



Comparative effectiveness of interventions to facilitate deprescription of benzodiazepines and other sedative hypnotics: systematic review and meta-analysis

Dena Zeraatkar,^{1,2} Sumanth Kumbargere Nagraj,³ Michael Ling,¹ Tanvir Jassal,¹ Sarah Kirsh,^{1,2} João Pedro Lima,^{1,2} Tyler Pitre,⁴ Rachel Couban,¹ Muizz Hussain,⁵ Siri Seterelv,³ Stijn Van de Velde,³ Katarzyna Gustavsson,^{6,7} Adam Wichniak,⁶ Carole E Aubert,^{8,9} Antoine Christiaens,^{10,11} Anne Spinewine,^{11,12} Thomas Agoritsas^{2,3,13}

For numbered affiliations see end of the article

Correspondence to: D Zeraatkar zeraatd@mcmaster.ca; (ORCID 0000-0003-4287-0541)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2025;389:e081336 <http://dx.doi.org/10.1136/bmj-2024-081336>

Accepted: 10 March 2025

ABSTRACT

OBJECTIVE

To review evidence from randomised trials assessing the effectiveness of strategies to deprescribe benzodiazepines and closely related sedative hypnotics (BSH).

DESIGN

Systematic review and meta-analysis of randomised controlled trials.

DATA SOURCES

MEDLINE, Embase, CINAHL, PsycInfo, and CENTRAL, searched from inception to August 2024, and reference lists of included studies and similar systematic reviews.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES

Eligible studies randomised adults using BSH for insomnia to interventions aimed at deprescribing BSH, strategies to implement these interventions in healthcare settings, or usual care or placebo.

METHODS

Reviewers worked independently and in duplicate to screen search results, extract data, and assess risk of bias. Similar interventions were grouped together, frequentist random effects meta-analysis was conducted, and the certainty of evidence was assessed using the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach.

RESULTS

The review identified 58 publications reporting on 49 unique trials with more than 39 000 patients. Interventions were classified into the following categories: tapering, patient education, physician education, combined patient and physician education, cognitive behavioural therapy, medication review, mindfulness, motivational interviewing, pharmacist led interventions, and drug assisted tapering and withdrawal. Low certainty evidence suggests that education of patients (144 (95% confidence interval 61 to 246) more per 1000 patients), medication review (104 (34 to 191) more), and a pharmacist led educational intervention (491 (234 to 928) more) may increase the proportion of patients who discontinue BSH compared with usual care. Moderate certainty evidence suggests that education of patients probably has little or no effect on physical function, mental health, and signs and symptoms of insomnia. No evidence was found regarding these other outcomes for medication review or for the pharmacist led educational intervention. No compelling evidence was found that other interventions may help patients to discontinue BSH. Moreover, no high or moderate certainty evidence was found that any of the interventions caused an increase in dropouts. Finally, low certainty evidence suggests that multicomponent interventions may be more effective at facilitating discontinuation of BSH than single component interventions.

CONCLUSION

The evidence on the effectiveness of interventions to discontinue BSH is of low certainty. Educating patients, doing medication reviews, and a pharmacist led educational intervention may increase the proportion of patients who discontinue BSH.

Introduction

Benzodiazepines and closely related sedative hypnotics (BSH) are commonly used to treat insomnia. Because these drugs may cause falls, cognitive impairment, and dependence, particularly in older and/or vulnerable people, guidelines advise healthcare providers to discontinue BSH.¹⁻²¹ Nevertheless, prolonged use of BSH remains common, with associated harms, healthcare burdens, and costs.²²⁻²⁶

Research has documented diverse reasons for prolonged use of BSH. Patients report that they are not aware of the hazards of BSH or alternative drug and non-drug options to manage insomnia.²⁷⁻³⁰

WHAT IS ALREADY KNOWN ON THIS TOPIC

Many patients use benzodiazepines and related sedative hypnotics (BSH) to treat insomnia despite risks such as falls, cognitive impairment, and dependence

Healthcare providers have reported limited knowledge of alternatives to BSH, a lack of clarity on its harms, and insufficient evidence on optimal strategies to support its discontinuation

WHAT THIS STUDY ADDS

Low certainty evidence suggests that patient education, medication reviews, and pharmacist led educational interventions may increase BSH discontinuation rates

Multicomponent interventions may be more effective than single component approaches, but further high quality research is needed

This low certainty body of evidence highlights the need for additional high quality research

When patients attempt to discontinue BSH, they may be deterred by symptoms of physiological and psychological dependence, anxiety, or poor sleep.³⁰⁻³² Healthcare providers report limited knowledge of alternatives to BSH, a lack of clarity on the harms of BSH, and insufficient evidence on optimal strategies to support the deprescription of BSH, defined as the planned, supervised, and safe process of dose reduction or complete cessation.^{29 33}

In response to these problems, this systematic review and meta-analysis identifies, pools, and appraises evidence from randomised trials assessing the effectiveness of interventions to facilitate deprescription of BSH and strategies for the implementation of such interventions in healthcare settings.

This review is part of the BMJ Rapid Recommendations project—a collaborative effort from the MAGIC Evidence Ecosystem Foundation (www.magicvidence.org) and *The BMJ* to produce trustworthy recommendations in response to practice changing evidence. This BMJ Rapid Recommendations project was triggered by BE-SAFE (Benzodiazepine and sedative hypnotic use to improve patient SAFETY and quality of care)—a project funded by the European Union's Horizon Europe research and innovation programme and the Swiss State Secretariat for Education, Research and Innovation to improve patient safety and prescription practices related to BSH for insomnia.³⁴ The parallel BMJ Rapid Recommendations guideline panel, comprised of patient partners, methodologists, general practitioners, internists, geriatricians, a psychiatrist, a public health physician, a clinical pharmacist, and a social worker, helped to define the research question and the scope of the review. The panel included only people free of financial conflicts of interest, and potential intellectual conflicts were balanced among panel members.

Methods

We developed a protocol for this systematic review and meta-analysis, which was reviewed and approved by the BE SAFE collaborators in February 2023 and registered online on Open Science Framework (<https://osf.io/8em4p/files/osfstorage>). We report our study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.³⁵

Eligibility criteria

Eligible studies randomised adult patients using BSH for insomnia to interventions aimed at deprescribing or discontinuation of BSH, strategies to implement these interventions, or standard care or placebo. Examples of eligible interventions include, but are not limited to, supervised tapering with or without drug support, education of patients about the harms of BSH, education of healthcare providers, behavioural interventions such as cognitive behavioural therapy, and system level interventions such as implementing new protocols or services in healthcare facilities. BSH

drugs include but are not limited to benzodiazepines (for example, diazepam, alprazolam) and non-benzodiazepine “z-drugs” (for example, zolpidem, zaleplon). We did not restrict eligibility on the basis of duration of BSH use.

We anticipated that some trials would recruit patients who used BSH for a mix of different conditions and that some trials may not report the primary reason for which patients were using BSH. We included such trials unless explicit evidence showed that 60% or more of patients were using BSH for conditions other than insomnia. However, we did subgroup analyses investigating differences in results between trials in which patients were explicitly using BSH for insomnia and other trials in which the reason for using BSH was not reported.

Eligible designs included parallel, cluster, or crossover-by-cluster randomised trials. We excluded crossover trials because successful discontinuation of BSH in the first phase of the trial would render patients ineligible for subsequent phases of crossover. We did not restrict eligibility on the basis of language or time of publication. We excluded observational studies; systematic, scoping, or narrative reviews; trials that investigated interventions that only targeted discontinuation of non-BSH drugs; and trials that investigated the effectiveness of societal interventions (for example, messaging in the media).

We also excluded trials with fewer than 20 patients per arm. We anticipated that smaller trials would contribute little to meta-analyses, be at higher risk of publication bias, and more often include unrepresentative samples and trial arms that were prognostically imbalanced.³⁶

Search strategy

We worked with an experienced research librarian to devise and implement a search of MEDLINE, Embase, CINAHL, PsycInfo, and CENTRAL, from inception to August 2024 (supplement 1). Our search combined terms related to BSH, discontinuation and deprescription, and randomised trials. We supplemented our search by soliciting the parallel BMJ Rapid Recommendations guideline panel for additional eligible trials, doing forwards and backwards searching, and reviewing references of similar systematic reviews.³⁷⁻⁴⁵

Screening

Following training and calibration exercises to ensure sufficient agreement, pairs of reviewers, working independently and in duplicate, screened the titles and abstracts of search records and subsequently the full texts of records deemed eligible at the title and abstract screening stage by using Covidence (<https://www.covidence.org>) online systematic review software. Reviewers resolved discrepancies by discussion or, when necessary, by adjudication with a third reviewer.

Data extraction

Following training and calibration to ensure sufficient agreement, pairs of reviewers, working independently and in duplicate, extracted data from eligible trials into a structured Excel spreadsheet. We extracted data on characteristics of trials (for example, trial design, country, setting), patients (for example, age, sex, duration of BSH use), and interventions (for example, type, duration, intensity) and outcomes of interest.

A formal outcome prioritisation exercise with the parallel BMJ Rapid Recommendations panel directed our selection of outcomes.^{46 47} Outcomes of interest included the proportion of patients who discontinue BSH, quality of life, physical function, mental health, cognitive function, signs and symptoms of insomnia, sleep efficiency, total sleep time, sleep onset latency, number of BSH prescriptions, drug-free nights, and dropouts of patients during the intervention.⁴⁸ We expected that trials would measure and report these outcomes by using several different instruments. We extracted data for all validated instruments measuring our outcomes of interest. When trials reported both self-reported and laboratory tested discontinuation of BSH, we prioritised laboratory test results for their objectivity.

We anticipated that the effects of interventions might vary according to the age of patients and the duration of BSH use. We intended to extract data stratified by these factors, when reported, but did not identify any cases in which trials reported results stratified according to these factors.

We preferentially extracted intention-to-treat data without any imputations at the longest reported time of follow-up at which randomised groups were still preserved. For dichotomous outcomes, we extracted numbers of participants and events in each arm and relative effects (for example, relative risks) and associated confidence intervals. For continuous outcomes, we extracted number of participants, means, and measures of variability in each arm. We preferentially extracted the mean change in continuous outcomes when reported or, when not reported, the mean measure at follow-up.⁴⁹ To facilitate meta-analyses, when trials reported medians or interquartile ranges for continuous outcomes, we converted these measures to means and standard deviations by using previously described methods.⁵⁰

For cluster randomised trials, we additionally extracted the intracluster correlation coefficient—a measure of the degree to which patients within the same cluster (for example, patients in the same nursing home) are similar to each other in their response to an intervention compared with individuals from other clusters.⁵¹ For interventions targeting the discontinuation of several different classes of inappropriate medications, we extracted outcome data only for patients using BSH. When trials only reported data in figures, we used WebPlotDigitizer to extract data (<https://automeris.io/WebPlotDigitizer/>).

For publications in languages other than English, we planned to use Google Translate ([https://translate](https://translate.google.com/)

[google.com/](https://translate.google.com/)) to perform extractions and ask a native language speaker to confirm the extractions. We did not, however, find any publications in any language other than English. Reviewers resolved disagreements by discussion, or, when necessary, by adjudication with a third reviewer.

Assessment of risk of bias

Following training and calibration, reviewers worked independently and in duplicate to assess risk of bias by using a modified version of the Cochrane endorsed Risk of Bias 2.0 tool (RoB 2.0)—the gold standard tool for assessing limitations in trials that may bias results.⁵² This tool assesses the risk of bias of trials across five domains: bias due to randomisation (for example, random sequence generation, allocation concealment), bias due to deviations from the intended intervention (for example, lack of blinding leading to imbalances in co-interventions across trial arms), bias due to missing outcome data (for example, attrition), bias due to measurement of the outcome (for example, unblinded outcome assessment), and selective reporting (for example, selective reporting of outcome measures on the basis of results). We additionally considered bias due to the timing of identification or recruitment of participants for cluster randomised trials and bias due to period and carryover effects for crossover-by-cluster trials.

To assess the risk of bias due to deviations from the intended intervention, we considered the effect of assignment rather than adherence to the intervention, as this effect is likely to be the observed effect in clinical settings and of the greatest interest to evidence users. We considered trials without masking of patients, healthcare providers, and investigators to be at high risk of bias for deviations from the intended intervention owing to the potential for differential care across trial arms. For example, we hypothesised that knowledge among healthcare providers that they are part of a randomised trial concerning the discontinuation of BSH could lead them to advise patients to stop using BSH regardless of the intervention to which participants have been randomised. Likewise, such information could make patients more likely to attempt to discontinue BSH independent of the intervention to which they have been randomised.

