Title: The clinical effects of sleep restriction therapy for insomnia: A meta-analysis of randomised controlled trials.

Authors: Leonie F Maurer (PhD)[1], Jules Schneider (PhD)[1], Christopher B Miller (PhD)[1,2], Colin A Espie (PhD)[1,2], Simon D Kyle (PhD)[1*]

1. Sleep and Circadian Neuroscience Institute, Nuffield Department of Clinical Neurosciences, University of Oxford, UK
2. Big Health Inc., San Francisco, USA & London, UK

Address of correspondence:

*Simon D. Kyle, PhD

Sleep and Circadian Neuroscience Institute, Nuffield Department of Clinical Neurosciences, University of Oxford

Dunn School of Pathology, South Parks Road,

Oxford, OX1 3RE

(e) simon.kyle@ndcn.ox.ac.uk

(t) +441865 618675

Word count: 4667

Disclosure: SDK and CAE are supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC) based at Oxford University Hospitals NHS Trust and the University of Oxford. This work was also supported by funding from NIHR-HTA (16/84/01; CI SDK). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. CAE is co-founder of and shareholder in Big Health Ltd, a company that specialises in the digital delivery of cognitive behavioural therapy for sleep improvement (the Sleepio programme). This study is in no way connected to Big Health Ltd or Sleepio. CBM is employed by Big Health Inc. and is salaried by the company.

Acknowledgements. We want to thank all authors who provided additional information to facilitate this meta-analysis.

Summary

Sleep restriction therapy (SRT) is an established treatment for insomnia that has been used in clinical practice for over 30 years. It is commonly delivered as part of multicomponent cognitive-behavioural therapy (CBT-I) but has also been linked to beneficial effects as a standalone intervention. In order to quantify the efficacy of SRT we performed a comprehensive meta-analysis of randomised controlled trials (RCTs) comparing SRT to minimally active or non-active control groups. Primary outcomes were self-reported insomnia severity and sleep diary metrics at post-treatment. Weighted effect sizes were calculated with Hedges' *g* and risk of bias was assessed by two independent raters with the Cochrane tool.

Our search yielded eight RCTs meeting inclusion/exclusion criteria. Random effects models revealed large treatment effects in favour of SRT versus control for insomnia severity measured with the insomnia severity index (*g*=-0.93; 95%CI=-1.15, -0.71), sleep efficiency (*g*=0.91; 95%CI=0.52, 1.31), sleep onset latency (*g*=-0.62; 95%CI=-0.84, -0.40), and wake-time after sleep onset (*g*=-0.83; 95%CI=-1.11, -0.55). No effects were found for total sleep time (*g*=0.02; 95%CI=-0.29, 0.34). Results should be interpreted in the context of the small number of comparisons (≤6 per outcome), high risk of bias (6 out of 8 studies met criteria for high risk), and heterogeneity in study design and SRT administration. Only a small number of studies provided outcomes at follow-up (n≤3), hindering assessment of long-term effects.

Sleep restriction therapy effectively improves insomnia severity and sleep continuity in the short term; more studies are needed to assess if effects are sustained at long-term follow-up (>3m). Post-treatment effect sizes appear as large as multicomponent CBT-I. To reduce risk of bias, future studies should consider testing the effects of SRT against control groups that are matched for ***non-specific*** treatment effects. Large-scale pragmatic trials are also needed to test if SRT is effective in clinical practice and to quantify effects on daytime functioning and quality of life.

Keywords: Meta-analysis, Insomnia, Sleep restriction therapy, Cognitive behavioural therapy, Sleep, Psychological intervention

Glossary of terms:

AASM	American Academy of Sleep Medicine
CAU	Care as usual
CBT-I	Cognitive behavioural therapy for insomnia
CI	Confidence interval
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
FU	Follow up
ISI	Insomnia severity index
m	month(s)
N_c	Number of comparisons
N_p	Number of participants
PT	Post treatment
RCT	Randomised controlled trial
RoB	Risk of Bias
SE	sleep efficiency
SH	Sleep hygiene
SOL	Sleep onset latency
SRT	Sleep restriction therapy
StdD	Standard deviation
StdE	Standard error
TIB	Time in bed
TST	Total sleep time
w	week(s)
WASO	Wake after sleep onset
WLC	Waitlist control

Introduction

While multicomponent cognitive-behavioural therapy for insomnia (CBT-I) has been assessed in a large number of RCTs, less is known about the efficacy of its constituent components [1, 2]. Testing the efficacy of specific therapies for insomnia is key to isolating and appraising mechanisms that drive clinical change and sets the stage for considering whether single-component interventions could offer a cost- and time efficient alternative. Sleep restriction therapy (SRT), a behavioural component of CBT-I, is often claimed to be the most effective standalone intervention in the treatment of insomnia [3]. Through systematic restriction of time spent in bed, SRT aims to initially match time in bed with reported sleep time in order to reduce wakefulness during the night and consolidate sleep [4]. The restricted sleep opportunity typically results in mild sleep deprivation during early implementation, which has been hypothesised to strengthen sleep drive and reduce arousal before and during sleep. Establishing consistent bed- and rise times has also been postulated to regularise exposure to zeitgebers, which in turn may stabilise the circadian rhythm, align it with the sleep homeostat, and aid both the predictability and consolidation of sleep [3, 4].

The effectiveness of single-component SRT has been assessed through multiple trial designs, including uncontrolled single-arm studies [4-10], comparisons with treatment as usual or minimally active control groups [11-17], and comparisons with active comparators (e.g. single component and full CBT-I [18-21]). The evidence for SRT was first systematically reviewed in 2014 [22]. From nine identified SRT studies, only four were randomised-controlled trials (RCTs) [11, 13, 15, 18]. While formal meta-analysis was not performed, weighted effect sizes indicated medium-to-large treatment effects for sleep-diary derived sleep onset latency (SOL; n=4), wake-time after sleep onset (WASO; n=3), and sleep efficiency (SE; n=3). Global insomnia measures, however, were not assessed at that time.

