Family Medicine and Community Health

Impact of different types of physical exercise on sleep quality in older population with insomnia: a systematic review and network meta-analysis of randomised controlled trials

Pakwan Bahalayothin, Kittiphon Nagaviroj, Thunyarat Anothaisintawee

To cite: Bahalayothin P, Nagaviroj K, Anothaisintawee T. Impact of different types of physical exercise on sleep quality in older population with insomnia: a systematic review and network meta-analysis of randomised controlled trials. *Fam Med Com Health* 2025;**13**:e003056. doi:10.1136/fmch-2024-003056

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/fmch-2024-003056).

Received 04 July 2024
Accepted 02 January 2025

Family Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Correspondence to
Dr Kittiphon Nagaviroj;
kittiphon.nag@mahidol.ac.th

ABSTRACT

Objective To measure the impact of each type of exercise on sleep quality and identify the exercise that enhances sleep quality the most.

Study selection Eligible randomised controlled trials that compare physical exercise to routine activities, usual care, non-physical activity, or health education to measure the Pittsburgh Sleep Quality Index.

Data source Studies retrieved from Medline, Embase, CINAHL, Scopus, ClinicalTrial.gov and ThaiJo from the database's inception to October 2022.

Data extraction and synthesis Two reviewers independently identified studies, collected data and assessed bias. In the absence of heterogeneity, a fixed effect model was used for pairwise meta-analysis. Alternatively, a random effect model was used. A two-stage network meta-analysis used the surface under the cumulative ranking curve (SUCRA) to compare exercise efficacy.

Main outcome Global Pittsburgh Sleep Quality Index (GPSQI) and subdomain score.

Results This review comprised 2170 people from 25 trials. Direct meta-analysis revealed significant improvement in GPSQI with combined exercise (unstandardised mean difference (USMD) −2.35, 95% CI−3.13 to −1.57, p<0.001, I^2=69.13%). GPSQI decreased considerably with aerobic activity (USMD −4.36, 95% CI −7.86 to −0.86, p=0.01, I^2=97.83%). For the network meta-analysis, strengthening, aerobic and combination exercise significantly lowered GPSQI (USMD−5.75, −3.76 and −2.54, respectively). Strength training improved GPSQI scores most effectively (SUCRA 94.6%).

Conclusion Exercise that strengthens muscles, rather than aerobic or combination exercises, is the most effective way to enhance sleep quality.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ According to the previous study, exercise is simple to implement, less expensive, and particularly effective for treating insomnia in the elderly.

WHAT THIS STUDY ADDS

⇒ Combination exercise and aerobic exercise are effective in improving sleep quality to a clinically significant level. The outcome of this study indicates that strengthening exercise has the highest efficacy among others, followed by aerobic exercise and combination exercise.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The results will help the multidisciplinary team determine which type of exercise is the most effective intervention to improve sleep quality beyond clinical significance.

INTRODUCTION

Sleep quality declines with age. Older people have more significant insomnia than younger people. Between 30% and 48% of seniors complain of sleepiness, while 12–20% have insomnia problems.[1] Family doctors see about 5.5 million patients every year for insomnia.[2] Strong evidence links sleeplessness to depression, anxiety and other mental health disorders.[3 4] Several studies have linked sleeplessness to metabolic syndrome, hypertension and heart disease.[5–7] Cognitive decline and prostate cancer risk are related to insomnia.[8 9] Insomnia also increases the likelihood of job impairment, absenteeism and poor performance, which costs the healthcare system and society.[10]

Older individuals should start non-pharmacological treatment due to hypnotic toxicity from a longer half-life.[11] CBT-I (cognitive behavioural therapy for insomnia) is the first-line treatment, although it is seldom used in clinical settings owing to the time and lack of qualified therapists.[12 13] However, exercise is simple, inexpensive and beneficial for alleviating insomnia in older adults. Exercise also reduces illness comorbidity (metabolic syndrome, stroke, depression, anxiety) and builds muscle.[12] Most studies have shown that exercise treatment reduces insomnia symptoms.[14–17] The usefulness of various forms of exercise in increasing sleep quality is still unclear, making it difficult to choose the best one.

The subjective sleep result is vital from the patient's perspective.[18] The Pittsburgh Sleep Quality Index (PSQI) is among the most used subjective sleep quality tools due to its validity and reliability.[19–24] The questionnaire is simple and utilised without medical assistance.

This systematic review aims to evaluate and compare the impact of various types of physical exercise on enhancing sleep quality in the elderly population. The main outcome of interest is the PSQI, which is used to determine the most effective therapeutic exercise option.

REVIEW QUESTIONS

1. How do individual types of exercise impact the PSQI?
2. What type of exercise is the best therapeutic way to improve the PSQI?

METHODS

Study design

This study was a systematic review and network meta-analysis performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[25] and the PRISMA Network Meta-Analysis Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-Analyses of Healthcare Interventions.[26]

Study selection

Two authors designed the search terms and search strategies for each database, as presented in the online supplemental material. Relevant studies were searched from the Medline, Embase, CINAHL, Scopus, ClinicalTrial.gov and ThaiJo web databases, encompassing published and unpublished research. The search begins from the date of database inception to October 2022.

The eligible randomised controlled trials (RCTs) were restricted to the Thai or English language based on the inclusion and exclusion criteria listed below. Comprehensive searching was conducted with the aid of Medical Subject Heading (MeSH) in Medline and entrée in Embase, along with free text terms. Researchers further investigated current randomised controlled studies in the US National Library of Medicine (www.clinicaltrials.gov).

Eligible papers were assessed by two authors (PB, KN) based on their titles and abstracts separately. When the information was insufficient, the full paper was reviewed for further details. When the two authors did not reach the same decision, discussion was conducted through meetings. If disagreement was not resolved by the two authors, consensus was reached by consulting a third party (TA). Microsoft Word, Microsoft Excel and Endnote software were used to organise the search, inclusion and extraction of papers.

Inclusion criteria

- RCTs compared physical exercise with a control group (eg, routine daily activities, usual care, non-physical activity, and health education).
- Participants in the study were adults aged 60 years and older.
- Participants were diagnosed with insomnia according to DSM-V (Diagnostic and Statistical Manual of Mental Disorders V) criteria (present sleep dissatisfaction for at least 3 days per week for at least 3 months, impact on daily functioning, adequate opportunity to sleep, inability to explain by other sleep-wake disorder, mental disorder, medical condition, or medication)[27] or have a GPSQI > 0.
- The PSQI was used as a standardised scale to measure sleep quality outcomes.

Exclusion criteria

- Duplicate studies.
- Reviews, conference abstracts and letters.
- Studies with incomplete data, such as protocol studies, ongoing studies, and insufficiently reported results that could not be converted to USMDs.

