Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits

Jennifer Glass, Krista L Lanctôt, Nathan Herrmann, Beth A Sproule, Usoa E Busto

Abstract

Objective: To quantify and compare potential benefits (subjective reports of sleep variables) and risks (adverse events and morning-after psychomotor impairment) of short term treatment with sedative hypnotics in older people with insomnia.

Data sources: Medline, Embase, the Cochrane clinical trials database, PubMed, and PsychLit, 1966 to 2003; bibliographies of published reviews and meta-analyses; manufacturers of newer sedative hypnotics (zaleplon, zolpidem, zopiclone) regarding unpublished studies.

Selection criteria: Randomised controlled trials of any pharmacological treatment for insomnia for at least five consecutive nights in people aged 60 or over with insomnia and otherwise free of psychiatric or psychological disorders.

Results: 24 studies (involving 2417 participants) with extractable data met inclusion and exclusion criteria. Sleep quality improved (effect size 0.14, P < 0.05), total sleep time increased (mean 25.2 minutes, P < 0.001), and the number of night time awakenings decreased (0.63, P < 0.001) with sedative use compared with placebo. Adverse events were more common with sedatives than with placebo: adverse cognitive events were 4.78 times more common (95% confidence interval 1.47 to 15.47, P < 0.001); adverse psychomotor events were 2.61 times more common (1.12 to 6.09, P > 0.05), and reports of daytime fatigue were 3.82 times more common (1.88 to 7.80, P < 0.001) in people using any sedative compared with placebo.

Conclusions: Improvements in sleep with sedative use are statistically significant, but the magnitude of effect is small. The increased risk of adverse events is statistically significant and potentially clinically relevant in older people at risk of falls and cognitive impairment. In people over 60, the benefits of these drugs may not justify the increased risk, particularly if the patient has additional risk factors for cognitive or psychomotor adverse events.

Introduction

Insomnia often affects the quality of life for older people. Acute episodes are usually treated with drugs. Between 5% and 33% of elderly people in North America and the United Kingdom are prescribed a benzodiazepine or a benzodiazepine receptor agonist (zolpidem, zopiclone, zaleplon) for sleep problems. Adverse events that are associated with sedative use, such as ataxia, falls, or memory impairment, are thought to be particularly detrimental for older people. Despite the widespread use of sedative hypnotics in older people, the risk-benefit relation is not known. This meta-analysis aims to study the benefits of sedative use, as determined by subjective reported changes in sleep variables, and the risks, as determined by adverse events.

Methods

Identification of studies

We searched Medline, Embase, the Cochrane clinical trials database, PubMed, and PsychLit from 1966 to 2003, using the keywords “elderly” or “aged” (Medline and Cochrane database only) and “sedatives” or “hypnotics” or “benzodiazepines” or “zolpidem” or “zaleplon” or “zopiclone” or “antihistamines” (PsychLit and Cochrane database only) or “diphenhydramine” or “sleep” (Cochrane database only) or “sleep disorders” (PsychLit and Cochrane database only). For each citation identified, we scanned titles or abstracts, or both, to identify randomised controlled trials that excluded patients under 60 years old. If studies included some patients who were under 60 years old, the mean age of participants had to be over 60. We searched bibliographies of published reviews and meta-analyses for relevant titles and asked Servier, Canada (manufacturer of zaleplon), Sanofi-Synthelabo, US (maker of zolpidem), and ICN Canada (maker of zopiclone) about unpublished studies.

Inclusion criteria

We considered published randomised controlled trials of sedative hypnotics in English that compared active treatment against placebo or another active comparator. The active treatment phases of included studies were double blind.

In the included studies, participants had a mean age of at least 60 years and met predetermined diagnostic criteria for insomnia. Any study that included diagnostic criteria that were defined a priori was accepted.

Investigators must have excluded patients with psychiatric disorders, concurrent use of drugs affecting the central nervous system, and severe or acute physical illnesses that might disrupt sleep. As outcomes were subjective reports made by patients, investigators must have deemed that participants were cognitively able to perform the assessments (for example, by reporting appropriate score on a scale such as the mini-mental state examination). Participants must have had a washout period after previous drug treatments, and studies with crossover designs need an appropriate washout period.

