criteria (Table 1). Of potential subjects, 43% passed the first screening step (a 30-minute telephone interview focusing on general medical history, use of tobacco and medications, and sleep). This was followed by a clinical interview and physical exam with medical history. Cognitive status (Folstein Mini-Mental Status, MMS), depression (Center for Epidemiological Studies Depression Scale, CES-D), stress (Life Events Scale, LES), anxiety (SCL-90 Anxiety Scale), and recent subjective sleep quality (Pittsburgh Sleep Quality Index, PSQI) were assessed during the interview (Table 2). Scores on specific PSQI components were used to identify and exclude subjects with significant subjective sleep difficulty while still retaining a meaningful range sleep quality. The seven PSQI components each have a possible value of 0–3, with a lower number indicating better sleep quality. Individuals were excluded for any one of the following: (1) Subjective Sleep Quality ≥2; (2) Sleep Latency=3; (3) Sleep Disturbance=3; or (4) Use of Sleeping Medications ≥2, or if there was any use at all of prescription sleeping medication.

Scores on the LES and the Anxiety scale of the SCL-90 were used as guidelines rather than absolute criteria for identifying and excluding individuals experiencing significant and disruptive stress and psychosocial instability (e.g., having a spouse recently diagnosed with a life-threatening illness). Answers to individual questions on those questionnaires, follow-up questions, and clinical impression contributed to the decision. Of excluded individuals, fewer than 5% were excluded for this reason.

Individuals with a score >16 on the CES-D depression questionnaire were excluded. The range of scores for the enrolled subjects in both the NERT and the ERT group was 0—16. As shown in Table 2, the mean CES-D scores were surprisingly low. About 79% of the sample had CES-D scores in the lowest quartile (total score of 0-4); 11%, 4%, and 6% of the sample had scores in the 2nd, 3rd, and 4th quartiles, respectively. These low depression scores may be due to health and medication exclusion criteria applied during the initial telephone screening interview (e.g., use of anti-depressants), together with exclusion of individuals with significant sleep problems (a criteria described in recruitment materials and applied as described above). Insomnia and distress/depression appear to have a bidirectional relationship.

Sixty-three percent of individuals who underwent the clinical interview and physical exam met the inclusion/exclusion criteria and were invited to participate in the study. The study protocol was approved by the University of Washington Institutional Review Board (Human Subjects Committee) and conformed to the Declaration of Helsinki. Informed written consent was obtained from all subjects.
Table 2—Demographic and other descriptive data for subject groups (mean±standard deviation; percents). There were no significant group differences except for use of estrogen.

<table>
<thead>
<tr>
<th></th>
<th>NERT (n=56)</th>
<th>ERT (n=37)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yrs)</td>
<td>70.9±5.9</td>
<td>69.7±6.6</td>
</tr>
<tr>
<td>BMI</td>
<td>25.9±3.7</td>
<td>25.8±3.6</td>
</tr>
<tr>
<td>Educ (yrs)</td>
<td>15.6±3.0</td>
<td>15.8±3.0</td>
</tr>
<tr>
<td>MMS</td>
<td>28.1±1.5</td>
<td>28.1±3.2</td>
</tr>
<tr>
<td>CES-D</td>
<td>2.4±3.7</td>
<td>3.4±3.6</td>
</tr>
<tr>
<td>SCL-90A</td>
<td>2.1±3.7</td>
<td>3.1±4.2</td>
</tr>
<tr>
<td>LES</td>
<td>8.2±5.9</td>
<td>10.4±6.6</td>
</tr>
<tr>
<td>PSQI</td>
<td>3.9±3.0</td>
<td>3.7±2.4</td>
</tr>
</tbody>
</table>

ERT duration

| 2-5 yrs        | ----- | 16.1% |
| 5-10 yrs       | ----- | 8.1%  |
| >10 yrs        | ----- | 75.7% |

BMI = body mass index (weight in kg/(height in meters)²); MMS = Mini-Mental Status score; CES-D = Center for Epidemiological Studies - Depression score; SCL-90A = SCL-90 Anxiety score; LES = Life Events Scale score; PSQI = Pittsburgh Sleep Quality Index.