Intervention

Any RCTs that compared physical exercise with a control group (no physical intervention, routine activity or sleep education) on the quality of sleep were included. The types of exercises featured in these papers were divided into five categories by the authors as follows:[28]

- Aerobic or endurance physical activity (A). This activity is characterised by its continuous use of large muscle groups and its rhythmic nature. This type of exercise depends on aerobic metabolism and ultimately produces ATP. Examples of aerobic exercise include cycling, dancing, hiking, jogging or long-distance running, swimming, gardening and brisk walking.[29]
- Strength exercise or resistance exercise (S). Exercises that increase muscle strength by making muscles work against a weight or force and using anaerobic metabolism,[30] such as lifting weights, arm curls, wall push-ups, and resistance machines or equipment.
- Balance exercise (B). Balance exercise refers to exercise in which an individual maintains their line of gravity within their base of support (BOS) or is described as having the ability to maintain equilibrium, for instance, balance walking, sideways walking, heel-to-toe walking, one-leg standing, and step-up.[31]
- Flexibility exercise (F). This is described as an activity that increases the movement of the joint through soft tissue around the joint to improve joint range of motion,[32] such as gymnastics and dance.[33]
- Combination exercise (A+S+B+F). Combination exercise is defined as a mixed type of aerobic, strengthening, balance and flexibility exercise, as previously described. In the event that any study does not declare the type of exercise or in such exercises as Tai Chi,

Qigong, Baduanjin yoga and Pilates, this type of exercise is defined as a combination exercise.

Patient and public involvement

Patients and the public were not involved in the production of this research.

Outcome

The primary outcome was the GPSQI, which is composed of 19 self-rated questions and five questions rated by a bed partner or roommate (if one is available), but only self-rated questions were included in the scoring. In scoring the PSQI, seven component scores were derived, each scored from 0 (no difficulty) to 3 (severe difficulty). All component scores were summed to produce a global score, which ranged from 0 to 21. Higher scores indicated worse sleep quality.[34]

Secondary outcomes were the seven domains of PSQI scores (subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbance, use of sleep medication, and daytime dysfunction).

Data management

All matched studies were extracted to the Endnote 20 reference software program. Duplicate or redundant studies were removed.

Data extraction

Two reviewers (PB, KN) independently extracted data from each included study, such as those on the characteristics of the study and participants, details on the interventions, and intended outcomes. The researchers attempted to contact the original authors through email to obtain any missing or essential information. Initially, any discrepancies were resolved through discussion, and if disagreement remained, a third party (TA) was consulted.

The data record form is displayed in online supplemental table S3.

Risk of bias assessment

The risk of bias was assessed independently by the two review authors using the revised Cochrane risk of bias tool for randomised trials (RoB 2.0).The RoB 2.0 for individually randomised trials assessed five main domains as follows:

1. Bias arising from the randomisation process
2. Bias due to deviations from intended interventions
3. Bias due to missing outcome data
4. Bias in the measurement of the outcome
5. Bias in the selection of reported results.

Each risk of bias domain was assessed as low, high or some concern. The research was determined to have a low risk of bias if all domains were assessed as low risk. However, if many categories are classified as some concerns, it may result in an overall high risk of bias.

In the event of any disagreement among reviewers, a third party was involved to establish consensus.

Statistical analysis

Direct meta-analysis

A pairwise meta-analysis was performed when more than two RCTs compared similar populations, interventions, comparators and outcomes. This study employed aggregate or summary data for the meta-analysis. The USMD of the PSQI was pooled between the intervention and control groups. The assessment of heterogeneity was conducted using the Cochrane Q test and the degree of heterogeneity. The criteria for heterogeneity included a p-value<0.1 from the Cochrane Q test or an I^2 value >25. In studies with considerable heterogeneity, random effects models pooled USMDs. Otherwise, USMD data were pooled using a fixed effect model.

Analysis of a subgroup or subset

Meta-regression analysis was used to investigate the source of heterogeneity in studies if presented. The meta-regression analysis considered each of the following covariates individually as contributing factors to the heterogeneity: region, setting, age, percentage of women, diagnostic criteria, exercise, time per session, frequency of exercise, duration of exercise, total exercise length, measurement method, and baseline Global Pittsburgh Sleep Quality Index (GPSQI)—total PSQI of the seven domains.

Network meta-analysis

A network meta-analysis was conducted to compare the efficacy of all available exercise interventions for insomnia. This network meta-analysis employed a two-stage methodology[35] to calculate the average differences between continuous outcomes for all treatments. The Surface Under the Cumulative Ranking Curve (SUCRA) was employed to rank exercise therapies based on the predicted probability of their efficacy. A clustered ranking plot was generated using the SUCRA values of the mean difference to assess and compare which therapy was most effective in terms of efficacy. When comparing treatments directly and indirectly, inconsistencies might occur between the two groups. A design-by-treatment technique was utilised to mitigate potential loops and design discrepancies.[36] If inconsistency was identified, the loop with an inconsistency factor greater than two was explored. The sensitivity analysis was performed by excluding studies with different characteristics.

The analyses were conducted using STATA version 18. A two-tailed p-value<0.05 was used to establish statistical significance, with the exception of the Q-test, where a p-value<0.10 was used.

Publication bias

Assessment of publication bias was carried out using a funnel plot and the Egger test. Publication bias was identified when the funnel plot exhibited asymmetry or when the p-value from the Egger test was less than 0.05.[37]

RESULTS

A total of 1643 studies underwent screening, of which 1543 studies were subsequently excluded by title, abstract and full text in some papers. Out of the 100 studies, four papers could not be retrieved. Of the remaining 96 studies, 71 studies were excluded for various reasons, including not being RCTs (n=4), showing only protocol RCTs (n=5), not reporting the results in English or Thai (n=1), having participants younger than 60 years old (n=50), having additional interventions in both the control and intervention groups (n=3), having additional interventions solely in the intervention group (n=1), and having additional interventions only in the control group (n=1). Four studies did not report PSQI scores as an outcome. One study reported only the percentage change in PSQI. Additionally, there was one duplicate study. The search and study selection process is illustrated in figure 1 using a flow diagram.

Risk of bias assessment

The risk of bias in the outcome was categorised according to the analysis approach, specifically the intention-to-treat group and the per-protocol group. Within the intention-to-treat group, 43.8% of studies had a low risk of bias, as