Interventions were pharmacological treatments for insomnia for at least five consecutive nights. Any sedative hypnotic currently used in clinical practice was included in the search, including over the counter medications such as antihistamines and prescription medications such as benzodiazepines and
zolpidem, zopiclone, and zaleplon. We excluded studies of barbiturates and chloral hydrate or chloral hydrate derivatives as these are not recommended for elderly people.²⁴³⁵

Study selection, data abstraction, and assessment of quality

Articles were selected on the basis of the inclusion criteria (JG) and verified by another investigator (UEB). Studies were included if data from at least one of the outcome variables could be extracted. Three investigators rated study quality using Jadad criteria,¹¹ of whom two were blinded with respect to authors, author affiliation, date, and source of publication. Method of randomisation and allocation concealment were evaluated at this time while assessing the quality of the study. Two investigators (JG, UEB) abstracted data; UEB was blind to journal, authors, and date of publication. Any discrepancies were resolved through consensus.

Outcome measures

Benefits were measured by the participants’ perceived change in sleep. The variables that we considered were sleep quality (soundness or depth of sleep); total sleep time (total amount of time participants perceive that they have slept, measured in minutes); sleep onset latency (measured in minutes) or ease of getting to sleep (qualitative measure, score on questionnaire); and number of awakenings during the night. Ratings were made by patients and not an observer. All variables were analysed separately.

To measure risks, we determined the total number of adverse events for treatment and placebo (including placebo run-in phases) and then categorised them as: cognitive adverse events (memory loss, confusion, disorientation); psychomotor-type adverse events (reports of dizziness, loss of balance, or falls); and morning hangover effects (residual morning sedation). We analysed morning impairment (as measured by performance tasks such as reaction time or hand-eye coordination tasks) separately from adverse events as a risk of sedative use.

Data synthesis

As no standard accepted instrument measures sleep quality, we used effect sizes of the change in scores. The equation for Cohen’s d was used to estimate effect size: \( \frac{M_1 - M_2}{\sigma_p} \) (where \( M_1 \) = mean sleep quality score for the treatment group, \( M_2 \) = mean sleep quality score for the control group, and \( \sigma_p \) = pooled standard deviations from either control or treatment or both groups). To minimise false positive results, we used the larger standard error, whether in the control group, treatment group, or all participants pooled. As more than one instrument had been used to measure psychomotor impairment, we also calculated effect sizes for morning-after performance. We assessed total sleep time, number of awakenings, and sleep onset latency as continuous variables.

If means were reported without variances, estimates were obtained from studies with similar methodology and patient population and a similar or smaller sample size.²² If the study did not have a placebo arm but did have a placebo run-in period, we used the run-in period as the control to preserve as much raw data as possible. We obtained common odds ratios for all adverse events. All results used random effects models and 95% confidence intervals. We used \( \chi^2 \) analysis to test heterogeneity for all combined results.

We calculated numbers needed to treat and to harm by using the inverse of the absolute risk reduction (efficacy data) and the absolute risk increase (adverse event data).¹³¹⁴

Results

Of 120 studies identified, 20 satisfied inclusion and exclusion criteria and had extractable subjective data.²⁶⁻²⁷ Four further studies reported on adverse events only and have been included in the assessment of risk (fig 1).²⁸⁻⁴¹

A total of 830 participants were treated with a benzodiazepine, 106 with zopiclone, 384 with zolpidem, 609 with zaleplon, 14 with diphenhydramine, and 468 with placebo (not including placebo run-ins) (table 1).