Clinical guidelines regarding SRT efficacy and application are mixed. The 2017 European guideline for the treatment of insomnia considers SRT to have 'good efficacy' [23]. Practice parameters from the American Academy of Sleep Medicine (AASM; 2006) recommended SRT as a 'guideline', not as a 'standard', intervention [24]. However, this suggestion was based on only two RCTs published at the time. In 2016 the American College of Physicians regarded the evidence for SRT as insufficient [25]. Since the last systematic review, several SRT studies have been published, and in particular have reported on global measures of insomnia severity. With more data available, and in consideration of the need to establish the efficacy of SRT, we set out to undertake the first formal meta-analysis. We sought to establish the impact of SRT on insomnia severity and self-reported sleep, and to perform a risk of bias assessment for each trial.

During the write-up of this review the AASM published an updated systematic evaluation of SRT as part of a GRADE process that included all currently available behavioural and psychological treatments for insomnia [2]. Based on evidence from 6 RCTs, a 'conditional recommendation' was given for the clinical use of SRT [26]. Overall, the evidence for SRT was considered low quality as only 4 RCTs [11, 19, 21, 27] provided data suitable for meta-analysis on *critical outcomes*. We note, however, that one of the four studies included in this meta-analysis did not actually implement SRT - but sleep compression instead [27] - and one did not meet strict criteria for an RCT because it was a partially randomized-preference trial [21]. Our view is that a dedicated appraisal of the entire RCT literature on SRT is needed to 1) quantitatively estimate short and long-term efficacy; 2) appraise risk of bias; and 3) make recommendations for the field in terms of next steps.

Methods

Search strategy

This meta-analysis was conducted following PRISMA guidelines and pre-registered on PROSPERO (ID:CRD42020185750)[28]. We searched databases EMBASE, PsycINFO, CINAHL, AMED, and PubMed for peer-reviewed research articles published in English between 1987 (when the first SRT study was published [4]) and July 2020. We also searched relevant sleep research journals such as Journal of Sleep Research, Sleep, Sleep Medicine and Behavioural Sleep Medicine. We included the following search terms: “insomnia” or “chronic insomnia” or “sleeplessness” or “sleep disorder*” or “sleep initiation” or “sleep maintenance” or “poor sleep” or “sleep problem” or “sleep disturbance" and “sleep restriction”, "time in bed restriction", "bedtime restriction", “sleep compression”, “behavio?ral treatment”, “behavio?ral intervention”, “behav* therap*” or “behav* modification”. The first author screened all titles and abstracts to identify studies that administered SRT as a single component therapy, within a randomized-controlled trial, recruiting adults with insomnia, and reporting on global insomnia measures or self-reported sleep outcomes. Full-text versions were obtained when abstracts indicated that inclusion/exclusion criteria were met regarding study population, design and intervention (at this stage full texts of studies investigating any type of behavioral therapy for insomnia were extracted unless its components were specified in the abstract and did not meet inclusion/exclusion criteria). In uncertain cases, inclusion/exclusion was discussed between the first and last authors (LFM, SDK).

Study selection

The following inclusion criteria were applied for study selection: 1) study participants were adults (≥ 18 years) and met criteria for insomnia disorder or insomnia symptoms, as verified by diagnostic criteria such as DSM-5 [29] or self-report questionnaires; 2) sleep restriction therapy was delivered as a standalone intervention (or in combination with sleep hygiene or

other minimally active strategies), involving the systematic restriction of time in bed for the treatment of insomnia, whereby an initial, curtailed sleep window was prescribed with the aim of extending it over subsequent weeks based on pre-specified sleep efficiency criteria (the intervention could be delivered in any format e.g., face to face, on the phone or online); 3) the control group was judged to be minimally active (e.g. sleep hygiene) or non-active (e.g. wait-list control), or had no prior evidence for improving sleep or insomnia (e.g. time in bed regularisation); 4) the study design was a randomised controlled trial; and 5) the study reported on global insomnia measures (e.g. Insomnia severity index) or sleep-diary derived sleep outcomes. Additionally, we applied the following exclusion criteria: 1) the main focus of the study was on a population that met criteria for any other co-morbid psychiatric, medical or sleep disorder besides insomnia; and 2) non-randomised controlled trials, cohort and case-control studies.

Data extraction

Details of the identified studies were recorded by the first and second author using a standardised data-extraction form. The form was constructed to extract the following study characteristics: 1) citation of the publication; 2) geographic location; 3) total number of participants that were randomised and the number that provided outcome data, as well as the number of female participants; 4) mean age (and standard deviation when applicable); 5) method of recruitment; 6) diagnostic criteria for insomnia; 7) medication status; 8) comorbidity; 9) inclusion and exclusion criteria for participation; 10) components, format, frequency and duration of the intervention and who delivered it; 11) type, frequency and duration of the control condition and who delivered it; 12) subjective sleep measurements used; 13) the duration of the sleep diary if applicable; and 14) the format and timing of post-treatment and follow-up.

Findings regarding the effect of SRT on self-reported sleep measures and global insomnia measures at post-treatment were extracted and coded independently by the first and second author. These included: insomnia severity (Insomnia severity index [ISI]); sleep-diary defined sleep onset latency (SOL), wake-time after sleep onset (WASO), sleep efficiency (SE), and total sleep time (TST). As a secondary aim, we extracted data at follow-up time points, and for sleep-diary derived time in bed (TIB), seep-diary derived sleep quality ratings, and attrition (defined as number of participants discontinuing with the intervention, and number of participants not providing outcomes at post-treatment).

Risk of bias assessment

The Cochrane Collaboration's tool for assessing risk of bias (RoB) [30] was used to assess risk of bias for each study. The first and second author independently appraised all papers using the RoB tool, assessing five domains of bias: randomisation process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. Disagreements between reviewers were resolved by involving a third author (SDK). The integrated algorithm of the Cochrane Collaboration's tool was then applied to judge the risk of bias for each domain and across all domains (overall bias) for each study, whereby a study is judged to be at low risk of bias if a low risk of bias was given across all domains. Similarly, a study is judged to raise 'some concerns' if at least one domain was evaluated to raise 'some concerns', and a high risk of bias is given if a study was judged to be at high risk of bias in at least one domain, or if it was judged to have some concerns for multiple domains.