Our modified version of the RoB 2.0 tool includes the same domains, but with revised response options (that is, definitely low risk of bias, probably low risk of bias, probably high risk of bias, and definitely high risk of bias) and, in response to evidence suggesting that the original RoB 2.0 tool has poor inter-rater reliability,⁵³ detailed guidance tailored to factors relevant for this review (for example, removing guidance for assessing risk of bias of adhering to the intervention, listing important co-interventions to consider). Reviewers resolved disagreements by discussion or by consultation with a third reviewer, if necessary.

Data synthesis and analysis

A group of content experts and methodologists, with feedback from the wider guideline panel, classified similar interventions together for synthesis, considering the objective of the intervention, the target stakeholder (for example, patients, physicians, nurses, pharmacists), class of therapy (for example, cognitive behavioural therapy, mindfulness based therapy), and whether the intervention involved gradual or abrupt reduction in BSH dose, educational materials, or meetings with professionals. For analyses investigating the effects of cognitive behavioural therapy against usual care, we included trials that compared cognitive behavioural therapy against no additional interventions, cognitive behavioural therapy against tapering, the combination of cognitive behavioural therapy and tapering against tapering alone, and the combination of cognitive behavioural therapy and tapering against no additional interventions. We anticipated that all cognitive behavioural therapy interventions would include some suggestion of tapering. However, we did subgroup analyses to investigate differences in the effects of cognitive behavioural therapy when administered alone or with tapering and when compared against tapering or usual care.

We expected that interventions comprising more than one component would be more effective than interventions with a single component. Thus, we also classified interventions as multicomponent or single component to facilitate comparisons between the two. Multicomponent interventions included the combination of audit and feedback with education of physicians, cognitive behavioural therapy and tapering, education of both patients and physicians, and complex interventions that involved the introduction of a new professional role or new healthcare professionals in the healthcare setting or the training of healthcare professionals to perform additional tasks outside of their main scope of practice. Conversely, we classified cognitive behavioural therapy as a single component intervention as it comprises a distinct class of therapy.⁵⁴ We classified the combination of tapering and education of patients and cognitive behavioural therapy alone or in combination with education of patients as single component interventions because tapering and cognitive behavioural therapy will ultimately always include an educational component. We also considered drug assisted tapering as a single component intervention.

We restricted comparisons of multicomponent and single component interventions to our most important outcome representing the effectiveness of interventions (the proportion of patients who discontinue BSH) and another outcome representing acceptability and safety of the intervention (the proportion of patients who discontinued the intervention).

When possible, for each unique comparison of interventions and outcome, we did frequentist random effects meta-analysis using the restricted maximum likelihood (REML) estimator—a conservative approach

to meta-analysis—using data from the last point of follow-up from each trial at which randomised groups were still preserved.⁵⁵ For comparisons for which meta-analysis was not possible, we planned to describe the range and distribution of results narratively, according to synthesis without meta-analysis (SWiM) guidelines.⁵⁶ We did not, however, encounter comparisons for which we could not do quantitative synthesis.

We found that trials measured and reported results for physical function, mental health, and quality of life by using several different instruments. We pooled the results of each unique measure separately. We avoided using standardised mean differences owing to their potential to be influenced by differences in variability across trial populations.⁵⁷ We also avoided converting units of one measure to another owing to potential differences in the range of the construct covered by each measure.⁵⁷

For cluster trials that appropriately accounted for correlations within clusters, we used the reported results in the analysis. Otherwise, we adjusted trial results by the design effect, which we calculated using either the intraclass correlation coefficient reported in the trial or, if one was not reported, the median intraclass correlation coefficient across all included trials.⁴⁹

For trials with multiple arms assessing the same intervention, we combined the number of participants and events for dichotomous outcomes and did fixed effect meta-analysis for the arms that assessed the same intervention for continuous outcomes. For trials with multiple arms assessing the same intervention that reported only relative effects or mean differences against a common comparator group, we pooled the arms assessing the same intervention together, using previously described methods to avoid double counting the comparator group.⁵⁸

We summarised heterogeneity by using the I^2 statistic and interpreted an I^2 value of 0% to 40% as not important, 30% to 60% as moderate heterogeneity, 50% to 90% as substantial heterogeneity, and 75% to 100% as considerable heterogeneity.⁴⁹ The I^2 value is prone to misinterpretation as even small degrees of clinically unimportant inconsistency may translate to high I^2 values if estimates from studies are precise.^{59 60} Hence, we also considered the absolute magnitude of differences in estimates across studies. For analyses with 10 or more studies, we had planned to test for publication bias by visually inspecting funnel plots and Eggers tests, but all analyses included fewer than 10 studies.⁶¹

Like conventional meta-analysis, which pools results from studies assessing the same or similar research questions, network meta-analysis also pools study results but allows comparisons among three or more interventions simultaneously. We opted against network meta-analyses owing to the sparsity of evidence across outcomes and outcome measures and to concerns with transitivity or joint randomness—a critical assumption of network meta-analysis.⁶² In particular, the types of interventions studied in

trials varied across settings. For example, studies conducted in the community often focused on less intensive interventions such as patient education and tapering, whereas trials conducted in specialised settings studied more intensive interventions, such as introducing new healthcare professionals or training healthcare professionals to perform additional tasks outside of their main scope of practice. Given the likely differences in patient populations across these settings, comparisons between these interventions violate the assumption of transitivity requisite for network meta-analysis.

We used the meta and metafor packages in R (version 4.1.2) for all analyses and uploaded our data, code, and forest plots on Open Science Framework (<https://osf.io/8em4p/files/osfstorage>).^{63 64}

Subgroup analyses

We anticipated that the effects of interventions may vary according to several a priori specified factors—risk of bias, care setting (inpatients versus outpatients versus nursing homes), age (<65 versus ≥65), country (high income versus low/middle income), duration of BSH use (<median or ≥median across trials), and method for ascertaining BSH discontinuation (for example, patient report, medical records, or toxicology). To test for differences in effects according to these factors, we planned to do subgroup analyses. We were unable to do subgroup analyses based on care setting because of insufficient variation in care settings across trials reporting on similar interventions, age because trials did not report results stratified by age and large overlap existed in the age of participants across trials, country because almost all trials were performed in high income countries, and duration of BSH use because all trials did not report results stratified by duration of BSH use and large overlap existed in the duration of BSH use across trials.

We also hypothesised that the effects of interventions may vary according to more specific characteristics of the intervention (for example, pace at which patients were instructed to taper BSH, mode of delivery of cognitive behavioural therapy). In such cases, we did post hoc subgroup analyses to look for differences based on these factors. When we encountered situations in which a single study could contribute to more than one subgroup, we split the shared control group between the two subgroups to avoid double counting patients.⁵⁸ We evaluated the credibility of subgroup effects by using the ICEMAN tool.⁶⁵ In situations in which we had credible evidence of a subgroup effect, we planned to present stratified results based on the subgroup. We did not, however, identify any credible evidence of subgroup effects.

Certainty of evidence and reporting

We assessed the certainty of evidence by using the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach.⁶⁶ For each comparison and outcome, we rated the certainty of evidence as either high, moderate, low, or very low on

the basis of considerations of risk of bias (that is, study limitations), inconsistency (that is, heterogeneity in results across trials), indirectness (that is, differences between the questions investigated in studies and the question of interest), imprecision (that is, random error), and publication bias (that is, the tendency for studies with statistically significant results or positive results to be published, published faster, or published in journals with higher visibility).

High certainty evidence suggests that the estimated effect (the results from a rigorous systematic review and meta-analysis) is likely to be close to the true effect. Conversely, low or very low certainty evidence suggests that the estimated effect may in fact be substantially different from the true effect.

To make judgments about indirectness, we considered the context in which the trial was conducted. For non-drug interventions, we anticipated that evidence primarily comprised of trials conducted before 2000 would be indirect owing to the potential for patients to react differently to behavioural interventions. For example, mailed educational materials may be less effective in the digital age, in which physical mail is more quickly discarded. We did not, however, encounter interventions and outcomes dominated by evidence from before 2000, so we did not need to rate down the certainty for this reason.

We judged imprecision on the basis of the plausibility of the intervention achieving or exceeding the minimally important effect, without considering the possible magnitude of effect.⁶⁷ We used minimally important differences, sourced from the literature and by consensus of the parallel BMJ Rapid Recommendations guideline panel to make these judgments. The final assessment of certainty was fully contextualised by the guideline panel to formulate recommendations.

To facilitate interpretation, we calculated the risk differences per 1000 patients for all dichotomous outcomes by multiplying the risk ratio obtained from meta-analyses with an estimate of the risk of the outcome in the control group. For the control group risk, we used the median risk reported in standard care groups across trials. The parallel BMJ Rapid Recommendations guideline panel provided input on the applicability of the control group risks.

We report results using GRADE simple language summaries (for example, describing high certainty evidence with declarative statements, moderate certainty evidence with “probably,” low certainty evidence with “may,” and very low certainty evidence with “very uncertain”).⁶⁸ In the main manuscript, for comparisons and outcomes with multiple measures of the same outcome, we prioritise reporting results from the measure with the highest certainty. If certainty ratings are comparable, we report results from the most commonly used measure.

Patient and public involvement

The parallel BMJ Rapid Recommendations guideline panel, which included three patient partners, provided

Table 1 | Characteristics of trials and participants. Values are numbers (percentages) unless stated otherwise

Characteristic	Trials (n=49)
Design:	
Parallel randomised trial	32 (65)
Cluster randomised trial	16 (33)
Crossover by cluster randomised trial	1 (2)
Country*:	
Argentina	1 (2)
Australia	3 (6)
Belgium	5 (10)
Brazil	1 (2)
Canada	8 (16)
Finland	2 (4)
Ireland	1 (2)
Israel	1 (2)
Japan	2 (4)
Netherlands	4 (8)
Spain	2 (4)
Sweden	3 (6)
Switzerland	1 (2)
United Kingdom	6 (12)
United States	11 (22)
Not reported	1 (2)
Registered	21 (43)
Funding*:	
Industry	7 (14)
Government	35 (71)
Institutional	11 (22)
Not-for-profit	11 (22)
None	2 (4)
Time published:	
1980-90	1 (2)
1991-2000	7 (14)
2001-10	19 (39)
2011-present	22 (45)
Median (IQR) duration of follow-up across studies, weeks	24 (14-50.5)
Percentage male across all trials	31.9
Weighted mean age, years	53.7
Median (IQR) No of participants	188 (70-528)
Targeted drugs:	
Benzodiazepines including or excluding other related hypnotics	39 (80)
Non-benzodiazepine hypnotics	3 (6)
Range of inappropriate drugs including benzodiazepines and related hypnotics	7 (14)
Interventions*	No of trials (No of participants)
Tapering	6 (1122)
Education of patients	10 (5570)
Education of physicians	7 (2291)
Education of patients and physicians	1 (2367)
Cognitive behavioural therapy	11 (1930)
Medication review	3 (2499)
Pharmacist led interventions	3 (23 117)
Mindfulness	1 (70)
Drug assisted tapering	9 (616)

IQR=interquartile range.

*A trial could be classified in more than one category.

shows study selection, and supplement 2 lists excluded studies and the reasons for exclusion. All trials were published in English.

Trial and patient characteristics

Table 1 shows characteristics of trials and patients. Trials investigated a range of interventions, including tapering^{73 74 78 79 82-84 86-88}; education of patients⁸⁹⁻⁹⁸; education of physicians⁹⁹⁻¹⁰⁵; the combination of education of patients and physicians¹⁰⁶; cognitive behavioural therapy^{71 72 76-79 84 107-114}; medication review^{70 81 115 116}; pharmacist led interventions including a multicomponent intervention involving pharmacists, a pharmacist led educational intervention, and the introduction of a clinical pharmacy service to nursing homes¹¹⁷⁻¹¹⁹; mindfulness and motivational interviewing¹²⁰; drug assisted tapering and withdrawal^{69 75 85 88 121-126}; and auricular acupuncture.¹¹⁰ Trials typically compared active interventions against no additional intervention or a low intensity and minimally disruptive educational intervention targeting another topic or class of drugs other than insomnia and BSH (for example, antihypertensive agents), which we classified as usual care. Cognitive behavioural therapy, sometimes combined with tapering, was frequently compared against tapering alone, which we also classified as usual care in comparisons involving cognitive behavioural therapy.

