SLEEP, 2025, 48, 1–19
https://doi.org/10.1093/sleep/zsaf021
Advance access publication 6 February 2025
Review

Review

A systematic review and meta-analysis of group-based trajectory modeling of sleep duration across age groups and in relation to health outcomes

Wei Wang[1,2], Sing-Hang Cheung[3], Shu Fai Cheung[4], Rong Wei Sun[5], C. Harry Hui[3], Ho Yin Derek Ma[1] and Esther Yuet Ying Lau[1,2,6,*]

[1]Department of Psychology, The Education University of Hong Kong, Tai Po, New Territories, Hong Kong SAR, China,
[2]Analytics\Assessment Research Centre, The Education University of Hong Kong, Tai Po, New Territories, Hong Kong SAR, China,
[3]Department of Psychology, The University of Hong Kong, Pokfulam, Hong Kong SAR, China,
[4]Department of Psychology, University of Macau, Taipa, Macau SAR, China,
[5]School of Arts and Humanities, Tung Wah College, Mong Kok, Kowloon, Hong Kong SAR, China and
[6]Centre for Psychosocial Health, The Education University of Hong Kong, Tai Po, New Territories, Hong Kong SAR, China

*Corresponding author. Esther Yuet Ying Lau, Room 22, 2/F., Block D1, Department of Psychology, The Education University of Hong Kong, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong SAR, China. Email: laueyy@eduhk.hk.

Abstract

Study Objectives: To shed light on understanding sleep duration trajectories (SDTs) using different classification methods and their outcomes, this study aimed to (1) identify common SDTs among different age groups, (2) investigate the alignment versus differences between SDTs identification by group-based trajectory modeling (GBTM) and clinical standards, and (3) examine the impacts of SDTs on health outcomes.

Methods: A systematic literature search from four databases yielded 34 longitudinal SDT studies with GBTM analyses spanning three or more data waves. Apart from the proportion meta-analysis, a three-level meta-analysis was conducted with 14 of the studies that examined the association between SDT groups and health outcomes. Assessment of study quality was performed using the Guidelines for Reporting on Latent Trajectory Studies checklist.

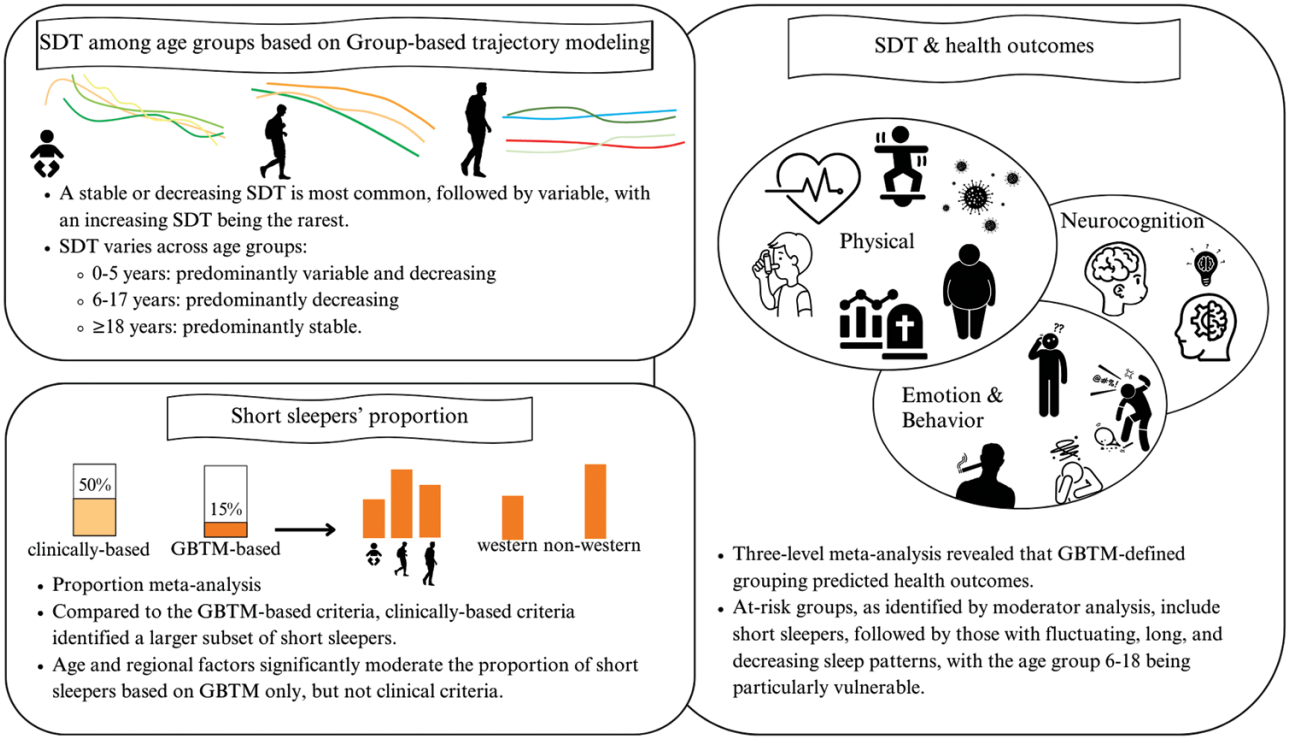
Results: Qualitative analysis identified four age-related SDT classes based on longitudinal trends: "persistent sleepers," "increase sleepers," "decrease sleepers," and "variable sleepers." Meta-analysis also showed differential proportions of "GBTM-defined shortest sleepers" across age groups and sample regions, as well as significant discrepancies in the prevalence of short sleep identified by clinical standards (=50% vs. 15% per GBTM). Overall, SDTs predicted emotional and behavioral outcomes, neurocognitive problems, and physical health ($OR = 1.538$, $p < 0.001$), in GBTM-defined "short," "fluctuating," "long," and "decreasing" sleepers as compared to the "adequate" group. The effects were stronger in adolescents and in datasets with more waves.

Conclusions: The identification of the GBTM-defined "short," "fluctuating," "long," and "decreasing" SDT groups and their associations with various health outcomes supported longitudinal investigations, as well as the development of interventions focusing on both the length and stability of sleep durations, especially in younger populations. Study registration: PROSPERO registration number CRD42023412201.

Key words: sleep duration trajectory; group-based trajectory modeling; three-level meta-analysis; proportion of short sleepers

Submitted for publication: September 17, 2024; **Revised:** December 17, 2024

Graphical Abstract



Short sleep is relatively common, with approximately one-third of both children and adults sleeping less than clinical recommendations [1–3]. Sleep duration has been associated with various consequences across life stages, including emotional and behavioral functioning such as prosocial behaviors [4], cognitive performance [5], and physical health [6]. However, some reported the impact of sleep duration, such as its cross-sectional U-shaped relationship with the metabolic syndrome was not supported in longitudinal studies [6]. To represent the intrapersonal patterns of sleep duration over time, sleep duration trajectories (SDTs) have emerged as a major development in sleep research.

Following this trend, many studies categorized SDTs based on professionals' consensus, such as the widely adopted recommendations by the National Sleep Foundation [7], in identifying short or long-sleepers. Previous reviews have comprehensively examined studies of clinical-based SDTs, concluding that insufficient sleepers were at a higher risk of adverse outcomes such as adiposity and mood deficits [8–10]. Clinical-based SDTs capture at-risk sleepers uniformly but may overlook populations' heterogeneity [11, 12]. Group-Based Trajectory Modeling (GBTM), also known as latent class growth modeling, identifies distinct subgroups with similar developmental trajectories across populations studied, accounting for heterogeneity in patterns of change over time [13, 14]. Beyond the static labels of more than just short or long sleepers, GBTM characterizes change vs. stability over time and enables the identification of classes such as persistently short sleepers, initially short sleepers, and those with decreasing sleep duration, etc. Two recent reviews analyzing GBTM-defined SDTs in children and adolescents, respectively, identified that younger membership in the persistently short sleep group was associated with various health outcomes [15, 16]. Moreover, lower socioeconomic status can predict persistently short sleepers, which is associated with a higher risk of cognitive and emotional deficits in children and adolescents [15], in contrast, to persistently long and naturally decreasing sleepers, who had lower risks of cognitive issues [16]. However, both reviews primarily discussed health risks without detailing prevalence rates for the persistently short groups.

Each with its own merits and constraints, the two classification methods should be compared to provide insight into their alignments versus discrepancies for a more accurate interpretation of findings and to inform study design that avoids overestimation or underestimation for specific purposes of future SDT studies. Furthermore, understanding the predictive capabilities of existing sleep duration standards—both clinically-defined and GBTM-defined—on health outcomes is crucial for validating and potentially adjusting these standards. Examining these classification methods as moderators could show which better captures the complexities of sleep patterns and their health associations. Additionally, there remains a significant gap in reviewing the existing SDT literature across a wide age range. To address these gaps, we aim to answer the following research questions:

1) How do SDTs vary among different age groups?
2) How do sleep duration standards based on GBTM align with professionals' consensus groups?
3) How is membership in specific SDT groups associated with health outcomes, and what are the moderators of the associations?

To achieve the above goals, we first qualitatively summarized the SDT classes among age groups. We then used a heatmap to visualize how SDT groups perform based on clinical standards and conducted a meta-proportion analysis to compare the proportions of clinical-based and GBTM-defined short sleepers. Finally, we employed multilevel meta-analysis to examine the

relationship between SDTs and health outcomes and conducted the moderator analysis to reveal whether different classification methods differ in their ability to predict health risks associated with various SDTs.

Methods

Selection strategy and eligibility criteria

Adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines, a systematic literature search was conducted across ProQuest, EBSCO, Scopus, and Web of Science. As GBTM of SDTs was a recent development of sleep research, the search was limited to articles published from January 2003 to March 2023. The search string employed was TS = (((sleep*) AND (trajector* OR "latent growth" OR "growth mixture" OR "mixture modeling"))). The detailed search strategy can be found in Table S1. We removed duplicates and filtered out studies that did not contain "sleep" in their title or abstract. In the initial screening phase, we reviewed the titles and abstracts of identified studies to eliminate those that did not provide information related to SDTs. For instance, studies examining the relationship between sleep and depression that primarily focused on depression trajectories rather than sleep trajectories were excluded. For studies that passed our initial screening, one researcher accessed the full text to screen the studies against the eligibility criteria. The primary ambiguities presented were methodological issues and the definition of specialized terms. For instance, some articles did not explicitly mention GBTM, yet their analyses and results included multiple sleep trajectory groups. Any ambiguities and issues were resolved with five other team members at a weekly meeting to reach a consensus collaboratively.

Only SDT-related articles with "sleep" in their title or abstract were retained. Eligibility was further determined based on study design and content: (1) quantitative and longitudinal studies in English investigating SDTs in generally healthy populations without diagnosed physical, psychological, or sleep conditions, or particular features such as pregnancy; (2) using GBTM techniques on at least three data waves; (3) classifying a minimum of two SDT groups (identified solely by sleep duration) with plots provided for visualization.