Strategy for data synthesis

Only outcomes provided by $n \geq 5$ studies were considered sufficient for interpretation of random-effects meta-analysis [31]. Effect sizes and 95% confidence intervals (CI) were

calculated for between-group comparisons at post-treatment. Effect size (Hedges' adjusted g [32]) was calculated with RevMan 5.4 (Cochrane Collaboration) [33] as the quotient of the difference between the mean of the SRT group against the mean of the control group, divided by the pooled weighted standard deviation. In this way, Hedges'g adjusts for small sample bias. Effect sizes from 0 to 0.32 can be interpreted as small, 0.33 to 0.55 as moderate, and 0.56 to 1.2 as large [34]. In the absence of standard deviations (StdD), standard errors (StdE) were transformed with StdD=StdE*SQRT(n), whereby n equals the sample size of the respective group. Random effects modelling was applied for the main analysis and heterogeneity of effects size was assessed with I^2 (where I^2=0 represents 0% heterogeneity, while greater values indicate higher observed heterogeneity; 25%=low, 50%=moderate, 75%=high). We calculated 95% confidence intervals (CI) around I^2 using the formula by Borenstein and colleagues [35].To aid interpretation, raw mean differences between groups and confidence intervals are also presented.

Results

Our primary database search yielded 2079 records. Additional search of sleep research journals identified a further 15 records. After removal of duplicates, 1057 records were screened, and 30 full texts extracted. Upon examination of full texts, we found that SRT instructions were often poorly described, making it impossible to judge whether the curtailed sleep window was adjusted based on pre-specified sleep efficiency criteria. We therefore decided to include studies as long as the authors described the intervention as sleep restriction therapy or provided information on therapeutic instructions that allowed us to conclude that SRT was implemented. Of the 30 full texts that were extracted, 21 were excluded because they did not meet inclusion/exclusion criteria. One study met all criteria but did not report on any of the relevant outcomes [13], leaving eight records for qualitative and quantitative synthesis (see Figure 1).

One study did not strictly meet our inclusion criteria for participant characteristics (age ≥18) because participants were permitted to be as young as 16 years of age [14]. Nonetheless the study was included because baseline characteristics clearly indicated that recruited participants were from the adult population (mean age SRT=55.4±12.7; mean age Control=51.8±13.4). One study [16] included two SRT arms that differed in delivery format (face-to-face and chat-based treatment). Since the majority of included studies delivered SRT face-to-face (see below), we extracted data for the face-to-face group only. This study included an 'imagination exercise' which was delivered in combination with SRT. Upon discussion between authors, this component was judged to be a 'minimally active treatment strategy'. A similar pragmatic decision was made for Taylor and colleagues, who combined SRT with advice on tapering and discontinuing hypnotic use (which was only triggered when sleep efficiency was ≥90%).

[INSERT FIGURE 1]

Study characteristics

Of the eight studies included in the analysis, four were conducted in the USA [11, 15, 18, 19], three in Europe [16, 17, 20], and one in New Zealand [14]. Studies were published between the years 2000 and 2020 and included a total of 533 participants (mean age range SRT=39.30 to 68.00; age range control=40.93 to 69.50) who were randomised to either SRT (n=277) or control (n=256). Percentage of female participants ranged from 41.7% to 100%. Half of the studies (n=4) recruited from the community and from primary care [15, 16, 19, 20], some from the community only (n=3) [11, 17, 18], and one from primary care only [14]. Diagnostic criteria for insomnia varied widely between studies and are summarised in Table 1. While the majority of studies (n=5) excluded sleep medication, three studies permitted them; whereby medication had to be maintained regularly on the same dosage [16], was recorded with no further instructions [20], or advice on its tapering and discontinuation was part of the intervention [15].

Length of SRT intervention varied across studies from two to eight weeks, and time point of post-treatment assessments varied between two weeks and three months (see Table 1). Time point of controlled follow-up assessments varied from three to six months. Out of all eight studies, only one implemented SRT as a standalone intervention [19], three in combination with sleep hygiene [11, 14, 17], one in combination with psychoeducation (including sleep hygiene) [20], one in combination with sleep hygiene and sleep education [18], one in combination with sleep hygiene and imagination exercises [16], and one in combination with sleep hygiene and advice on hypnotic tapering and discontinuation (once a sleep efficiency criterion was reached and for those who were regular hypnotic users) [15]. Sleep restriction therapy was typically delivered face-to-face [11, 14-17, 19], in combination with follow-up phone calls [17-19], in group settings [18], or via an internet platform [20]. Therapist background varied and included nurses that specialised in sleep medicine/mental health [15, 18, 19], a single GP researcher [14], graduate psychologists/psychology students [16, 20], and a doctoral researcher trained in CBT-I [17].

[INSERT TABLE 1]

Risk of bias

Out of all eight studies, none was classified with a *low* overall risk of bias in regard to the intention-to-treat effect. This was mainly due to high risk in the category ‘measurement of the outcome’. Ratings in this category were high because the outcome of interest was self-reported and, given the nature of behavioural interventions, participants were aware of the group they had been assigned to. However, in two trials, positioning of the study and matching for non-specific treatment effects (e.g. number of treatment sessions) helped to mask whether the allocated group was experimental or control [11, 17]. While six studies [14-16, 18-20] were rated with a high overall risk of bias, two [11, 17] were given an overall risk of bias of ‘some

concerns'. The risk of bias for each study and for each category of the Cochrane tool is summarised in Figure 2.

[INSERT FIGURE 2]

Meta-analysis

Sufficient outcome data were available for ISI, SE, SOL, WASO, and TST. Due to lack of follow-up data (n≤3 studies provided follow-up data for SRT and control) only results for post-treatment assessments are reported here. Similarly, n≤3 studies provided post-treatment outcomes for TIB and sleep quality. Analysis for follow-up data, TIB, and sleep quality can be found in the supplement and in Table S2. Results for standardised mean differences across sleep-diary derived sleep outcomes and ISI at post-treatment are summarised in Table 2 and illustrated in Figure 3/4. Average length of sleep diary at post-treatment was 1.83 weeks for SE, SOL and TST, and 2.00 weeks for WASO.