Subject Groups

Data presented here are from all female subjects in the larger study who (1) were 57 or older and at least five years post-menopause; (2) were not currently experiencing hot flashes or night sweats, and (3) who had a sleep efficiency (total sleep time/time in bed) of at least 70% on the baseline (no catheter) night. Further, they met the following criteria with regard to use of hormone replacement therapy. The NERT group (n=56) consisted of women who were not taking estrogen or progesterin in any form currently or in the past two years. The ERT group (n=37) contained women who were taking oral estrogen (without progesterin) five to seven days per week for at least 21 days per cycle, for at least two years. The mean duration of ERT could not be calculated, as many women endorsed a categorical answer (e.g., “more than 10 years”) to this question rather than giving a specific number of years. However, as indicated in Table 2, 76% had been on ERT for more than 10 years. Most ERT women (78%) were taking conjugated equine estrogens (Premarin; major estrogen component=estrone). The remaining women (22%) were taking an esterified estrogen combination (Estratab) whose principal estrogen component is also estrone. The estrogen dose for all but one estrogen-using woman was 0.625 mg/day; the single exception was a woman taking 0.3 mg/day. Although each ERT subject typically cited more than one reason why she began and continued to take estrogen, prevention of osteoporosis was the most frequently given reason. For those women who routinely discontinued estrogen for a few days each month, study participation always occurred on days when they were taking estrogen.

Because of the cross-sectional study design, it was important to characterize each group on variables that might influence sleep (Table 2). The two groups were equivalent on all demographic, psychosocial and medical variables measured. All subjects passed a rigorous health screening, and there were no group differences in overall health, specific medical problems, or use of prescription medications (other than estrogen).

General Procedures

Subjects were admitted to the University of Washington Medical Center General Clinical Research Center (GCRC) for a 65-hour stay. All meals were provided through the GCRC Nutrition Research Kitchen. The daily use of estrogen by ERT women was verified during the GCRC stay. All subjects went to bed at their customary bedtimes and slept until their customary rise time or until spontaneous awakening. Upon awakening each morning, they filled out a short questionnaire about the night’s sleep (Sleep Quality Questionnaire, SQQ). Napping was strongly discouraged.

Sleep and EEG Recording

Night 1 was an adaptation night. Apnea screening via pulse oximetry also occurred on this night. None of the subjects showed evidence of clinically significant sleep apnea. During the following day, subjects underwent a series of cognitive tests. On Nights 2 and 3, the sleep EEG was recorded. Standard procedures were used for sleep recordings, including EEG, electro-oculogram (EOG), and electromyogram (EMG). Pulse oximetry was not performed on these two nights. EEG electrodes were positioned for conventional sleep recordings at C3, C4, O1, and O2 (international 10-20 system of measurement) and were referenced to the contralateral mastoids. Data were recorded using a Grass 8-24 polygraph with filter settings of 0.1 and 70 Hz. Four channels (C3, C4, EMG, O2) were digitized simultaneously with recording, using a 12-bit digitizer with a voltage range of ±2.5 volts that was installed in a microcomputer. The sampling rate was 256 Hz, with data averaging and decimation to 128 Hz. Additional recording details are provided elsewhere. The night technician was blinded as to subjects’ estrogen status (NERT, ERT).

Catheterization for Blood Sampling

An indwelling catheter was placed in a forearm vein at 0800 in the morning after Night 2, by a phlebotomist blinded about the subject’s estrogen status. A blood sample (3 ml) was drawn through the catheter every 20 minutes for the next 24 hours, which included Night 3. These samples were assayed for growth hormone levels later, as part of the larger study.