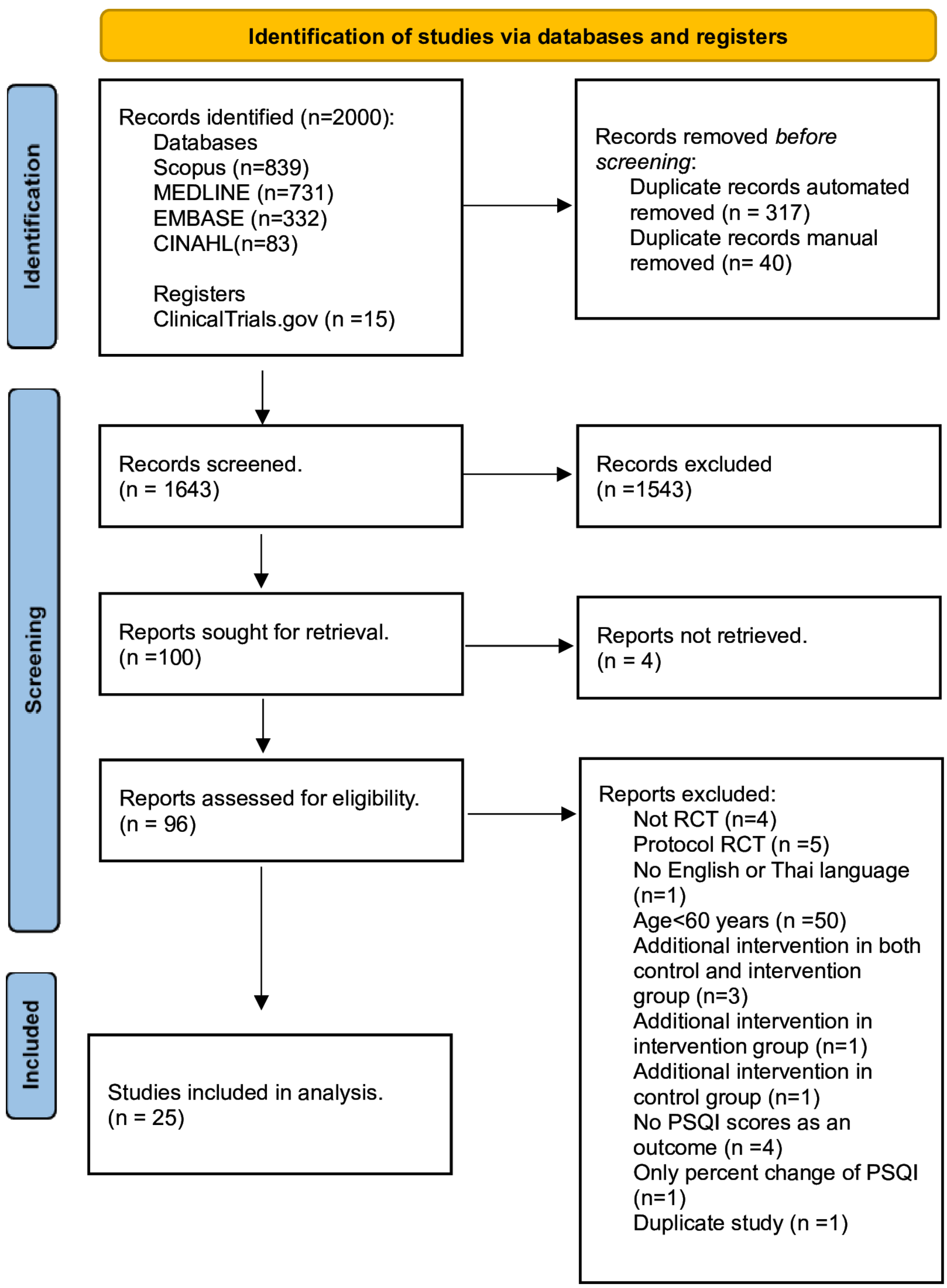

Figure 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for new systematic reviews, which included searches of databases, registers and other sources. PSQI, Pittsburgh Sleep Quality Index; RCT, randomised controlled trial.

shown in online supplemental figure S1. Subsequently, this was followed by studies with a high risk of bias at 31.3% and studies with some concern for bias at 25%. The main determinant of the high risk of bias was the measurement of outcomes by an unblinded assessor. Among the per-protocol group, 55.6% were categorised as low risk, whereas 22.2% were individually identified as having some concern and high risk (online supplemental figure S2). The primary contributory factors in concern and high-risk groups were unidentified randomisation techniques, concealment and measurement methods.

Study characteristics

This study incorporated a total of 25 studies undertaken between 1996 and 2021, encompassing 2170 participants. The most prevalent areas of study originated in Asia (56%), North America (16%), South America (16%) and European countries (12%). The majority of studies (80%) were conducted in community settings, whereas only 20% were carried out in nursing homes. The mean age of the participants was 70.38 (±4.56) years. The mean percentage of female participants was 71.85 (±21.86). The majority of patients were selected based on symptomatic criteria. Only one study had been carried out by recruiting participants based on the DSM-V criteria. Combination exercise, consisting of aerobic, strengthening, balancing or flexibility exercises, was the most prevalent form of exercise intervention, accounting for 48.15% of the total. Aerobic exercise alone accounted for 14.81%. The major level of exercise intensity was mild to moderate/moderate, accounting for 54.55% of the total reported studies. The average duration per session was approximately 53.69 (±11.79) minutes, and the number of sessions was three times per week (48%) and two times per week (35%). The average exercise duration per session was 150.65 (±70) minutes per week, and the exercise programme lasted for 14.4 weeks (±6.25). All participants in the study stated a compliance rate of 70% or above, as indicated by their self-reports (60%). The baseline GPSQI scores in the exercise and control groups were 8.37 (±2.38) and 8.27 (±2.16), respectively.

The characteristics of the included studies are outlined in online supplemental table S1.

Primary outcome

A total of 24 studies, with a sample size of 2045 participants, incorporated global PSQI scores as a measure of outcomes. Two treatment comparisons yielded a significant number of studies to conduct pairwise meta-analyses. The first comparison was aerobic exercise alone versus routine physical activity, with a total of four studies and 341 participants. The second comparison involved combination exercise versus routine physical activity, with a total of 13 studies and 1155 participants. A detailed summary of the global PSQI scores is presented in online supplemental table S2.

The direct meta-analysis demonstrated that the combination exercise had a significant impact on reducing GPSQI scores (USMD −2.35, 95% CI −3.13 to −1.57, p<0.001), but the studies had notable heterogeneity (I^2=69.13%) (figure 2). It was found that the variation in exercise intensity may be responsible for the observed