Quality of sleep

Number needed to treat was derived from four studies that in which participants reported any improvement in sleep quality (considered “successes”) compared with no improvement or a worsening in sleep quality (“failures”). On the basis of four studies (1072 participants), the number of patients who would need to be treated with a sedative (zaleplon, brotizolam or nitrazepam, or loprazolam) for one to have an improvement in sleep quality is 13 (95% confidence interval 6.7 to 62.9).²⁶⁻²⁷

Eight studies (719 participants) had extractable sleep quality score data (means and standard deviations) for any sedative versus placebo, and we used these to determine the magnitude of effect.²⁸⁻⁴¹ Reported sleep quality was significantly better with sedative use (mean effect size 0.14, 0.05 to 0.23; P < 0.005, fig 2). This effect size indicates a difference in mean scores on sleep quality for sedative versus placebo groups of 0.11. In the most heavily weighted study in the analysis,²⁹ this would correspond to mean scores of 3.8 in the placebo group and 3.7 in the sedative group on a seven point scale.

Seven studies (277 participants) were combined to see the effect of benzodiazepines only versus placebo on sleep quality measures.²⁶⁻²⁷ A significant improvement in sleep quality improved significantly (mean effect size 0.37, 0.01 to 0.73; fig
2). This effect size corresponds to a difference of 0.46 between the treatment means—scores of 2.7 for placebo and 3.1 for drug, on the five point scale used by the most heavily weighted study in this analysis.7

Three studies (339 participants) comparing benzodiazepines with benzodiazepine receptor agonists ( zaleplon, zolpidem, and zopiclone) found no significant difference in sleep quality (mean effect size 0.04, −1.11 to 1.19; test for heterogeneity P = 1.0).25 26 27

Two studies (116 participants) that reported sleep quality data for zopiclone versus placebo had a magnitude of effect of 0.41 (−0.76 to 1.58; test for heterogeneity P = 0.58).25 26 Data for zolpidem or zaleplon versus placebo were insufficient for inclusion.

Amount of sleep

In eight studies (601 participants) with extractable data, the increase in total sleep time was 34.2 minutes (16.2 to 52.8 minutes, P < 0.01; test for heterogeneity P = 0.13).20 21 22 23 24 25 26 27 31 34 37 Data were insufficient to analyse sleep onset latency or ease of getting to sleep.

Number of awakenings

In six studies (441 participants) with extractable data, the mean number of awakenings decreased by 0.63 (−0.48 to −0.77, P < 0.0001; test for heterogeneity P = 0.71).20 22 23 27 29 34 36 In six studies with benzodiazepines versus placebo (296 participants) the mean number of awakenings decreased by 0.60 (−0.41 to −0.78, P < 0.0001; test for heterogeneity P = 0.58).20 23 27 29 34 36

Adverse events

On the basis of all adverse events reported in 16 studies (2220 participants), the number needed to harm for sedative hypnotics compared with placebo is 6 (4.7 to 7.1),20 22 23 24 25 26 27 29 30 34 35 36 38 39 40 41 The most common adverse events were drowsiness or fatigue, headache, nightmares, and nausea or gastrointestinal disturbances. As severity of adverse events was reported in only one study, pooled estimates could not be determined.

On the basis of the 10 studies (712 participants), cognitive effects were significantly more common with sedative use than with placebo (odds ratio 2.25, 0.93 to 5.41, P = 0.07; test for heterogeneity P = 0.08, fig 3).20 23 24 25 26 27 30 36 39 41 Of
the 59 psychomotor effects that were reported, seven were serious events (six falls and one motor vehicle crash). Three resulted in broken bones: a fall resulting in a broken hip after five nights of 15 mg temazepam,13 a fall resulting in a broken femur after 0.125 mg triazolam,15 and one case of hip fracture while taking placebo. A motor vehicle crash after a single dose of temazepam 15 mg was reported.16

There were significantly more subjective reports of morning or daytime fatigue (seven studies, 829 participants) after treatment than after placebo (odds ratio 3.82, 1.88 to 7.80, P < 0.001). Studies of zolpidem, zopiclone and zolpidem (combined) versus benzodiazepine reported little difference (odds ratio 1.48, 0.75 to 2.93; test for heterogeneity P = 0.46).

