Sleep Medicine Reviews 89 (2026) 102326

Contents lists available at ScienceDirect

Sleep Medicine Reviews

journal homepage: www.elsevier.com/locate/smrv

Clinical Review

Thirty-five years of the Pittsburgh sleep quality index: a critical appraisal and a path forward for its contemporary use

Matteo Carpi [a,*], Sofia Fontoura Dias [b,c], Daniel Ruivo Marques [b,d,1], Claudio Liguori [e,f,1]

[a] Department of Human Neuroscience, Sapienza University of Rome, Rome, Italy
[b] Department of Education and Psychology, University of Aveiro, Aveiro, Portugal
[c] William James Center for Research (WJCR), University of Aveiro, Aveiro, Portugal
[d] CINEICC – Centre for Research in Neuropsychology and Cognitive Behavioral Intervention, Faculty of Psychology and Educational Sciences, University of Coimbra, Coimbra, Portugal
[e] Neurology Unit, University Hospital of Rome Tor Vergata, Rome, Italy
[f] Department of Systems Medicine, University of Rome Tor Vergata, Rome, Italy

ARTICLE INFO

Keywords:
Pittsburgh sleep quality index (PSQI)
Sleep quality
Sleep health
Clinical assessment
Psychometrics
Self-report scales

ABSTRACT

Since its publication in 1989, the Pittsburgh sleep quality index (PSQI) has become one of the most widely used self-report instruments in sleep medicine. Despite its enduring popularity, its extensive adoption has been accompanied by conceptual and methodological concerns regarding its psychometric foundations and the construct it purports to measure.

This review offers a critical appraisal of the PSQI nearly four decades after its introduction and outlines directions for its contemporary use. Five interrelated issues are examined: (i) the absence of a clearly specified construct model of "sleep quality"; (ii) the frequent use of the PSQI as a proxy for insomnia or sleep disturbance; (iii) the interpretive fragility of the conventional global score cutoff (>5); (iv) heterogeneity in component scoring logic and instability of factorial structure; and (v) the redundancy and recall vulnerability of retrospective sleep continuity estimates.

Drawing on the shift from "sleep quality" toward multidimensional sleep health, we propose reframing PSQI outputs within a theoretically grounded sleep health perspective. Accordingly, alternative strategies for scoring, reporting, and interpretation are discussed to preserve the value of this legacy instrument while improving its conceptual alignment with contemporary models of sleep assessment.

Practice points and research agenda

Practice points

1. The Pittsburgh sleep quality index should not be interpreted as a measure of insomnia severity or as a diagnostic tool for specific sleep disorders without careful consideration of its multifaceted content.
2. Reliance on the global score and conventional cutoff may overshadow clinically meaningful variability; examining component profiles and item-level responses can improve interpretation in both research and clinical settings.
3. Reframing questionnaire outputs within a multidimensional sleep health perspective may enhance their conceptual clarity and support more precise reporting and interpretation.

Research agenda

1. Future studies should formally evaluate alternative scoring approaches, including domain- and profile-based strategies, and assess their validity across diverse clinical and community samples.

* Corresponding author. Department of Human Neuroscience, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy.
E-mail address: matteo.carpi@gmail.com (M. Carpi).
[1] Co-last authors.

https://doi.org/10.1016/j.smrv.2026.102326
Received 24 March 2026; Received in revised form 19 May 2026; Accepted 14 June 2026
Available online 17 June 2026

Abbreviations

PSG	polysomnography
PSQI	Pittsburgh sleep quality index
REM	rapid eye movement
Ru-SATED	regularity, satisfaction, alertness, timing, efficiency, duration
SATED	satisfaction, alertness, timing, efficiency, duration
SWS	slow-wave sleep

2. The instrument's dimensional structure requires further examination using contemporary psychometric methods, with explicit consideration of construct specification and measurement models.
3. Longitudinal and interventional studies are needed to determine whether alternative scoring and profile-based interpretations improve sensitivity to change and clinical decision-making.

1. Introduction

In the intricate history of sleep research, the seemingly straightforward notion of "sleep quality" has long played the role of an elusive actor, standing in contrast to the more clearly delineated and objectively assessed concept of "sleep quantity". Despite its intuitive appeal and the understandable tendency to model it after successful constructs such as general health or quality of life, defining sleep quality inevitably runs into the intrinsically elusive nature of the construct itself. From the outset, the idea of "quality" appears to imply a subjective judgement, typically assessed via self-report, concerning the perceived characteristics of a sleep episode [1,2]. This, however, collides with the reduced or absent conscious awareness inherent to sleep, which renders individuals largely unaware of the physiological processes (i.e., objective sleep parameters) that are presumed to contribute to making sleep "good" or "poor" [3,4], along an implicit continuum that often overlaps with the equally ill-defined notion of "sleep disturbances".

