Pharmacologic Treatment of Insomnia Disorder: An Evidence Report for a Clinical Practice Guideline by the American College of Physicians

Timothy J. Wilt, MD, MPH; Roderick MacDonald, MS; Michelle Brasure, PhD, MSPH, MLIS; Carin M. Olson, MD, MS; Maureen Carlyle, MPH; Erika Fuchs, PhD, MPH; Imran S. Khawaja, MD; Susan Diem, MD, MPH; Erin Koffel, PhD; Jeannine Ouellette; Mary Butler, PhD; and Robert L. Kane, MD

Background: Pharmacologic interventions are often prescribed for insomnia disorder.

Purpose: To assess the benefits, harms, and comparative effectiveness of pharmacologic treatments for adults with insomnia disorder.


Study Selection: 35 randomized, controlled trials of at least 4 weeks’ duration that evaluated pharmacotherapies available in the United States and that reported global or sleep outcomes; 11 long-term observational studies that reported harm information; FDA review data for nonbenzodiazepine hypnotics and orexin receptor antagonists; and product labels for all agents.

Data Extraction: Data extraction by single investigator confirmed by a second reviewer; dual-investigator assessment of risk of bias; consensus determination of strength of evidence.

Data Synthesis: Eszopiclone, zolpidem, and suvorexant improved short-term global and sleep outcomes compared with placebo, although absolute effect sizes were small (low- to moderate-strength evidence). Evidence for benzodiazepine hypnotics, melatonin agonists, and antidepressants, and for most pharmacologic interventions in older adults, was insufficient or low strength. Evidence was also insufficient to compare efficacy within or across pharmacotherapy classes or versus behavioral therapy. Harms evidence reported in trials was judged insufficient or low strength; observational studies suggested that use of hypnotics for insomnia was associated with increased risk for dementia, fractures, and major injury. The FDA documents reported that most pharmacotherapies had risks for cognitive and behavioral changes, including driving impairment, and other adverse effects, and they advised dose reduction in women and in older adults.

Limitations: Most trials were small and short term and enrolled individuals meeting stringent criteria. Minimum important differences in outcomes were often not established or reported. Data were scant for many treatments.

Conclusion: Eszopiclone, zolpidem, and suvorexant may improve short-term global and sleep outcomes for adults with insomnia disorder, but the comparative effectiveness and long-term efficacy of pharmacotherapies for insomnia are not known. Pharmacotherapies for insomnia may cause cognitive and behavioral changes and may be associated with infrequent but serious harms.

Primary Funding Source: Agency for Healthcare Research and Quality.

For author affiliations, see end of text.
This article was published at www.annals.org on 3 May 2016.

Sleep problems are common and associated with a decline in overall and sleep-related health (1–3). Insomnia is more common in women and older adults (4). Aging is often accompanied by disrupted sleep and frequent and early waking (5, 6). Other conditions coexist with, are due to, or lead to poor sleep (7).

Insomnia symptoms are typically transient and may not cause distress or impaired activity. However, insomnia disorder as defined by the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, requires sleep problems that are chronic, persistent, and associated with daytime dysfunction. Insomnia disorder includes a predominant symptom of difficulty with sleep initiation, difficulty maintaining sleep, or early-morning waking with inability to return to sleep causing clinically significant distress or impairment in activities, occurring at least 3 nights per week. Current definitions require insomnia symptoms to have persisted for 3 months or more (8). Individuals must have adequate opportunity for sleep, and symptoms must not be better explained by another sleep disorder, effects of a substance, or other medical or mental conditions.

Treatment goals include meaningful improvements in sleep-associated distress or dysfunction (global outcomes). The Insomnia Severity Index (ISI) and the Pittsburgh Sleep Quality Index measure both problems and worry about sleep and accompanying distress or dysfunction. Sleep measures are based on specific sleep variables that can be assessed in a sleep laboratory with polysomnography or actigraphy or subjectively with patient-reported sleep diaries. These include sleep-onset latency (SOL), wake after sleep onset (WASO), total sleep time (TST), and sleep efficiency (percentage of time in bed sleeping). In general, insomnia disorder is treated by clinicians on the basis of patient-reported sleep-associated distress, not laboratory assessment. Additionally, polysomnography is not necessary.
indicated for the diagnosis or treatment of insomnia disorder (9).

Prescription medications are often used. Several are approved for insomnia, typically for short-term use, by the U.S. Food and Drug Administration (FDA). Although medication indications often focus on specific sleep variables, it is not known how frequently primary care physicians target medications to specific or global measures of insomnia or prescribe them long term.

We evaluated the efficacy, comparative effectiveness, and harms of pharmacologic interventions for insomnia disorder. An accompanying article assesses psychological and behavioral interventions (10). These articles summarize findings from a report conducted by the Agency for Healthcare Research and Quality/Minnesota Evidence-based Practice Center (11). The work served, in part, as the evidence base for an American College of Physicians’ clinical practice guideline.

METHODS

We developed and followed a standard protocol. The protocol was registered at www.crd.york.ac.uk/PROSPERO/ (registration number: CRD42014009908) in May 2014 and published at http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and -reports/?pageaction=displayproduct&productid=1888. Full details of the findings and evidence tables are available in the report (11).