Although subjects could not leave the GCRC during the 24-hour blood sampling period, they were strongly encouraged to remain as active as possible during daytime hours (e.g., walking around the unit). A slow heparinized (200 U/L) saline drip kept the catheter patent. A small electric heating pad set at the lowest setting was secured to the forearm as necessary, to ensure successful sampling. Nighttime samples were drawn remotely from the room next door through a long extension of the catheter. The study was completed at the end of the blood sampling period. Due to the requirements of the larger study, the order of the baseline night and catheter night could not be varied.

Sleep Scoring

Paper records were scored for sleep stages by a single human rater using standard techniques and 30-second epoch lengths. The rater was blind to subjects’ group assignments. Standard polysomnographic variables were then calculated, with sleep...
onset latency defined as the onset to stage 1 sleep. The number of awakenings is reported in two ways: (1) WAKES, defined as the number of transitions from a sleep stage to an epoch scored as “wake”; and (2) TAW1, a subset of WAKES, defined as the number of transitions from a sleep stage to a “wake” period lasting for two or more consecutive 30-sec epochs. All sleep-wake variables reported here are from these human-scored data, except for measures of stages 3 and 4 (slow-wave sleep, or SWS). Minutes of SWS were calculated by a computer algorithm (C STAGE;35) that quantifies and scores EEG epochs using EEG spectral analysis. This algorithm was designed to address some of the difficulties inherent in scoring SWS in older subject while using the conventional 75 µvolt amplitude criterion.35,36

Data Analysis

Unless otherwise stated, data are reported as means ± standard deviation of the mean (SD). All variables were examined prior to analyses for extreme outliers (>3 SD from the mean) and normality of distributions. There were several outliers for sleep latency, or for REM latency; the outliers were distributed equally among the two groups for sleep latency, whereas REM latency outliers were almost all in the ERT group. These and other significantly non-normal distributions were transformed (e.g., square root transformation) prior to analyses or were analyzed using non-parametric tests. There were also two outliers for total wake time (both NERTs) and two for minutes of stage 3 and 4 (one NERT, one ERT) which were retained for data analyses. The analyses for these two variables was performed both with and without these outliers, and with and without a nonparametric test. The results were the same in all cases. The results of the parametric tests including these outliers are reported here. Statistical significance was always construed as p ≤ .05 (two-tailed tests).

For a cross-sectional study such as this one, it is important to identify any group differences that might influence dependent measures. To maximize the possibility of identifying any such group differences, group comparisons on all demographic and psychosocial measures were performed using statistical tests that are relatively more sensitive to Type I error (erroneously rejecting the null hypothesis) and relatively less sensitive to Type II error (erroneously failing to reject the null hypothesis).

Polysomnographic sleep variables were examined by multivariate analysis of variance (MANOVA; SPSS-PC statistical software) followed by multiple post-hoc t-tests. Three series of comparisons were performed. The first series determined whether there were any baseline sleep differences between the two groups. The second series of analyses compared sleep on the baseline vs. the catheter nights, separately by group. In the third series of analyses, the percent change from the baseline night to the catheter night was compared for NERT vs. ERT women.

Three questions from the SQQ (i.e., the subjective sleep rating) were analyzed in a similar manner, with NERT vs. ERT comparisons made separately for each night, and baseline vs. catheter night comparisons made separately for each group. Correlations between subjective sleep quality and several polysomnographic sleep-wake variables were also examined.

As is normal for subjects of this age, there were occasional blood sampling problems during the 24-hour sampling period. Prior to performing the above analyses, technician and nursing notes were examined (while blind to hormone status) for each subject. There were no NERT vs. ERT differences with regard to which forearm vein was catheterized. For more than 90% of the subjects there were no blood sampling problems during the 24-hour sampling period. The remaining subjects who experienced some sampling difficu-

---

### Table 3—Polysomnographic sleep-wake variables (means±standard deviation) for the baseline night and catheter night for NERT women (not on estrogen replacement therapy; n=56) and ERT women (on estrogen replacement therapy for 2 years or more; n=37).

The unit of measure is minutes, except where indicated. WASO%=minutes of wake after sleep onset, as a percent of sleep period time (where sleep period time=time in bed minus sleep latency).