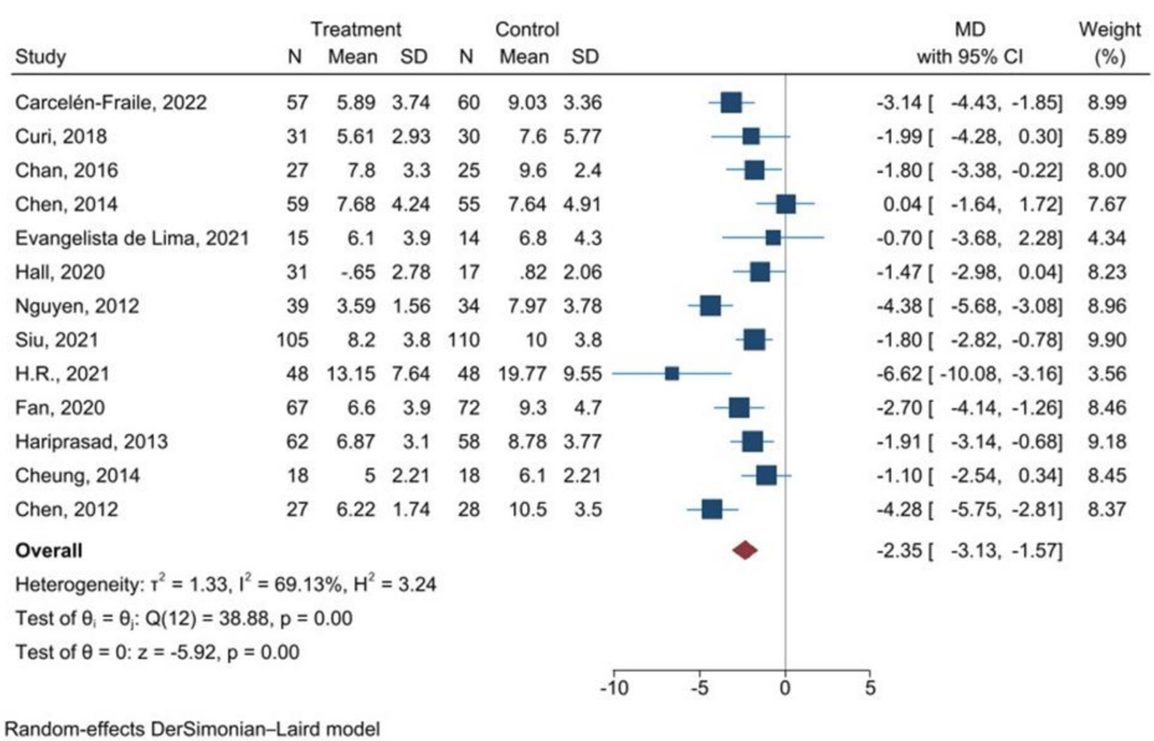

Figure 2 Pooled mean difference of Global Pittsburgh Sleep Quality Index (GPSQI) when comparing combination exercise with routine physical activity.

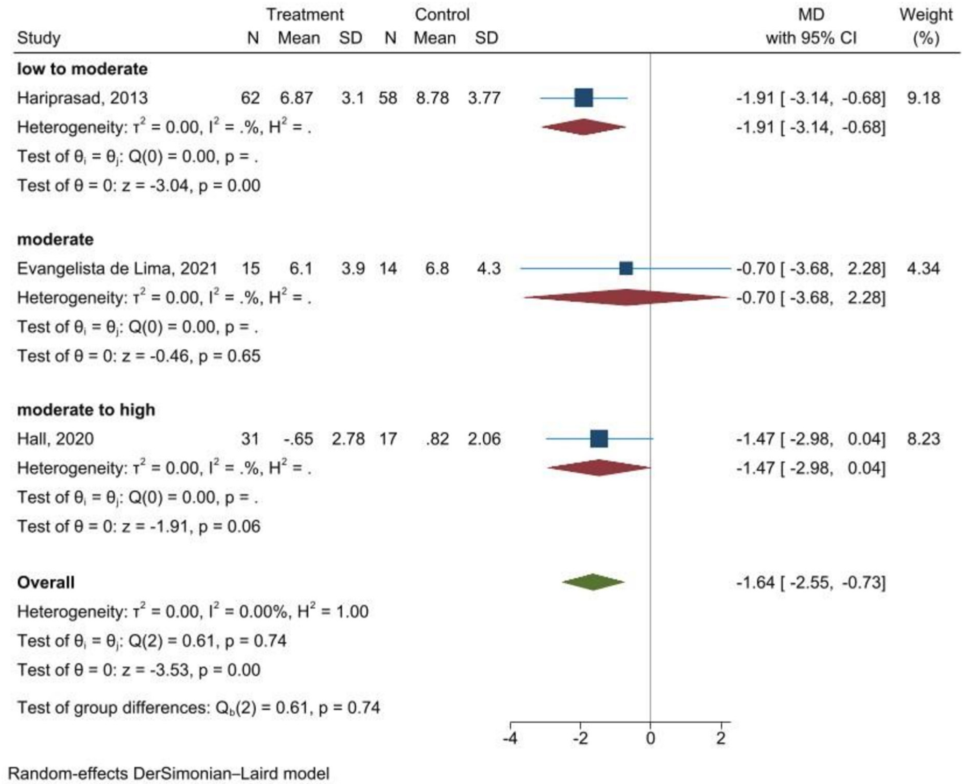

Figure 3 Pooled mean difference of Global Pittsburgh Sleep Quality Index (GPSQI) when comparing combination exercise with routine physical activity according to intensity.

heterogeneity, and this subsequently reduced the degree of heterogeneity to 0% (figure 3). Low to moderate-intensity exercise appeared to decrease USMDs the most (USMD −1.91, 95% CI −3.14 to −0.68).

The comparison between aerobic exercise alone and routine physical activity also showed a significant decrease in GPSQI scores (USMD −4.36, 95% CI −7.86 to −0.86, p=0.01). However, the heterogeneity was also remarkable (I^2=97.83%) (figure 4). The heterogeneity was reduced after accounting for time, duration, length and baseline GPSQI scores. The group with a session duration of less than 60 min had a substantial decrease in scores (USMD −2.07, 95% CI −2.55 to −1.58). Likewise, the group that had a session duration of 60 min or more also exhibited a notable decrease in scores (USMD −6.34, 95% CI −11.56 to −1.13), as depicted in online supplemental figure S3. There were significant decreases in scores among various subgroups, including those who performed more than 100 min of exercise per week (USMD −2.75, 95% CI −4.23 to −1.28), those who exercised for less than 10 weeks (USMD −2.75, 95% CI −4.23 to −1.28), those who exercised for 10 weeks or more (USMD −9.01, 95% CI −10.06 to −7.96), those with a baseline GPSQI score below 10 (USMD −2.75, 95% CI −4.23 to −1.28), along with those with a baseline GPSQI score of 10 or higher (USMD −9.01, 95% CI −10.06 to −7.96). The aforementioned

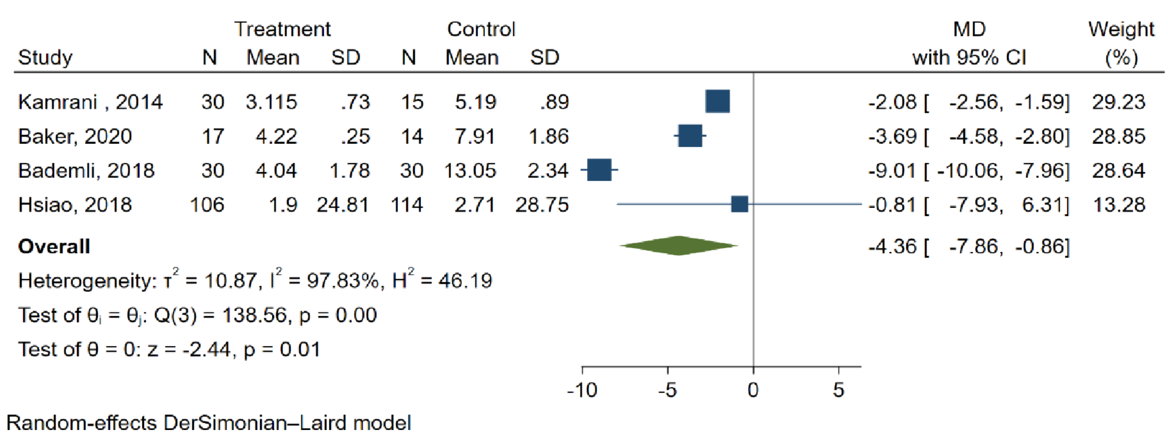

Figure 4 Pooled mean difference of Global Pittsburgh Sleep Quality Index (GPSQI) when comparing aerobic exercise with routine physical activity group.