Six studies (648 participants) that compared benzodiazepine receptor agonists and benzodiazepine reported little difference in numbers of adverse events (odds ratio 1.11, 0.59 to 2.07, P = 0.75; test for heterogeneity P = 0.07).22 23 26 28 36 40 Studies of zaleplon, zopiclone and zolpidem (combined) versus benzodiazepines found no significant difference in cognitive adverse events (four studies, 298 participants; odds ratio 1.12, 0.16 to 7.76, P = 0.91; test for heterogeneity P = 0.35).22 23 26 36 40

Publication bias

Funnel plot analyses indicated a possible publication bias on all reported psychomotor-type adverse events (six studies, 625 participants; odds ratio 1.48, 0.75 to 2.93; test for heterogeneity P = 0.46).22 25 26 36 40

Discussion

Treatment with sedative hypnotics improves the quality of sleep, evaluated subjectively. The effect size (0.14) is small according to the classification by Cohen (in which a small effect size is about 0.2).18 Sleep quality is a measure of efficacy for sleep medications and also a measure of patient satisfaction. Findings for both total sleep time and number of night time awakenings showed significant although small improvements in patients who took a sedative rather than placebo.

Risk of adverse events

The risk of adverse events was higher with sedative treatment. Most adverse events were reported to be reversible and not severe.13 23 26 36 47 51 53 54 61 Patients who took sedatives had a higher incidence of falls and motor vehicle crashes.

Numbers needed to treat versus numbers needed to harm

The number needed to treat for improved sleep quality was 13 and the number needed to harm for any adverse event was 6. This ratio indicates that an adverse event is more than twice as likely as enhanced quality of sleep. This ratio can be used as a rough indicator only, as more than double the number of participants contributed to the “harm” data than to the “effectiveness” data (2220 vs 1072). The effectiveness estimate is less precise (the confidence interval is wider).

### Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment (all) (N)</th>
<th>Placebo (N)</th>
<th>Effect size (SE)</th>
<th>Effect size (95% CI)</th>
<th>Weight (%)</th>
<th>Effect size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All treatments (effect size)</td>
<td>Bayer et al22</td>
<td>40</td>
<td>40</td>
<td>1.07 (6.83)</td>
<td>0.00</td>
<td>1.07 (-12.32 to 14.46)</td>
</tr>
<tr>
<td></td>
<td>Klimm et al24</td>
<td>72</td>
<td>72</td>
<td>0.57 (4.91)</td>
<td>0.01</td>
<td>0.57 (-9.06 to 10.19)</td>
</tr>
<tr>
<td></td>
<td>Murphy et al20</td>
<td>16</td>
<td>16</td>
<td>1.16 (4.50)</td>
<td>0.01</td>
<td>1.16 (-7.66 to 9.96)</td>
</tr>
<tr>
<td></td>
<td>Roger et al26</td>
<td>221</td>
<td>221</td>
<td>1.17 (3.30)</td>
<td>0.02</td>
<td>1.17 (-5.30 to 7.64)</td>
</tr>
<tr>
<td></td>
<td>Dehlin et al22</td>
<td>26</td>
<td>40</td>
<td>0.50 (0.60)</td>
<td>0.02</td>
<td>0.50 (-6.68 to 1.68)</td>
</tr>
<tr>
<td></td>
<td>Elle et al23</td>
<td>29</td>
<td>15</td>
<td>0.23 (0.58)</td>
<td>0.02</td>
<td>0.23 (-0.91 to 1.38)</td>
</tr>
<tr>
<td></td>
<td>Vlukal et al27</td>
<td>32</td>
<td>32</td>
<td>0.37 (0.20)</td>
<td>0.03</td>
<td>0.37 (-0.03 to 0.77)</td>
</tr>
<tr>
<td></td>
<td>Hedner et al25</td>
<td>268</td>
<td>136</td>
<td>0.13 (0.05)</td>
<td>0.13</td>
<td>0.13 (0.03 to 0.22)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>701</td>
<td>557</td>
<td>100.00</td>
<td>0.14 (0.05 to 0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test for heterogeneity: χ²=1.91, df=7, P=0.96, I²=0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test for overall effect: z=2.99, P=0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzodiazepines only (effect size)</td>
<td>Bayer et al22</td>
<td>40</td>
<td>40</td>
<td>1.07 (6.83)</td>
<td>0.07</td>
<td>1.07 (-12.32 to 14.46)</td>
</tr>
<tr>
<td></td>
<td>Klimm et al24</td>
<td>72</td>
<td>72</td>
<td>0.57 (4.91)</td>
<td>0.14</td>
<td>0.57 (-9.06 to 10.19)</td>
</tr>
<tr>
<td></td>
<td>Murphy et al20</td>
<td>16</td>
<td>16</td>
<td>1.16 (4.50)</td>
<td>0.16</td>
<td>1.16 (-7.66 to 9.96)</td>
</tr>
<tr>
<td></td>
<td>Roger et al26</td>
<td>221</td>
<td>221</td>
<td>1.17 (3.30)</td>
<td>0.30</td>
<td>1.17 (-5.30 to 7.64)</td>
</tr>
<tr>
<td></td>
<td>Dehlin et al22</td>
<td>26</td>
<td>40</td>
<td>0.50 (0.60)</td>
<td>0.50</td>
<td>0.50 (-6.68 to 1.68)</td>
</tr>
<tr>
<td></td>
<td>Elle et al23</td>
<td>29</td>
<td>15</td>
<td>0.23 (0.58)</td>
<td>0.92</td>
<td>0.23 (-0.91 to 1.38)</td>
</tr>
<tr>
<td></td>
<td>Vlukal et al27</td>
<td>32</td>
<td>32</td>
<td>0.37 (0.20)</td>
<td>80.38</td>
<td>0.37 (-0.03 to 0.77)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>436</td>
<td>431</td>
<td>100.00</td>
<td>0.37 (0.01 to 0.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test for heterogeneity: χ²=0.20, df=6, P=1.00, I²=0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test for overall effect: z=2.04, P=0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 2 Mean effect size (95% confidence intervals) for subjective improvements in sleep quality with any sedative treatment and benzodiazepines only compared with placebo for at least five nights in people aged 60 or older with insomnia
Comparison with the literature