Data Sources and Searches

We searched MEDLINE, Embase, Cochrane Central Register of Controlled Trials, and PSYCinfo from 2004 through September 2015 for relevant English-language literature (Table 1 of the Supplement, available at www.annals.org), reference lists of pertinent studies, FDA Web sites and product labels, and ClinicalTrials.gov.

Study Selection

Two investigators independently reviewed titles and abstracts of search results and screened the full text of potentially eligible English-language references. We included randomized, controlled trials (RCTs) of pharmacologic therapies available in the United States, regardless of the comparator (placebo, another medication, nonpharmacologic therapy), if they enrolled adults with insomnia disorder, provided at least 4 weeks of follow-up, and reported global or sleep outcomes. We excluded from analysis studies deemed to have high risk of bias. We included observational studies that reported harms if 1) the population included 100 or more adults with chronic insomnia without other major diagnoses, such as cancer and Parkinson disease; 2) the pharmacologic agents evaluated were FDA-indicated for insomnia and probably administered for sleep disorders; 3) study duration was at least 6 months; and 4) harms were reported by drug class. We also examined harms reported in the following: eligible trials; FDA Web sites for nonbenzodiazepine hypnotics, eszopiclone, zaleplon, zolpidem, and the orexin receptor antagonist suvorexant; FDA product labels for all pharmacologic therapies; and systematic reviews.

Data Extraction and Quality Assessment

One investigator extracted study data, which was independently checked by a second investigator. Our main outcome of interest was patient-reported global measure of effectiveness: improvement in sleep variables and daytime functioning and distress. The 7-item ISI (score range, 0 to 28) or the 7-component, 19-item Pittsburgh Sleep Quality Index (score range, 0 to 21) were commonly used instruments measuring both problems and worry about sleep and accompanying distress or dysfunction (Table 2 of the Supplement, available at www.annals.org). We also assessed sleep-specific parameters on the basis of patient-recorded sleep diaries, including SOL, TST, WASO, sleep efficiency, sleep quality, function, mood, and quality of life.

Risk of bias for RCTs was independently assessed by 2 investigators using an instrument developed for this project (12). We classified overall risk of bias for each study as low, moderate, or high according to a summary of bias risk across domains and confidence that the results were believable given the study’s limitations. We excluded studies with high risk of bias from analysis.

Data Synthesis and Analysis

We grouped studies and rated strength of evidence by drug. Evidence ratings within each comparison were determined as high, moderate, low, or insufficient (13) by at least 2 trained research associates and the principal investigator; final determinations were ascertained through consensus. For assessments of efficacy, we used established minimum important differences for main outcomes if they were available and statistical significance alone if minimum important differences were not established.

We pooled results for doses approved by the FDA if more than 1 study assessed that dose, used a similar comparator (such as placebo), involved a similar study population (for example, elderly patients), and assessed similar outcomes (such as the same sleep variable). Data were analyzed in RevMan 5.3 (Nordic Cochrane Center). DerSimonian and Laird random-effects models were used to calculate pooled risk ratios and absolute risk differences with 95% CIs for categorical outcomes. Weighted mean differences or standardized mean differences with 95% CIs were calculated for continuous outcomes. We assessed statistical heterogeneity with the Cochran Q test and measured magnitude with the I2 statistic (I2 > 75% indicates substantial heterogeneity) (14). We searched several databases to identify and reduce potential publication bias.

Role of Funding Source

This review was nominated to the Agency for Healthcare Research and Quality Effective Healthcare Program by an anonymous individual and funded by the Agency for Healthcare Research and Quality. Agency staff and key informants representing various perspectives offered suggestions as we developed and refined the scope of the review. A technical expert panel provided input to the protocol and reviewed the draft report. The American College of Physicians pro-

Downloaded From: http://annals.org/ by a University of New Mexico User  on 05/14/2016
vided support for manuscript preparation. The authors are solely responsible for its content.

RESULTS

Search results are shown in the Appendix Figure (available at www.annals.org). Thirty-five RCTs with acceptable risk of bias evaluated pharmacologic treatments (15–46) (Figure). Six were in older adults (42–47). We did not find evidence of publication bias in a search of ClinicalTrials.gov through September 2015. Eleven observational studies provided data for long-term harms (48–58). Drug doses used in older adults were lower than those used in general populations. Few antidepressant or benzodiazepine trials met inclusion criteria, primarily because of short treatment durations. Patients were typically diagnosed with insomnia disorder according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria, which require symptoms to be present for at least 1 month (current criteria require symptoms for at least 3 months).

The mean symptom duration was rarely reported. Additional enrollment criteria were based on thresholds for SOL, TST, WASO, or number of awakenings per night during a typical night in the month before enrollment. Only 2 trials based enrollment on, or provided baseline scores of, global measures. Most required TST of 6.5 hours or less or SOL of greater than 30 minutes. When WASO was included, thresholds ranged from 30 to 120 minutes. Common exclusion criteria included psychiatric disorders or substance abuse (Tables 3 and 4 of the Supplement, available at www.annals.org). Trials rarely assessed treatments for longer than 4 weeks. Most participants were women, white, and younger than 50 years. Most studies were industry-sponsored. Doses used in some trials exceeded FDA recommended doses, especially for women and older or debilitated adults (Tables 1 and 2).