Sixteen trials used a cluster randomised trial design,^{82 83 96 99-101 103-105 111 117-119} one a crossover by cluster design,^{76 77} and the remainder parallel randomised designs. The median duration of follow-up was 24 (interquartile range 14-50.5) weeks. All trials were conducted in high income countries, apart from two trials conducted in Argentina and Brazil.^{120 121} More than three quarters of trials were published since 2001. Four trials were conducted in nursing homes and the remainder in the community.^{86 100 116 118} More than half of all trials were funded by government organisations, followed by not-for-profit organisations or hospitals and universities.

Three trials targeted non-benzodiazepine hypnotics,^{69 92 110} seven a range of inappropriate drugs including BSH,^{95 102 104 106 115 117 118} and the remainder benzodiazepines including or excluding other related hypnotics. Trials included a median of 188 (interquartile range 70-528) patients. Approximately one third of all patients were male. Twenty one trials with 5848 patients reported that most of their participants (>60%) were using BSH for insomnia.^{69 71-77 82 83 85 87 90 93 96 107-114 120 121 124} The reasons for BSH use in other trials were not reported.

Risk of bias

Figure 2 shows the risk of bias of trials that compared tapering with usual care, and supplements 3-10 show risk of bias judgments for all trials and outcomes. Nearly all data were rated as being at high risk of bias, primarily owing to concerns about deviations from intended intervention that arose from lack of masking

input on the scope of the systematic review, including the outcomes of interest.

Results

Search results

Our search yielded more than 30 000 records. Of these, 58 publications, reporting on 49 unique trials with 39 336 patients, proved eligible. Eight trials were reported in more than one publication.⁶⁹⁻⁸⁵ Figure 1

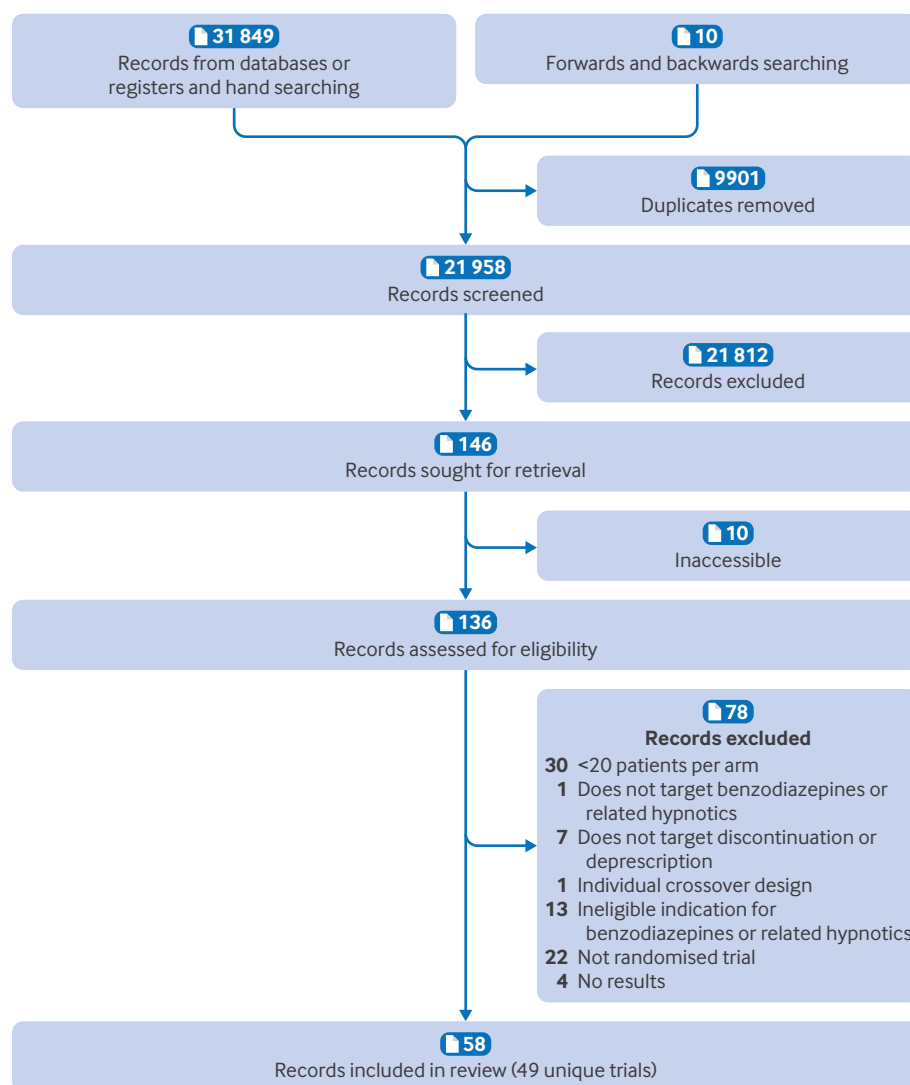


Fig 1 | Study selection

of patients and healthcare providers and potential for differential care across trial arms. Concerns about the randomisation procedures and allocation concealment also contributed to ratings of high risk of bias.

Summary of findings

Figure 3 and figure 4 summarise the results of this systematic review and meta-analysis. Low certainty evidence suggests that, compared with usual care, education of patients, medication review, and a pharmacist led educational intervention may increase the proportion of patients who discontinue BSH. However, we did not find evidence that these interventions improve physical function, mental health, or signs and symptoms of insomnia. Moderate certainty evidence suggests that education of patients probably has little or no effect on physical function, mental health, and signs and symptoms of insomnia. We did not find evidence assessing these outcomes for medication review or the pharmacist led educational intervention.

We did not find evidence that other interventions, including tapering, education of physicians, the combination of education of patients and physicians, cognitive behavioural therapy, mindfulness, other pharmacist led interventions, and drug assisted tapering or withdrawal may help patients to discontinue BSH. Moreover, we did not find high or moderate certainty evidence that any of the interventions caused an increase in dropouts. Finally, low certainty evidence suggests that multicomponent interventions may be more effective at facilitating deprescription of BSH than single component interventions. Below, we describe the results for each type of intervention.

Tapering

Six trials, with 1122 patients, reported on tapering.^{73 74 78 79 82-84 86-88} Five trials, with 839 patients, compared the effects of tapering BSH against usual care.^{82 84 86-88} One trial compared two different tapering strategies, one with written instructions and one with follow-up appointments with physicians, against

Trial		Randomisation	Deviations from intended intervention	Missing outcome data	Measurement of outcome	Selection of reported results	Carryover or period effects	Identification or recruitment of participants into clusters
Discontinuation of BSH								
Voshaar 2003, 2006		●	●	●	●	●		
Vicens 2014, 2018		●	●	●	●	●		●
Kosto 2023		●	●	●	●	●		
Dropouts during intervention								
Habraken 1997		●	●	●	●	●		
Vicens 2014, 2018		●	●	●	●	●		●
Physical function								
Voshaar 2003, 2006	SF-36 (physical functioning scale)	●	●	●	●	●		
Mental health								
Nakao 2006	Hamilton Rating Scale for Anxiety	●	●	●	●	●		
Nakao 2006	Hamilton Rating Scale for Depression	●	●	●	●	●		
Voshaar 2003, 2006	SF-36 (mental health scale)	●	●	●	●	●		
Vicens 2014, 2018	Hospital Anxiety and Depression Scale - anxiety subscale	●	●	●	●	●		●
Vicens 2014, 2018	Hospital Anxiety and Depression Scale - depression subscale	●	●	●	●	●		●
Cognitive function								
Voshaar 2003, 2006	Delayed recall 15 words test	●	●	●	●	●		
Quality of life								
Voshaar 2003, 2006	Health Utility Index	●	●	●	●	●		

● Low risk of bias ● Probably low risk of bias ● Probably high risk of bias ● High risk of bias

Fig 2 | Risk of bias of trials that compared tapering with usual care

usual care.⁸² In this trial, patients were instructed to reduce their dose of BSH by 10-25% at two to three week intervals.⁸² A second trial reduced BSH by 25% at one week intervals for three weeks followed by 12.5% at one week intervals for two weeks.⁸⁶ A third reduced BSH by 25% at two week intervals.⁸⁸ The fourth trial reduced BSH by 25% per week with the option to divide the final step to 12.5% over four days.⁸⁴ The final trial combined tapering with education.⁸⁷ Patients in the tapering arm received information on insomnia, information on the harms of prolonged BSH use, sleep hygiene recommendations, a tapering instruction table, and a letter for their family physician requesting discontinuation of BSH.⁸⁷ Patients in the usual care arm received only information on insomnia and the harms of BSH, along with the letter for their physician.⁸⁷ This trial does not describe the pace of tapering. Supplement 11 shows additional trial characteristics, and supplement 12 shows the summary of findings table.

Low certainty evidence suggests that tapering may have little or no effect on discontinuation of BSH, dropouts during the intervention, and cognitive function, and may worsen quality of life and physical

functioning. Although tapering seems to have a statistically significant effect on the proportion of patients who discontinue BSH, the absolute effect did not meet our threshold of clinical importance. Moderate certainty evidence suggests that tapering probably has little or no effect on symptoms of anxiety and depression.

We anticipated that the effects of tapering may differ according to the pace of tapering, reason for BSH use, and method for ascertaining discontinuation of BSH. We did not, however, find evidence of a difference in the effects of tapering based on pace or based on method for ascertaining discontinuation. Two trials reported including patients using BSH primarily for insomnia. An analysis restricted to these trials produced results that were consistent with the primary analysis.

The trial that compared written instructions for tapering against follow-up with physicians did not report a difference in discontinuation of BSH between the two tapering strategies.⁸² One trial compared masked tapering combined with an augmented cognitive behavioural therapy programme against unmasked tapering combined with standard cognitive behavioural therapy.^{73 74} The trial suggests that masked

Intervention	Discontinuation of BSH	Physical function	Mental health	Signs and symptoms of insomnia	Dropouts during intervention
Tapering v usual care	74 more (95% CI 25 more to 135 more) per 1000. Odds ratio 1.72 (95% CI 1.23 to 2.41) 3 trials (735 patients)	7.00 lower* (95% CI 19.00 lower to 5.00 higher) 1 trial (137 patients)	0 higher† (95% CI 0.87 lower to 0.87 higher) 1 trial (506 patients) 0 higher‡ (95% CI 0.62 lower to 0.62 higher) 1 trial (506 patients)		27 more (95% CI 67 fewer to 250 more) per 1000. Odds ratio 1.28 (95% CI 0.36 to 4.56) 2 trials (506 patients)
Education of patients v usual care	144 more (95% CI 61 more to 246 more) per 1000. Odds ratio 2.53 (95% CI 1.58 to 4.04) 8 trials (4055 patients)	0.10 higher§ (95% CI 1.42 lower to 1.62 lower) 1 trial (521 patients)	1.56 higher¶ (95% CI 0.12 lower to 3.23 higher) 1 trial (521 patients)	1.90 lower** (95% CI 2.97 lower to 0.83 lower) 1 trial (521 patients)	38 more (95% CI 24 fewer to 135 more) per 1000. Odds ratio 1.41 (95% CI 0.76 to 2.63) 2 trials (587 patients)
Education of physicians v usual care	607 more (95% CI 127 more to 828 more) per 1000. Odds ratio 18.78 (95% CI 2.31 to 152.69) 1 trial (109 patients)				
Education of patients and physicians v usual care	4 fewer (95% CI 26 fewer to 23 more) per 1000. Risk ratio 0.97 (95% CI 0.80 to 1.18) 1 trial (465 patients)				
Cognitive behavioural therapy v usual care	71 more (95% CI 21 fewer to 210 more) per 1000. Odds ratio 1.68 (95% CI 0.82 to 3.45) 8 trials (1472 patients)	8.32 lower§ (95% CI 19.60 lower to 2.96 higher) 1 trial (43 patients)	1.09 lower†† (95% CI 10.90 lower to 8.72 higher) 1 trial (43 patients)	1.56 lower** (95% CI 4.41 lower to 1.29 higher) 3 trials (168 patients)	23 more (95% CI 32 fewer to 107 more) per 1000. Odds ratio 1.24 (95% CI 0.68 to 2.24) 8 trials (1582 patients)
Medication review v usual care	104 more (95% CI 34 more to 191 more) per 1000. Odds ratio 2.04 (95% CI 1.31 to 3.17) 2 trials (736 patients)				
Mindfulness v usual care				4.82 lower** (95% CI 7.87 lower to 1.77 lower) 1 trial (52 patients)	94 more (95% CI 55 fewer to 417 more) per 1000. Odds ratio 2.07 (95% CI 0.47 to 9.02) 1 trial (70 patients)
Pharmacist led interventions					
Multicomponent intervention involving pharmacists v educational intervention for pharmacists	4 more (95% CI 62 fewer to 158 more) per 1000. Odds ratio 1.05 (95% CI 0.36 to 3.13) 1 trial (19 398 patients)				
Pharmacist led educational intervention v usual care	491 more (95% CI 234 more to 928 more) per 1000. Risk ratio 4.78 (95% CI 2.80 to 8.14) 1 trial (301 patients)				
Clinical pharmacy services in nursing homes					
High certainty		Definitely more effective than usual care	Definitely worse than usual care	Definitely no different than usual care	
Moderate certainty		Probably more effective than usual care	Probably worse than usual care	Probably no different than usual care	
Low certainty		May be more effective than usual care	May be worse than usual care	May be no different than usual care	
Very low certainty		We are very uncertain			