Although meta-analyses have examined the effects of benzodiazepines and the new benzodiazepine-receptor agonists, they have not examined their effects in older people. Nowell et al combined five studies that included people under 65 with insomnia and found a positive effect size of 0.02 (0.45 to 0.79) for benzodiazepines and zolpidem versus placebo for sleep quality scores.43 Smith et al found an effect size of 1.2 when pharmacotherapy was compared with placebo in their meta-analysis of studies including both younger and older adults, using effect size that was weighted using studies’ sample sizes.44 These studies describe greater overall effects of sedative on sleep quality than we found. Although our results for sleep quality were heavily weighted by one study that had a large sample size (n = 404) and small standard error,21 removing this study from the analysis resulted in a mean effect size that was still lower than those reported by previous meta-analyses (d = 0.37).

Similarly, in a meta-analysis of benzodiazepine use in adults with insomnia, Holbrook et al found a significant increase in adverse events with benzodiazepine use (odds ratio 1.8, 1.4 to 2.4). The increase in psychomotor-type side effects found with sedative use in our study (odds ratio 2.61) is similar to the increase in reports of dizziness and lightheadedness found in the Holbrook meta-analysis after benzodiazepine use (odds ratio 2.6, 0.7 to 10.3). Holbrook et al also report a significant increase in reports of daytime fatigue with benzodiazepine use (odds ratio 2.4, 1.8 to 3.4). This is lower than our reported odds ratio for subjective reports of daytime fatigue after night time sedative use (odds ratio 3.82). This may indicate that older people have similar or greater potential for risks such as adverse events than younger adults.

Loss of memory and confusion have been reported with older sedative hypnotics such as triazolam and newer sedatives.
such as zolpidem in patients of all ages.45 46 The risk for these events may be even higher in elderly patients. Although investigators in the studies we included stated that patients were not cognitively impaired, they did not always use validated scales such as the mini-mental state examination to support these claims. Any pre-existing cognitive impairment may exacerbate confusion or memory problems.31

An increased risk for adverse events is consistent with reported risks for falls and motor vehicle crashes or household incidents.46 Older people may be at greater risk for adverse effects because of pharmacokinetic considerations, such as reduced clearance of certain sedative hypnotics.47 48 There is also some evidence of pharmacodynamic differences such as increased sensitivity to peak drug effects.31

We found that the impairment of performance tasks the morning after sedatives are used, though significant, is small. This may indicate that, even when fatigue is reported, reaction time or hand-eye coordination after sedative use does not deviate greatly from normal.