Assessment of study quality

The Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) [17], a 21-item checklist that has been used in systematic reviews involving latent growth mixture models [18], was used to assess various assumptions and reporting standards in the included studies. This checklist helped us evaluate assumptions related to trajectory group number and shape, model selection criteria, data quality and fit statistics, group homogeneity, and handling of missing data. Higher GRoLTS scores indicate better adherence to reporting standards. A detailed assessment of the standards used for model estimation and selection in each article was also conducted (Table S2).

Qualitative synthesis methods

To address the first question "How do SDTs vary among different age groups?," we categorized participants into three age groups: 0–5, 6–17, and ≥18. This classification was based on optimizing sample sizes for statistical analysis while reflecting key developmental stages. The 0–5 age group represents early childhood, crucial for foundational skill development. The 6–17 age group encompasses both primary and secondary education years, where significant cognitive and social development occurs. By combining SDT groups from all eligible studies and identifying common classes based on the longitudinal trends, we did not solely rely on the labels provided in individual studies. We extracted key information, including age ranges, labels, numbers, proportions, and characteristics of SDT groups, measurement types, as well as SDT group plots from each study.

For the second research question concerning the alignment between clinically and GBTM-defined SDT groups, we created heatmaps based on specific sleep duration values at each time point, via the *ggplot2* and *tidyr* packages in R software. We extracted these values from the studies or their plots using the "Getdata graph digitizer" software [19]. The heatmaps visually represent the age (*x*-axis), the proportion of SDT groups (*y*-axis), and sleep duration disparities (color intensity) between clinically adequate standards and GBTM results. The color scale corresponds to sleep duration in relation to the recommended sleep duration standards established by the National Sleep Foundation [7]. This comprehensive visualization depicts the association between clinical and GBTM-defined SDT groups. To address the third question, we reported the relationships between GBTM-defined SDT groups and health-outcome variables.

Quantitative synthesis methods

We conducted four separate meta-analyses using R software. The first three meta-analyses were designed to estimate proportions for different SDT sleepers using the meta package in all included studies. The fourth was a multilevel meta-analysis using *metafor* and *dmetar* packages on 14 of the included studies that examined the association between SDT groups and health outcomes.

Meta-analyses of proportions.

To further address the second research question, we extracted the group with the shortest sleep duration consistently observed across all-time points as GBTM-defined shortest sleepers (GSS) from each study. We also identified two groups, clinically adequate sleepers (CAS) and clinically short sleepers (CSS), using the clinical-based classification method [7]. The CAS group included individuals who consistently maintained sleep duration within the recommended range for their respective age at all recorded time points, while the CSS group consisted of individuals who at one or more time points had sleep duration below the recommended range for their age.

We employed a random-effects model with weighted trajectory estimates based on data extracted from studies that reported results from a combined sample of 174 960 individuals. Cochran's Q test was used to examine heterogeneity among studies, and I^2 statistic was used to estimate heterogeneity severity. Then, subgroup analyses were conducted to examine differences among age groups, regions, and measurement types. Additionally, meta-regression analysis was performed to identify potential sources of heterogeneity among the studies, using publication year, number of final model trajectories, number of waves of sleep duration measurement, and quality assessment score of each article as continuous predictors.

A three-level meta-analysis.

To answer our third research question: "How is membership in specific SDT groups associated with health outcomes, and what are the moderators of the associations?," we selected 14 out of the 34 studies for a multilevel meta-analysis. We extracted essential information, including ORs, from these different SDT groups and their corresponding health outcomes, resulting in a total of 69 effect sizes. The reasons for excluding certain studies

and detailed information can be found in Table S3. Given the diversity of health outcomes examined in the 14 studies, three groups were identified in order to increase the number of effect sizes within each category, thereby enhancing the robustness and reliability of the analysis. They are: (1) emotional and behavioral outcomes, which include three studies on smoking behavior, non-suicidal self-harm in adolescents, and hyperactivity/inattention, emotional symptoms, conduct problems, and peer relationship problems in children; (2) neurocognitive outcomes, including three studies on cognitive and language development, fine and gross motor abilities, and receptive vocabulary in children, and cognitive decline in adults; and (3) physical outcomes, including new-onset asthma, obesity in adolescents and adults, balance problems, cardiovascular events, all-cause mortality and risk of hypertension in adults, and specific serum cytokines in children, as reported in eight studies.

To accommodate studies that provided multiple effect estimates for the same cohort, our meta-analysis adopted a three-level random-effects model by considering three sources of variance, including the sampling variance of the observed effect sizes (level 1), the variance between effect sizes from the same study (level 2), and the variance between different studies (level 3). Given that each study included multiple SDT groups and that each SDT group was associated with varied outcomes with multiple effect sizes, this model proved crucial in accounting for both the dependency and heterogeneity within and between studies.

This multilevel approach allowed us to capture the hierarchical structure of the data, providing a more accurate estimation of the overall effect. Our analysis followed the steps of the three-level meta-analytic model [20]. Firstly, we tested the overall association between SDTs and health outcomes. We then examined the significance of the within-study variance (level 2) and the between-study variance (level 3). Subsequently, we explored the distribution of the total variance across the three levels of the meta-analytic model. Finally, we conducted subgroup analyses to evaluate whether the associations between different SDT groups and health outcomes differed under three classification methods (i.e. class-based, clinical-based, and GBTM-based) and also investigated whether the associations between SDTs and health outcomes differed across age groups, regions, publication years, and study designs. To assess whether outcome type (physical, emotional and behavioral, neurocognitive) moderated the association, we performed a multivariate meta-regression, modeling random effects at both the study and effect size level to account for within- and between-study heterogeneity. Additionally, separate random-effects multivariate models were conducted for each outcome type to estimate pooled effect sizes independently.

We used funnel plot and Egger's regression tests to assess publication bias for the three-level meta-analysis, but not for the proportion meta-analyses, because these methods are reported to be inappropriate for the latter. In the case of proportion meta-analyses, the concept of a positive result is less defined, as the studies focus on estimating prevalence or proportions rather than comparing effect sizes. Barker et al. (2021) recommended against using traditional publication bias tests in proportion meta-analyses [21]. Instead, a qualitative assessment of potential publication bias is more appropriate, considering factors such as study design, population diversity, and geographic representation. In addition, to assess the consistency of the main results, we employed sensitivity analysis by systematically excluding each study to examine the robustness of the findings. This analysis was conducted separately for both the three-level meta-analysis and the meta-analysis of proportions. In all analyses, a significance level of $p < .05$ (two-tailed) was used.

Results

General description of included studies

A PRISMA flow chart illustrating the study selection process is presented in Figure 1. SDTs are defined as patterns of change in sleep duration over time, identified using GBTM. This method classifies individuals into distinct trajectory groups with specific labels. Initially, we identified 4784 articles across four databases. After removing duplicates and using Endnote to screen out studies without "sleep" in their title or abstract, 2011 articles remained. We then carried out a full-text review of 246 potentially eligible articles, ultimately including 34 articles published between 2009 and 2022 in our final analysis, with the majority (68%) published in or after 2019[a].

Sixteen studies primarily focused on children aged 0–5 years, a phase when daytime naps are necessary to meet children's sleep requirements [22]. Therefore, these studies presented various types of SDTs, including separate trajectories for daytime and nighttime sleep. Eight studies investigated populations with a baseline age between 6 and 17 years, with a stronger emphasis on nighttime SDTs; while ten studies focused on populations aged 18 or above, with one study examining both daytime and nighttime SDTs in older adults. The population in the 34 reviewed studies represented a diverse range of populations from various geographical regions, including America (44%), Asia (29%), and others. The majority (94%) of the studies employed self-report or parent-report to measure sleep with the rest using objective measures such as actigraphy. The number of SDT groups identified in these studies ranged from two to five. Seven studies spanned over approximately a decade, while eight collected sleep duration data over three waves, while the majority conducted more than three waves. Twenty-one and 20 studies analyzed predictors and outcomes for SDT groups, respectively (details in Table 1).

The GRoLTS checklist revealed substantial variations in the reporting quality of the articles, with scores ranging from 5 to 15 (average score of 8.6), with higher scores indicating more rigor (see details in Table S4). All articles outlined the model comparison tools used for model selection, with the Bayesian information criterion being the most common (79%). Akaike's information criterion was used in 15% of the articles, while 12% utilized the Mendel–Rubin–likelihood ratio test. Every article included the number or proportion of each subgroup and provided plots of the final trajectory result. Most articles stated the analytical software used and the control for covariates. However, only two studies provided within-wave mean and variance information, and only one clearly documented latent trajectory model choices. None of the articles provided plots of estimated average trajectories for each model (Figure 2).

SDT variations from qualitative analyses

From the original 167 SDT groups in 34 studies, we categorized them into four main classes based on patterns of within-person sleep duration variation over time: Class A (persistent sleepers) are individuals with stable sleep duration over time, typically within a range of 0.5 hours without significant fluctuations. Class B (increase sleepers) includes individuals who exhibit a consistent upward trend in sleep duration. Class C (decrease sleepers) are those who show a consistent decline in sleep duration over time, and class D (variable sleepers) is characterized by individuals