Insomnia severity index. Five studies [14, 17-20] provided post-treatment outcome data for the ISI (see Figure 3). Random effect models revealed large effects of SRT on the ISI when compared to control (N_c=5; *g*=-0.93 [95% CI= -1.15 to -0.71], *p*<.001). Heterogeneity between studies was low (I^2=9%). The ISI score was, on average, 4.20 points (95% CI=-5.25, -3.15) lower in the SRT group.

[INSERT FIGURE 3]

Sleep efficiency. Six studies [11, 15-19] reported sleep efficiency. Similar to the ISI, analysis showed large effect sizes in favour of SRT (N_c=6; *g*=0.91 [95% CI= 0.52 to 1.31], *p*<.001), yet with moderate heterogeneity between studies (I^2=65%, *p*=0.01).

Sleep onset latency. Post-treatment outcomes for SOL were provided by 6 studies [11, 15-19]. Compared to control, SRT had large effects on SOL (N_c=6; *g*=-0.62 [95% CI=-0.84 to -0.40], *p*<.001). Heterogeneity between studies was low (I^2=0%).

Wake-time after sleep onset. Five studies [11, 15, 17-19] provided post-treatment outcome data for WASO. Again, random effect models revealed large effects for SRT when compared to control (N_c=5; *g*=-0.83 [95% CI= -1.11 to -0.55], *p*<.001). Heterogeneity between studies was low (I^2=24%).

[INSERT FIGURE 4]

Total sleep time. Six studies [11, 15-19] reported on TST. Analysis showed no between-group effects (N_c=6; *g*=0.02 [95% CI= -0.29 to 0.34], *p*=.88). Heterogeneity between studies was moderate (I^2=50%).

[INSERT TABLE 2]

Attrition and acceptability. Reporting on attrition was often limited to the study flow chart and no further details were given on the exact nature of withdrawals or lost to follow-up. The reporting did not allow to retrieve information on the number of participants who completed each of the available treatment session, the number of participants who discontinued the intervention, or the number of participants who completed the full intervention but did not provide outcome data. Consequently, we only report overall completion rate for outcomes. Out of 278 participants who were randomised to SRT, 255 completed post-treatment assessments resulting in a low attrition rate of 8.27%. Across all control groups, 234 out of 256 participants completed post-treatment assessments, revealing a comparable attrition rate of 8.59%. To assess potential harms of SRT, we conducted further exploratory searches for adverse events and side effects. Out of all eight studies, only two reported on adverse events and measured side effects in a systematic manner [14, 17]. Both studies reported no serious adverse events in

the SRT group, and one study reported between-group differences in minor adverse events, with the SRT group reporting more negative effects of the intervention during week 2 compared to the control arm (time in bed regularisation) [17].

Discussion

We found evidence of large treatment effects for SRT versus control on insomnia severity and all self-reported sleep continuity variables at post-treatment (*g* range: 0.62 to 0.93). The most robust effect was found for the ISI (*g*=0.93). Together, these results indicate that SRT reliably improves sleep and insomnia symptoms. Indeed, effect sizes are comparable in magnitude to those reported for full CBT-I (e.g. ISI *g*=0.82) [1, 36] but conclusions are tentative at this stage given the difference in number of trials (e.g. ISI : SRT n=6 vs full CBT-I n=34 [1]) and the absence of adequately powered non-inferiority trials (although we note that head-to-head studies have typically failed to find clear differences [19, 20]). Consistent with meta-analyses of multicomponent CBT-I [36], no effect was found for TST (g=0.02), which ranged from *g*=-0.38 [11] to *g*=0.57 [18]. When interpreting variability in effect sizes for TST it should be kept in mind that studies differed in relation to treatment protocol [e.g. strict (TIB=TST) vs adjusted sleep opportunity restriction (TIB=TST+30min)] and time point of the post-treatment outcome (two to 12 weeks); factors that may influence both the direction and magnitude of group differences for TST [3].

Further analysis of attrition data revealed that outcome completion rates were comparable and low for SRT (8.27%) and control groups (8.59%). However, due to lack of specific reporting across studies, it was not possible to comment on treatment engagement and session completion rates. In line with CONSORT guidelines [37], future SRT studies should specify the number of participants who completed each treatment session and the total treatment protocol; and the reasons given for treatment discontinuation. A related point is that only two studies reported

on adverse events and side effects of SRT, outcomes that are commonly found to be neglected in CBT-I research [38] but important for adequate trial reporting [37]. This information will enable appraisal of SRT acceptability, which is known to be challenging for patients to implement [5, 39].

Other secondary outcomes, such as TIB and sleep quality, as well as sleep continuity and insomnia severity at follow-up, were considered exploratory due to the low number of studies ($n \leq 3$) and are addressed in the supplement only. There is clearly a need to perform studies with long-term follow-up of outcomes. A recently published meta-analysis on long-term treatment effects of CBT-I showed that although significant effects were observed for insomnia severity, sleep efficiency, and SOL at 3, 6, and 12-months follow-up, effect sizes appeared to decline over time [40]. Since SRT is part of CBT-I, similar declines might be expected. Only future controlled studies with long-term assessments will help determine whether initial SRT gains are sustained.

While not considered a clinically relevant sleep outcome, it is surprising that only few studies ($n=3$) reported on sleep-diary derived TIB, given the focus of SRT on restricted sleep opportunity. Previous research has shown that the restriction of TIB is central to the efficacy of SRT [17], and that therapeutic instructions on TIB reduction and titration varies widely across studies [41]. Consequently, adequate reporting of TIB may help interpret variability in treatment effects between studies (e.g. variability in TST at post-treatment). Additionally, reporting prescribed, self-reported, and objective TIB during treatment (where collected) would enable quantification of adherence to SRT and facilitate synthesis across studies. Adherence to SRT (and stimulus control) has previously been identified as a predictor of sleep outcomes [39].