Global outcomes were measured less often than were sleep variables. Minimum important differences were rarely used for global outcomes and were not established for sleep outcomes. We did not identify minimum important differences for sleep variables. Because most studies did not provide mean baseline global or sleep values, percentage improvement from baseline to follow-up frequently could not be calculated. A large placebo response was often noted. For example, in a trial evaluating zolpidem as needed, 24% of participants randomly assigned to placebo were rated as “much or very much improved” on the Clinician Global Impression (CGI) scale. Another trial evaluating extended-release zolpidem reported that 48% of individuals randomly assigned to placebo had a CGI rating of “much or very much improved.” In 2 trials of suvorexant, 42% of individuals randomly assigned to placebo had a “treatment response,” defined by a change from baseline to end of study of at least 6 points on the ISI (30). A large placebo response was also noted in SOL, TST, and WASO: 19-, 39-, and 30-minute improvements from baseline to follow-up frequently could not be calculated. A large placebo response was often noted. For example, in a trial evaluating zolpidem as needed, 24% of participants randomly assigned to placebo were rated as “much or very much improved” on the Clinician Global Impression (CGI) scale. Another trial evaluating extended-release zolpidem reported that 48% of individuals randomly assigned to placebo had a CGI rating of “much or very much improved.” In 2 trials of suvorexant, 42% of individuals randomly assigned to placebo had a “treatment response,” defined by a change from baseline to end of study of at least 6 points on the ISI (30). A large placebo response was also noted in SOL, TST, and WASO: 19-, 39-, and 30-minute improvements from baseline, respectively.

Nonbenzodiazepine Hypnotics

The FDA labeling information for nonbenzodiazepine hypnotics (eszopiclone, zaleplon, and zolpidem) warns of daytime memory and psychomotor impairment, abnormal thinking and behavioral changes, complex behaviors (such as sleep driving, which is driving
after taking a hypnotic but having no recollection of driving), and depression and suicidal thoughts and actions, the latter primarily in patients taking both sedatives and hypnotics (Table 5 of the Supplement, available at www.annals.org). Observational data indicated that hypnotic drugs, benzodiazepines and nonbenzodiazepines, were associated with dementia (hazard ratio, 4.6 points [CI, 1.5 to 8.5]) (49).

### Eszopiclone

Three RCTs with moderate risk of bias (n = 1929) analyzed the efficacy of eszopiclone, 2 to 3 mg daily (Tables 6 and 7 of the Supplement, available at www.annals.org) (15-17). One 12-week study showed that eszopiclone improved global outcomes compared with placebo (−9 versus −5 points) (16). Compared with placebo, the mean change in ISI score between groups (−4.6 points [CI, −5.3 to −3.9]) was less than the threshold used to define a "treatment response" versus a comparator (that is, a 7-point difference). At 12 weeks, eszopiclone more often improved ISI scores to less than 7 at end point, indicating remission or no clinically significant insomnia, compared with placebo (50% versus 19%) (low-strength evidence). Pooled results showed that eszopiclone reduced SOL by 19 minutes and increased TST by 45 minutes compared with placebo (moderate-strength evidence). However, mean SOL remained greater than 30 minutes and TST greater than 6.5 hours in all 3 trials. Eszopiclone reduced WASO by 11 minutes compared with placebo (low-strength evidence). One study in older adults (n = 388) found low-strength evidence that eszopiclone, 2 mg, increased the percentage of patients having a minimum important difference in global outcomes versus placebo (37% versus 24%) (43). The mean effect size on the ISI score was small. Compared with placebo, TST and WASO improved 30 and 22 minutes, respectively.

Trials showed that somnolence, unpleasant taste, and myalgias were higher with eszopiclone than placebo. Serious adverse events, reported in 1 trial (15), were also higher (3% versus 1%). Data reported in medical reviews for the FDA indicated higher incidences of memory impairment, psychiatric adverse effects, depression, anxiety, and accidental injury with eszopiclone compared with placebo (Table 8 of the Supplement, available at www.annals.org). In an open-label extension of an RCT evaluating eszopiclone, serious adverse events leading to study withdrawal occurred in 2% of individuals (53) (Table 9 of the Supplement, available at www.annals.org).