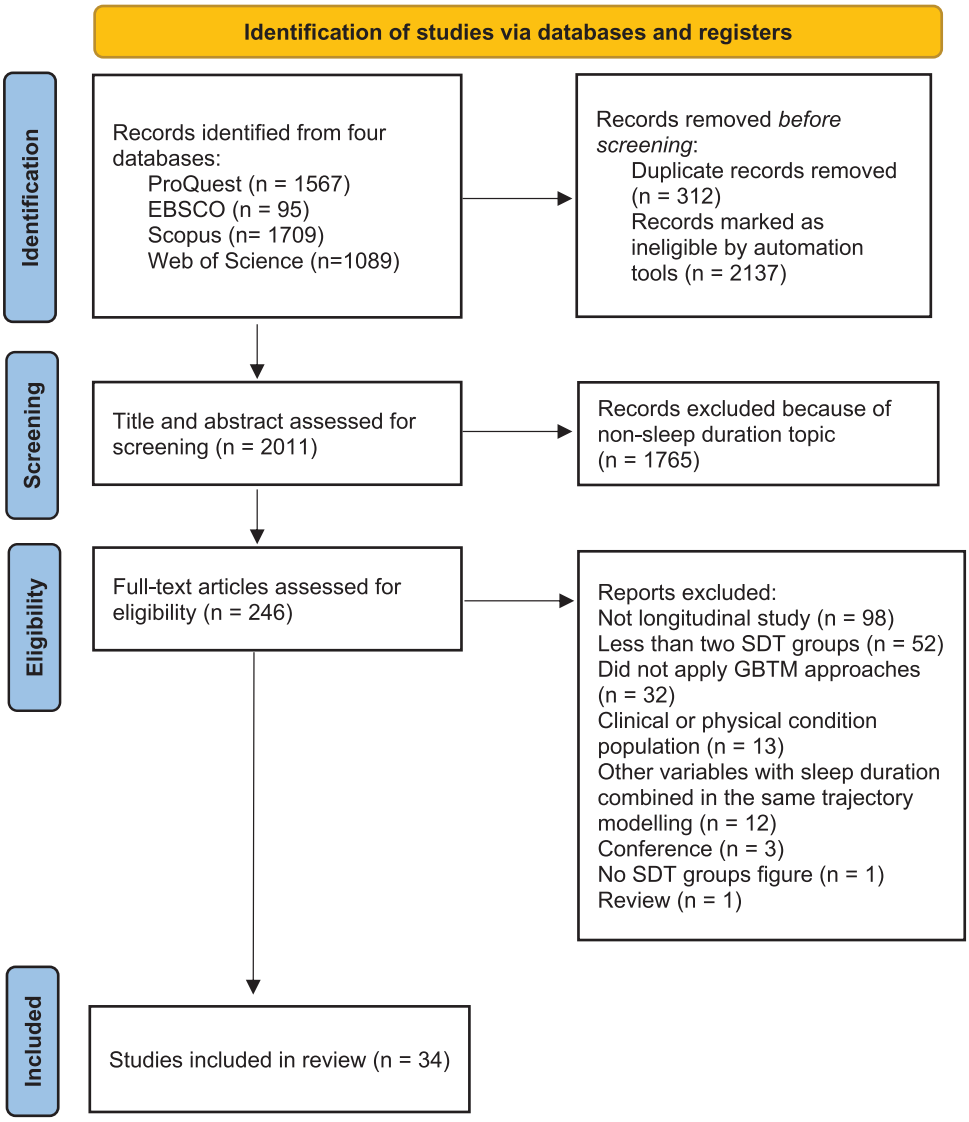


Figure 1. Study Selection Process. Flow chart including inclusion and exclusion criteria.Abbreviations: SDT, sleep duration trajectory; GBTM, group-based trajectory modeling.

with varying sleep patterns, with changes exceeding 0.5 hours and no consistent trend. Classes A and C were the most prevalent, each accounting for 34% of the total classes. Class D accounted for 27%, and class B was the least common, representing only 5% of the total.

Our analysis further revealed that the SDTs showed distinct distributions of SDT classes across different age groups (Figure S1). For infants and toddlers (0–5 years), more sleepers fell into classes C and D. Sleep duration in this age group tended to initially increase before decreasing [30, 33, 37, 46, 49, 52, 57], while some trajectories displayed an initial decrease followed by an increase in the 6 months to 5.5 years age range [41–43, 49, 53]. Throughout childhood and adolescence, a consistent declining trend in sleep duration was common [26, 27, 40, 44, 50]. This phase of life was predominantly represented by Class C (decreased sleepers). In contrast, adulthood presented a more stable pattern in sleep duration, with most SDT groups belonging to class A (persistent sleepers) [24, 29, 40, 45, 48, 50, 51, 55, 56]. Notably, labels given to class A were different across studies, due to their specific research focus (see Table S5 for details).

Relationship between clinical sleep duration standards and GBTM-defined SDT groups

To understand the relationship between clinical standards for sleep duration and GBTM-defined SDT groups, we conducted a meta-analysis and created several heatmap representations. The meta-analysis included 175 748 participants from 34 selected studies, while the heatmaps visually depicted the SDT patterns and their correlation with clinical sleep duration standards. Figure 3, A–C illustrate three example heatmaps from three studies, showcasing the three studies with the largest sample sizes in each respective age group. The complete set of heatmaps for all studies can be found in Figure S2.

The findings revealed that a random pooled effect estimate identified 36% (95% *CI*: [0.2360; 0.4878]) of study participants (N = 59 910) as CAS (see Figure S3), suggesting that their sleep durations consistently fell within the recommended range of adequate sleep. Another, 50% (95% *CI*: [0.3881; 0.6217]) of participants (N = 109 932) were identified as CSS (see Figure S4), indicating sleep durations below the recommended threshold at one or more than one wave. The GSS, which was formed by extracting

Table 1. Dataset Characteristics

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Abdul Jafar et al (2021) [23]	Singapore	Chinese (54.3%), Malay (28.4%), Indian (17.3%); 46.3% female	3 months	3–54 months	7	15	Y	N	Parent-report	Total	62	27.0%	Short variable
											75	28.0%	Moderate consistent
											62	26.0%	Long consistent
											42	19.0%	Long variable
										Nighttime	73	28.0%	Short variable
											99	41.0%	Moderate consistent
											76	30.6%	Long consistent
										Daytime	67	25.0%	Short consistent
											84	34.0%	Moderate consistent
											54	20.0%	Long consistent
											38	21.0%	Long variable
Bakour et al (2020) [24]	US	Non-Hispanic white (67.4%); Non-hispanic black (15.1%), Hispanic (11.7%), Other (5.8%); 49.9% female	15.5 ± 0.2 years	12–32 years	4	7	N	Y	Self-report	Night	1471	14.2%	Consistently short
											8300	80.1%	Consistently adequate
											145	1.4%	Adequate-increasing
											446	4.3%	Long-decreasing
Cao et al (2022) [25]	China	Chinese (unspecified%); 47.25% female	1 month	1–24 months	4	8	N	Y	Parent-report	Total	187	14.4%	Short persistent
											235	15.5%	Intermediate and decreased slowly
											709	38.1%	Decreased to short
											577	32.0%	Long and decreased
										Nighttime	456	27.8%	Increased
											1104	61.8%	Long persistent
											148	10.5%	Decreased and then increased
										Daytime	162	9.4%	Decreased slowly
											846	49.6%	Short
											700	41.0%	Long and decreased
Chang et al (2018) [26]	Taiwan	Taiwan Chinese (unspecified%); 49.16% female	11.7 ± 0.5 years 0.45 years	12–18 years	7	9	Y	Y	Self-report	Time in bed	506	20.0%	Short declining
											1659	66.0%	Typical sleep
											345	14.0%	Long sleep

Table 1. Continued

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Fang et al (2021) [27]	China	Chinese (unspecified%); 41.1% female	8.1 ± 0.9 years	8–14 years	5	7	Y	Y	Parent-report	Nighttime	227	11.5%	Rapidly decreasing
											1199	60.8%	Moderately decreasing
											547	27.7%	Persistent sleeping >=8 h/day
Fatima et al (2020) [28]	UK	White (97.3%), Ethnic Minorities (2.72%); 51.9% female	50–62 years	37–73 years	3	9	Y	N	Self-report	Total	11 330	27.6%	Poor
											16 770	40.8%	Healthy
											12 994	31.6%	Borderline poor
Gilmour et al (2013) [29]	Canada	Canadians (100%; no other info reported); 56.54% female	46.9 years	47–55 years	5	11	Y	N	Self-report	Nighttime	873	11.1%	Short
											4372	49.4%	Low-normal
											3249	37.0%	High-normal
											179	2.4%	Long
Gui et al (2022) [30]	China	Chinese (unspecified%); 49.38% female	42 days	42 days—36 months	8	9	Y	Y	Parent-report	Nighttime	36	17.6%	Increasing
											196	76.3%	Stable
											11	6.1%	Decreasing
										Total	51	21.5%	Short
											150	59.9%	Medium
											42	18.6%	Long
										Daytime	31	14.4%	Short
											155	60.4%	Medium
											57	25.2%	Long
Guimaraes et al (2021) [31]	Canada	White (100%); 45.1% female	9.6 ± 0.9 years	9–17 years	3	11	Y	N	Accele-rometry in baseline; Parent-report in follow ups	Nighttime	160	66.8%	Shallow decrease
											80	33.2%	Steep decrease
Hu et al (2022) [32]	China	Chinese (unspecified%); 52.9% female	>45 years	60–67 years	4	8	N	Y	Self-report	Nighttime	2480	32.4%	Dominant short
											1405	18.4%	Dominant healthy-long
											1875	24.5%	Long decreasing
											1895	24.8%	Short increasing
										Daytime	2530	48.9%	Short increasing
											1379	18.0%	Stable normal
											2530	33.1%	Long decreasing
Jansen et al (2022) [33]	US	US Caucasian (100%); 46% female	2 ± 0.1 years	2–5.5 years	6	7	Y	N	Parent-report	Nighttime	25	4.5%	Short duration
											201	37.3%	Steady 9 h
											333	58.2%	Longer, slightly decreasing

Table 1. Continued

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Kline et al (2021) [34]	US	White (49.6%), Black (24.6%), Chinese (11%), Japanese (11%), Hispanic (3.9%); 100% female	45–55 years	45–68 years	4	10	Y	Y	Self-report	Total	602	63.0%	Persistent insufficient
											1025	37.0%	Persistent sufficient
Machado et al (2021) [35]	Brazil	White (66%), Black (15.7%), Other (18.3%); 53.3% female	11 years	11–22 years	3	8	Y	N	Self-report	Total	51; 43	3.4%; 2.4%	Increase and maintenance women: increase and decrease
											713; 464	45%; 25.6%	Fast reduction and maintenance
											818; 1304	51.6%; 72.0%	Constant reduction
Machado et al (2021) [36]	Brazil	Brazilian (unspecified%); 53.7% female	11 years	11–22 years	3	8	N	Y	Self-report	Total	46; 37	3.4%; 2.4%	Men: increase and maintenance women: increase and decrease
											608; 401	45%; 25.6%	Fast reduction and maintenance
											696; 1127	51.6%; 72.0%	Constant reduction
Magee, Gordon & Caput (2014) [37]	Australia	Australian (unspecified%); 51.9% female	0–1 years	0–7 years	4	9	Y	Y	Parent-report in wave 1–3; self-report in wave 4	Nighttime	74	2.5%	Poor
											339	11.6%	Persistent short
											1190	40.6%	Typical
											1323	45.2%	Initially short
McMahon et al (2018) [38]	Colombia	European American (69%), African American (12%), Hispanic/ Latino (3%), Asian (11%), Native American (3%), mixed race (3%); 51% female	28 ± 4 years	21–35 years	5	9	Y	N	Actigraphy	Total	82	21.0%	Short
											172	44.0%	Intermediate
											137	35.0%	Long
Myllyntausta et al (2020) [39]	Finland	Finnish (unspecified%); 87% female	63.3 ± 1.1 years	63–68 years	5	8	Y	N	Actigraphy	Nighttime	227	54.0%	Shorter mid-range with increase
											139	33.0%	Longer mid-range with increase
											55	13.0%	Constantly short