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>NERT Catheter</th>
<th>Baseline</th>
<th>ERT Catheter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lights out (decimal hrs)</td>
<td>22.72±6.7</td>
<td>22.94±5.2a</td>
<td>22.64±6.0</td>
<td>22.87±6.0a</td>
</tr>
<tr>
<td>Lights on (decimal hrs)</td>
<td>6.40±.74</td>
<td>6.04±.82a</td>
<td>6.33±.60</td>
<td>6.09±.66</td>
</tr>
<tr>
<td>Time in bed (TIB)</td>
<td>457.4±47.5</td>
<td>422.8±55.6a</td>
<td>457.3±42.</td>
<td>428.8±38.4a</td>
</tr>
<tr>
<td>Sleep latency (SLAT)</td>
<td>12.2±8.9</td>
<td>20.0±18.5a</td>
<td>18.0±11.4</td>
<td>17.3±12.0</td>
</tr>
<tr>
<td>Total sleep time (TST)</td>
<td>399.2±43.8</td>
<td>316.8±77.1a</td>
<td>389.1±51.6</td>
<td>335.5±43.2a</td>
</tr>
<tr>
<td>Total wake time (TWT)</td>
<td>57.7±27.4</td>
<td>105.5±54.9a</td>
<td>67.7±27.0</td>
<td>92.8±43.2a</td>
</tr>
<tr>
<td>WASO (TWT - SLAT)</td>
<td>46.0±25.2</td>
<td>86.0±48.5a</td>
<td>50.2±25.2</td>
<td>75.7±42.7a</td>
</tr>
<tr>
<td>WASO%</td>
<td>10.2±5.2</td>
<td>21.2±11.6a</td>
<td>11.3±5.3</td>
<td>18.3±9.6a</td>
</tr>
<tr>
<td>Sleep efficiency (SE; %)</td>
<td>87.4±5.2</td>
<td>74.3±14.8a</td>
<td>84.9±6.0</td>
<td>78.4±9.6a</td>
</tr>
<tr>
<td>Total wakes (WAKES; #)</td>
<td>18.1±6.7</td>
<td>17.1±8.2</td>
<td>18.2±6.0</td>
<td>17.1±6.6</td>
</tr>
<tr>
<td>Wakes $1 minute (TAW1; #)</td>
<td>7.8±3.7</td>
<td>9.3±5.2a</td>
<td>7.4±3.6</td>
<td>8.5±4.8</td>
</tr>
<tr>
<td>REM latency (REML)</td>
<td>69.6±31.9</td>
<td>89.8±62.3a</td>
<td>71.6±45.6</td>
<td>74.3±42.6</td>
</tr>
<tr>
<td>Minutes REM (REM)</td>
<td>78.0±21.5</td>
<td>60.3±28.9a</td>
<td>77.2±22.8</td>
<td>65.6±24.0a</td>
</tr>
<tr>
<td>REM as % of TST (%REM)</td>
<td>19.5±4.8</td>
<td>19.0±6.7</td>
<td>19.8±5.1</td>
<td>19.3±6.2</td>
</tr>
<tr>
<td>Minutes Stage 2 (%STg2)</td>
<td>254.4±38.6</td>
<td>201.8±57.8a</td>
<td>244.1±49.8</td>
<td>202.4±43.2</td>
</tr>
<tr>
<td>STg 2 as % of TST (%STg2)</td>
<td>63.9±8.4</td>
<td>63.3±9.5</td>
<td>62.4±8.5</td>
<td>60.3±10.9</td>
</tr>
<tr>
<td>Minutes SWS (SWS)</td>
<td>71.6±32.6</td>
<td>51.0±28.9a</td>
<td>69.6±49.2</td>
<td>63.0±45.6</td>
</tr>
<tr>
<td>SWS as % of TST (%SWS)</td>
<td>17.8±7.5</td>
<td>16.9±9.9</td>
<td>18.3±13.6</td>
<td>19.0±13.8</td>
</tr>
</tbody>
</table>