Our results appear to differ from the recent AASM appraisal in relation to magnitude of between-group effects on sleep outcomes, with larger effect sizes reported here. Mean differences between SRT and control are about twice the size for SOL (11.44min vs 6.42min) and WASO (23.25min vs 11.67min) when compared to the results from the AASM task force. While we identified eight studies that met inclusion criteria, only four were submitted to meta-analysis by the task force; with an overlap of just two studies between the two meta-analyses. We believe that our appraisal is strengthened by the following points: 1) we found and included more of the published evidence; 2) our analysis only comprised fully randomised-controlled trials; and 3) we only considered studies implementing SRT (rather than sleep compression; although we acknowledge that these treatments ***may*** share similar mechanisms). We therefore consider our review to be the most comprehensive synthesis of SRT efficacy to date.

Methodological considerations

Two key limitations of our meta-analysis relate to the number of included studies (just n=8) and overall low study quality. None of the eight studies were classified as having a low risk of bias and 6 out of 8 studies were classified as high risk of bias. The high risk of bias can mostly be explained by the nature of behavioural intervention and our choice of outcome; that is, participants were typically aware of the group they had been assigned to, and outcomes of interest were self-reported. This resulted in a high-risk score on the measurement of outcome, driving the overall risk of bias. To understand the active ingredients within SRT and to minimise bias associated with the nature of self-reported outcomes, future studies should consider choosing comparison groups that control for non-specific factors (e.g. therapist time, treatment expectations, attention, monitoring of sleep-wake pattern) [42]. This meta-analysis has shown that SRT is effective compared to minimally active or non-active control groups. In line with recommendations for psychotherapy trial design [43, 44], we believe the field is now

ready for fine-grained research studies that focus on the active ingredients and mechanisms of SRT.

Similar to meta-analyses of CBT-I [1, 36], included studies differed in terms of diagnostic criteria for insomnia, method of recruitment, medication status, nature of control conditions, delivery methods, therapist expertise, duration of treatment, number of treatment sessions, time point of outcome measures, and treatment instructions given alongside SRT. Interestingly, variability between studies did not appear to have a substantial effect as indicated by small heterogeneity in effect sizes for several outcomes, including the ISI. However, large 95% confidence intervals for the I^2 demonstrates uncertainty around the heterogeneity estimates and should therefore be interpreted with caution. Robust treatment effects with little heterogeneity across a variety of study design features *may* indicate the potency of SRT; that is, the greatest contribution to the observed treatment effect may be participant implementation of SRT instructions rather than other factors, such as therapist type or format. Another limitation is that our review included studies that combined SRT with sleep hygiene or other minimally active strategies. Indeed, our results showed that all but one of the included studies combined SRT with additional advice (mainly sleep hygiene). Consequently, this review could not determine the effectiveness of the 'pure' SRT intervention. However, it should also be noted that there is no 'standard' SRT protocol implemented across studies [41]. Our review of included studies revealed variation in the reporting and implementation of SRT: out of eight studies, only three [14, 15, 17] described in detail how the initial sleep window was calculated *and* how the sleep window was subsequently titrated.

Studies included in this meta-analysis were mostly conducted within research settings and recruited thoroughly screened participants with primary insomnia. More research is needed in 'real-world' insomnia populations, applying limited exclusion criteria, in order to generate

evidence suitable for clinical practice. Such a study is underway [45]. If pragmatic trials show that SRT is effective, it might be an ideal candidate for delivery across primary care settings as an alternative to pharmacological interventions. Finally, while this meta-analysis focused on self-reported insomnia symptoms and sleep outcomes - revealing encouraging results - our reading of the published literature suggests that studies rarely report on outcomes beyond sleep, such as daytime functioning or quality of life. Given the importance of these outcome to patients [46], future studies should incorporate appropriate measures to assess the generalised effects of sleep improvement.

Practice points

- Compared to control groups, SRT is an effective intervention for insomnia as determined by post-treatment assessments.
- Large treatment effects were found for self-reported insomnia severity, sleep onset latency, wake-time after sleep onset and sleep efficiency, but not for total sleep time.
- There is limited outcome data at long-term follow-up.

Research agenda

Future studies on SRT for insomnia should aim to:

- Include long-term follow-up of clinical outcomes (beyond 3 months)
- Specify the number of participants who completed each treatment session and the full intervention protocol; and the reasons given for treatment discontinuation
- Carefully design intervention and control arms to match groups on *nonspecific* factors and reduce response bias.
- Conduct large-scale pragmatic trials to test SRT in clinical practice.
- Compare the clinical efficacy and cost-effectiveness of SRT vs. *full* CBT-I in adequately powered non-inferiority trials

References

[1] A. Van Straten, T. Van der Zweerde, A. Kleiboer, P. Cuijpers, C.M. Morin, J. Lancee, Cognitive and behavioral therapies in the treatment of insomnia: A meta-analysis, *Sleep Med Rev* 38 (2018) 3-16.

*[2] J.D. Edinger, J.T. Arnedt, S.M. Bertisch, C.E. Carney, J.J. Harrington, K.L. Lichstein, et al., Behavioral and psychological treatments for chronic insomnia disorder in adults: an American Academy of Sleep Medicine systematic review, meta-analysis and GRADE assessment, *J Clin Sleep Med* 17(2) (2021) 263-298.

[3] L.F. Maurer, C.A. Espie, S.D. Kyle, How does sleep restriction therapy for insomnia work? A systematic review of mechanistic evidence and the introduction of the Triple-R model, *Sleep medicine reviews* 42 (2018) 127-138.

[4] A.J. Spielman, P. Saskin, M.J. Thorpy, Treatment of chronic insomnia by restriction of time in bed, *Sleep* 10(1) (1987) 45-56.

[5] S.D. Kyle, K. Morgan, K. Spiegelhalder, C.A. Espie, No pain, no gain: an exploratory within-subjects mixed-methods evaluation of the patient experience of sleep restriction therapy (SRT) for insomnia, *Sleep medicine* 12(8) (2011) 735-747.

[6] S.D. Kyle, C.B. Miller, Z. Rogers, A.N. Siriwardena, K.M. MacMahon, C.A. Espie, Sleep Restriction Therapy for Insomnia is Associated with Reduced Objective Total Sleep Time, Increased Daytime Somnolence, and Objectively Impaired Vigilance: Implications for the Clinical Management of Insomnia Disorder, *Sleep* 37(2) (2014) 229-237.