Table 1. Continued

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Patte, Qian & Leatherdale (2017) [40]	Canada	White (70.6%), Black (2.4%), Asian (4.9%), Off-reserve Aboriginal (1.1%), Latin American/ Hispanic (1%), Others/missing (20%); 53.9% female	Grade 9	Grade 9–12	3	12	Y	N	Self-report	Total	688	9.3%	Short
											3233	43.7%	Low-normal
											1975	26.7%	High-normal
											651	8.8%	Long
Plancoulaine et al (2018) [41]	France	French (unspecified%); 46.8% female	2 years	2–5.5 years	3	8	Y	N	Parent-report	Nighttime	59	4.9%	Short
											576	47.8%	Medium-low
											448	37.2%	Medium-high
											54	4.5%	Long
											67	5.6%	Changing
Radmanish et al (2022) [42]	France	French (unspecified%); 47.4% female	2 years	2–5.5 years	3	9	N	Y	Parent-report	Nighttime	59	4.9%	Short
											576	47.8%	Medium-low
											448	37.2%	Medium-high
											54	4.5%	Long
											67	5.6%	Changing
Reynaud et al (2021) [43]	France	French (unspecified%); 43.7% female	2 years	2–5.5 years	3	7	N	Y	Parent-report	Nighttime	59	4.9%	Short
											576	47.8%	Medium-low
											448	37.2%	Medium-high
											54	4.5%	Long
											67	5.6%	Changing
Seeger et al (2011) [44]	Canada	French-speaking (unclear); 49.9% female	10 years	10–13 years	4	8	Y	Y	Parent-report	Time in bed	316	14.5%	Short
											1486	68.2%	10.5-hour
											376	17.3%	11-hour
Seegers et al (2016) [45]	Canada	Quebec (unclear); 53% female	2.5 years	2.5–10 years	8	8	Y	Y	Parent-report	Nighttime	72	6.0%	Short persistent
											47	3.9%	Short increasing
											628	52.7%	10-h
											445	37.3%	11-h
Smithson et al (2018) [46]	Canada	Caucasian (70.6%), others (29.4%); 48.4% female	3 months	3–24 months	8	10	N	Y	Parent-report	Nighttime	68	9.7%	Short
											276	39.2%	Intermediate
											359	51.1%	Long
										Total	126	17.9%	Short
											259	36.9%	Intermediate
											148	21.1%	Decline to short
											169	24.1%	Long
										Daytime	207	29.5%	Short
											242	34.40%	Decrease to short
											167	23.8%	Intermediate
											86	12.2%	Decrease to intermediate

Table 1. Continued

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Tham et al (2021) [47]	Singapore	Chinese (60.3%), Malay (26.8%), Indian (12.9%); 47.8% female	3 months	3–54 months	7	9	N	N	Parent-report	Total	91	27.0%	Short variable
											97	28.0%	Moderate consistent
											90	26.0%	Long consistent
											67	19.0%	Long variable
										Nighttime	100	28.0%	Short variable
											146	41.0%	Moderate consistent
											110	31.0%	Long consistent
										Daytime	87	25.0%	Short consistent
											118	34.0%	Moderate consistent
											69	20.0%	Long consistent
											73	21.0%	Long variable
Touchette et al (2009) [48]	Canada	White (92.1%),Black African (2.3%), Native Amerindian (0.3%), Arab (1.3%), Asian (0.3%), other (3.7%); 48.8% female	2.5 years	2.5–5 years	5	6	Y	Y	Parent-report	Nighttime	104	5.1%	Short-persistent
											100	4.9%	Short-increasing
											1077	52.4%	10-hour persistent
											776	37.7%	11-hour persistent
Touchette et al (2013) [49]	Canada	Quebec (unclear); 51.4% female	6 months	6–48 months	4	10	Y	N	Parent-report	Nighttime	49	4.9%	Short-persistent
											46	4.6%	Short-increasing
											474	47.6%	10-hour
											426	42.8%	11-hour
										Daytime	42	4.3%	Rapidly decreasing
											744	75.7%	Normally decreasing
											197	20.0%	Slowly decreasing
Wang et al (2020) [50]	China	Chinese (100%); 23.8% female	52.5 ± 11.8 years	53–57 years	3	10	N	Y	Self-report	Nighttime	879	1.7%	Low-stable
											3384	6.4%	Low-increasing
											8074	15.4%	Normal-decreasing
											40 262	76.5%	Normal-stable
Wang et al (2020) [51]	China	Chinese (100%); 53% female	46.3 ± 12.5 years	20–70 years	3	7	N	Y	Self-report	Total	4521	61.1%	Stable
											1835	24.8%	Decreasing
											1041	14.1%	Increasing

Table 1. Continued

Study	Region	Demographic information	Baseline age	Age/ grade span	Waves	GRoLTS score	Pre	OC	Measure-ment type	SDT kind	# each group	% each group	Labels of SDT
Xavier et al (2020) [52]	Brazil	White mother (73.2%), Others (26.8%); 48.1% female	3 months	3–48 months	4	8	Y	N	Parent-report	Total	348	9.1%	Short
											2757	72.1%	Typical
											719	18.8%	Initially longer
Yong et al (2019) [53]	France	French (unspecified%); 40% female	2 years	2–5.5 years	3	6	Y	N	Parent-report	Nighttime	59	4.9%	Short
											576	47.8%	Medium-low
											448	37.2%	Medium-high
											54	4.5%	Long
											67	5.6%	Changing
Zheng et al (2021) [54]	Australia	Australian mother (78.9%), Others (21.1%); 47.2% female	4 months	4–60 months	5	5	N	Y	Parent-report	Nighttime	80	15.1%	Short stable
											126	23.9%	Catchup long
											322	61.0%	Long stable
										Total	88	16.6%	Short
											300	56.9%	Mid
											140	26.5%	Long-total
										Daytime	318	60.2%	Short
											155	29.4%	Mid
											55	10.4%	Long
Zhu et al (2021) [55]	China	Chinese (Unspecified%); 51.1% female	>65 years	76–84 years	4	7	N	Y	Self-report	Total	478	11.0%	Short stable
											3343	76.9%	Medium stable
											526	12.1%	Long increased
Zitser et al (2020) [56]	UK	Not reported (Unspecified%); 19.09% female	42.3 ± 5.03 years	42–69 years	5	9	N	Y	Self-report	Nighttime	29	4.7%	5 hours
											228	37.2%	A quadratic 6 hours
											278	45.4%	A quadratic 7 hours
											78	12.7%	8 hours

Abbreviations: Pre, Predictor; OC, Outcome; SDT, Sleep duration trajectory; Y, Yes; N, No.
Demographic note: Most studies reported a balanced sex/gender distribution, although a few were gender-specific based on the study's focus. While some studies provided ethnically diverse samples, many were region-specific, predominantly from North America, Europe, and Asia. However, detailed reporting of racial and ethnic composition was limited, restricting comprehensive subgroup analyses.

the shortest SDT group among all-time points from each study, accounted for 15% (95% *CI*: [0.1122; 0.1870]) of the participants (N = 25 568; Figure S5). No outlier studies that could significantly affect the primary results were identified in the sensitivity analyses conducted on the three proportion meta-analyses (see Figure S6–8 for sensitivity analysis forest plots for CAS, CSS, and GSS proportions).

In our meta-regression analysis, we found that the number of waves in the studies did not significantly affect the proportions of CAS, CSS, and GSS. However, the number of sleep trajectory groups within each study was associated with a decrease in the proportion of both GSS (β = –0.0815, 95% *CI*: [–0.1209, –0.0422], $p < .001$) and CAS (β = –0.1922, 95% *CI*: [–0.3420, –0.0425], $p < .005$). Furthermore, publication year significantly moderated the proportion of GSS (β = 0.0121, 95% *CI*: [0.0015, 0.0227], $p = .0268$), suggesting a temporal increase in this group over time. This finding may reflect a worsening trend in population sleep patterns, where shorter sleep durations are becoming more prevalent. Additionally, this trend reflected improvements in the sensitivity and precision of GBTM to detect short SDTs in more recent studies. Detailed results for moderating factors within the GSS are provided in Table 2, while those for CAS and CSS can be found in Table S6a and Table S7a, respectively.

Subgroup analyses revealed significant differences in the age range and sampling regions among GSS. Specifically, in the 0–5 years age group, 10.20% (95% *CI*: [0.0588; 0.1452]) were defined as GSS, while in the 6–18 years age group, the proportion was 21.42% (95% *CI*: [0.1229; 0.3054]), and among adults (over 18 ages), it was 17.21% (95% *CI*: [0.0886; 0.2555]). Additionally, the proportion of GSS in non-western regions (19.46%, 95% *CI*: [0.1356; 0.2537]) was higher than that of western regions (11.81%, 95% *CI*: [0.0707; 0.1655]; Table 3). Subgroup analyses for CAS and CSS can be found in Table S6b and Table S7b.

Association between SDTs and health outcomes

The qualitative synthesis of the results can be seen in Table S5. The three-level meta-analysis indicated a significant overall effect size estimate of *OR* = 1.54 (df = 68, $p < .001$, 95% *CI* [1.18;

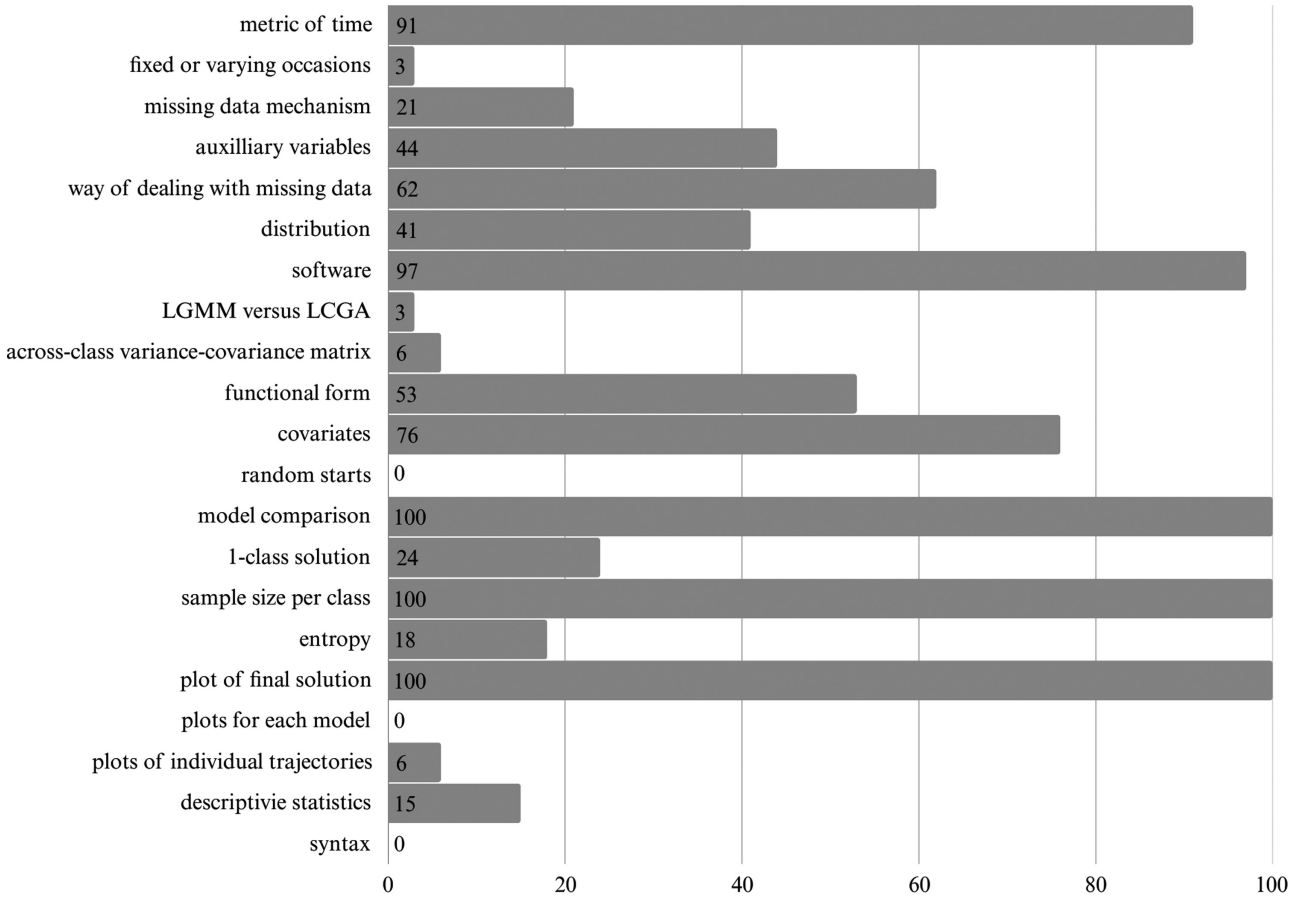


Figure 2. Articles (%) fulfilling individual Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) items. LGMM, latent growth mixture modeling; LGCA, latent class growth analysis. Total number of articles examined for systematic review is *N* = 34. The exact percentages are given in each bar.