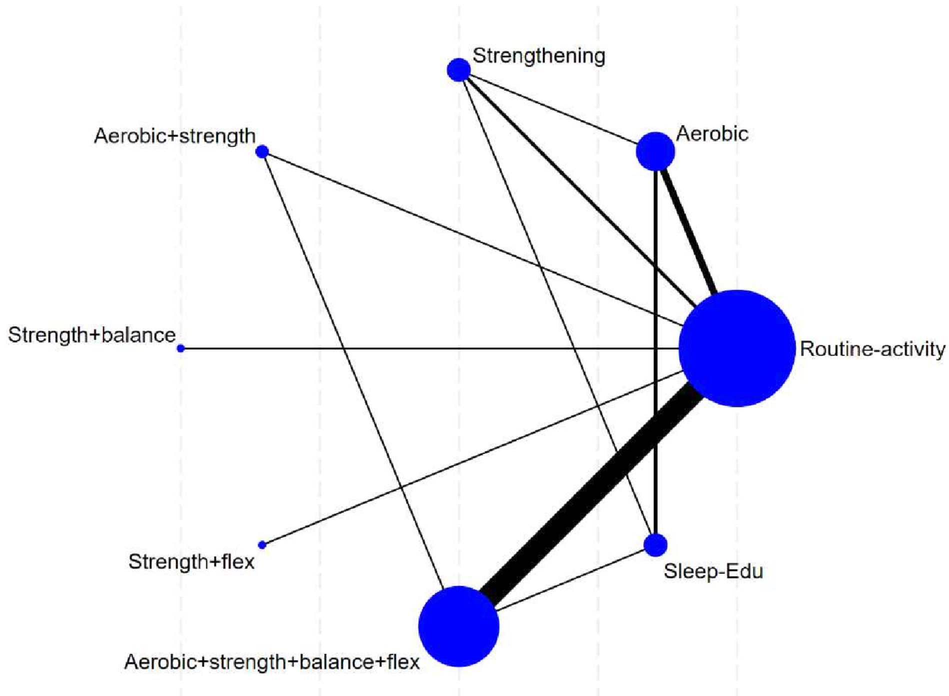

Figure 5 Network diagrams of Global Pittsburgh Sleep Quality Index (GPSQI) in all comparisons.

findings are shown in online supplemental figures S4 to S6.

Network meta-analysis

The network meta-analysis incorporated 24 out of 25 studies, as one study[38] did not report the standard deviation of GPSQI scores. The studies included in the analysis comprised seven different interventions: aerobic exercise, strengthening, a combination of aerobic exercise and strengthening, a combination of strengthening and balancing exercises, a combination of strengthening and flexibility exercises, a combination of aerobic exercise, strengthening, balance exercises and flexibility exercises, as well as sleep education. The data indicated that there was no significant inconsistency across all trials, as indicated by a χ^2 value of 3.78 and a p-value =0.58. The network graph showed that the pairs of combination verus routine, aerobic versus routine, and strengthening versus routine consisted of the largest number of studies. Meanwhile, the graph showed that the largest sample size was the routine activity group, followed by a combination of exercise and aerobic exercise, as shown in figure 5. Aerobic exercise demonstrated a substantial decrease in GPSQI scores when compared with regular physical activity (USMD−3.76, 95%CI −5.67 to −1.85), although the meta-analysis of the aerobic exercise group was characterised by marked heterogeneity (I^2=97.83%). Strengthening exercises resulted in a substantial decrease in GPSQI scores compared with regular activity (USMD −5.75, 95% CI −8.06 to −3.45), combined exercise (USMD −3.21, 95% CI −5.73 to −0.70), and a combination of aerobic and strengthening exercise (USMD−4.28, 95% CI −8.35 to −0.22). The combination exercise, as compared with normal activity, resulted in a substantial reduction in GPSQI scores (USMD −2.54, 95% CI −3.65 to −1.43). Nevertheless, the meta-analysis of the combination exercise group was characterised by notable heterogeneity (I^2=69.13%). Sleep education resulted in a substantial decrease in GPSQI scores compared with routine activity (USMD −4.63, 95% CI −7.12 to −2.15), as seen in table 1. Strengthening exercise was the most effective approach for reducing GPSQI, followed by sleep education and engaging in aerobic exercise, as seen in table 2 and online supplemental figure S7. The adjusted funnel plot demonstrated an overall symmetrical distribution of data across all studies, with the exception of one study[39] suggesting a potential risk of publication bias as shown in online supplemental figure S8.

Secondary outcome

The secondary outcome of this study covered all seven domains of PSQI, comprising subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbance, usage of sleep medicine, and daytime dysfunction.

This research included only trials with combination exercise in the meta-analysis because sufficient numbers of eligible trials were not available for the other exercise types. The meta-analysis indicated the beneficial impact of combination exercise on reducing the five domains of PSQI scores in comparison to the routine physical activity group. The domains mentioned were sleep quality (USMD −0.43, 95% CI -0.67 to −0.2, p<0.001, I^2=54.2%), sleep latency (USMD −0.44, 95% CI -0.76 to −0.12, p=0.01, I^2=70.59%), sleep duration (USMD −0.43, 95% CI −0.6 to −0.26, p<0.001, I^2=45.09%), sleep efficiency (USMD −0.39, 95% CI −0.58 to −0.2, p<0.001, I^2=50.85%), and

Table 1 Estimation of mean difference (95% CI) comparison of GPSQI in upper triangle

Routine activity	−4.63 (−7.12 to −2.15)*	−2.54 (−3.65 to −1.43)*	−1.22 (−5.29 to 2.85)	−1.57 (−5.80 to 2.66)	−1.47 (−4.83 to 1.89)	−5.75 (−8.06 to −3.45)*	−3.76 (−5.67 to −1.85)*
	Sleep education	2.10 (−0.52 to 4.71)	3.41 (−1.36 to 8.19)	3.06 (−1.84 to 7.97)	3.17 (−0.98 to 7.31)	−1.12 (−3.72 to 1.48)	0.87 (−1.33 to 3.08)
		A+S+B+F	1.32 (−2.90 to 5.54)	0.97 (−3.41 to 5.34)	1.07 (−2.29 to 4.43)	−3.21 (−5.73 to −0.70)*	−1.22 (−3.38 to 0.93)
			S+F	−0.35 (−6.22 to 5.52)	−0.25 (−5.53 to 5.03)	−4.53 (−9.21 to 0.15)	−2.54 (−7.04 to 1.96)
				S+B	0.10 (−5.30 to 5.51)	−4.18 (−9.00 to 0.64)	−2.19 (−6.83 to 2.45)
					A+S	−4.28 (−8.35 to −0.22)*	−2.29 (−6.15 to 1.56)
						S	1.99 (−0.44 to 4.42)
							A