### Zaleplon

Two 4-week RCTs with moderate risk of bias compared zaleplon (5 to 20 mg) with placebo (n = 973) (Tables 6 and 7 of the Supplement) (18, 19). Neither reported global outcomes. Zaleplon, 5 mg, did not improve SOL or other sleep variables. Although a statistically significant 10-minute improvement in SOL was reported with the 10-mg dose, mean SOL in the zaleplon group remained greater than 30 minutes. Total sleep

---

**Table 1. Nonbenzodiazepine and Other Drugs Approved for Insomnia**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Trade Name</th>
<th>Indications</th>
<th>Recommended Dosage</th>
<th>Dosage in Older/Debilitated Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nonbenzodiazepine “Z” drugs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eszopiclone*</td>
<td>Lunesta</td>
<td>Improve sleep onset and maintenance</td>
<td>1 mg; maximum 3 mg</td>
<td>Not to exceed 2 mg</td>
</tr>
<tr>
<td>Zaleplon†</td>
<td>Sonata</td>
<td>Short-term use to improve sleep onset</td>
<td>10 mg</td>
<td>5 mg</td>
</tr>
<tr>
<td>Zolpidem†</td>
<td>Ambien</td>
<td>Short-term use to improve sleep onset</td>
<td>Men: 5-10 mg§</td>
<td>5 mg</td>
</tr>
<tr>
<td>Zolpidem ER‡</td>
<td>Ambien CR</td>
<td>Improve sleep onset and maintenance</td>
<td>Women: 5 mg§</td>
<td>6.25 mg</td>
</tr>
<tr>
<td>Zolpidem SL¶</td>
<td>Edluar</td>
<td>Insomnia (short-term); improve sleep onset</td>
<td>Men: 5-10 mg</td>
<td>5 mg</td>
</tr>
<tr>
<td>Zolpidem SL**</td>
<td>Intermezzo</td>
<td>As-needed use to improve sleep onset after middle-of-the-night awakenings (≥4 h of bedtime remaining before awakening)</td>
<td>Women: 1.75 mg</td>
<td>1.75 mg</td>
</tr>
<tr>
<td><strong>Orexin receptor antagonist</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suvorexant††</td>
<td>Belsomra</td>
<td>Improve sleep onset and/or maintenance</td>
<td>10 mg; maximum 20 mg</td>
<td>Lowest effective dose advised</td>
</tr>
<tr>
<td><strong>Melatonin receptor agonist</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramelteon‡‡</td>
<td>Rozerem</td>
<td>Improve sleep onset</td>
<td>8 mg</td>
<td>8 mg</td>
</tr>
<tr>
<td><strong>Tricyclic antidepressant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doxepin§§</td>
<td>Silenor</td>
<td>Improve sleep maintenance</td>
<td>6 mg; 3 mg if appropriate</td>
<td>3 mg; maximum 6 mg</td>
</tr>
</tbody>
</table>

ER = extended release; SL = sublingual.

* www.lunesta.com/PostedApprovedLabelingText.pdf.
‡ www.accessdata.fda.gov/drugsatfda_docs/label/2006/019908s027lbl.pdf.
¶‡‡ http://general.takedapharm.com/content/file.aspx?filetypecode=rozerempi&cacheRandomizer=d4d90170-f7cb-4a9c-bc5a-81f1eb8b5318.
www.edluar.com/EDLUAR-Pi.pdf.
‡ www.accessdata.fda.gov/drugsatfda_docs/label/2008/019908s027lbl.pdf.
¶ www.edluar.com/EDLUAR-Pi.pdf.
‡‡ http://general.takedapharm.com/content/file.aspx?filetypecode=rozerempi&cacheRandomizer=d4d90170-f7cb-4a9c-bc5a-81f1eb8b5318.
evidence. Global outcomes were not reported or had insufficient power compared with placebo (44). Other sleep and wake variables were reported that zolpidem, 5 mg, reduced SOL by 18 minutes compared with placebo (22). One study in older adults (45) showed zolpidem, 10 mg as needed, reduced SOL by 15 minutes and increased TST by 48 minutes compared with placebo on nights when medication was taken (moderate-strength evidence) (25, 26). Sleep-onset latency remained greater than 30 minutes, and there was no difference in WASO. Another RCT reported smaller and nonsignificant changes in sleep variables (24).

### Zolpidem

Six RCTs with moderate risk of bias (n = 859) lasting 4 to 32 weeks compared zolpidem with placebo (Tables 6 and 7 of the Supplement) (18–23). No trial reported global outcomes. Effect on sleep variables was mixed. Zolpidem (5 to 10 mg) reduced SOL by 15 minutes compared with placebo in pooled results from 4 trials (18, 21–23) (moderate-strength evidence). Mean SOL remained greater than 30 minutes. In the 1 small longer-term trial (n = 91), Randall and colleagues reported no statistically significant difference in SOL versus placebo (22). In two 4-week trials not pooled, SOL results were mixed (18, 20). Zolpidem improved the proportion of participants “getting a better night’s sleep” versus placebo (69% versus 49%) (moderate-strength evidence) (18–20). Pooled trials showed that zolpidem, 5 to 10 mg, increased TST by 23 minutes compared with placebo (21–23). In trials not pooled, zolpidem did not consistently improve TST (18–20). The 1 longer-term trial reported that zolpidem was not statistically significantly different from placebo for TST or WASO (22). One study in older adults (n = 166) reported that zolpidem, 5 mg, reduced SOL by 18 minutes compared with placebo (44). Other sleep and global outcomes were not reported or had insufficient evidence.