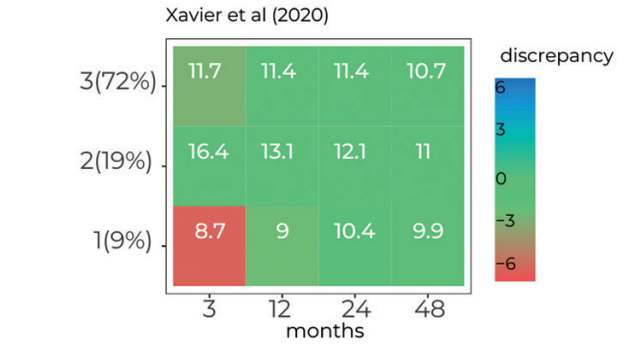
1.90]), suggesting that SDTs predicted levels of risk for the health outcomes examined. The estimated variance components were $\tau^2_{\text{Level 3}} = 0.26$ ($SD = 0.51$) and $\tau^2_{\text{Level 2}} = 0.34$ ($SD = 0.59$). The I^2_{Level3} = 56.51% of the total variation can be attributed to between-cluster, and I^2_{Level2} = 42.67% to within-cluster heterogeneity. We found that the three-level model provided a significantly better fit compared to a two-level model with level three heterogeneity constrained to zero ($\chi^2 = 17238.01$, $p < .001$). There was no significant publication bias (Egger's test $t = -0.53$, $p = .60$; Figure S9), and the sensitivity analysis confirmed the significant association between SDT groups and outcomes (*OR* with 95% *CI* ranging from 1.45 [1.19; 1.72] to 1.59 [1.18; 2.00]), demonstrating that the results were robust and reliable.

In the subgroup analysis, we found that only the "short" groups under the GBTM-based classification method, compared with "adequate" group, exhibited the most substantial average effect (*OR* = 1.54), followed by, "fluctuating" SDT group (*OR* = 1.47), "long" SDT group (*OR* = 1.42), and "decreasing" SDT group (*OR* = 1.32; Table 4).

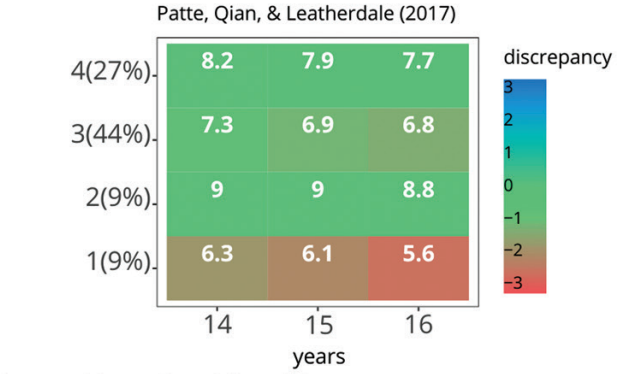
The moderating analysis revealed several important findings. Firstly, studies comprising more waves were more likely to show higher OR of SDTs' impact on health outcomes (F (1,67) = 5.23, $p = .025$). Secondly, age range substantially influenced the effect size (F (2,66) = 3.33, $p = .042$). Compared to adults (baseline age over 18 years old), the younger age group (6–18 years old) showed a greater propensity for serious adverse effects (mean *OR* = 2.13). However, outcome type was not a significant moderator of effect sizes ($F(2,66) = 1.69$, $p = .193$), indicating that the differences among physical, emotional and behavioral, and neurocognitive outcomes were not statistically significant. This suggested that while individual outcomes may exhibit different pooled effect sizes, these differences did not explain the overall heterogeneity of the data. The separate random-effects multivariate models for each outcome type, which were used to estimate their pooled effect sizes independently, are shown in Table S8.

Discussion

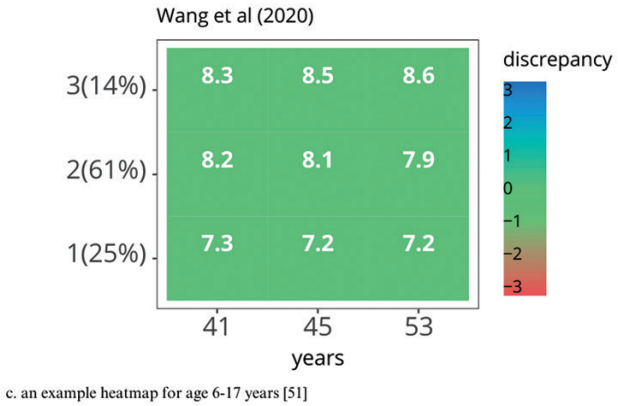
This study aimed to explore the SDT patterns across age groups throughout life, how they align with clinical sleep standards, and their associations with health outcomes. The results revealed significant variations in SDTs across different ages, with each group exhibiting unique patterns of sleep duration changes. While there was some agreement between GBTM-defined SDT groups and clinical criteria, particularly in the significant numerical difference between the proportions of short sleepers identified by clinical criteria versus GBTM, highlighting the need for more refined methods to accurately capture the complexity of sleep patterns. The meta-proportion analysis showed a higher proportion of participants classified as CSS compared to GSS. Moderator analyses indicated that age and region significantly influenced the proportion of GSS, but not CSS. The three-level meta-analysis emphasized the predictive role of SDTs in health outcomes, with the "short" SDT group, among adolescents, showing the most



a. an example heatmap for age 0-5 years [52]



b. an example heatmap for age 6-17 years [40]



c. an example heatmap for age 6-17 years [51]

Figure 3. (A) an example heatmap for ages 0–5 years [52]. (B) an example heatmap for ages 6–17 years [40]. (C) an example heatmap for age over 18 years [51]. Footnote: The Y-axis represents original GBTM groups, with numbers (1, 2, 3, etc.) indicating the SDT group number and their respective percentages. "1" consistently represents the shortest GBTM group. The X-axis indicates the time points under study. Color coding is as follows: green signifies sleep duration within the recommended range for the given age, the more intense the red, the shorter the sleep, and the more intense the blue, the longer the sleep. Actual sleep duration hours are displayed within the grid for clarity. By using a heatmap to illustrate each GBTM-defined group's sleep duration at different time points, researchers can quickly identify which trajectories consistently fall short of the recommended sleep durations and may therefore be at higher health risk. This visual representation not only enhances the interpretability of the data but also highlights the within-population variability that may be overlooked when clinical-based standards are used alone.

substantial adverse effects. These findings highlight the intricate nature of SDT patterns, underscoring the importance of continuous dynamic monitoring and personalized interventions for sleep duration.

Age-related heterogeneity in SDTs

Our study revealed age-related differences in SDTs. Unlike the linear decline suggested in previous reviews [58, 59], we found complex and varied patterns. Infants and young children (0–5 years) showed the most variable extended sleep durations. As sleep cycles stabilize, variations are most prevalent during early childhood, a period demanding extensive sleep due to rapid growth [60]. Adolescents (6–17 years) primarily fell under class C (decreased sleeper), possibly due to academic pressures, extracurricular commitments [61], and circadian rhythm shifts [62]. Adults predominantly maintained stable sleep patterns, though individual sleep needs varied. The least common pattern, increased sleepers (Class B), was observed in adults over 60, likely due to lifestyle changes and retirement [63].

Proportion variations in sleep duration patterns: GBTM insights and influencing factors

Our study discovered a notable discrepancy in the proportions of individuals classified as short sleepers, with CSS (50%) and GSS (15%). This large numerical difference highlights the variation in short sleeper identification depending on the methods used. The heatmap visualized GBMT-defined group's sleep duration and highlighted the alignment with clinical sleep performance. By traditional clinical standards, individuals are classified as "CSS" if their sleep duration falls short of the recommended amount at one or more time points. In contrast, only individuals consistently exhibiting the shortest sleep durations across all studied time points are identified as "GSS." This discrepancy between 50% and 15% underlines the complexity involved in defining and interpreting short sleep durations, suggesting that the universally applied clinical sleep duration standards may not fully capture the intricacies of individual sleep patterns. This distinction between CSS and GSS is crucial, as it highlights how the measurement approach can influence the understanding of sleep health. Our findings suggested that many individuals classified as short sleepers by clinical standards may not consistently exhibit short sleep durations. Conversely, albeit less common, some individuals who consistently maintain shorter sleep durations, and thus fall into the GSS category, may not be identified as short sleepers based on clinical standards. This could potentially lead to the misclassification of an individual's sleep health status. To avoid this, sleep patterns across multiple time points should be considered.

Notably, clinical standards and GBTM standards may measure different aspects of sleep. Clinical classifications typically focus on adherence to recommended sleep durations at predefined thresholds, providing a clear benchmark for sleep duration. While GBTM captures longitudinal patterns and variability in sleep duration over time, accounting for within-group heterogeneity and longitudinal changes that clinical standards do not address. To improve accuracy in assessing sleep health, it is essential to consider sleep patterns across multiple time points rather than relying solely on static classifications. By integrating both clinical and trajectory-based approaches, a more comprehensive understanding of sleep health and its implications for overall well-being can be achieved.

The meta-regression analysis indicated that the number of trajectory groups in each study significantly influenced the proportions of GSS and CAS. This finding suggests that the methodological choices made in trajectory modeling, such as the criterion for subgroup size, can substantially impact the interpretation of sleep pattern proportions [34, 42]. It is, therefore, important to

Table 2. Results for Random Effects Meta-regression of Moderators for GBTM-Defined Shortest Sleepers

Variable	Univariate coefficient	z value	P value	95%CI	Estimated tau ²	R ²
Publication year*	0.01	2.32	.027	[0.00; 0.02]	0.01	11.74%
No. of trajectory groups	−0.08	−4.22	.000	[−0.12; −0.04]	0.01	33.87%
Waves	0.00	0.22	.828	[−0.02; 0.03]	0.01	0.00%
Quality score	0.02	1.71	.097	[−0.00; 0.04]	0.01	5.29%

*Analysis using the data collection year as a moderator was also conducted, yielding results consistent with those using the publication year ($F(1, 32) = 6.10$, $p = .019$). Although both years were analyzed as moderators, we present the publication year results in this table because it is consistently reported across all studies and is more precise than the estimated data collection year.