Unsurprisingly, attempts to arrive at a shared definition of sleep quality have unfolded against the backdrop of the classic subjective–objective dichotomy in sleep research. Efforts have focused both on identifying the intrinsic, measurable sleep characteristics underlying a "good" sleep episode and on characterizing the nature of its subjective experience [1]. In line with this perspective, in 2017 the National Sleep Foundation convened a panel of experts to identify evidence-based indicators of good sleep quality across the lifespan, with the explicit aim of translating scientific knowledge into recommendations for the general public [2], actually reframing the problem from one of definition to one of measurement. After reviewing 277 studies, the task force agreed that so-called sleep continuity metrics (such as sleep latency, number of prolonged awakenings during nocturnal sleep, wake after sleep onset, and sleep efficiency) represent robust indicators of sleep quality across the lifespan. By contrast, and perhaps counter to commonsense expectations, no consensus was reached regarding sleep architecture or daytime functioning. For example, the panel concluded that a higher proportion of rapid eye movement (REM) sleep does not necessarily indicate better sleep quality, and it did not reach a shared position on the relationship between slow-wave sleep (SWS) and sleep quality. Notably, it was also acknowledged that self-reported sleep quality ratings may naturally reflect non-sleep-related phenomena and conditions, including mood, pain, and general health status [1,2], reinforcing the notion that what is commonly labeled as "sleep quality" is not fully captured by the experience of sleep per se, but rather emerges from the interplay of sleep-related processes with both trait-like and state-dependent factors. As a consequence, sleep quality remains an inherently complex construct that resists straightforward synthesis, distinct from objective descriptions of sleep and intrinsically shaped by the multifaceted nuances of individual experience.

Yet, despite this conceptual instability, sleep quality continues to occupy a central position in contemporary sleep research. The construct remains widely used in epidemiological studies, where it is frequently examined as an independent predictor of a broad range of health outcomes [5,6], with recent years witnessing a sustained and even increasing volume of publications explicitly referring to sleep quality in their aims and conclusions. In fact, this trend is also bibliographically evident: a simple PubMed search for "sleep quality" (all fields), conducted from database inception to 2025, shows a steady increase in indexed publications over the last ten years, with a recent rise from 4811 records in 2024 to 6066 in 2025.

However, in spite of these persistent uncertainties, a single legacy instrument has been routinely used to operationalize "sleep quality" since 1989, namely the Pittsburgh sleep quality index (PSQI). Originally developed by Daniel Buysse and colleagues to assess sleep disturbances in individuals with psychiatric diagnoses [7], the PSQI is a compact self-report questionnaire that has become familiar to sleep researchers and clinicians worldwide. It has been translated into more than 60 languages and comprises 18 items probing sleep-related experiences, habitual sleep patterns, and common sleep problems over the preceding month [7]. Through a scoring procedure that yields seven component scores (subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbance, use of sleep medication, and daytime dysfunction), a global score intended to summarize overall "sleep quality" is obtained.

Favored by its ability to respond to a clear assessment need shared by both sleep researchers and clinicians, the PSQI rapidly established itself as one of the reference instruments for evaluating sleep quality. Over time, the construct of sleep quality has become closely intertwined with the questionnaire itself, which is now widely used in large-scale studies of the general population [8,9] and commonly regarded as a screening tool for "sleep dysfunction" across clinical and non-clinical settings [10]. As of February 2026, the original 1989 article has been cited 13,728 times in PubMed and over 43,000 times on Google Scholar, with the number of studies mentioning or using the PSQI continuing to increase steadily year by year.

As the PSQI approaches four decades of use, the present perspective aims to take stock of its role in sleep research and clinical practice. Specifically, the paper offers a critical appraisal of the instrument, revisiting its merits and recurrent misuses in light of ongoing debates surrounding the concept of sleep quality, and situating the PSQI within the current research landscape by integrating perspectives from sleep medicine and psychometrics. Inspired by classical philosophical argumentation, the review is structured into two complementary sections: a *pars destruens*, which examines five core issues related to the PSQI that are well represented in the literature, and a *pars construens*, outlining constructive proposals for how the potential of the PSQI may be more coherently grounded and integrated within contemporary sleep research and assessment practices.