αp<.05, baseline versus catheter night, for same group; βp<.05, NERT versus ERT women, for same night
culties were distributed equally between the two groups. There were two types of sampling problems. (1) The catheter lost its patency, due to infiltration, venous spasm, or unknown cause. This typically occurred after several difficult or sluggish samples. This was the most common type of sampling difficulty. When this occurred, the catheter was not replaced. (2) One or a few samples were difficult to draw, requiring the night nurse to enter the subject’s room and manipulate the catheter and/or the subject’s arm. There was no difference among the groups in the relative frequency of each type of problem. Sleep-wake analyses were repeated after excluding those subjects. These analyses yielded the same results as when subjects with sampling difficulties were included; therefore, the results described below are for the larger, more inclusive sample (ERT=37, NERT=56).

RESULTS

Demographic and Psychosocial Variables

There were no group differences in age, body mass index, education, cognitive level, depression, anxiety, stress, or subjective sleep quality (Table 2). The mean PSQI scores for each group were <5, which is the criterion used to identify “good” vs. “poor” sleepers.32 Approximately 20% of subjects in both groups had PSQI scores above 5 (“poor” sleepers). The highest score in each group was 12, out of the maximum 21 possible points. The Subjective Sleep Quality component of the PSQI was also analyzed (non-parametrically), and showed no difference for NERT (median=0; percents of group having a score of 0, 1, or 2 were 60.7%, 35.7%, and 3.6%) vs. ERT (median=0; percents with 0, 1, or 2 were 67.6%, 32.4%, and 0%).

NERT vs. ERT on Baseline Night

The first series of analyses compared the sleep of ERT vs. NERT women on the baseline night. As shown in Table 3, there were no NERT–ERT differences for any sleep variables except sleep latency and sleep efficiency. ERT women averaged a longer sleep latency and a lower sleep efficiency than NERT women. The group difference in sleep efficiency was primarily a consequence of the group difference in sleep latency. When the sleep efficiency analysis was repeated, using sleep latency as a covariate, the group difference in sleep efficiency was no longer evident. In other words, the amount of sleep as a percent of time in bed (i.e., sleep efficiency) was different between the groups, because of the group difference in sleep latency. However, the amount of sleep as a percent of sleep period time (time in bed minus sleep latency) was not different for the two groups.

The exact p-value for the sleep latency comparison was p=.008. As the data displayed a significantly non-normal distribution, a Mann-Whitney (non-parametric) test was subsequently used to repeat the analysis (p=.011). To assess the reliability of this finding, a power analysis was performed. The power of the non-parametric test data was 90%, at p=.05. There were no out-

Figure 1—Polysomnographic sleep-wake variables for the catheter night are depicted as median percent change from the baseline night, for each group. Medians rather than means are depicted because they are a more accurate representation of central tendency for these percent change variables. Abbreviations are defined in Table 3. The NERT group is significantly different compared to the ERT group (p<.05, indicated by **) for: TWT, SLAT, SE, SWS, and %SWS.
liers in either group that contributed disproportionately to the outcome, and the variances of the two groups did not differ. The median sleep latency for NERT women was 10.7 minutes; the median sleep latency for ERT women was 14.5 minutes.

To further explore this issue, the sleep latency component of the PSQI was then examined. This sleep quality questionnaire contains numerous specific questions about subjects’ sleep at home during the preceding month. Two questions assess sleep latency; the answers are coded and combined to yield a PSQI Sleep Latency score with a range of 0–3 and a higher score indicating a longer sleep latency. The NERT group averaged a longer sleep latency at home: 0.74±0.12 (SEM), with a median score of 1. The ERT group averaged 0.39±0.01, with a median score of 0 (Mann-Whitney test, p=.047).