[7] C.B. Miller, S.D. Kyle, N.S. Marshall, C.A. Espie, Ecological momentary assessment of daytime symptoms during sleep restriction therapy for insomnia, *Journal of sleep research* 22(3) (2013) 266-272.

[8] C.B. Miller, S.D. Kyle, C.J. Gordon, C.A. Espie, R.R. Grunstein, A.E. Mullins, et al., Physiological markers of arousal change with psychological treatment for insomnia: a preliminary investigation, *PloS one* 10(12) (2015).

[9] A. Vallières, T. Ceklic, C.H. Bastien, C.A. Espie, A preliminary evaluation of the physiological mechanisms of action for sleep restriction therapy, *Sleep disorders* 2013 (2013).

[10] H. Whittall, M. Pillion, M. Gradisar, Daytime sleepiness, driving performance, reaction time and inhibitory control during sleep restriction therapy for Chronic Insomnia Disorder, *Sleep Medicine* 45 (2018) 44-48.

*[11] L. Friedman, K. Benson, A. Noda, V. Zarcone, D.A. Wicks, K. O'Connell, et al., An actigraphic comparison of sleep restriction and sleep hygiene treatments for insomnia in older adults, *Journal of Geriatric Psychiatry and Neurology* 13(1) (2000) 17-27.

[12] L. Friedman, D.L. Bliwise, J.A. Yesavage, S.R. Salom, A preliminary study comparing sleep restriction and relaxation treatments for insomnia in older adults, *Journals of Gerontology* 46(1) (1991) P1-P8.

[13] A. Fernando III, B. Arroll, K. Falloon, A double-blind randomised controlled study of a brief intervention of bedtime restriction for adult patients with primary insomnia, *Journal of primary health care* 5(1) (2013) 5-10.

*[14] K. Falloon, C.R. Elley, A. Fernando 3rd, A.C. Lee, B. Arroll, A. Fernando, 3rd, Simplified sleep restriction for insomnia in general practice: a randomised controlled trial, *British Journal of General Practice* 65(637) (2015) e508-15.

*[15] D.J. Taylor, W. Schmidt-Nowara, C.A. Jessop, J. Ahearn, Sleep restriction therapy and hypnotic withdrawal versus sleep hygiene education in hypnotic using patients with insomnia, *Journal of Clinical Sleep Medicine* 6(2) (2010) 169-175.

*[16] A. Gieselmann, R. Pietrowsky, The effects of brief chat-based and face-to-face psychotherapy for insomnia: a randomized waiting list controlled trial, *Sleep Medicine* 61 (2019) 63-72.

*[17] L.F. Maurer, C.A. Espie, X. Omlin, M.J. Reid, R. Sharman, D. Gavriloff, et al., Isolating the role of time in bed restriction in the treatment of insomnia: a randomized, controlled, dismantling trial comparing sleep restriction therapy with time in bed regularization, *Sleep* 43(11) (2020).

*[18] D.R. Epstein, S. Sidani, R.R. Bootzin, M.J. Belyea, Dismantling multicomponent behavioral treatment for insomnia in older adults: a randomized controlled trial, *Sleep* 35(6) (2012) 797-805.

*[19] C.L. Drake, D.A. Kalmbach, J.T. Arnedt, P. Cheng, C.V. Tonnu, A. Cuamatzi-Castelan, et al., Treating chronic insomnia in postmenopausal women: a randomized clinical trial comparing cognitive-behavioral therapy for insomnia, sleep restriction therapy, and sleep hygiene education, *Sleep* 42(2) (2019) zsy217.

*[20] T. Krieger, A. Urech, S.B. Duss, L. Blattler, W. Schmitt, H. Gast, et al., A randomized controlled trial comparing guided internet-based multi-component treatment and internet-based guided sleep restriction treatment to care as usual in insomnia, *Sleep Medicine* 62 (2019) 43-52.

[21] S. Sidani, D.R. Epstein, M. Fox, L. Collins, Comparing the Effects of Single-and Multiple-Component Therapies for Insomnia on Sleep Outcomes, *Worldviews on Evidence-Based Nursing* 16(3) (2019) 195-203.

*[22] C.B. Miller, C.A. Espie, D.R. Epstein, L. Friedman, C.M. Morin, W.R. Pigeon, et al., The evidence base of sleep restriction therapy for treating insomnia disorder, *Sleep medicine reviews* 18(5) (2014) 415-424.

[23] D. Riemann, C. Baglioni, C. Bassetti, B. Bjorvatn, L. Dolenc Groselj, J.G. Ellis, et al., European guideline for the diagnosis and treatment of insomnia, *Journal of sleep research* 26(6) (2017) 675-700.

[24] T. Morgenthaler, M. Kramer, C. Alessi, L. Friedman, B. Boehlecke, T. Brown, et al., Practice parameters for the psychological and behavioral treatment of insomnia: An update. An American Academy of Sleep Medicine Report, *Sleep* 29(11) (2006) 1415-1419.

[25] A. Qaseem, D. Kansagara, M.A. Forciea, M. Cooke, T.D. Denberg, Management of chronic insomnia disorder in adults: a clinical practice guideline from the American College of Physicians, *Annals of internal medicine* 165(2) (2016) 125-133.

[26] J.D. Edinger, J.T. Arnedt, S.M. Bertisch, C.E. Carney, J.J. Harrington, K.L. Lichstein, et al., Behavioral and psychological treatments for chronic insomnia disorder in adults: an American Academy of Sleep Medicine clinical practice guideline, *J Clin Sleep Med* 17(2) (2021).

[27] B.W. Riedel, K.L. Lichstein, W.O. Dwyer, Sleep compression and sleep education for older insomniacs: Self-help versus therapist guidance, *Psychology and Aging* 10(1) (1995) 54.

[28] D. Moher, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, et al., Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, *Systematic reviews* 4(1) (2015) 1.

[29] A.P. Association, Diagnostic and statistical manual of mental disorders, *BMC Med* 17 (2013) 133-137.