2. Scope and procedural approach

Consistent with this introductory outline, this article is conceived as a theory-driven perspective and narrative critical review rather than a systematic review or meta-analysis. Hence, its aim is not to provide an exhaustive synthesis of all psychometric and clinical studies of the PSQI, but to reassess selected conceptual, methodological, and interpretive issues that have become particularly relevant after nearly four decades of widespread use. To support this argument and encourage debate, we combined: (i) a critical reading of key conceptual, psychometric, and clinical literature on the PSQI and, more broadly, sleep quality and self-report sleep assessment; (ii) focused audits of recent literature to illustrate contemporary interpretive practices; (iii) an illustrative simulation of possible component-score configurations to examine structural properties of the scoring system; and (iv) heuristic examples mapping PSQI item and component information onto multidimensional sleep health domains. While some of these elements follow a structured approach, within the context of this review they are intended to support a conceptual appraisal and generate directions for future research, not to replace formal validation studies, systematic reviews, or meta-analyses.

3. Conceptual and methodological limitations of the PSQI

3.1. The PSQI construct of "sleep quality" is underspecified

The original validation study of the PSQI was not immune to the broader "sleep quality" conundrum. Drawing on their clinical expertise in psychiatric sleep medicine, Buysse and colleagues developed a standardized instrument to investigate sleep complaints that were highly prevalent among patients with psychiatric conditions [7]. In doing so, they explicitly acknowledged the difficulty of defining their construct of interest, which in 1989 was arguably even more nebulous than today. As they frankly acknowledged, "the exact elements that compose sleep quality, and their relative importance, may vary between individuals" and "because sleep quality is largely subjective, sleep laboratory measures may correlate with perceived sleep quality, but they cannot define it" [7, p. 194].

Faced with this definitional ambiguity, and unable to rely on some of the most recent conceptual and technical developments in clinimetric and psychometric theories at the time the PSQI was developed, the authors took a pragmatic and clinically grounded stance: rather than resolving the theoretical status of sleep quality, they proceeded to operationalize it. In this sense, the PSQI shifted the emphasis from construct definition to measurement, echoing a broader tendency in clinical research whereby instruments are sometimes developed in advance of a fully articulated theoretical framework (a dynamic that has been critically discussed, for instance, in the measurement of depression [11]). Therefore, by its very origin, the PSQI sought to quantify a construct that remained both loosely specified and potentially heterogeneous, spanning perceived restfulness, sleep continuity metrics, symptoms of sleep disorders, and aspects of daytime functioning, and the selection and weighting of these elements relied substantially on clinical judgement and experience with patients presenting sleep complaints.

Unsurprisingly, a similar logic is evident in more recent developments, such as the Brief Version of the PSQI (B-PSQI), which reduced the instrument to six items selected to preserve screening efficiency and measurement properties, with reported sensitivity and specificity comparable to the full version for identifying poor sleepers [12]. Although this methodological refinement applies contemporary psychometric techniques and effectively enhances the instrument's efficiency, it does not fundamentally alter the underlying conceptual ambiguity, leaving the construct of sleep quality still not fully specified.

Notably, one of the primary aims of the original PSQI was discrimination rather than exhaustive construct assessment. The instrument was designed to distinguish "good" from "poor" sleepers relative to clinical criteria, effectively separating healthy sleepers from patients with depression or sleep complaints [7]. However, as will be discussed in subsequent sections, this emphasis on classification entails trade-offs. Instruments optimized for case discrimination are not necessarily optimized for construct validity or explanatory depth, and an instrument that efficiently separates groups may not provide a transparent account of what, precisely, is being measured [13].

A further core issue emerges when the pragmatic origins of the PSQI are retrospectively interpreted through a more naturalistic perspective. From a pragmatic standpoint, sleep quality functions as a practically useful summary label for a range of indicators. From a naturalistic or realist standpoint, however, sleep quality may be implicitly treated as a coherent latent entity that causally gives rise to its observed indicators [14,15]. In measurement terms, the former view aligns more closely with a formative logic (the components jointly define the total score, much like education, income, and occupation define socioeconomic status), whereas the latter corresponds to a reflective model (a latent construct gives rise to component manifestations) [14], as schematically illustrated in Fig. 1. Hence, the problem arises when the PSQI total score is tacitly interpreted as if it indexed a well-defined causal latent variable, with such reification potentially leading to conceptually vague inferences. For example, when neuroimaging studies investigate the neural correlates of PSQI-measured sleep quality [16], it becomes unclear whether the target is a unified biological process or a heterogeneous aggregation of subjective complaints. Similarly, when randomized clinical trials report changes in PSQI total score as evidence that an intervention improves sleep quality [17–19], the interpretation depends critically on which aspect of sleep-related functioning is actually modified. Without a clearly specified construct model, the meaning of such changes remains partially indeterminate.