**Sleep During Baseline vs. Catheter Nights**

The second series of analyses compared sleep variables from the baseline night vs. the catheter night, separately for each group. The sleep of NERT women was significantly and negatively affected by the blood sampling procedure (Table 3). On the catheter night, NERT women spent less time in bed, took longer to fall asleep, and spent more time awake and less time in REM, stage 2 and SWS sleep (in absolute minutes, though not as a percent of total sleep time). In contrast, the sleep of ERT women appeared to be less affected by the catheter. As with NERT women, ERT women spent less time in bed on the catheter night, took longer to fall asleep, and spent more time awake and less time in REM and stage 2 sleep. However, SWS and several other measures (e.g., number of wakes) were unaffected.

**Sleep Disruption: NERT vs. ERT**

The effect of the catheter and blood sampling procedure on sleep was compared for the two groups by examining the percent change from baseline to catheter night for each group (Figure 1). There were significantly larger percent changes from baseline to catheter night in the sleep of NERT women compared with ERT women in sleep latency, total wake time, sleep efficiency, minutes of SWS, and WASO as a percent of sleep period time. The percent change from baseline for all remaining sleep-wake variables was also numerically smaller for ERT compared to NERT women, but there was large variance associated with most of these variables (e.g., TAW1) and these group comparisons were not statistically significant. Statistical significance was defined as p≤.05. However, because of the multiple t-test comparisons, it should be noted that all of the exact p-values for each group’s baseline vs. catheter comparisons were less than .005, except for the following: TAW1 (NERT group only, p=.02); REML (NERT group, p=.02). Lights out (ERT group, p=.013); minutes of REM (ERT group, p=.01).

**Subjective Sleep Quality**

Analyses of subjective sleep ratings from the morning SQQ yielded results parallel to the polysomnographic analyses (Table 4). There were no NERT-ERT differences in subjective sleep measures on the baseline night. However, the impact of the catheter night again was more significant for the NERT group. The subjective sleep quality rating (from the 1—9 scale) was significantly lower for NERT women on the catheter night compared to the NERT baseline night and also compared to ERT rat- ings for the catheter night. This is also illustrated by the median percent change on the catheter night from baseline (18.3% for NERT women and 0.0% for ERT women, p=.05). The sleep quality ratings were subsequently grouped together for further analysis as follows: significantly better than or the same as at home (rating=5—9); and worse than at home (rating=1—4). For the adaptation night, there were no significant differences in the percent of NERT women who said they slept significantly worse (28.6%) or better (10.7%) than at home, compared with ERT women (32.0%, 10.8%). Similar results were obtained for the baseline night (NERT: 14.3% worse, 14.3% better; ERT: 13.5% worse, 16.2% better). However, after the catheter night, 50.0% of NERT women and 24.3% of ERT women said they slept significantly worse than at home (p<.05 for NERT baseline-catheter comparison; p<.05 for NERT-ERT comparison on catheter night). Fewer NERT women said they slept better that night than at home (8.9%), compared with ERT women (18.9%). NERT women also reported they were more likely on the catheter night to lie awake in bed trying to sleep; this was not true of ERT women.

The percent change (from baseline to catheter nights) in subjective sleep quality rating was significantly correlated with the percent change in several polysomnographic sleep-wake variables, for both groups. A decreased sleep quality rating on the catheter night relative to the baseline night was correlated with increased TWT (r = .40 and .32 for NERT and ERT groups, respectively), and decreased TST (.31, .43), sleep efficiency (.37, .39).
.38), and percent REM sleep (.41, .31) on the catheter night relative to the baseline night (p<.05, all comparisons).

DISCUSSION

These results indicate that, in this cross-sectional study, the use of chronic oral ERT had little effect on the sleep of older post-menopausal women without hot flashes, except in the presence of a challenge to sleep (i.e., frequent nocturnal blood sampling). ERT appeared to ameliorate the disruptive effect of this stressor on both objectively assessed and subjectively assessed sleep.