Epidemiological evidence shows an increased risk of car crashes in elderly people who are using long acting sedatives (nitrazepam or flurazepam, for example) but not with short acting sedatives (such as triazolam or temazepam).32 Impairment depends on dose and time since dosing.33 The small effects of morning after impairment that we found may be due to the heterogeneous half lives and doses of the drugs tested, as well as the variable times of testing after dosing.

Limitations of this study

Interpretation of this meta-analysis must take into account that all sedatives or all benzodiazepines were grouped together for analyses, irrespective of differences in half life, potency, or dosage. As well no standard method of collecting subjective sleep variables is available. The studies used various measures: ordinal scales (three, five, or seven point), visual analogue scales, and combined scales. Though subjective reports create more variable results than objective measures, we focused on subjective ratings of sleep variables as the primary outcome measure because consumption of healthcare resources is driven by subjective report rather than objective measures of sleep.

Another potential source of variability was the health status of the participants in the studies. Some were community dwelling ambulatory patients attending a health clinic and others were inpatients on a geriatric ward. Differences in environment or health status may affect how people respond to subjective assessments. Furthermore, although studies were double blind, the hypnotic effects of sedative hypnotics may be cues that compromise blinding, which may in turn affect subjective reports.

In some cases, such as when studies did not have placebo arms, placebo run-in scores were used as control scores (see table). Sleep has improved over time during clinical studies with self monitoring and recording patterns, and all study tips for improving sleep hygiene have been given.41 43 Two studies excluded placebo responders, leading to an enriched sample.41 43 Using placebo run-in scores or enriched samples might lead to underestimated placebo effects and inflated estimates of treatment effects.

Dependence and habituation are some of the major concerns with sedative use, but were not addressed directly in any of the studies. These factors also weigh on the risk-benefit relation, but their role could not be determined in this study.

What is already known on this topic

Benzodiazepines and newer benzodiazepine receptor agonists are thought to be efficacious for sleep disturbances in elderly people.

They are associated with risks that are particularly detrimental in elderly people, such as ataxia, cognitive effects, and falls.

Little is known about how the risks and benefits compare for non-prescription sedative hypnotics

What this study adds

In people over 60, the benefits associated with sedative use are marginal and are outweighed by the risks, particularly if patients are at high risk for falls or cognitive impairment.

Conclusions

Although the improvements in sleep variables obtained from prescription sedative hypnotics are statistically significant, the effect size is small, and the clinical benefits may be modest at best. The added risk of an adverse event may not justify these benefits, particularly in a high risk elderly population. These factors should be considered when sedative hypnotics are prescribed for older patients. Non-pharmacological therapies such as cognitive behaviour therapy have been shown to be as efficacious as pharmacotherapy for insomnia in older people.44 45 Because fewer risks are associated with behavioural therapies,44 45 they may be a viable treatment alternative in a healthy elderly population with no cognitive impairment.

Contributors: JG wrote the protocol with input from all authors, did the initial searches and study selection, data abstraction, and data analyses including graphics, and wrote the paper. She is guarantor. UEB helped write the protocol and establish the inclusion and exclusion criteria, evaluated study quality, carried out data abstraction, and edited the final manuscript. KLL helped write the protocol, assisted and reviewed data analyses and statistical analyses, graphing, and interpretation of results, and edited the final report. NH helped write the protocol, gave clinical input into the study design, inclusion and exclusion criteria, evaluated study quality, and edited the final manuscript. BAS provided input on the protocol, including clinical input, commented on interim results, and reviewed and edited results (including graphs and tables) and the final manuscript. Funding: None. This study was part of a PhD thesis for JG.

Competing interests: None declared.

Ethical approval: Not required.