*significant change of mean difference comparison of GPSQI.
A, aerobic; A+S, aerobic+strengthening; A+S+B+F, aerobic+strengthening+balance+flexibility; GPSQI, Global Pittsburgh Sleep Quality Index; S, strengthening; S+B, strengthening+balance; S+F, strengthening+flexibility.

sleep medication (USMD −0.53, 95% CI −0.85 to −0.2, p=0.00, I^2=56.58%). Nevertheless, the combination exercise group was not able to decrease their PSQI scores in the domains of sleep disturbance (USMD −0.15, 95% CI −0.39 to 0.08, p=0.2, I^2=74.93%) and daytime dysfunction (USMD −0.29, 95% CI −0.62 to 0.04, p=0.09, I^2=80.35%) compared with the group that performed regular physical activity. The aforementioned findings are shown in online supplemental figure S9A to S9G.

After performing a subgroup analysis on the sleep quality domain, it revealed that both the duration of exercise per session (<60 min and ≥60 min) and the weekly exercise duration (≤100 min and >100 min) resulted in improved sleep quality scores in the combination exercise group. Online supplemental figures S10A, S10B illustrate these findings.

In the domain of sleep latency, the subgroup analysis based on the baseline GPSQI score, percentage of female participants, time per session, and weekly exercise duration likewise revealed a decrease in sleep latency scores among the combination exercise group (online supplemental figures S11A-S11D). The duration of exercise per week (≤100 min and >100 min) both demonstrated the improvement in sleep efficiency scores (online supplemental figure S12). The subgroup analysis, which examined the female percentage, time per exercise session, duration per week, and baseline GPSQI scores of the sleep duration and sleep medication domains, revealed significant improvement in these scores among participants in the combination exercise group (online supplemental figure S13A–D,S15A–D). The length of the study may account for the variations in the sleep disturbance domain (I^2=37.02%): studies lasting 1–12 weeks exhibited a reduction in sleep disturbance scores (USMDs −0.35, 95% CI −0.46 to −0.23), whereas studies lasting 13–30 weeks showed an increase in scores (USMD 0.15, 95% CI −0.08 to 0.38) (online supplemental figure S14).

The daytime dysfunction domain in older people aged 60–70 who undertook combination exercise compared with routine physical activity was improved; however, the benefit was not seen in people older than 70 years (online supplemental figure S16A). Similar outcomes were observed in the domain of daytime dysfunction. For studies lasting 1–12 weeks, the combination exercise group showed an improvement in daytime dysfunction scores (USMD −0.43, 95% CI −0.7 to −0.15). However, for studies lasting 13–30 weeks, there was no significant difference between the two groups (USMD −0.09, 95% CI −0.86 to 0.69) (online supplemental figure S16B).

DISCUSSION

Impact of individual exercise on PSQI

Our systematic review revealed that only the combination exercise and aerobic exercise groups had enough studies to perform the direct meta-analysis, which showed both types of exercise significantly help reduce GPSQI scores with USMDs of −2.35 (95% CI −3.13 to −1.57, p<0.001,

Table 2 Surface under the cumulative ranking curve of GPSQI

Treatment	SUCRA	PrBest	Mean rank
Routine physical activity	9.6	0.0	7.3
Aerobic exercise	67.3	2.1	3.3
Strengthening exercise	94.6	72.5	1.4
Aerobic+strengthening exercise	32.7	1.0	5.7
Strengthening+balance exercise	35.0	3.2	5.6
Strengthening+flexibility exercise	32.7	2.9	5.7
Aerobic+strengthening+balance+flexibility exercise	46.7	0.0	4.7
Sleep education	81.4	18.3	2.3

GPSQI, Global Pittsburgh Sleep Quality Index; PrBest, Probability best; SUCRA, The surface under the cumulative ranking curve.

I^2=69.13%) and −4.36 (95% CI −7.86 to −0.86, p=0.01, I^2=97.83%), respectively (figures 2 and 4). Congruent results in subjective sleep quality scores and objective sleep index (total sleep time by polysomnography, PSG) have also been published in other systematic reviews.[16 17 40–43] Even though there was a high level of heterogeneity between studies, when controlling for covariates, almost all the data indicated significantly reduced GPSQI scores.

By contrast, a recent systematic review and network meta-analysis compared the efficacy of different exercise regimens to improve sleep quality in older people.[44] The results showed that combining muscle endurance training with walking was more effective in improving sleep quality than regimens involving sleep hygiene, pilates, walking alone, health education, or resistance training. Factors such as including participants over the age of 65, including non-pharmacological therapies like CBT for insomnia (CBT-I), and not restricting the study to those employing PSQI as the main outcome may have contributed to different outcomes from this study. The study indicated that face-to-face CBT-I was more effective at enhancing sleep quality than other activities and standard care, which could have influenced the different results from this study as well.

In the combination exercise group, the results showed that only low to moderate intensity exercise reduced GPSQI significantly. Nevertheless, this finding was mainly based on only one large RCT with a high risk of bias, and most of the intensity data were missing in the combination group.[45] For the aerobic exercise group, exercising for over 60 min—at least 100 min per week for more than 10 weeks—demonstrated the most significant score reduction among others. The length of exercise is consistent with previous studies on the effect of physical activity on sleep outcome[16] that show an exercise program of 12 weeks to 6 months is the most effective strategy to improve sleep outcome. Although the duration of exercise was shorter for the aerobic exercise group in our systematic review, it significantly improved sleep quality.

Regarding the secondary outcome, the combination exercise helped to significantly reduce subjective symptoms of insomnia in most of the subdomain scores except sleep disturbance and daytime dysfunction domains, which did not significantly improve. This outcome was inconsistent with one previous study, which was about the effect of Tai Chi.[15] This study found that Tai Chi significantly reduced all seven subdomain scores. The explanation behind this discordance could arise from the variation of exercise characteristics in this review, and half of the studies in the combination exercise group (7 of 13 publications) were classified as high risk of bias.

The best therapeutic type of exercise to improve sleep quality

Our network meta-analysis of 24 studies indicated strengthening exercise as the most effective in improving GPSQI, followed by sleep education, aerobic exercise and combination exercise. The ranking was based on the SUCRA presented in table 2 and online supplemental figure S7. The strengthening exercise not only reduced GPSQI the most compared with routine activity, but also considerably decreased the score compared with the combination of aerobic and strengthening exercise groups and the combination exercise group. The unexpected discrepancy in this outcome might be attributed to the intensity of exercise among the strengthening exercise group, in which high-intensity exercise represented approximately 50% of studies compared with other groups with very limited information on intensity data. The relationship between the intensity of exercise and sleep quality was explored in previous studies but was still unclear,[16 46] even though it was reported that moderate intensity physical activity might be more effective than vigorous exercise in enhancing sleep quality for young people and older individuals. Moreover, another study indicated that a high physical load at work or physical activity was associated with difficulties initiating sleep and may be a risk factor for insomnia.[47] In contrast, another study found that high-intensity exercise did not disrupt and might even improve subsequent nocturnal sleep.[48]

Although sleep education is a common approach to managing insomnia in older individuals, many studies in our analysis failed to clearly outline the specific procedure used, thus introducing confounding variables into the outcomes of our study. It is also worth mentioning that when analysing data in the combination exercise group, it is important to interpret the results cautiously due to the high risk of bias among studies in this group. Nevertheless, a previous study found that the minimal clinically significant difference (MCID) for the PSQI is 3 points.[49] According to this cut-off value, strengthening exercise, sleep education, and aerobic exercise all significantly improve sleep quality at the clinical level.