### As-Needed Zolpidem

Three RCTs with moderate risk of bias (n = 607) compared zolpidem, 10 mg as needed, with placebo (24–26). No trial lasted longer than 12 weeks. Results were mixed. Low-strength evidence from 1 RCT showed that as-needed zolpidem improved global outcomes (CGI score of “much or very much improved”) versus placebo (54% versus 24%) (24). Two trials (n = 355) showed zolpidem, 10 mg as needed, reduced SOL by 15 minutes and increased TST by 48 minutes compared with placebo on nights when medication was taken (moderate-strength evidence) (25, 26). Sleep-onset latency remained greater than 30 minutes, and there was no difference in WASO. Another RCT reported smaller and nonsignificant changes in sleep variables (24).

### Sublingual Zolpidem

One 4-week RCT with moderate risk of bias (n = 295) compared low-dose sublingual zolpidem, 3.5 mg as needed, with placebo in participants with difficulty returning to sleep after middle-of-the-night awakenings (27). No global outcomes were reported. Sublingual zolpidem reduced SOL after middle-of-the-night awakenings by 18 minutes compared with placebo (low-strength evidence).

### Extended-Release Zolpidem

One RCT with low risk of bias (n = 1018) compared extended-release zolpidem, 12.5 mg taken at least 3 nights per week, with placebo over 24 weeks (28). The CGI rating of “much or very much improved” favored zolpidem ER over placebo (85% versus 48%). Evidence was low for most sleep variables, although effect sizes were small.

### Adverse Effects of Zolpidem

Trials showed that withdrawals due to adverse effects, but not any specific adverse effect or overall withdrawals, were greater with zolpidem than placebo (6% versus 3%). Some adverse effects were noted with greater frequency in trials that evaluated as-needed, clinically significant improvements in SOL and TST were reported with zaleplon, 20 mg, compared with placebo in both trials, although this exceeds the FDA-recommended dose.

### Adverse Effects

Trials showed that withdrawals due to adverse effects, but not any specific adverse effect or overall withdrawals, were greater with zolpidem than placebo (6% versus 3%). Some adverse effects were noted with greater frequency in trials that evaluated as-needed.

### Table 2. Benzodiazepine Drugs Approved for Insomnia

<table>
<thead>
<tr>
<th>Drug</th>
<th>Trade Name</th>
<th>Indications</th>
<th>Recommended Dosage</th>
<th>Dosage in Older/Debilitated Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estazolam*</td>
<td>Prosom</td>
<td>Short-term use to improve sleep onset and reduce nocturnal awakenings and/or early morning awakenings</td>
<td>1 mg: 2 mg if appropriate</td>
<td>0.5 mg</td>
</tr>
<tr>
<td>Flurazepam†</td>
<td>Dalmane</td>
<td>Improve sleep onset and reduce nocturnal awakenings and/or early morning awakenings</td>
<td>30 mg: 15 mg if appropriate</td>
<td>15 mg</td>
</tr>
<tr>
<td>Quazepam‡</td>
<td>Doral</td>
<td>Short-term use to improve sleep onset and reduce nocturnal awakenings and/or early morning awakenings</td>
<td>7.5 mg: 15 mg if appropriate</td>
<td>Lowest effective dose advised</td>
</tr>
<tr>
<td>Temazepam§</td>
<td>Restoril</td>
<td>Short-term use to improve sleep onset and maintenance and/or to reduce early morning awakenings</td>
<td>15 mg: 7.5 or 30 mg if appropriate</td>
<td>7.5 mg</td>
</tr>
<tr>
<td>Triazolam</td>
<td></td>
<td>Halcin</td>
<td>Short-term use to improve sleep onset</td>
<td>0.25 mg: 0.125 or 0.5 mg if appropriate</td>
</tr>
</tbody>
</table>

‡ www.drugs.com/mtm/quazepam.html.
§ www.drugs.com/pro/temazepam.html.
|| www.drugs.com/pro/halcin.html.
sublingual, or extended-release zolpidem. Differences were small and not considered serious. Data reported in medical reviews for the FDA indicated that zolpidem resulted in higher incidence of psychiatric adverse effects and memory and driving impairment (Table 8 of the Supplement). In observational studies (Table 9 of the Supplement), zolpidem was also associated with risk for fractures (adjusted odds ratio, 1.72 [CI, 1.37 to 2.16]) (50). Two studies showed inconsistent results for mortality risk (50, 55). One study showed that zolpidem was associated with risk for major head injury or fracture requiring hospitalization (adjusted hazard ratio, 1.67 [CI, 1.19 to 2.34]) (48). In another, zolpidem and temazepam were associated with incident cancers (51). One open-label extension study of zolpidem, 20 mg, noted that 19% of patients withdrew from the study with adverse effects (55).

**Orexin Receptor Antagonists**

The FDA labeling information for the orexin receptor antagonist suvorexant warns of cognitive and behavioral changes, such as amnesia, anxiety, hallucinations, and other neuropsychiatric symptoms; complex behaviors, such as sleep driving; worsening of depression, including suicidal thinking in individuals with depression; daytime impairments; sleep paralysis; and hypnagogic/hypnopompic hallucinations (Table 5 of the Supplement).