Fig 3 | Summary of results of interventions excluding drug assisted tapering. *SF-36 physical functioning scale (range 0-100 with higher scores indicating better function). †Hospital Anxiety and Depression Scale anxiety subscale (range 0-21 with lower scores indicating better mental health). ‡Hospital Anxiety and Depression Scale depression subscale (range 0-21 with lower scores indicating better mental health). §SF-36 physical component score (range 0-100 with higher scores indicating better function). ¶SF-12 mental component score (range 0-100 with higher scores indicating better mental health). **Insomnia Severity Index (range 0-28 with lower scores indicating fewer signs and symptoms of insomnia). ††SF-36 mental component score (range 0-100 with higher scores indicating better mental health). CI=confidence interval

Intervention	Discontinuation of BSH	Physical function	Mental health	Signs and symptoms of insomnia	Dropouts during intervention
Drug assisted tapering and withdrawal					
Melatonin assisted taper v tapering alone	39 fewer (95% CI 133 fewer to 130 more) per 1000. Odds ratio 0.77 (95% CI 0.30 to 1.96) 1 trial (83 patients)				0 fewer (95% CI 128 fewer to 586 more) per 1000. Odds ratio 1.00 (95% CI 0.06 to 16.48) 1 trial (92 patients)
Paroxetine assisted taper v tapering alone	300 more (95% CI 2 more to 595 more) per 1000. Odds ratio 3.96 (95% CI 1.01 to 15.52) 1 trial (45 patients)		5.2 lower* (95% CI 9.98 lower to 0.42 lower) 1 trial (45 patients) 3.2 lower† (95% CI 5.74 lower to 0.66 lower) 1 trial (45 patients)		
Ramelteon assisted taper v tapering alone	27 fewer (95% CI 120 fewer to 129 more) per 1000. Odds ratio 0.84 (95% CI 0.36 to 1.95) 1 trial (101 patients)				38 fewer (95% CI 87 fewer to 48 more) per 1000. Odds ratio 0.69 (95% CI 0.33 to 1.43) 1 trial (135 patients)
Propranolol v abrupt withdrawal					200 more (95% CI 85 fewer to 468 more) per 1000. Odds ratio 2.27 (95% CI 0.67 to 8.33) 1 trial (40 patients)
Dothiepin assisted taper v tapering alone					37 fewer (95% CI 124 fewer to 114 more) per 1000. Odds ratio 0.78 (95% CI 0.34 to 1.82) 1 trial (87 patients)
Lormetazepam v abrupt withdrawal	199 more (95% CI 60 fewer to 535 more) per 1000. Odds ratio 2.66 (95% CI 0.65 to 11.11) 1 trial (40 patients)				300 fewer (95% CI 443 fewer to 5 more) per 1000. Odds ratio 0.25 (95% CI 0.06 to 1.02) 1 trial (40 patients)
Flumazenil, oxazepam v oxazepam	300 more (95% CI 13 fewer to 548 more) per 1000. Odds ratio 3.50 (95% CI 0.94 to 12.97) 1 trial (40 patients)				
Bromazepam v diazepam	76 fewer (95% CI 379 fewer to 143 more) per 1000. Odds ratio 0.70 (95% CI 0.20 to 2.51) 1 trial (45 patients)				32 more (95% CI 166 fewer to 342 more) per 1000. Odds ratio 1.17 (95% CI 0.32 to 4.25) 1 trial (45 patients)
Bromazepam v lorazepam	87 more (95% CI 201 fewer to 295 more) per 1000. Odds ratio 1.44 (95% CI 0.44 to 4.74) 1 trial (46 patients)				130 fewer (95% CI 319 fewer to 160 more) per 1000. Odds ratio 0.57 (95% CI 0.17 to 1.91) 1 trial (46 patients)
Diazepam v lorazepam	162 more (95% CI 131 fewer to 338 more) per 1000. Odds ratio 2.05 (95% CI 0.59 to 7.15) 1 trial (45 patients)				161 fewer (95% CI 338 fewer to 132 more) per 1000. Odds ratio 0.49 (95% CI 0.14 to 1.70) 1 trial (45 patients)

High certainty	Definitely more effective than usual care	Definitely worse than usual care	Definitely no different than usual care
Moderate certainty	Probably more effective than usual care	Probably worse than usual care	Probably no different than usual care
Low certainty	May be more effective than usual care	May be worse than usual care	May be no different than usual care
Very low certainty	We are very uncertain		

Fig 4 | Summary of results for drug assisted tapering and withdrawal interventions. *Hamilton Rating Scale for anxiety (range: 0 to 56 with lower scores indicating less anxiety). †Hamilton Rating Scale for depression (range: 0 to 56 with lower scores indicating less depression). CI=confidence interval

tapering may increase the proportion of patients who discontinue BSH compared with unmasked tapering with cognitive behavioural therapy but may have little or no effect on signs and symptoms of insomnia compared with unmasked tapering.

Education of patients

Ten trials, with 5570 patients, investigated the effects of educating patients.⁸⁹⁻⁹⁸ Nine trials, with 4709 patients, compared education of patients against usual care.⁸⁹⁻⁹⁶ Educational content described adverse effects of BSH and beneficial reasons for reducing BSH, tapering schedules, stories of peers who have successfully stopped or reduced BSH, self-assessment tools, and suggestions for alternative treatments to manage insomnia. One trial included educational materials that also had a link to a cognitive behavioural therapy website.⁹³

Seven trials investigated the effects of single letters or distribution of written educational materials such as brochures,⁸⁹⁻⁹⁴ one trial investigated the effects of the distribution of multiple letters,⁸⁹ four trials investigated the effects of written educational materials coupled with follow-up by physicians, pharmacists, or researchers,⁹¹⁻⁹³ and one trial investigated the effects of counselling by researchers or healthcare providers.⁹⁸ Supplement 13 shows additional trial characteristics, and supplement 14 shows the summary of findings table.

Low certainty evidence suggests that education of patients may increase the proportion of patients who discontinue BSH and may have little or no effect on dropouts. Moderate certainty evidence also suggests that education of patients probably has little or no effect on physical function, mental health, signs and symptoms of insomnia, total sleep time, sleep onset latency, and sleep efficiency, although the results for these outcomes came from only a single trial and may not reflect the effects of all educational interventions targeting patients.⁹⁰

We anticipated that the effect of educational interventions may differ according to the mode of the educational intervention, reason for BSH use, and method for ascertaining BSH discontinuation. We did not find evidence that the effect estimates varied according to the mode of the educational intervention or method for ascertaining BSH discontinuation. Three trials reported including patients using BSH primarily for insomnia.⁹⁰ An analysis restricted to these trials produced results that were consistent with the primary analysis.

One trial compared a single generic letter, a single tailored letter, and three tailored letters.⁹⁷ The results of the trial suggest that tailored letters may be more effective at achieving discontinuation than generic letters but did not report a difference between distributing a single tailored letter and three tailored letters. Another trial compared patient education through a passive letter or physician led consultations with patients against usual care.⁹¹ The trial did not report a statistically significant difference

between the two education strategies. A third trial compared two educational materials that differed in their content and approach against usual care.⁹⁰ The YAWNS-1 and YAWNS-2 packages were similar but had some differences. YAWNS-1 included a flexible dose reduction schedule and emphasised behaviour change techniques and cognitive behavioural therapy for insomnia. The YAWNS-2 package used a fixed dose reduction schedule, described risks of BSH, and described strategies for improving sleep. The results of the trial suggest that YAWNS-1 may be preferable to YAWNS-2.

Education of physicians

Seven trials, with more than 2000 patients, reported on interventions targeted at educating physicians.⁹⁹⁻¹⁰⁵ Six trials, with more than 2000 patients, compared education of physicians against usual care.¹⁰⁰⁻¹⁰⁵ The educational interventions typically focused on the indications for BSH use, the effects of BSH in older people, strategies for counselling patients on reducing or discontinuing BSH, and alternative treatments for sleep. Some interventions focused on reducing physicians' overall prescribing of BSH rather than educating physicians to encourage current users to discontinue BSH.

One trial reported on a multisession educational programme for nurses, nursing assistants, and physicians in nursing homes.¹⁰⁰ Two trials reported on educational visits to physicians,¹⁰¹ one of which also involved follow-up calls.¹⁰¹

Four trials investigated the combination of physician education with audit and feedback.⁹⁹ One trial compared four unique strategies: audit and feedback alone, audit and feedback combined with a letter from the medical regulatory authority body describing best practices and advice on how to reduce medications in patients receiving higher doses of BSH and links to external prescribing resources, audit and feedback combined with the letter and a phone call from a pharmacist, or audit and feedback combined with the letter and a phone call from a physician.⁹⁹ Another trial compared audit and feedback combined with in-person education of physicians on doing medication reviews against usual care.¹⁰⁴ This trial also provided incentive payments to physicians to do medication reviews.¹⁰⁴ One trial compared audit and feedback combined with a two hour educational in-person workshop for physicians against usual care.¹⁰⁵ The last trial compared sending written educational bulletins to physicians every two months for six months, also with audit and feedback, against usual care.¹⁰³ Supplement 15 shows additional trial characteristics, and supplement 16 shows the summary of findings table.

Only one trial reported on discontinuation of BSH. Owing to problems of serious risk of bias and very serious imprecision, the evidence on the effects of education of physicians on the proportion of patients who discontinue BSH is very uncertain. Several trials reported the number of prescriptions or average dose

of BSH in medical practices instead of discontinuation. These trials either reported no difference or a small reduction in prescriptions of BSH following education of physicians.¹⁰¹⁻¹⁰⁵

We hypothesised that the effect of education of physicians may differ according to the educational strategy, reason for BSH use among patients, and method for ascertaining BSH discontinuation. Insufficient evidence precluded subgroup analyses to investigate these factors. The trial comparing audit and feedback—alone or combined with a letter from the medical regulatory authority or a phone call—did not report differences between the four educational strategies on the number of patients receiving high doses of BSH.⁹⁹

Education of patients and physicians

One trial, with 2367 patients, compared an intervention that combined the education of patients and physicians against usual care.¹⁰⁶ The trial targeted drugs acting on the central nervous system, including benzodiazepines, opioids, muscle relaxants, tricyclic antidepressants, and first generation antihistamines. Primary care practices within an integrated healthcare system were randomised to receive the intervention or usual care. Patient education comprised medication specific educational brochures, handouts with tips on managing symptoms with non-drug approaches, and a brochure on ways to prevent falls. Physician education comprised information on the target drug class, evidence based treatment alternatives, and practice support for deprescribing. Supplement 17 shows additional trial characteristics, and supplement 18 shows the summary of findings table.

Moderate certainty evidence suggests that a programme comprised of distributing educational materials to patients and healthcare providers probably has little or no effect on discontinuation of BSH compared with usual care. This comparison, however, was informed by a single trial, so the evidence may not be broadly applicable to all educational interventions targeting patients and physicians.

We hypothesised that the effect of education of patients and physicians may differ according to the educational strategy, reason for BSH use among patients, and method for ascertaining BSH discontinuation. Insufficient evidence precluded subgroup analyses to investigate these factors.

Cognitive behavioural therapy

Eleven trials, with 1930 patients, investigated cognitive behavioural therapy.^{71 72 76-79 84 107-114} Nine trials, with 1762 patients, compared cognitive behavioural therapy alone or in combination with tapering against usual care or tapering alone.^{76 84 107-109 111-114} Cognitive behavioural therapy typically involved principles of cognitive behavioural therapy for insomnia,⁵⁴ such as education about stimulus control, sleep restriction, sleep hygiene, and dysfunctional beliefs about sleep. As most cognitive behavioural interventions included

some suggestion of tapering, the analysis comparing cognitive behavioural therapy against usual care also includes trials that compared cognitive behavioural therapy in combination with tapering against tapering alone, trials that compared cognitive behavioural therapy in combination with tapering against no other interventions, and trials that compared cognitive behavioural therapy alone against tapering alone.