3.2. The component scoring logic is heterogeneous, and the factor structure is unstable

The unclear construct definition underlying the PSQI extends to its individual components and to the implied latent structure of the instrument. Beyond the diversity of item formats employed, the scoring logic is not homogeneous across components [7]. While some components are directly item-like (e.g., subjective sleep quality and use of sleep medication), corresponding exactly to the respondent's 0–3 rating, others are algorithmic composites, derived from summative procedures and mapped onto 0–3 categories using predefined thresholds (e.g., sleep latency and sleep efficiency categories based on specific cut-points that resemble objective measurement criteria). Furthermore, the weighting of items and components appears to reflect clinical intuitions and pragmatic decisions rather than an explicitly articulated construct model, resulting in blurred conceptual boundaries across domains. For example, the sleep disturbance component is obtained by summing the frequency ratings (0–3) of nine distinct sleep-related problems (including nightmares, breathing difficulties, and feeling too hot or too cold) and subsequently recoding the total into a single 0–3 score based on predefined thresholds. No explicit rationale is provided in the original scoring description regarding the equivalence or differential weighting of these heterogeneous complaints. Similarly, the daytime dysfunction component combines an item assessing the likelihood of dozing during daytime activities with another item investigating motivation or enthusiasm in doing things. Beyond potential overlap with affective symptoms, this scoring choice leaves the broader content of daytime dysfunction partially specified, with domains such as fatigue or cognitive difficulties not being directly represented. Although these features are transparent in the scoring procedure, they have rarely been examined critically, and the components are typically treated as fixed units, with little attention to their constituting elements.

Unsurprisingly, the psychometric consequences of this heterogeneity become evident in studies investigating the dimensional structure of the PSQI [10,20]. Given the instrument's structure and mixed item formats, factor-analytic investigations typically use the seven components, rather than the original items, as indicators of latent factors, and such studies have yielded non-unique solutions, with one-, two-, and three-factor models reported across different populations [10].

While a comprehensive review of the PSQI's dimensionality is beyond the scope of the present paper and has been addressed in previous systematic work [10,20], a brief overview of selected findings illustrates the variability in proposed structures. In the original validation study, the seven components were combined into a single global score based on clinical reasoning, but no formal factor analysis was conducted [7]. Subsequent research explicitly examining dimensionality has produced divergent results.