Three RCTs with moderate risk of bias (n = 2811) compared suvorexant with placebo in mixed general and older populations (29, 30) (Tables 6 and 7 of the Supplement). Most participants were women, were white, and had moderate insomnia. Two trials lasted 3 months (30), and 1 (n = 781) lasted 1 year (29). The 1-year trial evaluated 30 to 40 mg of suvorexant, doses not approved by the FDA (29). The mean difference in ISI scores at 3 months was -1.2 points. Moderate-strength evidence from the 3-month trials of 15 and 20 mg showed that compared with placebo, suvorexant improved response to therapy (55% versus 42%), defined as 6-point or greater improvement from baseline in ISI score (28). Short-term therapy with suvorexant, 15 to 20 mg, reduced SOL by 6 minutes compared with placebo, but SOL remained greater than 30 minutes. Suvorexant, 15 to 20 mg, improved TST by 16 minutes and WASO by 5 minutes. Strength of evidence was moderate for all sleep variables.

Within the two 3-month trials, the specific adverse effect due to suvorexant, 15 and 20 mg, was somnolence; rates of somnolence were 7% for suvorexant and 3% for placebo (30). Withdrawal for any reason and withdrawals due to adverse effects did not significantly differ between suvorexant, 15 to 20 mg, and placebo over the short term (30). No difference between groups in the proportions of participants reporting at least 1 adverse effect was seen (moderate-strength evidence). Data reported in medical reviews for the FDA indicated few serious adverse effects, and none were more common than with placebo for doses approved for use in the United States (20 mg for nonelderly patients and 15 mg for elderly patients) (Table 10 of the Supplement, available at www.annals.org).

**Melatonin and Melatonin Agonists**

The FDA labeling information for the melatonin receptor agonist ramelteon (Tables 6 and 7 of the Supplement) warns of new cognitive or behavioral abnormalities; such complex behaviors as sleep driving; and, in primarily depressed patients, exacerbation of depression and suicidal ideation (Table 5 of the Supplement).

**Melatonin**

One RCT with moderate risk of bias (n = 791) compared melatonin, 2 mg prolonged release, with placebo (31, 32). Evidence on global and sleep outcomes was insufficient, although effect sizes were small. The type or frequency of adverse effects did not differ between groups.

**Ramelteon**

Five RCTs with moderate risk of bias (n = 3124) (2 of which were unpublished) compared ramelteon (4 to 16 mg) with placebo (33–36). All but 1 were short term (4 to 5 weeks); the exception lasted 6 months (34). No trial reported global outcomes. Ramelteon did not reduce sleep variables compared with placebo (low-strength evidence). Ramelteon improved SOL by 10 minutes but did not improve TST in a single study of older adults (44). Global outcomes and other sleep variables were not reported.

Trials showed no differences between ramelteon and placebo in the type or frequency of adverse effects, although overall withdrawals were slightly greater with ramelteon. Observational data from 2 open-label studies (n = 1403) reported longer-term harms related to ramelteon (57, 58) (Table 9 of the Supplement). Adverse effects associated with ramelteon were common but rarely were severe or required study withdrawal. Study withdrawal for any reason occurred in 58% of older adults administered ramelteon (57).

**Benzodiazepine Hypnotics**

Observational data indicated that benzodiazepines were associated with dementia (Table 9 of the Supplement) (49). The effect was greatest for higher-dose hypnotics and for benzodiazepine drugs with half-lives exceeding 24 hours (diazepam, flurazepam, and chlordiazepoxide).

One small trial with moderate risk of bias (n = 39) evaluating temazepam reported improvement in SOL and TST compared with placebo (37) (Table 11 of the Supplement, available at www.annals.org). The proportion of patients achieving sleep efficiency of 80% or greater and SOL of 30 minutes or less was similar between groups. Strength of evidence was insufficient. Two trials with moderate risk of bias (n = 208) studying temazepam in older adults reported no improvement in TST (43, 47). One trial (n = 168) reported improvement in SOL (low-strength evidence) (43), and a smaller
trial \( (n = 40) \) reported improvement in WASO \( (47) \). The evidence for all outcomes except SOL was insufficient.

No trials reported significant differences in overall withdrawals or withdrawals due to adverse effects between temazepam and placebo (insufficient-strength evidence) \( (37, 43, 47) \). Two trials did not report specific adverse effects \( (37, 47) \). In 1 trial, the proportion of participants with at least 1 adverse effect was similar between the temazepam and placebo groups \( (43) \). However, incidences of drowsiness and fatigue were higher with temazepam than with placebo.

### Antidepressants

Two RCTs with moderate risk of bias \( (n = 276) \) of 4 weeks’ duration compared doxepin with placebo \( (\text{Table 11 of the Supplement}) \) \( (38, 39) \). One trial \( (n = 229) \) reported improvements in TST and WASO with 3- and 6-mg doses compared with placebo \( (39) \). Strength of evidence was low. Another small RCT \( (n = 47) \) did not report sleep variables and found insufficient evidence on global outcomes with high doses of doxepin (25 mg to 50 mg) \( (38) \). Two trials in older adults evaluated doxepin \( (n = 494) \) over 4 to 12 weeks \( (45, 46) \). Improvement in ISI scores favored doxepin, 1 to 6 mg, compared with placebo. The mean difference in ISI scores was small. Doxepin improved sleep variables (low- to moderate-strength evidence). Adverse effects and study withdrawals did not significantly differ between participants receiving doxepin and those receiving placebo.