Four trials offered cognitive behavioural therapy in person and in groups,^{84 108 113 114} three trials in person and individually,^{76 107 112} one trial using booklets,¹⁰⁹ and another trial using an interactive e-learning program.¹¹¹ The intensity of cognitive behavioural therapy varied from five sessions over 10 weeks to 10 sessions over 10 weeks. The duration of sessions varied between 50 minutes and two hours. Supplement 19 shows additional trial characteristics, and supplement 20 shows the summary of findings table.

Owing to concerns about risk of bias, inconsistency, and imprecision, we are very uncertain of the effects of cognitive behavioural therapy on the proportion of patients who discontinue BSH. Moderate certainty evidence suggests that cognitive behavioural therapy probably has little or no effect on signs and symptoms of insomnia, sleep onset latency, and sleep efficiency. Low certainty evidence also suggests that cognitive behavioural therapy may improve quality of life, may have little or no effect on dropouts, mental health, cognitive function, and total sleep time, and may worsen physical function.

We anticipated that the effects of cognitive behavioural therapy may differ according to whether it was administered alone or in combination with tapering, whether it was compared with no other intervention or tapering, the mode of delivery, the number of sessions, the reason for BSH use, and the method for ascertaining BSH discontinuation. We did subgroup analyses investigating the effects of these factors and did not find evidence that these factors influenced effect estimates.

One trial compared cognitive behavioural therapy against auricular acupuncture and did not find a statistically significant difference in discontinuation of BSH between the two groups.¹¹⁰ Another trial compared cognitive behavioural therapy alone or in combination with two different blinded tapering strategies (reduction of BSH by 10% or 20% every two weeks).⁷¹ More patients in the blinded tapering groups discontinued BSH than in the cognitive behavioural therapy alone group. Finally, a trial that compared cognitive behavioural therapy alone, cognitive behavioural therapy coupled with tapering, and tapering alone did not find a statistically significant difference between the three interventions in the number of patients that discontinued BSH.¹¹³

Medication review

Three trials, with 2499 patients, compared medication reviews with usual care.^{70 81 115 116} In one trial, for older adults with multimorbidity and polypharmacy,

a physician and a pharmacist jointly did a medication review supported by clinical decision software followed by shared decision making between the patient and physician.^{70 81} In the second trial, the intervention was implemented in nursing homes and involved interdisciplinary case conferences with the general practitioner, pharmacist, and nurse doing medication reviews for each nursing home resident.¹¹⁶ In this trial, medication review was also combined with education of healthcare providers, which included e-learning and a face-to-face workshop covering pharmacotherapy for older people. The last trial implemented a medication review, done by pharmacists and nurses, for home healthcare patients.¹¹⁵ Supplement 21 shows additional trial characteristics, and supplement 22 shows the summary of findings table.

Low certainty evidence suggests that medication review may increase the proportion of patients who discontinue BSH. We hypothesised that the effect of medication review may differ according to the reason for BSH use among patients and the method for ascertaining BSH discontinuation. Insufficient evidence precluded subgroup analyses to investigate these factors.

Mindfulness intervention

One trial, with 70 patients, compared eight sessions of a group mindfulness based relapse prevention intervention coupled with motivational interviewing against motivational interviewing alone.¹²⁰ Supplement 23 shows additional trial characteristics, and supplement 24 shows the summary of findings table.

We are very uncertain of the effects of mindfulness on dropouts during intervention. Low certainty evidence suggests that mindfulness may have little or no effect on signs and symptoms of insomnia.

We anticipated that the effects of mindfulness based interventions may differ according to the mode of delivery, number of sessions, reason for BSH use, and method for ascertaining BSH discontinuation. Insufficient evidence, however, precluded subgroup analyses based on these factors.

Pharmacist led interventions

Three trials, with 23 117 patients, reported on pharmacist led interventions.¹¹⁷⁻¹¹⁹ Supplement 25 shows additional trial characteristics, and supplements 26-28 show summary of findings tables.

One of these trials compared an intervention that encouraged pharmacists to distribute educational materials to patients and physicians against usual care.¹¹⁷ Materials for patients included drug specific brochures outlining why the drug might be inappropriate, alternative treatment options, and a visual tapering protocol. Materials for physicians provided the rationale for deprescribing, evidence on drug harms, credible sources for recommendations, and safer alternatives. Low certainty evidence suggests that this intervention may increase the proportion of patients who discontinue BSH.

The second trial compared an intensive support programme in which pharmacists received training to work with physicians to identify long term BSH users and send letters encouraging patients to discontinue BSH against a programme in which pharmacists received a manual to perform the same task.¹¹⁹ Low certainty evidence suggests that the intensive programme may have little or no effect on BSH discontinuation.

The last trial investigated the effects of introducing clinical pharmacy services in nursing homes compared with usual care.¹¹⁸ The trial reported moderate certainty evidence that the clinical pharmacy service probably reduced the number of prescriptions for benzodiazepines in nursing homes compared with usual care.

Insufficient evidence precluded any subgroup or sensitivity analyses.

Drug assisted tapering or withdrawal

Nine trials, with 616 patients, reported on drug assisted tapering or withdrawal.^{69 75 85 88 121-126} Two trials, with 140 patients, compared melatonin assisted tapering against tapering without melatonin^{75 121}; one trial, with 66 patients, compared tapering with paroxetine against tapering alone and usual care⁸⁸; one trial, with 135 patients, compared tapering with ramelteon against tapering alone⁶⁹; one trial, with 40 patients, compared abrupt withdrawal with propranolol against abrupt withdrawal alone¹²⁶; and one trial, with 87 patients, compared tapering with dothiepin against tapering alone.¹²⁵

Three trials involved replacing patients' regular BSH with an alternative BSH or placebo.¹²²⁻¹²⁴ In the first trial, 40 patients' BSH was replaced with either lormetazepam or placebo for one week, after which the drug was discontinued.¹²⁴ In another trial, 40 patients were randomised to either discontinue BSH and receive oxazepam for the first three nights coupled with flumazenil for eight nights or receive oxazepam for the first three nights without flumazenil.¹²² In the last trial, 68 patients were randomised to be switched to diazepam, lorazepam, or bromazepam, after which each of the drugs was tapered over 10 weeks.¹²³ Supplement 29 shows additional trial characteristics, and supplements 30-39 show summary of findings tables.

Tapering with paroxetine compared with tapering alone may improve anxiety and depression, but the evidence is of low certainty. The evidence on the effects of tapering with melatonin compared with tapering alone, tapering with paroxetine compared with tapering alone, tapering with ramelteon compared with tapering alone, switching to lormetazepam before withdrawal compared with withdrawal alone, flumazenil and oxazepam supported withdrawal compared with oxazepam alone, and switching to bromazepam, lorazepam, or diazepam before withdrawal on discontinuation of BSH is very uncertain. Insufficient evidence precluded subgroup analyses.

Multicomponent interventions

Eleven trials, with 20 250 patients, compared multicomponent with single component interventions.^{71 72 76-79 84 99 107-109 112 113 119 120} Supplement 40 shows the summary of findings table. Low certainty evidence suggests that multicomponent interventions may increase the proportion of patients who discontinue BSH compared with single component interventions. We did not find compelling evidence that multicomponent interventions led to more dropouts during the intervention.

Discussion

Main findings

Our systematic review and meta-analysis summarised all the evidence from randomised trials assessing the effectiveness of interventions to discontinue or facilitate deprescription of BSH in healthcare settings. We reviewed 49 unique randomised trials, with more than 30 000 patients, and found that education of patients, medication reviews, and a pharmacist led educational intervention may increase the proportion of patients who discontinue BSH.

We did not find evidence that other interventions increase the proportion of patients who discontinue BSH, including tapering, education of physicians, the combination of education of patients and physicians, cognitive behavioural therapy, mindfulness, other pharmacist led interventions, or drug assisted tapering or withdrawal. The evidence for these interventions was either very low certainty or did not suggest an important effect.

Interpretation of findings

The current body of evidence is overall uncertain, primarily because of concerns about risk of bias and imprecision. For many interventions, such as medication reviews and pharmacist led interventions, no evidence is available on several outcomes important to patients, including physical function, mental health, and signs and symptoms of insomnia. Trials also seldom reported outcomes beyond one year of follow-up. Initial successes in discontinuing BSH might not be sustained in the long term.

Many of the interventions described in this review can be implemented in various ways. For example, tapering strategies can vary in the degree of follow-up by healthcare providers and the pace of tapering. We had insufficient evidence to confidently conclude whether the effects of interventions may vary according to their specific characteristics. Even for comparisons for which we were able to do subgroup analyses to investigate the influence of other intervention characteristics on effect estimates, the effect estimates within subgroups were imprecise.

Our results may not generalise to all patients. Fewer than half of trials reported most patients to be predominantly using BSH for insomnia. Whether the effects of interventions for deprescribing BSH may be different depending on the reason for its use is unclear. Moreover, the evidence came primarily from patients

in their 50s, despite the use of BSH being highly prevalent and likely more problematic in older adults. Nearly all trials excluded patients with significant mental health conditions, so whether these results also apply to those with comorbid mental health diagnoses is uncertain. Poor reporting and insufficient evidence also precluded investigation of how factors such as the habitual dose of BSH, degree of dependence, and patient's adherence may influence the effects of interventions.

Additionally, the effects of interventions facilitated by healthcare providers may vary depending on healthcare providers' knowledge, skills, motivation, and self-efficacy. Healthcare providers who choose to participate in clinical trials may be more knowledgeable and skilled than others and may hold stronger beliefs about the effectiveness of the interventions that are under study.

Not all interventions to reduce BSH may be feasible across all health systems. Many trials examined costly, time intensive interventions that overburdened health systems may struggle to support. For instance, although pharmacist led education of patients and physicians seemed promising, pharmacists' scope and nature of practice varies across jurisdictions and settings and is often constrained by limited time and high workloads. Health systems may mitigate this challenge by offering incentives to support the implementation of such interventions.

We are uncertain of the optimal triggers and contraindications for BSH deprescription. Patients who enrol in clinical trials and adhere to the intervention may be more willing to discontinue BSH than are those in the general population. However, this willingness likely applies equally to patients randomised to the control group, making it unlikely to meaningfully bias our results.

Finally, key organisational aspects of interventions that may contribute to their success remain uninvestigated. Except for one trial that offered financial incentives to physicians for doing medication reviews,¹⁰⁴ the effectiveness of changes in financial and governance structures has not been studied. These factors, along with implementation strategies and the context in which interventions are implemented, may affect their success.

Relation to previous literature and guidance

This systematic review builds on previous reviews by providing the most up-to-date evidence. Previous reviews have been restricted to select interventions,^{37 39 44 45 127} have not adequately differentiated between various types of interventions,^{18 44} have been restricted to older adults,^{37 127} have not provided quantitative synthesis of study results,^{18 40 45} or have not assessed the certainty of the body of evidence.^{37 39 40 45} This review adheres to the highest established standards for producing trustworthy, transparent, and actionable systematic reviews: we assessed the risk of bias of studies, provided quantitative estimates of the effects of interventions, and evaluated the certainty of evidence

by using the latest GRADE guidance.¹²⁸ Our findings are largely consistent with other reviews. We show that although certain interventions may be effective in increasing the proportion of patients who discontinue BSH, the evidence is overall of low certainty.

Our findings on the effectiveness of tapering contradicts existing guidance that recommends tapering.¹²⁹ We show that although tapering increases the number of patients who discontinue BSH, this effect is modest and may not be clinically important. This may partly be attributed to the effects of tapering alone not being sustained in the long term. In our review, we prioritised extracting and synthesising effects at the longest reported point of follow-up. Trials typically reported a reduction in the effects of tapering over time.^{78 79 82-84}

Our findings on the effectiveness of cognitive behavioural therapy also contrast with those of a previous review that found it to be effective.⁴⁵ This discrepancy can be attributed to differences in eligibility criteria and the time point of interest. For example, a previous review reports results for the most positive time point for a trial investigating cognitive behavioural therapy instead of the longest reported point of follow-up.¹¹³ At the longest reported point of follow-up, the data suggest that initial cessation of BSH was not later sustained.