[30] J.P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M.J. Page, et al., Cochrane handbook for systematic reviews of interventions, John Wiley & Sons2019.

[31] D. Jackson, R. Turner, Power analysis for random-effects meta-analysis, *Research Synthesis Methods* 8(3) (2017) 290-302.

[32] L.V. Hedges, Distribution theory for Glass's estimator of effect size and related estimators, *journal of Educational Statistics* 6(2) (1981) 107-128.

[33] R.M.R.C.p.V.C. The, T.C.C. Nordic Cochrane Centre, 2014.

[34] M.W. Lipsey, D.B. Wilson, The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis, *American psychologist* 48(12) (1993) 1181.

[35] M. Borenstein, L.V. Hedges, J.P. Higgins, H.R. Rothstein, Introduction to meta-analysis, John Wiley & Sons2009.

[36] L. Mitchell, L. Bisdounis, A. Ballesio, X. Omlin, S.D. Kyle, The impact of cognitive behavioural therapy for insomnia on objective sleep parameters: A meta-analysis and systematic review, *Sleep medicine reviews* (2019).

[37] D. Moher, K.F. Schulz, D.G. Altman, The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials, *The Lancet* 357(9263) (2001) 1191-1194.

[38] H.E. Condon, L.F. Maurer, S.D. Kyle, Reporting of Adverse Events in Cognitive Behavioural Therapy for Insomnia: A Systematic Examination of Randomised Controlled Trials, *Sleep Medicine Reviews* (2020) 101412.

[39] N. Vincent, S. Lewycky, H. Finnegan, Barriers to engagement in sleep restriction and stimulus control in chronic insomnia, *Journal of consulting and clinical psychology* 76(5) (2008) 820.

[40] T. van der Zweerde, L. Bisdounis, S.D. Kyle, J. Lancee, A. van Straten, Cognitive behavioral therapy for insomnia: A meta-analysis of long-term effects in controlled studies, *Sleep Medicine Reviews* 48 (2019) 101208.

[41] S.D. Kyle, M.R. Aquino, C.B. Miller, A.L. Henry, M.R. Crawford, C.A. Espie, et al., Towards standardisation and improved understanding of sleep restriction therapy for insomnia disorder: A systematic examination of CBT-I trial content, *Sleep Med Rev* 23 (2015) 83-8.

[42] D.J. Buysse, S. Ancoli-Israel, J.D. Edinger, K.L. Lichstein, C.M. Morin, Recommendations for a Standard Research Assessment of Insomnia, *Sleep* 29(9) (2006) 1155-1173.

[43] D.C. Mohr, B. Spring, K.E. Freedland, V. Beckner, P. Arean, S.D. Hollon, et al., The Selection and Design of Control Conditions for Randomized Controlled Trials of Psychological Interventions, *Psychotherapy and Psychosomatics* 78(5) (2009) 275-284.

[44] I. Michopoulos, T.A. Furukawa, H. Noma, S. Kishimoto, A. Onishi, E.G. Ostinelli, et al., Different control conditions can produce different effect estimates in psychotherapy trials for depression, *Journal of Clinical Epidemiology* 132 (2021) 59-70.

[45] S.D. Kyle, C. Madigan, N. Begum, L. Abel, S. Armstrong, P. Aveyard, et al., Primary care treatment of insomnia: study protocol for a pragmatic, multicentre, randomised controlled trial comparing nurse-delivered sleep restriction therapy to sleep hygiene (the HABIT trial), *BMJ Open* 10(3) (2020) e036248.

[46] S.D. Kyle, C.A. Espie, K. Morgan, "… Not Just a Minor Thing, It Is Something Major, Which Stops You From Functioning Daily": Quality of Life and Daytime Functioning in Insomnia, *Behavioral Sleep Medicine* 8(3) (2010) 123-140.

Figures and Tables

Figure legends:

Figure 1. Prisma flow diagram.

Figure 2. Risk of bias for each category by study. Green circles are indicating low risk of bias, yellow circle express 'some concerns' and red circles are representing a high risk of bias. D1 (Domain 1)=Randomisation process; D2 (Domain 2)=Deviations from intended interventions; D3 (Domain 3)=Missing outcome data; D4 (Domain 4)=Measurement of the outcome; D5 (Domain 5)=Selection of reported results.

Figure 3. Overview of standardised between-group effect sizes on the ISI at post-treatment. Standardised means (black diamonds) are polarity adjusted such that values on the left-hand side (>0) favour sleep restriction therapy while values on the right-hand side (<0) favour control. Whiskers represent 95% confidence intervals of the standardised mean. Green diamond and whiskers represent weighted effects across all studies. ISI=Insomnia severity index.

Figure 4. Overview of standardised between-group effect sizes on sleep-diary derived sleep outcomes at post-treatment. Top left: SE=Sleep efficiency, top right: SOL=Sleep onset latency, bottom left: WASO=Wake after sleep onset, bottom right: TST=Total sleep time. Standardised means (black diamonds) are polarity adjusted such that values on the left-hand side (>0) favour SRT and values on the right-hand side (<0) favouring control. Whiskers represent 95% confidence intervals of the standardised mean. Green diamonds and whiskers represent weighted effects across all studies.