Strengths and limitations

This study has several strengths, as it extensively searched both published and unpublished databases to reduce potential research bias. The research focused on the older population, making it directly applicable to the demographic most affected by insomnia. The review included people with symptoms of insomnia without being restricted by specific diagnostic criteria, which could hinder the study selection process. Although no cut-off number was specified in this review, almost all baseline mean GPSQI scores were over 5, with only one study[50] scoring below 5. The study does not have a floor effect on sleep category scores (a score >5 refers to poor sleep quality)[34] and instead reports outcomes as mean differences and 95% CIs. The exercise interventions recruited in this review varied in type, intensity and duration, which could have enhanced the capacity of this study to determine which form of exercise should be applied. Finally, the current study implemented network meta-analysis as a statistical strategy, which served as a head-to-head comparison and ranked the best therapeutic exercise based on the efficacy of improving subjective sleep quality.

There are limitations in this study that could be addressed in future reviews. First, the heterogeneity of this study is high, mainly because of variations in the characteristics of exercise. Exercises of the same type may have varying structures that can impact the outcome. Another factor that could impact the outcome is intensity. Several studies lack data on the intensity of exercise, despite our efforts to communicate with the authors to address this issue. Several studies focused only on comorbidity as a restriction to physical activity, but other diseases, including cardiovascular disease, respiratory disease and uncontrolled hypertension, are also indirectly affected by exercise. Therefore, clinical applications should take this into account or may require further investigation in older individuals with specific medical conditions. Some exercise categories lack sufficient studies to illustrate the impact of those types of exercise on sleep quality due to low statistical power. Some exercises may be challenging for older people because of complicated techniques, restricted physical capabilities, and the requirement for an experienced instructor. Finally, the subjective sleep quality may not fully align with objective sleep outcomes such as polysomnography; this limitation might reduce the generalisability of the results.

CONCLUSION

This study shows that exercise, particularly strengthening exercise and aerobic exercise, is beneficial for enhancing subjective sleep quality at a clinically significant level compared with normal activities, which is consistent with previous studies. Strengthening has the highest efficacy, followed by aerobic exercise and combination exercise. Nevertheless, caution should be applied when interpreting this study because of the diverse exercise characteristics, the small number of studies, and the high risk of bias among studies.

Implications for future practice

This study aimed to compare the impact of exercise compared with routine activities on sleep quality and determine the most effective form of exercise. The results will help evaluate if exercise could be used as a supplemental treatment for insomnia and identify the most effective type of exercise. The outcomes of this study indicate that strengthening exercise is the most efficacious among others, followed by aerobic exercise and combination exercise. Nevertheless, all these types of exercise improve sleep quality beyond significant differences.

Implications for future research

Future studies could incorporate more evidence to enhance statistical power in certain exercise categories. The study's results could incorporate both subjective and objective measures to validate the sleep outcome. Adopting a prespecified intensity scale, such as the Borg Rating of Perceived Exertion (RPE) Scale, the Balance Intensity Scale (BIS), and the Talk test, could reduce bias and variation in exercise intensity. Additionally, objective outcomes such as maximum heart rate (MHR) and oxygen consumption (VO2) may be tracked by a smartwatch or other electronic devices to reduce bias and variation in exercise intensity and duration.

Contributors PB and KN are guarantors. PB contributed to drafting the study protocol, evaluating the quality of included studies, extracting data, interpreting data, and drafting the manuscript. KN contributed to revising the study proposal, evaluating the quality of included studies, extracting data, interpreting the data, and editing the final manuscript.TA contributed to the study concept and design, data analysis, data interpretation, and editing of the final manuscript. AI applications like QuillBot and Grammarly were used for language proofreading.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was exempted from Human Research Ethics Committee, Faculty of Medicine Ramathibodi Hospital, Mahidol University due to nature of study (systematic review and network meta-analysis). The ID number is 2484. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon request. The data used for analysis are available in the supplementary tables. The entire set of data will be provided upon request from the authors.

ORCID iDs
Pakwan Bahalayothin http://orcid.org/0009-0005-7500-0608
Kittiphon Nagaviroj http://orcid.org/0000-0002-5803-5244
Thunyarat Anothaisintawee http://orcid.org/0000-0003-1002-8536