### Comparative Effectiveness of Pharmacologic Interventions

#### Nonbenzodiazepine Hypnotics Versus Benzodiazepine Hypnotics

One 4-week RCT with moderate risk of bias \( (n = 223) \) compared zolpidem, 10 mg, with temazepam, 20 mg \( (\text{Table 11 of the Supplement}) \) \( (40) \). Global CGI, SOL, and WASO were similar between groups. Total sleep time was increased 27 minutes with zolpidem compared with temazepam. One 4-week trial with moderate risk of bias among older adults \( (n = 166) \) compared zolpidem, 8 mg, with temazepam, 15 mg \( (43) \). Groups had similar SOL and TST. Overall, evidence was mostly insufficient. Specific adverse effects, overall withdrawals, and the proportion of participants reporting at least 1 adverse effect were not reported according to treatment group.

#### Nonbenzodiazepines Versus Nonbenzodiazepine Hypnotics

Two short-term RCTs with moderate risk of bias provided insufficient evidence about the comparative effectiveness of zaleplon versus zolpidem \( (\text{Table 11 of the Supplement}) \). Both compared zaleplon with placebo but also included a zolpidem group \( (18, 19) \). Neither trial reported global outcomes. Overall withdrawals and the proportion of participants reporting at least 1 adverse effect did not differ between the zaleplon and zolpidem groups. One trial reported a higher incidence of dizziness with zolpidem, 10 mg, than with zaleplon, 5 mg \( (14\% \text{ versus } 3\%) \) \( (19) \).

### Cognitive–Behavioral Therapy versus Drugs

Four trials with moderate risk of bias compared cognitive behavioral therapy (CBT) with a commonly used sleep medication (zolpidem \( [2 \text{ studies}] \) or temazepam \( [2 \text{ studies}] \)) or compared combination treatment with either alone \( (23, 37, 41, 47) \) \( (\text{Table 11 of the Supplement}) \). Only 1 study \( (\text{zolpidem combined with CBT versus CBT alone; } n = 163) \) reported the percentage of responders or remitters on the basis of global outcomes \( (41) \). Overall, evidence was insufficient for global outcomes and most sleep variables. Differences were small and not statistically significant. Overall withdrawals did not differ between the drug and CBT groups. The proportion of participants reporting at least 1 adverse effect was not reported.

### Discussion

Two nonbenzodiazepine hypnotics, eszopiclone and zolpidem, as well as the orexin receptor antagonist suvorexant, improved short-term global outcomes and sleep variables for adults with insomnia disorder. However, the absolute mean effect versus placebo was small, and we found evidence for harms. Evidence for benzodiazepine hypnotics, melatonin agonists, and antidepressants in general populations and for most interventions in older adults was insufficient or low strength. Comparative effectiveness evidence was limited to a few small, short-term studies, which precludes meaningful comparisons between and across categories of pharmacologic agents as well as comparisons with CBT. Authors infrequently reported sleep quality, efficiency, quality of life, function, or mood. When reported, results were mixed. When positive, effects were typically small. We looked for, but did not find, evidence of publication bias.

Trial participants were typically middle-aged, free of comorbid conditions, white, and female. They had to meet diagnostic criteria for insomnia disorder (typically at least 1 month in duration) and report sleep variables that surpassed severity thresholds for SOL, WASO, TST, night-time awakenings, or daytime problems associated with sleep. It is not clear whether the patients enrolled in these studies are similar, or if the corresponding findings applicable, to most patients with insomnia presenting to primary care clinics or those meeting current symptom criteria and duration for insomnia disorder.

We included RCTs lasting at least 4 weeks and prioritized patient-reported global outcomes rather than patient-reported assessment of sleep variables because insomnia disorder is currently defined as sleep problems persisting at least 3 months and causing daytime distress or dysfunction \( (59) \). We did not include laboratory assessment of sleep variables (actigraphy and polysomnography) because primary care clinicians typically make diagnostic and treatment decisions on the basis of patient-reported sleep-related concerns,
Pharmacologic Treatment of Insomnia Disorder

not sleep laboratory findings. Furthermore, insomnia disorder is also not an indication for polysomnography. In addition, it is not known how many minutes change in SOL, TST, or WASO indicate clinically meaningful improvement. We also found few data indicating that baseline sleep variables can accurately guide prescribing. Few studies evaluated long-term or as-needed therapy. Physicians should be aware that many medications used regularly and long term for insomnia disorders have an FDA indication only for short-term, as-needed use, in part because hypnotics are associated with dementia, fractures, major injuries, and possibly cancer. The FDA labels warn about cognitive and behavioral changes, including impaired driving, and other adverse effects that may be serious or life-threatening. Dose reduction is advised in women and older/debilitated adults, partly because drugs taken at bedtime remain at levels high enough to interfere with morning driving. The FDA-recommended doses are often lower than those used in some studies. Most individuals at the end of study continued to have global outcomes and sleep measures exceeding thresholds used for study enrollment, indicating that medications do not typically result in remission. We suggest that physicians target treatment decisions and assess clinical effectiveness according to global outcomes that encompass sleep variables as well as daytime dysfunction and distress, including adverse effects of pharmacologic treatments, such as next-morning impairment and sleep driving.