Table 1. Baseline characteristics

Authors (year)	Geographic location	Number randomised per group	% females	Mean age± STDV	Method of recruitment	Diagnostic criteria for insomnia	Medication status	SRT rules**	Time point PT/FU
Drake et al. (2019) [19]	Michigan USA	SRT=52 SH=50	100.0%	SRT= 56.76±5.39 SH= 57.24±5.55	Recruited from primary care/sleep clinic/community (via newspaper advertisements)/ database of prior sleep studies	WASO≥1hr on ≥3nights/w & DSM-5 insomnia disorder with onset/exacerbation during peri-/postmenopausal period, verified per clinical interview & PSG-defined WASO≥45min	Medication affecting sleep was excluded, hormone therapy was accepted	TIB calculation: not specified Minimum TIB: not specified SE criteria: not specified	PT: 2w (SRT)/ 6w (SH) FU: 6m
Epstein et al. (2012) [18]	Arizona USA	SRT=44 WLC=50	SRT= 56.8% WLC= 64.0%	SRT= 68.00±8.25 WLC= 69.50±8.34	Media advertisements in newspapers, radio, and television (community)	SOL or WASO≥45 min≥3 nights/w (verified by 14 days of sleep diary) & insomnia duration≥6m & impaired daytime functioning resulting from insomnia	Medication affecting sleep was excluded (verified through urinalysis)	TIB calculation: TST Minimum TIB: 5hrs SE criteria: not specified	PT:6w
Falloon et al. (2015) [14]	Auckland New Zealand	SRT=46 SH=51	SRT= 85.0% SH= 71.0%	SRT= 55.40±12.70 SH= 51.80±13.40	Primary care patients from 14 GP practices	Primary insomnia lasting >6m & difficulty sleeping on ≥3nights/w as verified by sleep diary	Prescribed sleep medication & hypnotic medication use within 2w prior to baseline were excluded	TIB calculation: TST+50% of TIB awake Minimum TIB: 5hrs SE Criteria: if SE < 85%, then TIB=TST+30min; if excessively sleepy, then TIB=TIB+30min	PT:3m FU:6m

Friedman et al. (2000) [11]	California USA	SRT=16 SH=11	SRT= 56.25% SH= 81.81%	SRT= 65.10±8.60 SH= 61.90±7.10	Community advertisement	2w of sleep diary: SE<80%, SOL>30 min, or TST<6hr, or>30min of WASO ≥5nights/2w	Participants were required to be free of sleeping medication for ≥3w & refrained from medication use during treatment	TIB calculation: TST Minimum TIB: not specified SE criteria: n/a	PT:4w FU:3m
Gieselmann et al. (2019) [16]	Düsseldorf Germany	SRT=27 WLC=22	SRT= 48.0% WLC= 53.0%	SRT= 39.30±14.47 WLC= 42.74±11.73	Flyers/Posters on university campus, medical practices & pharmacies; study homepage and social media	Research criteria for primary insomnia & SOL > 30 min on ≥ 3 nights/1w for the last 6m	Sleep/psychotropic medications were permitted but patients were asked to maintain current pattern of usage	TIB calculation: TST Minimum TIB: not specified SE criteria: if SE>85%, then TIB=TIB+15min	PT:4w
Krieger et al. (2019) [20]	Bern Switzerland	SRT=41 CAU=21	SRT= 68.3% CAU= 81.0%	SRT= 46.59±17.52 CAU= 45.24±12.40	Newspaper advertisements, online postings, flyers, and physician referrals	Acute or chronic insomnia according to ICSD-3	Medication was measured but not excluded	TIB calculation: Not specified Minimum TIB: 6hrs SE criteria: if SE≥90%, then TIB=TIB+30min	PT:8w
Maurer et al. (2020) [17]	Oxford UK	SRT=27 TBR=29	SRT= 70.4% TBR= 69.0%	SRT= 40.63±9.13 TBR= 40.93±9.24	Community advertisement (social medial, posters, newspaper, radio)	DSM-5 chronic insomnia disorder verified by the Sleep condition indicator and interview	Medication was excluded	TIB calculation: TST Minimum TIB: 5hrs SE criteria: if SE≥90%, then TIB=TIB+15min	PT:4w FU:3m
Taylor et al. (2010) [15]	Texas USA	SRT=24 SH*=22	SRT= 41.7% SH= 68.2%	SRT= 56.75±15.29 SH= 50.36±14.10	Recruited from sleep medicine practice, primary care physician practices, and from the community via newspaper and radio advertising	Chronic insomnia (≥6m), defined as persistent problems initiating/maintaining sleep & SE<85% verified by sleep diary over 2w	Advice on hypnotic tapering/discontinuation was part of the intervention and therefore regular hypnotic use was inclusion criterion	TIB calculation: TST+10% Minimum TIB: 5hrs SE criteria: if SE≥90%, then TIB=TIB+15min & advice on hypnotic tapering	PT:8w

Notes. SRT=Sleep restriction therapy, SH=Sleep hygiene, WLC=Waitlist control, STDV=Standard deviation, CAU=Care as usual, w=week(s), m=month(s), TIB=Time in bed, TST=Total sleep time, SOL=Sleep onset latency, WASO=Wake after sleep onset, PSG=Polysomnography, PT=Post treatment, FU=Follow up. Follow-up time points are only provided for those studies that reported follow-up data on both groups: SRT and control. *Both groups received sleep hygiene advice pre-randomisation. Participants in the sleep hygiene group were offered SRT after their post-treatment assessment at week 8. ** SRT rules comprise of the initial time in bed calculation that determines the first sleep window, the minimum of time spent in bed at any given time during the intervention, and the sleep efficiency criteria that is used to titrate the sleep window (upwards and downwards) in the subsequent weeks of intervention. For simplicity, we only present the SE criteria for upwards titration.

Table 2. Effect sizes for outcomes at post-treatment between SRT and control group

Outcome	N_c	N_p	Mean difference (95% CI)	I^2 (95% CI)	Hedges' *g* (95% CI)	*p*
ISI	5	397	-4.20 [-5.25, -3.15]	9% (0-82)	-0.93 [-1.15, -0.71]	**<.001**
SE (%)	6	345	9.85 [7.29, 12.41]	65% (0-85)	0.91 [0.52, 1.31]	**<.001**
SOL (min)	6	345	-11.44 [-16.82, -6.05]	0% (0-74)	-0.62 [-0.84, -0.40]	**<.001**
WASO (min)	5	306	-23.25 [-32.47, -14.03]	24% (0-69)	-0.83 [-1.11, -0.55]	**<.001**
TST (min)	6	345	0.00 [-18.37, 18.37]	50% (0-80)	0.02 [-0.29, 0.34]	0.88

Notes. N_c= Number of comparisons, N_p=Number of participants, ISI=Insomnia severity index, SE=Sleep efficiency, SOL=Sleep onset latency, WASO=Wake after sleep onset, TST=Total sleep time. I^2 refers to heterogeneity between studies for standardised mean differences. *P* values in bold represent statistically significant group differences for standardised mean differences. Raw mean differences are presented to aid interpretation.