REFERENCES

1 Roth T, Coulouvrat C, Hajak G, *et al*. Prevalence and perceived health associated with insomnia based on DSM-IV-TR; International Statistical Classification of Diseases and Related Health Problems, Tenth Revision; and Research Diagnostic Criteria/International Classification of Sleep Disorders, Second Edition criteria: results from the America Insomnia Survey. *Biol Psychiatry* 2011;69:592–600.
2 Matheson E, Hainer BL. Insomnia: Pharmacologic Therapy. *Am Fam Physician* 2017;96:29–35.
3 Jansson-Fröjmark M, Lindblom K. A bidirectional relationship between anxiety and depression, and insomnia? A prospective study in the general population. *J Psychosom Res* 2008;64:443–9.
4 Jaussent I, Bouyer J, Ancelin M-L, *et al*. Insomnia and daytime sleepiness are risk factors for depressive symptoms in the elderly. *Sleep* 2011;34:1103–10.
5 Schwartz S, McDowell Anderson W, Cole SR, *et al*. Insomnia and heart disease: a review of epidemiologic studies. *J Psychosom Res* 1999;47:313–33.
6 Troxel WM, Buysse DJ, Matthews KA, *et al*. Sleep symptoms predict the development of the metabolic syndrome. *Sleep* 2010;33:1633–40.
7 Palagini L, Bruno RM, Gemignani A, *et al*. Sleep loss and hypertension: a systematic review. *Curr Pharm Des* 2013;19:2409–19.
8 Yaffe K, Falvey CM, Hoang T. Connections between sleep and cognition in older adults. *Lancet Neurol* 2014;13:1017–28.
9 Sigurdardottir LG, Valdimarsdottir UA, Mucci LA, *et al*. Sleep Disruption Among Older Men and Risk of Prostate Cancer. *Cancer Epidemiol Biomarkers Prev* 2013;22:872–9.
10 Kucharczyk ER, Morgan K, Hall AP. The occupational impact of sleep quality and insomnia symptoms. *Sleep Med Rev* 2012;16:547–59.
11 NIH State-of-the-Science Conference Statement on manifestations and management of chronic insomnia in adults. *NIH Consens State Sci Statements* 2005;22:1–30.
12 Maness DL, Khan M. Nonpharmacologic Management of Chronic Insomnia. *Am Fam Physician* 2015;92:1058–64.
13 Trauer JM, Qian MY, Doyle JS, *et al*. Cognitive Behavioral Therapy for Chronic Insomnia. *Ann Intern Med* 2015;163:191–204.
14 Rubio-Arias JÁ, Marín-Cascales E, Ramos-Campo DJ, *et al*. Effect of exercise on sleep quality and insomnia in middle-aged women: A systematic review and meta-analysis of randomized controlled trials. *Maturitas* 2017;100:49–56.
15 Du S, Dong J, Zhang H, *et al*. Taichi exercise for self-rated sleep quality in older people: a systematic review and meta-analysis. *Int J Nurs Stud* 2015;52:368–79.
16 Vanderlinden J, Boen F, van Uffelen JGZ. Effects of physical activity programs on sleep outcomes in older adults: a systematic review. *Int J Behav Nutr Phys Act* 2020;17:11.
17 Wu YHT, He WB, Gao YY, *et al*. Effects of traditional Chinese exercises and general aerobic exercises on older adults with sleep disorders: A systematic review and meta-analysis. *J Integr Med* 2021;19:493–502.
18 Buysse DJ, Hall ML, Strollo PJ, *et al*. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. *J Clin Sleep Med* 2008;4:563–71.
19 Wang L, Wu Y-X, Lin Y-Q, *et al*. Reliability and validity of the Pittsburgh Sleep Quality Index among frontline COVID-19 health care workers using classical test theory and item response theory. *J Clin Sleep Med* 2022;18:541–51.
20 Backhaus J, Junghanns K, Broocks A, *et al*. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res* 2002;53:737–40.
21 Sitasuwan T, Bussaratid S, Ruttanaumpawan P, *et al*. Reliability and validity of the Thai version of the Pittsburgh Sleep Quality Index. *J Med Assoc Thai* 2014;97 Suppl 3:S57–67.
22 Marques D, Gomes AA, Meiavia A, *et al*. Reliability and initial validation of the Pittsburgh Sleep Quality Index, European Portuguese version: A preliminary study in a sample of higher education students. *Sleep Med* 2013;14:e140.
23 Zheng B, Li M, Wang KL, *et al*. Analysis of the reliability and validity of the Chinese version of Pittsburgh sleep quality index among medical college students. *Beijing Da Xue Xue Bao Yi Xue Ban* 2016;48:424–8.
24 Spira AP, Beaudreau SA, Stone KL, *et al*. Reliability and validity of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older men. *J Gerontol A Biol Sci Med Sci* 2012;67:433–9.
25 Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
26 Hutton B, Salanti G, Caldwell DM, *et al*. The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations. *Ann Intern Med* 2015;162:777–84.
27 Administration. SAaMHS. *Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health*. National Library of Medicine, 2016.
28 Services USDoHaH. Physical activity guidelines for americans 2022. 2022. Available: https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf
29 Patel H, Alkhawam H, Madanieh R, *et al*. Aerobic vs anaerobic exercise training effects on the cardiovascular system. *World J Cardiol* 2017;9:134–8.
30 Sundell J. Resistance Training Is an Effective Tool against Metabolic and Frailty Syndromes. *Adv Prev Med* 2011;2011:984683.
31 project. TOP. balance. 2022. Available: https://www.physio-pedia.com/Balance#Introduction
32 Hampton L. Flexibiity. 2022. Available: https://www.physio-pedia.com/Flexibility#cite_note-1
33 Medicine CoFMaHOiYFaNBIo. *Health-Related Fitness Measures for Youth: Flexibility*. National Academic Press, 2012.
34 Buysse DJ, Reynolds CF III, Monk TH, *et al*. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213.
35 Lu G, Welton NJ, Higgins JPT, *et al*. Linear inference for mixed treatment comparison meta-analysis: A two-stage approach. *Res Synth Methods* 2011;2:43–60.
36 Jackson D, Boddington P, White IR. The design-by-treatment interaction model: a unifying framework for modelling loop inconsistency in network meta-analysis. *Res Synth Methods* 2016;7:329–32.
37 Sterne JAC, Sutton AJ, Ioannidis JPA, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
38 Brandão GS, Gomes GSBF, Brandão GS, *et al*. Home exercise improves the quality of sleep and daytime sleepiness of elderlies: a randomized controlled trial. *Multidiscip Respir Med* 2018;13:2.
39 Bademli K, Lok N, Canbaz M, *et al*. Effects of Physical Activity Program on cognitive function and sleep quality in elderly with mild cognitive impairment: A randomized controlled trial. *Perspect Psychiatr Care* 2019;55:401–8.
40 Montgomery P, Lilly J. Insomnia in the elderly. *BMJ Clin Evid* 2007;2007:2302.
41 Yang PY, Ho KH, Chen HC, *et al*. Exercise training improves sleep quality in middle-aged and older adults with sleep problems: a systematic review. *J Physiother* 2012;58:157–63.
42 D'Aurea CVR, Frange C, Poyares D, *et al*. Physical exercise as a therapeutic approach for adults with insomnia: systematic review and meta-analysis. *Einstein (Sao Paulo)* 2022;20:eAO8058.
43 Bonardi JMT, Lima LG, Campos GO, *et al*. Effect of different types of exercise on sleep quality of elderly subjects. *Sleep Med* 2016;25:122–9.
44 Hasan F, Tu YK, Lin CM, *et al*. Comparative efficacy of exercise regimens on sleep quality in older adults: A systematic review and network meta-analysis. *Sleep Med Rev* 2022;65:101673.

45 Hariprasad V, Sivakumar P, Koparde V, *et al*. Effects of yoga intervention on sleep and quality-of-life in elderly: A randomized controlled trial. *Indian J Psychiatry* 2013;55:364.
46 Tseng TH, Chen HC, Wang LY, *et al*. Effects of exercise training on sleep quality and heart rate variability in middle-aged and older adults with poor sleep quality: a randomized controlled trial. *J Clin Sleep Med* 2020;16:1483–92.
47 Dubinina E, Korostovtseva LS, Rotar O, *et al*. Physical Activity Is Associated With Sleep Quality: Results of the ESSE-RF Epidemiological Study. *Front Psychol* 2021;12:705212.
48 Thomas C, Jones H, Whitworth-Turner C, *et al*. High-intensity exercise in the evening does not disrupt sleep in endurance runners. *Eur J Appl Physiol* 2020;120:359–68.
49 Hughes CM, McCullough CA, Bradbury I, *et al*. Acupuncture and reflexology for insomnia: a feasibility study. *Acupunct Med* 2009;27:163–8.
50 Hsiao CY, Chen KM, Tsai HY, *et al*. Self-Perceived Health and Sleep Quality of Community Older Adults after Acupunch Exercises. *Am J Geriatr Psychiatry* 2018;26:511–20.