Baseline Sleep

It was somewhat surprising to observe that there was a significant group difference in only two of the polysomnographic sleep variables on the baseline night: sleep latency and sleep efficiency. (However, our analyses clearly indicate that the sleep efficiency difference was secondary to the sleep latency difference. When sleep latency was subtracted, the amount of sleep as a percent of sleep period time was not different for the two groups.) Moreover, it was the NERT women whose sleep was better according to these two measures. This contrasts with most other polysomnographic studies of estrogen effects, which have reported improved sleep following ERT.6,10,12,27 There are substantial and meaningful differences between those studies and the one reported here. The present study was comprised of older (mean = 70 yrs; range=57-80) women, and study criteria excluded the participation of women with hot flashes, nocturnal diaphoresis, insomnia, and depression. The studies reporting positive effects of ERT on sleep have been based on much younger women who had been amenorrheic for much shorter times on average (as little as three months).10 In addition, the ERT treatment in those studies was relatively short, ranging from four to eight weeks, in contrast to the long-term ERT of the women described here (76% had been taking estrogen for 10 years or more).

However, the most important difference between this study and those studies lies in another characteristic of the subject populations: the presence of hot flashes, nocturnal diaphoresis (“night sweats”), and/or disturbed sleep. Such symptoms were an inclusion criterion in three of the four studies, and were experienced by a substantial majority of women in the fourth study. Moreover, some studies specifically recruited subjects from menopause clinics. In the present study, women were recruited from the general population. Significant sleep difficulty was an exclusion criterion for the study. None of the women were experiencing hot flashes or nocturnal diaphoresis.

Three more recent studies are consistent with this analysis. Polo-Kantola et al11 found no effect of short-term ERT on standard polysomnographic variables. Their relatively large subject population (n=62) consisted of women drawn from the general population, with a wide variety of self-reported mild to severe subjective menopausal symptoms. However, the authors reported that the measured baseline occurrence of those symptoms was actually low. Antonijevic et al18 also found no estrogen effects on all-night polysomnographic variables, in women not experiencing hot flashes, though estrogen decreased wakefulness and increased REM sleep during the first two sleep cycles of the night. Interpretation of this small study (n=11) is complicated by the inclusion of two different ERT conditions. Half the women were on ERT for a mean of five years before the study began. For those subjects, assessment of estrogen effects on sleep were made first, followed by a baseline (non-ERT) assessment after a two-week “wash-out” period. The other women in the study were not on ERT when the study began. They underwent a baseline assessment first; the estrogen treatment assessment occurred after two weeks of ERT. That is, half of the subjects experienced ERT for about five years, the other half of the subjects experienced ERT for only two weeks. Finally, Purdie et al39 studied relatively young (mean age = 54) self-reported symptomatic women drawn from a menopause clinic, yet found no change in polysomnographic variables following hormone treatment. As with Polo-Kantola et al11 the objective baseline assessments showed that their subjects’ menopausal symptoms were quite mild (e.g., only about 10 minutes of wakefulness during the night). Also, theirs is the only published study to date that used hormone replacement therapy (HRT; estrogen plus a progestin) rather than ERT (estrogen alone). The possible sleep effects of adding a progestin to ERT clearly deserve further study.

The shorter sleep latency of the NERT group relative to the ERT group is puzzling, particularly as the PSQI questionnaire revealed a longer sleep latency at home for the NERT group. Discrepancies between subjective and objective measures of sleep have been previously reported by many investigators, particularly for insomnia patients (reviewed in 49). However, it is possible that at least some NERT women were experiencing chronic difficulty initiating sleep at home such that they fell asleep sooner in the controlled laboratory sleep environment. Also, NERT women may have found the preceding adaptation night more disruptive than ERT women and consequently exhibited some recovery on the baseline recording night, though the lack of a NERT-ERT difference in the subjective sleep ratings for the adaptation night is inconsistent with that notion. Finally, though this latency finding is statistically significant, the difference between the two groups was less than four minutes (medians) which may or may not have clinical significance.

Sleep Disturbance Consequent to Nocturnal Blood Sampling

Sleep maintenance and sleep architecture were disrupted in both groups by the nocturnal blood sampling protocol, as indicated by a longer sleep latency, increased nocturnal wakefulness, and decreased time in REM and stage 2 sleep. This is consistent with most other studies41,42 and with our previously reported finding that this protocol negatively affected sleep in a different population of older men and women.20 One study43 failed to find an impact of this procedure on sleep, but their subjects consisted of a small group (n=8) of young men. Here, as in previous studies (e.g., 20,42), the subjects consisted entirely or largely of women, and there were large individual differences in the degree of disruption, ranging from subjects who experienced no apparent disruption to others whose sleep was severely disrupted.