Table 3. Results for Subgroup Effect Size Analysis of GBTM-Defined Shortest Sleepers

Subgroups	No. of studies	Pooled prevalence (%)	95%CI	P value in between group comparison
Overall	34	0.15	[0.11; 0.19]	
Age group				.020
0–5	16	0.10	[0.06; 0.15]	
6–17	8	0.21	[0.12; 0.31]	
≥18	10	0.17	[0.09; 0.26]	
Region sampled				.031
Western	20	0.12	[0.07; 0.17]	
Non-western	14	0.19	[0.14; 0.25]	
Measurement type				.694
Subjective†	32	0.15	[0.11; 0.19]	
Objective‡	2	0.18	[0.00; 1.00]	

†Subjective measurement type includes self or parent-report sleep duration.
‡Objective measurement includes actigraphy.

consider methodological differences when drawing conclusions from such studies. Furthermore, our analysis also showed a significant increase in the proportion of GSS over time. This trend aligns with prior reviews linking the rising prevalence of short sleep to modern lifestyle factors such as work-related stress [64] and frequent mobile phone use [65].

Our subgroup analysis revealed significant age-related disparities in the proportion of GSS. The 6–18 age group displayed the highest proportion of short sleepers, consistent with previous research indicating that students often sacrifice sleep due to academic demands, electronic device usage, or lack of awareness about the importance of adequate sleep [66]. Furthermore, non-western regions displayed significantly higher proportions of GSS compared to western regions, likely due to cultural differences in sleep perceptions and behaviors. Specifically, in some non-western cultures, it is common to sacrifice sleep for work or study, a practice often viewed as a sign of dedication [67, 68]. Additionally, the correlation between longer working hours and short sleep might also explain this finding [69]. However, studies employing different measurement methods (subjective vs. objective) did not differ in the proportions of the GSS. This is conceivably due to only two studies including objective measurements [38, 39]. The small sample size may have also affected our ability to detect differences.

Sleep trajectories and health outcomes

Our meta-analysis revealed a significant overall effect size (OR = 1.54), indicating that SDT group membership has a strong association with health risks [70]. Unlike previous meta-analyses focusing primarily on specific sleep patterns [55, 71], our analysis includes multiple SDTs and health-outcome variables, offering a more comprehensive understanding of the relationship between SDTs and health outcomes.

Our analysis of SDT types indicated that only the GBTM classification method, but not the class-based or clinical-based classification methods, revealed significant subgroup differences in health outcomes. This finding might be attributed to the fact that the GBTM grouping is based on labels assigned in the original studies, which took into consideration the specific characteristics of SDTs within their respective populations. In contrast, the clinical-based classification relies on traditional standards, and the class-based grouping is primarily based on the overall trend of SDTs. Both classifications consider overall SDT trends in the entire population, potentially overlooking specific subgroup traits.

Additionally, the "fluctuating" group showed worse outcomes than the "adequate" group. However, we must interpret this result cautiously, as it is based on only two studies that used the same cohort to label the SDT group. The "short" labeled SDT groups had the highest average *OR* compared to other SDT groups defined by GBTM. Beyond "short" sleepers, we also examined other patterns. For example, the "decreasing" group, which includes the "rapidly decreasing group" in which adolescents significantly reduced their sleep duration from 9 to 6.5 hours within 6 years, were at higher risk compared to healthy sleepers [27]. The higher risk observed in the "decreasing" group underscores the potential health risks tied to sudden changes in sleep duration, especially during important growth periods like adolescence. Conversely, the "optimized" group, which included groups with increases in short sleep [32, 45, 54], as well as the "long and decreased" group [25], exhibited lower risks, suggesting that modifying sleep duration in both directions could potentially lead to optimal health outcomes.

Table 4. Results of (a) Subgroup and (b) Moderator Analyses for the Association of Sleep Duration Trajectories and Adverse Health Outcomes

Variables	*Studies	†ES	Intercept [95% *CI*]	*β* [95% *CI*]	Mean *OR*	*F* (df1, df2)	*P*	Level 2 variance	Level 3 variance
SDT subgroups									
Class-based‡						*F*(3,65) = 0.58	.633	0.40***	0.27***
A (persist sleepers)	7	26	1.58[1.13; 2.03]***		1.41				
B (increase sleepers)	3	4	1.29[0.52; 2.06]***	−0.29[−1.08; 0.49]	1.19				
C (decrease sleepers)	7	25	1.46[1.01; 1.91]***	−0.12[−0.53; 0.28]	1.33				
D (changing sleepers)	5	14	1.75[1.23; 2.27]***	0.17[−0.28; 0.61]	1.41				
Clinical standard-based‖						*F*(2,66) = 0.32	.725	0.31***	0.28***
Clinically adequate	5	13	1.54[1.06; 2.02]***		1.36				
Clinically short	13	40	1.57[1.20; 1.93]***	0.03[−0.38; 0.44]	1.55				
Clinically long	6	16	1.40[0.92; 1.89]***	−0.14[−0.64; 0.37]	0.92				
GBTM-based¶						*F*(5,64) = 3.39	.009	0.44***	0.19***
Adequate	3	4	0.80[0.12; 1.48]**		1.16				
Optimized	5	7	0.84[0.20; 1.48]**	0.41[−0.69; 0.77]	0.85				
Short	11	21	1.77[1.33; 2.20]***	1.07[0.34; 1.59]***	1.54				
Fluctuating	2	8	2.07[1.50; 2.64]***	1.27[0.52; 2.03]***	1.47				
Decreasing	7	15	1.70[1.22; 2.18]***	0.90[0.20; 1.60]***	1.32				
Long	6	14	1.81[1.32; 2.31]***	1.02[0.30; 1.73]***	1.42				
Moderators									
Region sampled						*F*(1,67) = 0.08	.774	0.39	0.26
Western	7	45	1.59[1.06; 2.12]***	0.11[−0.65; 0.87]					
Non-western	7	24	1.48[0.94; 2.02]***	−0.11[−0.87; 0.65]					
Baseline age						*F*(2,66) = 3.33	.042	0.21***	0.27***
Over 18	5	13	1.13[0.61; 1.66]***		1.22				
6–18	4	13	2.10[1.54; 2.66]***	0.97[0.21; 1.74]**	2.13				
0–5	5	43	1.44[0.94; 1.94]***	0.31[−0.41; 1.03]	1.31				
Publication year	14	69	1.56[1.10; 2.01]***	0.17[−0.02; 0.24]		*F*(1,67) = 2.62	.111	0.63***	0.21***
Quality	14	69	2.55[0.25; 4.85]**	−0.13[−0.42; 0.16]		*F*(1,67) = 0.78	.379	0.35	0.26
No of waves	14	69	0.50[−0.44; 1.45]	0.24[0.03; 0.45]**		*F*(1,67) = 5.23	.025	0.22***	0.27***
Research duration	14	69	1.57[0.73; 2.40]***	−0.00[−0.12; 0.11]		*F*(1,67) = 0.01	.94	0.38***	0.26***
Sleep duration kind						*F* (1, 67) = 1.92	.170	0.33***	0.26***
Total	3	12	1.03[0.22; 1.84]***						
Night	11	57	1.66[1.26; 2.05]***	0.62[−0.28 1.53]					
Outcome type						*F*(2,66) = 1.69	.193	0.32***	0.26***
Physical	8	34	1.36[0.90; 1.83]***		1.26				
Emotional behavior	3	22	2.12[1.40; 2.85]***	0.76[−0.10; 1.62]	1.68				
Neurocognitive	3	13	1.34[0.54; 2.14]***	−0.02[−0.95; 0.91]	1.13				

*Studies = number of studies;
†ES = number of effect sizes extracted from studies; β = the regression coefficient from the moderator analysis, indicating the estimated change (log odds ratio) associated with one unit of the moderator or compared to the reference category (for categorical moderators); Mean *OR* = mean effect size (odds ratio); *F* = the results of the omnibus test based on the *F* distribution; Level 2 variance = variance between effect sizes extracted from the same study; Level 3 variance = variance between studies.
‡: Class-based SDT groups were defined by the patterns of the SDT that we concluded under Result 3.2.
‖: Clinical standard-based SDT groups were defined by comparing the actual sleep duration to the recommended sleep duration [1]. Meeting the recommended sleep duration at all-time points is considered clinically adequate. Failing to meet the recommended sleep duration at any time point is considered clinically short. Sleeping more than the recommended sleep duration at any time point is considered clinically long.
¶GBTM-based SDT groups were defined by the original labels given in the studies.
Adequate includes dominant healthy-long [32], 10.5hrs [44], and 10hrs [45]; **optimized** includes long decreasing [25], short increasing [32, 45], low increased [50], and catch up long [54] as these labels reflect the sleep time tends to be optimized; **short** includes consistently short [24], short persistent [25, 45], short [26, 42–44], persistent insufficient [34], low stable [50], short total [54], and short stable [55]; **decreasing** includes decreasing [24], intermediate and decreased slowly [25], medium-low sleep, moderately decreasing [27], normal decreasing [50], and rapidly decreasing [27]; **fluctuating** is the group changes observed in opposite directions at different time points [42, 43] and **long** includes original long [26, 42, 43], increasing [51], and long increased [55] labels provided in individual studies. The labels provided by included studies and the comparison for three classifications can be found in Table S9.

Furthermore, our study demonstrated that age range can moderate the relationship between SDT groups and outcomes. Compared to their peers who had sufficient sleep, adolescents who experienced a declining trajectory in sleep patterns faced an elevated risk of developing emotional and behavioral problems, such as a higher likelihood of non-suicidal self-harm [27] and smoking [26]. This finding can be interpreted through the reward system theory. The theory proposes that poor sleep, particularly among adolescents in a critical stage of brain development, can disrupt the brain's reward system, thereby increasing the likelihood of engaging in behaviors that offer immediate gratification [72, 73]. However, our study could not analyze the interaction between age groups and outcome types due to limited effect sizes. Future research with larger effect sizes could allow for this nuanced analysis.