Strengths and limitations of study

The strengths of our systematic review and meta-analysis include a rigorous and comprehensive search for eligible randomised trials, screening and extraction of data in duplicate, the application of GRADE methods to evaluate the certainty of evidence, and our focus on outcomes important to patients.

Despite our rigorous search of the literature, we might have missed eligible trials, particularly publications in languages other than English. We mitigated this problem by also reviewing the references of similar systematic reviews and soliciting experts about eligible trials that may not have come up in our search.³⁷⁻⁴⁵ We assessed the certainty of evidence by using the GRADE approach. Although the GRADE approach presents a comprehensive framework for systematically and transparently considering all factors that may bear on the certainty of evidence, its application is ultimately subjective, and other authors may come to different conclusions about the certainty of evidence.¹³⁰ We anticipated that the effects of interventions may vary according to several characteristics of patients and interventions such as age and duration of BSH use, but we were unable to do subgroup analyses owing to insufficient data. A group of content experts and methodologists, with feedback from the wider guideline panel, classified similar interventions together for synthesis. In making these classifications, we considered the objective of the intervention, the target stakeholder, the class of therapy, and whether the intervention involved gradual or abrupt reduction in BSH dose, educational materials, or meetings with

professionals. Other researchers might classify the interventions differently.

Implications of findings

This systematic review and meta-analysis presents a range of interventions that may be useful in reducing prolonged use of BSH. We found the body of evidence to be overall of low certainty, highlighting the need for additional research.

Many professional societies and authoritative organisations advise against the prolonged use of BSH, citing evidence suggesting that these drugs are associated with adverse events.^{1 2 18} This systematic review informs parallel BMJ Rapid Recommendations issued by the BE-SAFE consortium. Its dedicated panel appraised this evidence to issue a conditional recommendation in favour of deprescription.³⁴

Decision makers also need to consider practitioners' workload.⁴⁸ Although certain interventions may be effective, some of these interventions may reduce the amount of time that healthcare providers have available for other medical problems and so may not be feasible in overburdened settings. We also found other interventions that could be implemented with little disruption to existing healthcare services and workflows. Such interventions, including sending letters and booklets to patients, are unlikely to significantly affect the workload of healthcare providers.

Future trials that investigate the effects of pragmatic interventions (that is, those that could be implemented in the real world without massive interference with other healthcare services or demanding unrealistic resources or time from healthcare providers) are still needed. These trials should recruit sufficient participants to allow for more precise effect estimates and use sufficient safeguards against bias. For example, as the nature of the interventions will likely preclude masking of patients and healthcare providers, investigators may still reduce potential for bias by matching interventions with regard to the degree of interaction between patients and healthcare providers to reduce expectancy effects and use strict treatment and follow-up protocols to reduce deviations from the intended intervention.

Conclusion

The evidence to guide patients and clinicians on the effectiveness of strategies to discontinue BSH is uncertain. Low certainty evidence suggests that educating patients, medication review, and enabling pharmacists to educate patient and physicians may increase the proportion of patients who discontinue BSH. We did not find compelling evidence that tapering, education of physicians, the combination of education of patients and physicians, cognitive behavioural therapy, mindfulness, other pharmacist led interventions, and drug supported tapering and withdrawal strategies are effective. Finally, low certainty evidence suggests that multicomponent interventions may be more effective at facilitating

discontinuation of BSH than single component interventions.

AUTHOR AFFILIATIONS

¹Department of Anesthesia, McMaster University, Hamilton, ON, Canada

²Department Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

³MAGIC Evidence Ecosystem Foundation, Oslo, Norway

⁴Division of Respiriology; Department of Medicine, University of Toronto, Toronto, ON, Canada

⁵Department of Medicine, McMaster University, Hamilton, ON, Canada

⁶Institute of Psychiatry and Neurology, Third Department of Psychiatry and Sleep Medicine Center, Warsaw, Poland

⁷Department of Science and Evaluation, Medical Research Agency, Warsaw, Poland

⁸Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

⁹Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland

¹⁰Fund for Scientific Research – FNRS, Belgium

¹¹Clinical Pharmacy and Pharmacoevidence Research Group, Louvain Drug Research Institute, Louvain, Belgium

¹²CHU UCL Namur, Pharmacy department, Yvoir, Belgium

¹³Division General Internal Medicine, Department of Medicine, University Hospitals of Geneva, Geneva, Switzerland

This systematic review was conducted as a part of the BE-SAFE project, funded by the European Union's Horizon Europe research and innovation programme and the Swiss State Secretariat for Education, Research and Innovation. The overarching goal of BE-SAFE is to improve patient safety by reducing gaps in knowledge and practice related to the reduction of benzodiazepines and sedative hypnotics used for sleep difficulties in Europe. We thank the BE-SAFE Publication Committee for reviewing the manuscript.

Contributors: DZ, TA, SKN, AS, AC, SVV, and SS conceived the study. RC developed the search strategy. ML, TJ, SK, TP, TA, SVV, SS, JPL, and SKN screened studies for eligibility and collected data. DZ and TP did analyses. DZ wrote the first draft of the manuscript. All co-authors interpreted the data, substantively contributed to the manuscript, and reviewed and approved the final draft of the manuscript. TA supervised the study. DZ and TA are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This project has received funding from the European Union's Horizon Europe research and innovation programme under the grant agreement No 101057123 and from the Swiss State Secretariat for Education, Research and Innovation under the contract No 22.00116. The funders had no role in considering the study design or in the collection, analysis, and interpretation of data, the writing of the report, or the decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from the European Union's Horizon Europe research and innovation programme and from the Swiss State Secretariat for Education, Research and Innovation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The study uses publicly available data and is thus exempt from need for ethics review.

Data sharing: All data are available on Open Science Framework (<https://osf.io/8em4p/files/osfstorage>).

Transparency: The lead authors (the manuscript's guarantors) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Dissemination to participants and related patient and public communities: This systematic review is intended to inform forthcoming BMJ Rapid Recommendations.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>.

- 1 National Institute for Health and Care Excellence. Guidance on the use of zaleplon, zolpidem and zopiclone for the short-term management of insomnia. 2004. <https://www.nice.org.uk/guidance/ta77/chapter/1-Guidance>
- 2 By the 2023 American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2023 updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. *J Am Geriatr Soc* 2023;71:2052-81. doi:10.1111/jgs.18372
- 3 Bakken MS, Engeland A, Engesaeter LB, Ranhoff AH, Hunskaar S, Ruths S. Risk of hip fracture among older people using anxiolytic and hypnotic drugs: a nationwide prospective cohort study. *Eur J Clin Pharmacol* 2014;70:873-80. doi:10.1007/s00228-014-1684-z
- 4 Barbone F, McMahon AD, Davey PG, et al. Association of road-traffic accidents with benzodiazepine use. *Lancet* 1998;352:1331-6. doi:10.1016/S0140-6736(98)04087-2
- 5 Barker MJ, Greenwood KM, Jackson M, Crowe SF. Cognitive effects of long-term benzodiazepine use: a meta-analysis. *CNS Drugs* 2004;18:37-48. doi:10.2165/00023210-200418010-00004
- 6 Buffett-Jerrott SE, Stewart SH. Cognitive and sedative effects of benzodiazepine use. *Curr Pharm Des* 2002;8:45-58. doi:10.2174/1381612023396654
- 7 Clegg A, Young JB. Which medications to avoid in people at risk of delirium: a systematic review. *Age Ageing* 2011;40:23-9. doi:10.1093/ageing/afq140
- 8 Billioti de Gage S, Bégaud B, Bazin F, et al. Benzodiazepine use and risk of dementia: prospective population based study. *BMJ* 2012;345:e6231. doi:10.1136/bmj.e6231
- 9 Donnelly K, Bracchi R, Hewitt J, Routledge PA, Carter B. Benzodiazepines, Z-drugs and the risk of hip fracture: A systematic review and meta-analysis. *PLoS One* 2017;12:e0174730. doi:10.1371/journal.pone.0174730
- 10 Hill KD, Wee R. Psychotropic drug-induced falls in older people: a review of interventions aimed at reducing the problem. *Drugs Aging* 2012;29:15-30. doi:10.2165/11598420-000000000-00000
- 11 Katzman MA, Bleau P, Blier P, et al. Canadian Anxiety Guidelines Initiative Group on behalf of the Anxiety Disorders Association of Canada/Association Canadienne des troubles anxieux and McGill University. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry* 2014;14(Suppl 1):S1. doi:10.1186/1471-244X-14-S1-S1
- 12 Khong TP, de Vries F, Goldenberg JS, et al. Potential impact of benzodiazepine use on the rate of hip fractures in five large European countries and the United States. *Calcif Tissue Int* 2012;91:24-31. doi:10.1007/s00223-012-9603-8
- 13 Meuleners LB, Duke J, Lee AH, Palamara P, Hildebrand J, Ng JQ. Psychoactive medications and crash involvement requiring hospitalization for older drivers: a population-based study. *J Am Geriatr Soc* 2011;59:1575-80. doi:10.1111/j.1532-5415.2011.03561.x
- 14 Neutel CI, Skurtveit S, Berg C. What is the point of guidelines? Benzodiazepine and z-hypnotic use by an elderly population. *Sleep Med* 2012;13:893-7. doi:10.1016/j.sleep.2011.12.014
- 15 Orriols L, Philip P, Moore N, et al. CESIR Research Group. Benzodiazepine-like hypnotics and the associated risk of road traffic accidents. *Clin Pharmacol Ther* 2011;89:595-601. doi:10.1038/clpt.2011.3
- 16 Paterniti S, Dufouil C, Alperovitch A. Long-term benzodiazepine use and cognitive decline in the elderly: the Epidemiology of Vascular Aging Study. *J Clin Psychopharmacol* 2002;22:285-93. doi:10.1097/00004714-200206000-00009
- 17 Picton JD, Marino AB, Nealy KL. Benzodiazepine use and cognitive decline in the elderly. *Am J Health Syst Pharm* 2018;75:e6-12. doi:10.2146/ajhp160381
- 18 Pottie K, Thompson W, Davies S, et al. Deprescribing benzodiazepine receptor agonists: Evidence-based clinical practice guideline. *Can Fam Physician* 2018;64:339-51.
- 19 Cumming RG, Le Couteur DG. Benzodiazepines and risk of hip fractures in older people: a review of the evidence. *CNS Drugs* 2003;17:825-37. doi:10.2165/00023210-200317110-00004
- 20 Riemann D, Espie CA, Altena E, et al. The European Insomnia Guideline: An update on the diagnosis and treatment of insomnia 2023. *J Sleep Res* 2023;32:e14035. doi:10.1111/jsr.14035
- 21 Zhong G, Wang Y, Zhang Y, Zhao Y. Association between benzodiazepine use and dementia: a meta-analysis. *PLoS One* 2015;10:e0127836. doi:10.1371/journal.pone.0127836