Other researchers have summarized adverse events of drugs used for insomnia in studies not meeting our inclusion criteria. For example, Kripke found increased incidence of depression and skin cancer among individuals using nonbenzodiazepine hypnotics and ramelteon (60, 61). Using pooled RCT data, Joya and colleagues found increased infection incidence with eszopiclone, ramelteon, zaleplon, and zolpidem (62). One review that assessed observational studies and case reports found that eszopiclone and zaleplon were associated with mild to moderate adverse effects, whereas zolpidem was associated with serious adverse effects, including amnesia, vertigo, confusion, and diplopia (63). Others found an association of benzodiazepine and nonbenzodiazepine hypnotic use with fractures (54), daytime somnolence (64), and motor vehicle accidents (65). Glass and colleagues showed that use of sedative-hypnotics compared with placebo in older patients with insomnia resulted in a 5-fold increase in memory loss, confusion, and disorientation; a 3-fold increase in dizziness, loss of balance, or falls; and a 4-fold increase in residual morning sedation, although absolute rates were low (66). Weich and associates conducted a retrospective cohort study using data from the United Kingdom General Practice Research Database (67). Anxiolytic and hypnotic drugs were associated with increased all-cause mortality. Use of certain medications for insomnia may lead to tolerance and dependence. Findings of the observational data mentioned above, however, must be interpreted with caution because of possible unmeasured confounders.

Future research to improve our understanding of insomnia disorder treatments should address the limitations of evidence that we found. The following are needed: more thorough reporting of the baseline duration and severity of symptoms; use of global outcomes; reporting of percentage of remitters or responders; establishment of minimum important differences in sleep variables; head-to-head drug comparisons within and across classes, as well as versus nonpharmacologic therapies; treatment durations of 1 year or more; observational studies evaluating long-term harm and assessment in patients, including older adults, with insomnia typically seen in primary care clinics.

In conclusion, eszopiclone, zolpidem, and suvorexant improved global outcomes and sleep variables for carefully selected adults with insomnia disorder. The clinical significance and general applicability of beneficial effects, the comparative effectiveness and long-term efficacy of agents, and the frequency and severity of adverse effects, especially among older adults, are not clearly established. A large placebo response was reported. Observational studies, including FDA data, suggest an association between hypnotics with frequent but serious harms.

From Minneapolis Veterans Affairs Center for Chronic Disease Outcomes Research, Minneapolis Veterans Affairs Health Care System, University of Minnesota School of Medicine, Minnesota Agency for Healthcare Research and Quality Evidence-based Practice Center, University of Minnesota School of Public Health, and Hennepin County Medical Center, Minneapolis, Minnesota.

Grant Support: By Agency for Healthcare Research and Quality (AHRQ): contract no. HHSA 290-2012-00016-I.

Disclosures: Dr. Wilt reports grants (funding for the evidence report) from AHRQ during the conduct of the study. Mr. MacDonald reports grants from AHRQ and personal fees from the American College of Physicians during the conduct of the study. Dr. Brasure reports a contract with AHRQ during the conduct of the study. Dr. Diem reports grants from AHRQ during the conduct of the study. Authors not named here have disclosed no conflicts of interest. Disclosures can also be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-1781.


Requests for Single Reprints: Timothy J. Wilt, MD, MPH, Minneapolis Veterans Affairs Health Care System, One Veterans Drive, Minneapolis, MN 55417; e-mail, Tim.wilt@va.gov.
Pharmacologic Treatment of Insomnia Disorder

Current author addresses and author contributions are available at www.annals.org.

References

Pharmacologic Treatment of Insomnia Disorder


60. Kripke DF. Greater incidence of depression with hypnotic use than with placebo. BMC Psychiatry. 2007;7:42. [PMID: 17711589]


Appendix Figure. Evidence search and selection.

- **Bibliographic database searches** (n = 3766 references)
  - **Title and abstract review excluded** (n = 3207 references)
  - **Pulled for full-text review** (n = 559 references)
    - **Excluded (n = 451 references)**
      - Excluded population: 101
      - Not RCT: 106
      - Not insomnia disorder: 74
      - Inadequate duration: 61
      - Not available in English: 48
      - Not available in U.S.: 22
      - No outcomes of interest: 19
      - Not valid comparison: 4
      - Not double-blinded (pharmaceutical): 3
      - Not peer-reviewed: 13
    - **Eligible (n = 108 references)**
      - **Hand-search** (n = 31 references)
        - **Harms search (pharmaceutical interventions)** (n = 12 references; 9 unique observational studies)
      - **Eligible references (n = 139)**
        - By intervention type:
          - Psychological: 6 references; 70 RCTs for the general adult population and older adults
          - Pharmacologic: 41 references; 39 RCTs for the general adult population and older adults
          - Complementary and alternative medicine: 14 references; 4 systematic reviews; 9 RCTs
          - Head-to-head drug comparison or combination treatment: 15 references; 15 RCTs

Intervention type totals do not equal total references because several trials were used in the analysis for 2 different types of interventions. RCT = randomized, controlled trial.