Limitations and implications

To our knowledge, this is the first systematic review and meta-analysis to assess the agreement between professionals' consensus standards and GBTM-defined SDT groups. It also focuses on the relationship between SDTs and health outcomes. This review has several strengths, including its coverage of healthy individuals across all age groups as well as its inclusion of longitudinal studies. However, despite our efforts to identify all relevant studies from four databases, the non-inclusion of unpublished and gray literature can lead to an overestimation of the effect size [74]. Secondly, while all included studies reported the number or percentage of sleep trajectory groups, the level of detail and focus on the characteristics of these groups varied significantly. Besides, since not all studies provided sleep data at each time point, we had to rely on figures and plots for those studies, which could introduce inaccuracies in data extraction or potential misinterpretation of the graphical data, even with the help of professional software. Furthermore, variations existed in how different studies operationalized sleep duration, namely time in bed, nocturnal sleep time, and in a few studies, objective sleep time. Most studies relied on subjective measures, such as parental reports, introducing inherent biases. This predominance of subjective methods limited our ability to analyze the influence of measurement type, potentially affecting the classification of SDTs and their relationship to health outcomes. Furthermore, we used broad age categories (0–5, 6–17, and ≥18), which may overlook nuanced developmental differences. This categorization was chosen to ensure sufficient sample sizes and statistical power, but it may limit the granularity of insights into specific developmental stages. Future research could benefit from more detailed age stratification to capture these subtleties. Finally, while we documented the potential confounders adjusted for in the studies assessing SDT and health outcomes in the footnote of Table S5, our meta-analysis was limited in its ability to examine other potential moderating factors that might influence the relationship between SDTs and health. We identified several common covariates across the studies, including age, sex, and educational level. Notably, even though studies controlled for different covariates based on their research objectives and populations, seven out of eight studies that reported both unadjusted and adjusted models showed no significant differences between models, suggesting minimal confounding effects on the relationship between SDT and health outcomes. We initially intended to evaluate the influence of sociodemographic characteristics on these outcomes. However, moderator analysis was not feasible due to a lack of available data on these covariates in several included studies. Future research may provide detailed sociodemographic data to enable a more thorough examination of their potential impact on the relationship between SDT and health outcomes.

Despite these limitations, this study offers important insights. Firstly, the utilization of GBTM for studying sleep patterns while emphasizing the need for standardization in GBTM methods is advisable because different standards can yield varying results (particularly with model indicators) affecting the number of trajectory groups identified. Since the number of groups acted as a moderator in our analysis, standardization is key for both consistency and reliability. Secondly, from a clinical perspective, our findings highlight the importance of considering individual and group variations in sleep patterns rather than relying solely on clinical sleep duration standards. Finally, our study revealed disparities in SDTs among different age groups and how these variations can impact health, potentially facilitating the use of interventions that target both age and SDT groups.

Future research should focus on several key areas. Firstly, careful consideration of group numbers and adherence to transparent reporting practices are vital for accurately interpreting GBTM results. Researchers should prioritize checking and reporting key assumptions underlying GBTM, such as within-class homogeneity and the handling of missing data. Standardizing these practices will enhance the reliability and comparability of GBTM analyses across studies. Additionally, efforts to standardize GBTM methods can improve the alignment between clinical sleep duration standards and individual sleep patterns, thereby improving the accuracy of sleep health assessments and facilitating cross-study comparisons. Secondly, to improve clarity and consistency in future studies, we recommend that researchers adopt standardized definitions and metrics for reporting sleep trajectory characteristics. This could include explicit reporting of both average sleep duration and its variability (e.g. SD), as well as providing data at multiple time points when possible. Such practices would facilitate a more comprehensive understanding of sleep patterns and their implications for health. Finally, studies exploring factors such as genetics, lifestyle, cultural variations, and environmental factors that shape sleep patterns across different age and population groups are warranted.

Conclusion

This research provided comprehensive insights into the variations in SDTs across different age groups and their associations with health outcomes. Each age group exhibited unique SDT characteristics, and the relationship between SDT membership and adverse health outcomes was influenced by age. Our analyses also notably unveiled a significant discrepancy between the SDT grouping and the prevailing clinical standards of sleep duration. This discrepancy underlines the pressing need for more personalized and targeted sleep interventions, especially for those within the identified high-risk populations. In conclusion, our findings provide a solid foundation for future research and significantly enrich the current understanding of the role of SDTs in health outcomes, which in turn serves as an important guide for the formulation of targeted interventions to enhance sleep health.

Supplementary material

Supplementary material is available at *SLEEP* online.

Funding

The work in this paper was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (EdUHK#18611419; EdUHK#18609423).

Acknowledgment

Particular thanks are extended to Dr. Marc Zahradnik for his thoughtful copyediting.

Disclosure Statement

Financial disclosure: none. *Nonfinancial disclosure:* none.

Data Availability

Data supporting this review's findings are available from the corresponding author upon request.

References

1. Pankowska MM, Lu H, Wheaton AG, *et al*. Prevalence and geographic patterns of self-reported short sleep duration among US adults, 2020. *Prev Chronic Dis*. 2023;**20**:E53. doi:10.5888/pcd20.220400
2. Chaput JP, Wong SL, Michaud I. Duration and quality of sleep among Canadians aged 18 to 79. *Health Rep*. 2017;**28**(9):28–33.
3. Wheaton AG, Claussen AH. Short sleep duration among infants, children, and adolescents aged 4 months–17 years—United States, 2016–2018. *MMWR Morb Mortal Wkly Rep*. 2021;**70**(38):1315–1321. doi:10.15585/mmwr.mm7038a1
4. Lau EYY, Lam YC, Lee JCK. Well-slept children and teens are happier and more hopeful with fewer emotional problems. *Child Indic. Res*. 2021;**14**(5):1809–1828. doi:10.1007/s12187-021-09823-2
5. Lo JC, Groeger JA, Cheng GH, Dijk DJ, Chee MWL. Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. *Sleep Med*. 2016;**17**:87–98. doi:10.1016/j.sleep.2015.08.021
6. Xie J, Li Y, Zhang Y, *et al*. Sleep duration and metabolic syndrome: an updated systematic review and meta-analysis. *Sleep Med Rev*. 2021;**59**:101451. doi:10.1016/j.smrv.2021.101451
7. Hirshkowitz M, Whiton K, Albert SM, *et al*. National Sleep Foundation's updated sleep duration recommendations: final report. *Sleep Health*. 2015;**1**(4):233–243. doi:10.1016/j.sleh.2015.10.004
8. Chaput JP, Gray CE, Poitras VJ, *et al*. Systematic review of the relationships between sleep duration and health indicators in the early years (0–4 years). *BMC Public Health*. 2017;**17**:91–107.
9. Short MA, Booth SA, Omar O, Ostlundh L, Arora T. The relationship between sleep duration and mood in adolescents: a systematic review and meta-analysis. *Sleep Med Rev*. 2020;**52**:101311. doi:10.1016/j.smrv.2020.101311
10. Van Veen MM, Lancel M, Şener O, Verkes RJ, Bouman EJ, Rutters F. Observational and experimental studies on sleep duration and aggression: a systematic review and meta-analysis. *Sleep Med Rev*. 2022;**64**:101661. doi:10.1016/j.smrv.2022.101661
11. Chaput JP, Dutil C, Sampasa-Kanyinga H. Sleeping hours: what is the ideal number and how does age impact this? *Nat Sci Sleep*. 2018;**10**:421–430. doi:10.2147/NSS.S163071
12. Matricciani LA, Olds TS, Blunden S, Rigney G, Williams MT. Never enough sleep: a brief history of sleep recommendations for children. *Pediatrics*. 2012;**129**(3):548–556. doi:10.1542/peds.2011-2039
13. Nagin DS. Group-based trajectory modeling: an overview. In: Piquero AR, Weisburd D, editors. *Handbook of quantitative criminology*. NY: Springer; 2010: 53–67.
14. Nguena Nguefack HL, Pagé MG, Katz J, *et al*. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. *Clin Epidemiol*. 2020;**12**:1205–1222.
15. Wang Y, Li B, Zhang C, Buxton OM, Redline S, Li X. Group-based sleep trajectories in children and adolescents: a systematic review. *Sleep Med Rev*. 2024;**75**:101916. doi:10.1016/j.smrv.2024.101916
16. Tham EK, Jafar NK, Koh CT, Goh DY, Broekman BF, Cai S. Sleep duration trajectories and cognition in early childhood: a systematic review. *Sleep Med Rev*. 2024;**74**:101912. doi:10.1016/j.smrv.2024.101912
17. Van De Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-checklist: guidelines for reporting on latent trajectory studies. *Struct Equ Model*. 2017;**24**:451–467.
18. Lounassalo I, Salin K, Kankaanpää A, *et al*. Distinct trajectories of physical activity and related factors during the life course in the general population: a systematic review. *BMC Public Health*. 2019;**19**:1–12.
19. Zein H, Tran V, Ghazy A, *et al*. *How to extract data from graphs using plot digitizer or getdata graph digitizer*. 2015. doi:10.13140/RG.2.2.17070.72002
20. Assink M, Wibbelink CJ. Fitting three-level meta-analytic models in R: a step-by-step tutorial. *Quant Methods Psychol*. 2016;**12**(3):154–174.
21. Barker TH, Migliavaca CB, Stein C, *et al*. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesisers of evidence. *BMC Med Res Methodol*. 2021;**21**:1–9.
22. Galland BC, Taylor BJ, Elder DE, Herbison P. Normal sleep patterns in infants and children: a systematic review of observational studies. *Sleep Med Rev*. 2012;**16**(3):213–222. doi:10.1016/j.smrv.2011.06.001
23. Abdul Jafar NK, Tham EK, Pang WW, *et al*. Association between breastfeeding and sleep patterns in infants and preschool children. *Am J Clin Nutr*. 2021;**114**(6):1986–1996. doi:10.1093/ajcn/nqab297
24. Bakour C, Schwartz SW, Wang W, *et al*. Sleep duration patterns from adolescence to young adulthood and the risk of asthma. *Ann Epidemiol*. 2020;**49**:20–26. doi:10.1016/j.annepidem.2020.07.003
25. Cao S, Liu J, Li Y, Xu S, Xia W. Associations of sleep duration with neurocognitive development in the first 2 years of life. *J Sleep Res*. 2022;**31**(5):e13554. doi:10.1111/jsr.13554
26. Chang LY, Chang HY, Wu WC, Lin LN, Wu CC, Yen LL. Dual trajectories of sleep duration and cigarette smoking during adolescence: relation to subsequent internalizing problems. *J Abnorm Child Psychol*. 2018;**46**(8):1651–1663. doi:10.1007/s10802-018-0414-x
27. Fang J, Wan Y, Zhang X, Su P, Tao F, Sun Y. Sleep duration trajectory during the transition to adolescence and subsequent risk of non-suicidal self-harm. *Eur Child Adolesc Psychiatry*. 2021;**31**:1–9. doi:10.1007/s00787-021-01768-9
28. Fatima Y, Bucks RS, Mamun AA, *et al*. Sleep trajectories and mediators of poor sleep: findings from the longitudinal analysis of 41,094 participants of the UK Biobank cohort. *Sleep Med*. 2020;**76**:120–127. doi:10.1016/j.sleep.2020.10.020
29. Gilmour H, Stranges S, Kaplan M, *et al*. Longitudinal trajectories of sleep duration in the general population. *Health Rep*. 2013;**24**(11):14–20.
30. Gui Y, Deng Y, Sun X, *et al*. Early childhood sleep trajectories and association with maternal depression: a prospective cohort study. *Sleep*. 2022;**45**(5). doi:10.1093/sleep/zsac037
31. Guimarães R de F, Mathieu ME, Reid RE, Henderson M, Barnett TA. Physical activity, screen time, and sleep trajectories from childhood to adolescence: the influence of sex and body