- 22 Brandt J, Janzen D, Alessi-Severini S, et al. Risk of long-term benzodiazepine and Z-drug use following the first prescription among community-dwelling adults with anxiety/mood and sleep disorders: a retrospective cohort study. *BMJ Open* 2021;11:e046916. doi:10.1136/bmjopen-2020-046916
- 23 Davies J, Rae TC, Montagu L. Long-term benzodiazepine and Z-drugs use in England: a survey of general practice [corrected]. *Br J Gen Pract* 2017;67:e609-13. doi:10.3399/bjgp17X691865
- 24 Huerta C, Abbing-Karahagopian V, Requena G, et al. Exposure to benzodiazepines (anxiolytics, hypnotics and related drugs) in seven European electronic healthcare databases: a cross-national descriptive study from the PROTECT-EU Project. *Pharmacoepidemiol Drug Saf* 2016;25(Suppl 1):56-65. doi:10.1002/pds.3825
- 25 Kurko TA, Saastamoinen LK, Tähkäpää S, et al. Long-term use of benzodiazepines: Definitions, prevalence and usage patterns - a systematic review of register-based studies. *Eur Psychiatry* 2015;30:1037-47. doi:10.1016/j.eurpsy.2015.09.003
- 26 Panneman MJ, Goetsch WG, Kramarz P, Herings RM. The costs of benzodiazepine-associated hospital-treated fall injuries in the EU: a Pharmostudy. *Drugs Aging* 2003;20:833-9. doi:10.2165/00002512-200320110-00004
- 27 Anthierens S, Habraken H, Petrovic M, Deveugele M, De Maeseneer J, Christiaens T. First benzodiazepine prescriptions: qualitative study of patients' perspectives. *Can Fam Physician* 2007;53:1200-1.
- 28 Lynch T, Ryan C, Cadogan CA. 'I just thought that it was such an impossible thing': A qualitative study of barriers and facilitators to discontinuing long-term use of benzodiazepine receptor agonists using the Theoretical Domains Framework. *Health Expect* 2022;25:355-65. doi:10.1111/hex.13392
- 29 Parr JM, Kavanagh DJ, Young RM, McCafferty K. Views of general practitioners and benzodiazepine users on benzodiazepines: a qualitative analysis. *Soc Sci Med* 2006;62:1237-49. doi:10.1016/j.socscimed.2005.07.016
- 30 Sirdifield C, Chipchase SY, Owen S, Siriwardena AN. A systematic review and meta-synthesis of patients' experiences and perceptions of seeking and using benzodiazepines and Z-drugs: towards safer prescribing. *Patient* 2017;10:1-15. doi:10.1007/s40271-016-0182-z
- 31 Fang SY, Chen CY, Chang IS, Wu ECH, Chang CM, Lin KM. Predictors of the incidence and discontinuation of long-term use of benzodiazepines: a population-based study. *Drug Alcohol Depend* 2009;104:140-6. doi:10.1016/j.drugalcdep.2009.04.017
- 32 Pétursson H. The benzodiazepine withdrawal syndrome. *Addiction* 1994;89:1455-9. doi:10.1111/j.1360-0443.1994.tb03743.x
- 33 Evrard P, Damiaens A, Patey AM, Grimshaw JM, Spinewine A. Barriers and enablers towards benzodiazepine-receptor agonists deprescribing in nursing homes: A qualitative study of stakeholder groups. *Explor Res Clin Soc Pharm* 2023;9:100258. doi:10.1016/j.rcsop.2023.100258
- 34 Agoritsas T, Kumbargere Nagraj S, Callegari E, et al. Deprescription of Benzodiazepine and Sedative Hypnotics in Insomnia Disorder: A Clinical Practice Guideline. BE-SAFE Deprescription recommendations. 2023. <https://app.magicapp.org/#/guideline/jDePyl>
- 35 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi:10.1136/bmj.n71
- 36 Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann N Y Acad Sci* 1993;703:135-46, discussion 146-8. doi:10.1111/j.1749-6632.1993.tb26343.x
- 37 Dou C, Rebane J, Bardal S. Interventions to improve benzodiazepine tapering success in the elderly: a systematic review. *Aging Ment Health* 2019;23:411-6. doi:10.1080/13607863.2017.1423030
- 38 Gould RL, Coulson MC, Patel N, Highton-Williamson E, Howard RJ. Interventions for reducing benzodiazepine use in older people: meta-analysis of randomised controlled trials. *Br J Psychiatry* 2014;204:98-107. doi:10.1192/bjp.bp.113.126003
- 39 Lynch T, Ryan C, Hughes CM, et al. Brief interventions targeting long-term benzodiazepine and Z-drug use in primary care: a systematic review and meta-analysis. *Addiction* 2020;115:1618-39. doi:10.1111/add.14981
- 40 Mokhar A, Topp J, Härter M, et al. Patient-centered care interventions to reduce the inappropriate prescription and use of benzodiazepines and z-drugs: a systematic review. *PeerJ* 2018;6:e5535. doi:10.7717/peerj.5535
- 41 Rasmussen AF, Poulsen SS, Oldenburg LK, Vermehren C. The Barriers and Facilitators of Different Stakeholders When Deprescribing Benzodiazepine Receptor Agonists in Older Patients-A Systematic Review. *Metabolites* 2021;11:254. doi:10.3390/metabo11040254
- 42 Reeve E, Ong M, Wu A, Jansen J, Petrovic M, Gnjdijc D. A systematic review of interventions to deprescribe benzodiazepines and other hypnotics among older people. *Eur J Clin Pharmacol* 2017;73:927-35. doi:10.1007/s00228-017-2257-8
- 43 Ribeiro PRS, Schlindwein AD. Benzodiazepine deprescription strategies in chronic users: a systematic review. *Fam Pract* 2021;38:684-93. doi:10.1093/fampra/cmab017
- 44 Soni A, Thiagarajan A, Reeve J. Feasibility and effectiveness of deprescribing benzodiazepines and Z-drugs: systematic review and meta-analysis. *Addiction* 2023;118:7-16. doi:10.1111/add.15997
- 45 Takaesu Y, Utsumi T, Okajima I, et al. Psychosocial intervention for discontinuing benzodiazepine hypnotics in patients with chronic insomnia: A systematic review and meta-analysis. *Sleep Med Rev* 2019;48:101214. doi:10.1016/j.smrv.2019.101214
- 46 Bayliss EA, Albers K, Gleason K, et al. Recommendations for outcome measurement for deprescribing intervention studies. *J Am Geriatr Soc* 2022;70:2487-97. doi:10.1111/jgs.17894
- 47 Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011;64:395-400. doi:10.1016/j.jclinepi.2010.09.012
- 48 Johansson M, Guyatt G, Montori V. Guidelines should consider clinicians' time needed to treat. *BMJ* 2023;380:e072953. doi:10.1136/bmj-2022-072953
- 49 Chandler J, Cumpston M, Li T, Page MJ, Welch V. *Cochrane handbook for systematic reviews of interventions*. Wiley, 2019.
- 50 Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;14:135. doi:10.1186/1471-2288-14-135
- 51 Donner A, Koval JJ. The estimation of intraclass correlation in the analysis of family data. *Biometrics* 1980;36:19-25. doi:10.2307/2530491
- 52 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898
- 53 Minozzi S, Cinquini M, Gianola S, Gonzalez-Lorenzo M, Banzi R. The revised Cochrane risk of bias tool for randomized trials (RoB 2) showed low interrater reliability and challenges in its application. *J Clin Epidemiol* 2020;126:37-44. doi:10.1016/j.jclinepi.2020.06.015
- 54 Edinger JD, Means MK. Cognitive-behavioral therapy for primary insomnia. *Clin Psychol Rev* 2005;25:539-58. doi:10.1016/j.cpr.2005.04.003
- 55 Veroniki AA, Jackson D, Viechtbauer W, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* 2016;7:55-79. doi:10.1002/jrsm.1164
- 56 Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* 2020;368:l6890. doi:10.1136/bmj.l6890
- 57 Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH. Pooling health-related quality of life outcomes in meta-analysis-a tutorial and review of methods for enhancing interpretability. *Res Synth Methods* 2011;2:188-203. doi:10.1002/jrsm.46
- 58 Rücker G, Cates CJ, Schwarzer G. Methods for including information from multi-arm trials in pairwise meta-analysis. *Res Synth Methods* 2017;8:392-403. doi:10.1002/jrsm.1259
- 59 Mittlböck M, Heinzl H. A simulation study comparing properties of heterogeneity measures in meta-analyses. *Stat Med* 2006;25:4321-33. doi:10.1002/sim.2692
- 60 Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008;8:79. doi:10.1186/1471-2288-8-79
- 61 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34. doi:10.1136/bmj.315.7109.629
- 62 Brignardello-Petersen R, Tomlinson G, Florez I, et al. GRADE Working Group. Grading of recommendations assessment, development, and evaluation concept article 5: addressing intransitivity in a network meta-analysis. *J Clin Epidemiol* 2023;160:151-9. doi:10.1016/j.jclinepi.2023.06.010
- 63 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153-60. doi:10.1136/ebmental-2019-300117
- 64 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1-48. doi:10.18637/jss.v036.i03
- 65 Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ* 2020;192:E901-6. doi:10.1503/cmaj.200077
- 66 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD
- 67 Zeng L, Brignardello-Petersen R, Hultcrantz M, et al. GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol* 2021;137:163-75. doi:10.1016/j.jclinepi.2021.03.026