The sleep of NERT women was clearly more disrupted than the sleep of ERT women, as indicated by both objective and subjective sleep measures. Baseline vs. catheter night comparisons performed separately for each group showed that most polysomnographic variables were affected by nocturnal blood sampling in NERT women. In contrast, far fewer measures were affected in ERT women.

It is interesting to compare these results with those that we
described previously for older men and women, in which sleep was significantly more disrupted on the catheter night for the women compared with the men. The women we described in that previous report were a heterogeneous group with regard to estrogen use: 60% were not taking estrogen, 28% were on ERT, and 12% were taking estrogen and progesterin. The sleep of ERT women described here is similar to the sleep we observed for older men in that previous study, particularly with regard to the lack of significant baseline vs. catheter differences in SWS and sleep latency.

Group differences in baseline sleep might lead to group differences in the impact of nocturnal blood sampling on sleep, especially on SWS. This does not explain the results observed here. There were no group differences in any baseline sleep-wake variables except sleep latency, including variables such as SWS for which there were large NERT-ERT differences in the effect of nocturnal blood sampling.

Group differences in age and psychosocial variables or in blood sampling difficulties might also result in differential impact of the catheter night on sleep. Again, there was no evidence of such group differences. NERT and ERT women were of the same average age and educational level, and they experienced the same low incidence of blood sampling difficulties on the catheter night. Individuals with significant depression, stress, or anxiety were excluded from the study during the screening process, and there were no group differences on these measures. There were also no group differences in body mass index, general health, incidence of specific illnesses, or the use of prescription and non-prescription medications (except for estrogen). Women taking estrogen have been described as likely to have more frequent contacts with health care systems, and as possibly healthier, than women not taking estrogen. That is unlikely in this study because of the rigorous screening process and the lack of group differences just described. In short, there are no apparent differences between the NERT and ERT groups other than estrogen use that might account for the results observed here.

Estrogen and Stress Reactivity

One mechanism by which estrogen might influence susceptibility to sleep disruption is through its effects on stress reactivity. Low estrogen states such as the postmenopausal years are associated with enhanced cardiovascular and hormonal responses to stress, whereas ERT seems to lower stress reactivity. Intravenous catheterization is a mild stressor as indicated by cortisol levels and the current study, and anecdotal conversations with these subjects and the subjects in our earlier study. The differential impact of nocturnal blood sampling on the sleep of NERT vs. ERT women reported here is consistent with this hypothesis about the relationship between estrogen, stress reactivity, and susceptibility to sleep disruption. These data are also consistent with reports that the sleep of older women (heterogeneous with regard to estrogen use) is more susceptible to disruption by environmental factors and/or stressors than the sleep of older men.

The use of ERT has increased tremendously during the past two decades, with an estimated one in six postmenopausal women taking estrogen in 1992. This has been accompanied by a change in emphasis from relatively short-duration ERT for the treatment of hot flashes and other perimenopausal discomfort to long-term ERT after menopause, as a prophylactic measure against osteoporosis, cardiovascular disease, and possibly cognitive decline. The results described here suggest that another benefit of long-term ERT in older, non-symptomatic, postmenopausal women may be decreased susceptibility to sleep disruption under some circumstances. As sleep concerns and complaints increase dramatically with age in women, and the sleep of older women seems more susceptible to environmental disruption, this possibility deserves further study. Moreover, there are several other mechanisms by which estrogen might plausibly influence sleep and possibly health status among the women of the French GAZEL cohort. Maturitas 1995;20:113-20.

REFERENCES

41. Adam K. Sleep is changed by blood sampling through an indwelling venous catheter. Sleep 1982;5:154-8.