weight status. *J Phys Act Health*. 2021;**18**(7):767–773. doi:10.1123/jpah.2020-0389
32. Hu Z, Tian Y, Song X, Hu K, Yang A. Associations between incident asthma with comorbidity profiles, night sleep duration, and napping duration trajectories: a 7-year prospective study. *Int J Public Health*. 2022;**67**:1604939. doi:10.3389/ijph.2022.1604939
33. Jansen EC, Zhao W, Jones AD, *et al*. Early childhood diet in relation to toddler nighttime sleep duration trajectories. *Nutrients*. 2022;**14**(15):3059. doi:10.3390/nu14153059
34. Kline CE, Colvin AB, Pettee Gabriel K, *et al*. Associations between longitudinal trajectories of insomnia symptoms and sleep duration with objective physical function in postmenopausal women: the Study of Women's Health Across the Nation. *Sleep*. 2021;**44**(8). doi:10.1093/sleep/zsab059
35. Machado AKF, Wendt A, Baptista Menezes AM, Gonçalves H, Wehrmeister FC. Sleep duration trajectories from adolescence to emerging adulthood: findings from a population-based birth cohort. *J Sleep Res*. 2021;**30**(3):e13155.
36. Machado AKF, Wendt A, Menezes AMB, *et al*. Associations between sleep duration trajectories from adolescence to early adulthood and working memory, schooling and income: a prospective birth cohort study from Brazil. *Sleep Med*. 2021;**86**:40–47.
37. Magee CA, Gordon R, Caputi P. Distinct developmental trends in sleep duration during early childhood. *Pediatrics*. 2014;**133**(6):e1561–e1567. doi:10.1542/peds.2013-3806
38. McMahon DM, Burch JB, Wirth MD, *et al*. Persistence of social jetlag and sleep disruption in healthy young adults. *Chronobiol Int*. 2018;**35**(3):312–328. doi:10.1080/07420528.2017.1405014
39. Myllyntausta S, Pulakka A, Salo P, *et al*. Changes in accelerometer-measured sleep during the transition to retirement: the Finnish Retirement and Aging (FIREA) study. *Sleep*. 2020;**43**(7). doi:10.1093/sleep/zsz318
40. Patte KA, Qian W, Leatherdale ST. Sleep duration trends and trajectories among youth in the COMPASS study. *Sleep Health*. 2017;**3**(5):309–316. doi:10.1016/j.sleh.2017.06.006
41. Plancoulaine S, Reynaud E, Forhan A, Lioret S, Heude B, Charles M-A; EDEN mother–child cohort study group. Night sleep duration trajectories and associated factors among preschool children from the EDEN cohort. *Sleep Med*. 2018;**48**:194–201. doi:10.1016/j.sleep.2018.03.030
42. Radmanish M, Khalfallah O, Glaichenhaus N, *et al*. Sleep duration trajectories associated with levels of specific serum cytokines at age 5: a longitudinal study in preschoolers from the EDEN birth cohort. *Brain Behav Immun Health*. 2022;**21**:100429. doi:10.1016/j.bbih.2022.100429
43. Reynaud E, Forhan A, Heude B, Charles MA, Plancoulaine S. Night-sleep duration trajectories and behavior in preschoolers: results from a prospective birth cohort study. *Behav Sleep Med*. 2021;**19**(4):445–457. doi:10.1080/15402002.2020.1773467
44. Seegers V, Petit D, Falissard B, *et al*. Short sleep duration and body mass index: a prospective longitudinal study in preadolescence. *Am J Epidemiol*. 2011;**173**(6):621–629. doi:10.1093/aje/kwq389
45. Seegers V, Touchette E, Dionne G, *et al*. Short persistent sleep duration is associated with poor receptive vocabulary performance in middle childhood. *J Sleep Res*. 2016;**25**(3):325–332. doi:10.1111/jsr.12375
46. Smithson L, Baird T, Tamana SK, *et al*.; CHILD Study Investigators. Shorter sleep duration is associated with reduced cognitive development at two years of age. *Sleep Med*. 2018;**48**:131–139. doi:10.1016/j.sleep.2018.04.005
47. Tham EK, Xu HY, Fu X, *et al*. Variations in longitudinal sleep duration trajectories from infancy to early childhood. *Sleep Health*. 2021;**7**(1):56–64. doi:10.1016/j.sleh.2020.06.007
48. Touchette E, Côté SM, Petit D, *et al*. Short nighttime sleep-duration and hyperactivity trajectories in early childhood. *Pediatrics*. 2009;**124**(5):e985–e993. doi:10.1542/peds.2008-2005
49. Touchette E, Dionne G, Forget-Dubois N, *et al*. Genetic and environmental influences on daytime and nighttime sleep duration in early childhood. *Pediatrics*. 2013;**131**(6):e1874–e1880. doi:10.1542/peds.2012-2284
50. Wang YH, Wang J, Chen SH, *et al*. Association of longitudinal patterns of habitual sleep duration with risk of cardiovascular events and all-cause mortality. *JAMA Netw Open*. 2020;**3**(5):e205246. doi:10.1001/jamanetworkopen.2020.5246
51. Wang Y, Hou W, Siddiqi SM, Sun C, Han T, Yang J. Association of sleep trajectory in adulthood with risk of hypertension and its related risk factors: the China Health and Nutrition Survey. *J Clin Sleep Med*. 2020;**16**(4):515–521. doi:10.5664/jcsm.8254
52. Xavier MO, Bielemann RM, Carpena MX, Barros FC, Matijasevich A, Santos IS. Sleep duration trajectories from age 3 to 48 months in The Pelotas (Brazil) 2004 Birth Cohort Study. *Paediatr Perinat Epidemiol*. 2020;**34**(1):60–69. doi:10.1111/ppe.12628
53. Yong CY, Reynaud E, Forhan A, *et al*.; EDEN study group. Cord-blood vitamin D level and night sleep duration in preschoolers in the EDEN mother-child birth cohort. *Sleep Med*. 2019;**53**:70–74. doi:10.1016/j.sleep.2018.09.017
54. Zheng M, Hesketh KD, Wu JH, Heitmann BL, Downing K, Campbell KJ. Nighttime sleep duration trajectories were associated with body mass index trajectories in early childhood. *Pediatr Obes*. 2021;**16**(7):e12766.
55. Zhu Q, You Y, Fan L, Fan H. Associations between sleep duration patterns and cognitive decline trajectories in older Chinese adults. *Aging Clin Exp Res*. 2021;**33**:3057–3063. doi:10.1007/s40520-021-01851-w
56. Zitser J, Anatürk M, Zsoldos E, *et al*. Sleep duration over 28 years, cognition, gray matter volume, and white matter microstructure: a prospective cohort study. *Sleep*. 2020;**43**(5). doi:10.1093/sleep/zsz290
57. Zhang A, Fang J, Wan Y, Su P, Tao F, Sun Y. Joint trajectories of life style indicators and their links to psychopathological outcomes in adolescence. *BMC Psychiatry*. 2021;**21**:1–10.
58. Boulos MI, Jairam T, Kendzerska T, Im J, Mekhael A, Murray BJ. Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. *Lancet Respir Med*. 2019;**7**(6):533–543. doi:10.1016/S2213-2600(19)30057-8
59. Evans MA, Buysse DJ, Marsland AL, *et al*. Meta-analysis of age and actigraphy-assessed sleep characteristics across the lifespan. *Sleep*. 2021;**44**(9). doi:10.1093/sleep/zsab088
60. Cai S, Tham EKH, Xu H-Y, *et al*. Trajectories of reported sleep duration associate with early childhood cognitive development. *Sleep*. 2023;**46**(2). doi:10.1093/sleep/zsac264
61. Bauducco SV, Özdemir M, Gradisar M, Boersma K, Bayram-Özdemir S. Trajectories of insomnia symptoms and insufficient sleep duration in early adolescents: associations with school stress. *Sleep Adv*. 2022;**3**(1):zpac018. doi:10.1093/sleepadvances/zpac018
62. Guidolin M, Gradisar M. Is shortened sleep duration a risk factor for overweight and obesity during adolescence? A review of the empirical literature. *Sleep Med*. 2012;**13**(7):779–786. doi:10.1016/j.sleep.2012.03.016
63. Coutrot A, Lazar AS, Richards M, *et al*. Reported sleep duration reveals segmentation of the adult life-course into three phases. *Nat Commun*. 2022;**13**(1):7697. doi:10.1038/s41467-022-34624-8
64. Mao Y, Raju G, Zabidi MA. Association between Occupational Stress and sleep quality: a systematic review. *Nat Sci Sleep*. 2023;**15**:931–947. doi:10.2147/nss.s431442

65. Brautsch LA, Lund L, Andersen MM, Jennum PJ, Folker AP, Andersen S. Digital media use and sleep in late adolescence and young adulthood: a systematic review. *Sleep Med Rev.* 2023;**68**:101742. doi:10.1016/j.smrv.2022.101742
66. Perez Zarate R, Colman O, Blake SC, *et al*. Factors that influence sleep behaviors of high school students: findings from a semi-rural community in Georgia. *J Sch Health*. 2023;**93**(5):411–419. doi:10.1111/josh.13312
67. Cheung BY, Takemura K, Ou C, Gale A, Heine SJ. Considering cross-cultural differences in sleep duration between Japanese and Canadian university students. *PLoS One*. 2021;**16**(4):e0250671. doi:10.1371/journal.pone.0250671
68. Li S, Arguelles L, Jiang F, *et al*. Sleep, school performance, and a school-based intervention among school-aged children: a sleep series study in China. *PLoS One*. 2013;**8**(7):e67928. doi:10.1371/journal.pone.0067928
69. Afonso P, Fonseca M, Pires JF. Impact of working hours on sleep and mental health. *Occup. Med*. 2017;**67**(5):377–382. doi:10.1093/occmed/kqx054
70. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat - Simul Comput*. 2010;**39**(4):860–864. doi:10.1080/03610911003650383
71. Pomeroy A, Pagan Lassalle P, Kline CE, Heffernan KS, Meyer ML, Stoner L. The relationship between sleep duration and arterial stiffness: a meta-analysis. *Sleep Med Rev.* 2023;**70**:101794. doi:10.1016/j.smrv.2023.101794
72. Lewis RG, Florio E, Punzo D, Borrelli E. The Brain's reward system in health and disease. *Adv Exp Med Biol*. 2021;**1344**:57–69. doi:10.1007/978-3-030-81147-1_4
73. Mullin BC, Phillips ML, Siegle GJ, Buysse DJ, Forbes EE, Franzen PL. Sleep deprivation amplifies striatal activation to monetary reward. *Psychol Med*. 2013;**43**(10):2215–2225. doi:10.1017/S0033291712002875
74. Schmucker C, Bluemle A, Briel M, *et al*. A protocol for a systematic review on the impact of unpublished studies and studies published in the gray literature in meta-analyses. *Syst Rev*. 2013;**2**:1–7.