- 68 Santesso N, Glenton C, Dahm P, et al, GRADE Working Group. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020;119:126-35. doi:10.1016/j.jclinepi.2019.10.014
- 69 ClinicalTrials.gov. Facilitation of Zolpidem (≥ 10 mg) Discontinuation Through Use of Ramelteon in Subjects With Chronic Insomnia. 2010 <https://clinicaltrials.gov/study/NCT00492232>
- 70 Blum MR, Sallevelt BTGM, Spinewine A, et al. Optimizing therapy to prevent avoidable hospital admissions in multimorbid older adults (OPERAM): cluster randomised controlled trial. *BMJ* 2021;374:n1585. doi:10.1136/bmj.n1585
- 71 Edinger JD, Wamboldt F, Johnson RL, et al. Use of a blinded hypnotic tapering strategy to promote hypnotic discontinuation. *Sleep (Basel)* 2021;48:zsab270doi:10.1093/sleep/zsab270.
- 72 Edinger JD, Wamboldt FS, Johnson RL, et al. Adherence to behavioral recommendations of cognitive behavioral therapy for insomnia predicts medication use after a structured medication taper. *J Clin Sleep Med* 2023;19:1495-503. doi:10.5664/jcsn.10616
- 73 Fung CH, Alessi C, Martin J, et al. 0399 Masked Tapering and Augmented CBT Improves Benzodiazepine Receptor Agonist Discontinuation in a Randomized Trial. *Sleep (Basel)* 2024;47(Supplement_1):A171-2doi:10.1093/sleep/zsae067.0399.
- 74 Fung CH, Alessi C, Martin JL, et al. Masked Taper With Behavioral Intervention for Discontinuation of Benzodiazepine Receptor Agonists: A Randomized Clinical Trial. *JAMA Intern Med* 2024;184:1448-56. doi:10.1001/jamainternmed.2024.5020
- 75 Lhteenmki R, Puustinen J, Vahlberg T, et al. Melatonin for sedative withdrawal in older patients with primary insomnia: a randomized double-blind placebo-controlled trial. *Br J Clin Pharmacol* 2014;77:975-85. doi:10.1111/bcp.12294
- 76 Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the management of long-term hypnotic drug use: a pragmatic randomised controlled trial. *Br J Gen Pract* 2003;53:923-8.
- 77 Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the regulation of long-term hypnotic drug use. *Health Technol Assess* 2004;8:iii-iv, 1-68. doi:10.3310/hta8080
- 78 Oude Voshaar RC, Gorgels WJ, Mol AJ, et al. Long-term outcome of two forms of randomised benzodiazepine discontinuation. *Br J Psychiatry* 2006;188:188-9. doi:10.1192/bjp.bp.105.012039
- 79 Oude Voshaar RC, Krabbe PF, Gorgels WJ, et al. Tapering off benzodiazepines in long-term users: an economic evaluation. *Pharmacoeconomics* 2006;24:683-94. doi:10.2165/00019053-200624070-00007
- 80 Puustinen J, Lhteenmki R, Nurminen J, et al. Long-term persistence of withdrawal of temazepam, zopiclone, and zolpidem in older adults: a 3-year follow-up study. *BMC Geriatr* 2018;18:142. doi:10.1186/s12877-018-0829-9
- 81 Sibille FX, de Saint-Hubert M, Henrard S, et al. Benzodiazepine receptor agonists use and cessation among multimorbid older adults with polypharmacy: secondary analysis from the OPERAM trial. *Drugs Aging* 2023;40:551-61. doi:10.1007/s40266-023-01029-1
- 82 Vicens C, Bejarano F, Sempere E, et al. Comparative efficacy of two interventions to discontinue long-term benzodiazepine use: cluster randomised controlled trial in primary care. *Br J Psychiatry* 2014;204:471-9. doi:10.1192/bjp.bp.113.134650
- 83 Vicens C, Sempere E, Bejarano F, et al. Efficacy of two interventions on the discontinuation of benzodiazepines in long-term users: 36-month follow-up of a cluster randomised trial in primary care. *Br J Gen Pract* 2016;66:e85-91. doi:10.3399/bjgp16X683485
- 84 Voshaar RCO, Gorgels WJ, Mol AJ, et al. Tapering off long-term benzodiazepine use with or without group cognitive-behavioural therapy: three-condition, randomised controlled trial. *Br J Psychiatry* 2003;182:498-504. doi:10.1192/bjp.182.6.498
- 85 Zee P, Wang-Weigand S, Roth T. Facilitation of zolpidem discontinuation using ramelteon in subjects with chronic insomnia. AAAP 21st Annual Meeting Poster Abstracts 2011; Chicago, IL.
- 86 Habraken H, Soenen K, Blondeel L, et al. Gradual withdrawal from benzodiazepines in residents of homes for the elderly: experience and suggestions for future research. *Eur J Clin Pharmacol* 1997;51:355-8. doi:10.1007/s002280050213
- 87 Kosto A, Lev D, Reiss N, Meged-Book T, Press Y. Discontinuation of benzodiazepines and Z-drugs in hospitalised population at the age of 60 and above. An open-label randomized controlled trial. *Int J Geriatr Psychiatry* 2023;38:e6012. doi:10.1002/gps.6012
- 88 Nakao M, Takeuchi T, Nomura K, Teramoto T, Yano E. Clinical application of paroxetine for tapering benzodiazepine use in non-major-depressive outpatients visiting an internal medicine clinic. *Psychiatry Clin Neurosci* 2006;60:605-10. doi:10.1111/j.1440-1819.2006.01565.x
- 89 Cormack MA, Sweeney KG, Hughes-Jones H, Foot GA. Evaluation of an easy, cost-effective strategy for cutting benzodiazepine use in general practice. *Br J Gen Pract* 1994;44:5-8.
- 90 Gardner DM, Turner JP, Magalhaes S, Rajda M, Murphy AL. Patient Self-Guided Interventions to Reduce Sedative Use and Improve Sleep: The YAWNS NB Randomized Clinical Trial. *JAMA Psychiatry* 2024;81:1187-97. doi:10.1001/jamapsychiatry.2024.2731
- 91 Heather N, Bowie A, Ashton H, et al. Randomised controlled trial of two brief interventions against long-term benzodiazepine use: outcome of intervention. *Addict Res Theory* 2004;12:141-54doi:10.1080/1606635310001634528.
- 92 Kuntz JL, Kouch L, Christian D, Hu W, Peterson PL. Patient education and pharmacist consultation influence on nonbenzodiazepine sedative medication deprescribing success for older adults. *Perm J* 2019;23:18-161. doi:10.7812/TPP/18-161
- 93 Mak SS, Alessi CA, Kaufmann CN, et al. Pilot RCT testing A mailing about sleeping pills and cognitive behavioral therapy for insomnia: impact on benzodiazepines and Z-drugs. *Clin Gerontol* 2024;47:452-63. doi:10.1080/07317115.2022.2130849
- 94 Navy HJ, Weffald L, Delate T, Patel RJ, Dugan JP. Clinical pharmacist intervention to engage older adults in reducing use of alprazolam. *Consult Pharm* 2018;33:711-22. doi:10.4140/TCP.n.2018.711.
- 95 Salonoja M, Salminen M, Aarnio P, Vahlberg T, Kivel S-L. One-time counselling decreases the use of benzodiazepines and related drugs among community-dwelling older persons. *Age Ageing* 2010;39:313-9. doi:10.1093/ageing/afp255
- 96 Tannenbaum C, Martin P, Tamblyn R, Benedetti A, Ahmed S. Reduction of inappropriate benzodiazepine prescriptions among older adults through direct patient education: the EMPOWER cluster randomized trial. *JAMA Intern Med* 2014;174:890-8. doi:10.1001/jamainternmed.2014.949
- 97 Ten Wolde GB, Dijkstra A, van Empelen P, van den Hout W, Neven AK, Zitman F. Long-term effectiveness of computer-generated tailored patient education on benzodiazepines: a randomized controlled trial. *Addiction* 2008;103:662-70. doi:10.1111/j.1360-0443.2008.02141.x
- 98 Tinetti ME, Baker DI, McAvay G, et al. A multifactorial intervention to reduce the risk of falling among elderly people living in the community. *N Engl J Med* 1994;331:821-7. doi:10.1056/NEJM199409293311301
- 99 Ashworth N, Kain N, Wiebe D, Hernandez-Ceron N, Jess E, Mazurek K. Reducing prescribing of benzodiazepines in older adults: a comparison of four physician-focused interventions by a medical regulatory authority. *BMC Fam Pract* 2021;22:68. doi:10.1186/s12875-021-01415-x
- 100 Avorn J, Soumerai SB, Everitt DE, et al. A randomized trial of a program to reduce the use of psychoactive drugs in nursing homes. *N Engl J Med* 1992;327:168-73. doi:10.1056/NEJM199207163270306
- 101 de Burgh S, Mant A, Mattick RP, Donnelly N, Hall W, Bridges-Webb C. A controlled trial of educational visiting to improve benzodiazepine prescribing in general practice. *Aust J Public Health* 1995;19:142-8. doi:10.1111/j.1753-6405.1995.tb00364.x
- 102 Midlv P, Bondesson A, Eriksson T, Nerbrand C, Hglund P. Effects of educational outreach visits on prescribing of benzodiazepines and antipsychotic drugs to elderly patients in primary health care in southern Sweden. *Fam Pract* 2006;23:60-4. doi:10.1093/fampra/cmi105
- 103 Pimlott NJ, Hux JE, Wilson LM, Kahan M, Li C, Rosser WW. Educating physicians to reduce benzodiazepine use by elderly patients: a randomized controlled trial. *CMAJ* 2003;168:835-9.
- 104 Pit SW, Byles JE, Henry DA, Holt L, Hansen V, Bowman DA. A Quality Use of Medicines program for general practitioners and older people: a cluster randomised controlled trial. *Med J Aust* 2007;187:23-30. doi:10.5694/j.1326-5377.2007.tb01110.x
- 105 Vicens C, Leiva A, Bejarano F, et al. Evaluation of a multicomponent intervention consisting of education and feedback to reduce benzodiazepine prescriptions by general practitioners: The BENZORED hybrid type 1 cluster randomized controlled trial. *PLoS Med* 2022;19:e1003983. doi:10.1371/journal.pmed.1003983
- 106 Phelan EA, Williamson BD, Balderson BH, et al. Reducing Central Nervous System-Active Medications to Prevent Falls and Injuries Among Older Adults: A Cluster Randomized Clinical Trial. *JAMA Netw Open* 2024;7:e2424234. doi:10.1001/jamanetworkopen.2024.24234
- 107 Ayabe N, Okajima I, Nakajima S, et al. Effectiveness of cognitive behavioral therapy for pharmacotherapy-resistant chronic insomnia: a multi-center randomized controlled trial in Japan. *Sleep Med* 2018;50:105-12. doi:10.1016/j.sleep.2018.05.038
- 108 Baillargeon L, Landreville P, Verreault R, Beauchemin J-P, Grgoire J-P, Morin CM. Discontinuation of benzodiazepines among older insomniac adults treated with cognitive-behavioural therapy combined with gradual tapering: a randomized trial. *CMAJ* 2003;169:1015-20.
- 109 Belleville G, Guay C, Guay B, Morin CM. Hypnotic taper with or without self-help treatment of insomnia: a randomized clinical trial. *J Consult Clin Psychol* 2007;75:325-35. doi:10.1037/0022-006X.75.2.325

- 110 Bergdahl L, Broman J-E, Berman AH, Haglund K, von Knorring L, Markström A. Auricular acupuncture versus cognitive behavioural therapy in the discontinuation of hypnotic drug usage, and treatment effects on anxiety, depression and insomnia symptoms— a randomised controlled study. *Eur J Integr Med* 2017;16:15-21doi:10.1016/j.eujim.2017.10.002.
- 111 Coteur K, Henrard G, Schoenmakers B, et al. Blended care to discontinue benzodiazepine receptor agonists use in patients with chronic insomnia disorder: a pragmatic cluster randomized controlled trial in primary care. *Sleep* 2023;46:zsac278. doi:10.1093/sleep/zsac278
- 112 Lichstein KL, Nau SD, Wilson NM, et al. Psychological treatment of hypnotic-dependent insomnia in a primarily older adult sample. *Behav Res Ther* 2013;51:787-96. doi:10.1016/j.brat.2013.09.006
- 113 Morin CM, Bastien C, Guay B, Radouco-Thomas M, Leblanc J, Vallières A. Randomized clinical trial of supervised tapering and cognitive behavior therapy to facilitate benzodiazepine discontinuation in older adults with chronic insomnia. *Am J Psychiatry* 2004;161:332-42. doi:10.1176/appi.ajp.161.2.332
- 114 Sandlund C, Hetta J, Nilsson GH, Ekstedt M, Westman J. Improving insomnia in primary care patients: A randomized controlled trial of nurse-led group treatment. *Int J Nurs Stud* 2017;72:30-41. doi:10.1016/j.nurstu.2017.03.007
- 115 Meredith S, Feldman P, Frey D, et al. Improving medication use in newly admitted home healthcare patients: a randomized controlled trial. *J Am Geriatr Soc* 2002;50:1484-91. doi:10.1046/j.1532-5415.2002.50402.x
- 116 Strauven G, Annys P, Vandaal E, et al. Cluster-controlled trial of an intervention to improve prescribing in nursing homes study. *J Am Med Dir Assoc* 2019;20:1404-11. doi:10.1016/j.jamda.2019.06.006
- 117 Martin P, Tamblyn R, Benedetti A, Ahmed S, Tannenbaum C. Effect of a pharmacist-led educational intervention on inappropriate medication prescriptions in older adults: the D-PRESCRIBE randomized clinical trial. *JAMA* 2018;320:1889-98. doi:10.1001/jama.2018.16131
- 118 Roberts MS, Stokes JA, King MA, et al. Outcomes of a randomized controlled trial of a clinical pharmacy intervention in 52 nursing homes. *Br J Clin Pharmacol* 2001;51:257-65. doi:10.1046/j.1365-2125.2001.00347.x
- 119 van de Steeg-van Gompel CH, Wensing M, De Smet PA. Implementation of a discontinuation letter to reduce long-term benzodiazepine use—a cluster randomized trial. *Drug Alcohol Depend* 2009;99:105-14. doi:10.1016/j.drugalcdep.2008.07.009
- 120 Barros VV, Opaleye ES, Demarzo M, et al. Effects of mindfulness-based relapse prevention on the chronic use of hypnotics in treatment-seeking women with insomnia: a randomized controlled trial. *Int J Behav Med* 2022;29:266-77. doi:10.1007/s12529-021-10002-4
- 121 Cardinali DP, Gvozdenovich E, Kaplan MR, et al. A double blind-placebo controlled study on melatonin efficacy to reduce anxiolytic benzodiazepine use in the elderly. *Neuro Endocrinol Lett* 2002;23:55-60.
- 122 Gerra G, Zaimovic A, Giusti F, Moi G, Brewer C. Intravenous flumazenil versus oxazepam tapering in the treatment of benzodiazepine withdrawal: a randomized, placebo-controlled study. *Addict Biol* 2002;7:385-95. doi:10.1080/1355621021000005973
- 123 Murphy SM, Tyrer P. A double-blind comparison of the effects of gradual withdrawal of lorazepam, diazepam and bromazepam in benzodiazepine dependence. *Br J Psychiatry* 1991;158:511-6. doi:10.1192/bjp.158.4.511
- 124 Petrovic M, Pevernagie D, Mariman A, Van Maele G, Afschrift M. Fast withdrawal from benzodiazepines in geriatric inpatients: a randomised double-blind, placebo-controlled trial. *Eur J Clin Pharmacol* 2002;57:759-64. doi:10.1007/s00228-001-0387-4
- 125 Tyrer P, Ferguson B, Hallström C, et al. A controlled trial of dothiepin and placebo in treating benzodiazepine withdrawal symptoms. *Br J Psychiatry* 1996;168:457-61. doi:10.1192/bjp.168.4.457
- 126 Tyrer P, Rutherford D, Huggett T. Benzodiazepine withdrawal symptoms and propranolol. *Lancet* 1981;1:520-2. doi:10.1016/S0140-6736(81)92861-0
- 127 Melo TAR, Bezerra CO, Fernandes BD, Rotta I, Reis WCT, Aguiar PM. Pharmacists' contribution to benzodiazepine deprescribing in older outpatients: a systematic review and meta-analysis. *Int J Clin Pharm* 2023;45:1037-49. doi:10.1007/s11096-023-01637-2
- 128 Mpundu-Kaambwa C, Chen G, Dalziel K, Devlin N, Ratcliffe J, Quality Of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) project team. International guidelines for self-report and proxy completion of paediatric health-related quality of life measures: a protocol for a systematic review. *BMJ Open* 2021;11:e052049. doi:10.1136/bmjopen-2021-052049
- 129 National Institute for Health and Care Excellence. Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults. 2022. <https://www.nice.org.uk/guidance/ng215>
- 130 Murad MH, Mustafa R, Morgan R, Sultan S, Falck-Ytter Y, Dahm P. Rating the quality of evidence is by necessity a matter of judgment. *J Clin Epidemiol* 2016;74:237-8. doi:10.1016/j.jclinepi.2015.11.018

Web appendix: Supplementary materials