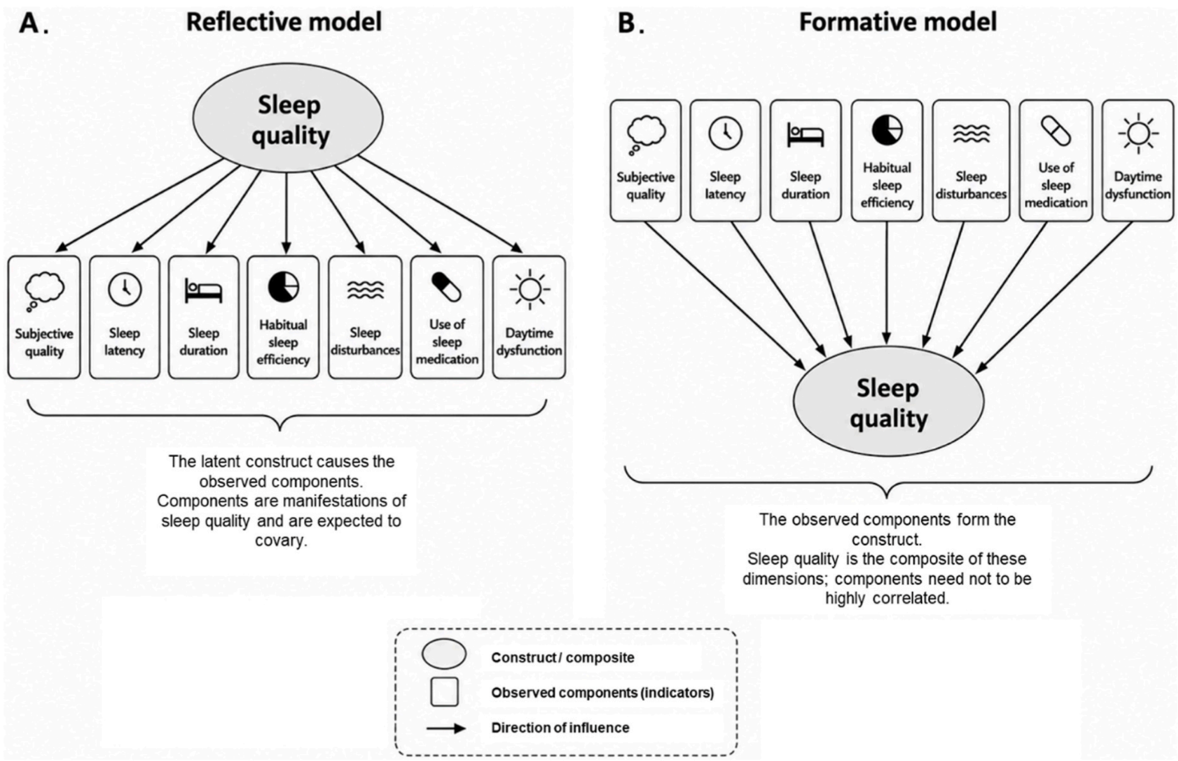

Fig. 1. Simplified representation of reflective and formative measurement models as applied to the Pittsburgh sleep quality index (PSQI). In a reflective model (Panel A), an underlying construct is assumed to give rise to the observed component scores. In a formative model (Panel B), the observed components combine to form the overall construct or composite score. The interpretive meaning of the PSQI global score differs under these two models.

Cole et al. conducted exploratory and confirmatory factor analyses in a community sample of older adults ($n = 417$) [21]. Exploratory factor analysis suggested a two-factor solution ("sleep efficiency" and "perceived sleep quality"), although the sleep medication component showed weak loading. Confirmatory analysis indicated better fit for a correlated three-factor model, comprising "sleep efficiency" (sleep duration and sleep efficiency), "perceived sleep quality" (subjective sleep quality, sleep latency, and sleep medication), and "daily disturbances" (sleep disturbances and daytime dysfunction).

In contrast, confirmatory factor analysis in a sample of Spanish adolescents and young adults ($n = 216$) supported a satisfactory one-factor solution [22]. In a large clinical sample of cancer-free women previously diagnosed with breast cancer ($n = 1174$), Otte et al. reported acceptable fit for a one-factor model, although a two-factor correlated solution ("sleep efficiency" and "perceived sleep quality") demonstrated superior fit [23]. More recently, an Australian study of 889 secondary school students supported a one-factor structure in both exploratory and confirmatory analyses, while also identifying residual correlations between related components (e.g., sleep duration and sleep efficiency; sleep efficiency and sleep latency) that required modeling [24].

Although this overview is deliberately selective, it illustrates the instability of the PSQI's factor structure across populations and contexts. Structural variability does not necessarily invalidate the instrument, but it complicates the assumption that the PSQI global score reflects a stable, unidimensional construct [10]. This observation further supports caution in relying exclusively on the global score and a fixed cutoff, particularly in longitudinal or interventional studies where changes in total score are interpreted as reflecting change in a coherent latent construct.

3.3. The PSQI is not a measure of insomnia severity or clinical sleep disturbance

Beyond the widely discussed conceptual ambiguity surrounding "sleep quality", the use of the PSQI in clinical contexts is widespread [10]. Here, a fundamental misunderstanding persists, namely the use of the PSQI as a measure of insomnia severity or, more broadly, clinical sleep disturbance, rather than as a measure of sleep quality. In this context, "sleep disturbance" is typically used as a broad umbrella term referring to clinically relevant complaints or disruptions of sleep, including difficulties with sleep initiation or maintenance, sleep-related breathing symptoms, parasomnia-like events, or daytime consequences [25,26]. Unlike insomnia disorder, however, it does not correspond to a single diagnostic construct with a unified symptom pattern, chronicity criterion, or severity metric.

The original purpose of discriminating sleep-disturbed patients likely contributed to these shifts, along with the scoring system in which higher PSQI scores correspond to poorer sleep quality. Implicitly equating poor sleep quality with clinical sleep problems is appealing and seemingly straightforward, yet it overlooks several critical elements.

This issue is particularly evident in chronic insomnia, where subjective sleep complaints (indeed including subjective sleep quality and sleep satisfaction) constitute the core of the diagnosis according to current diagnostic systems and clinical guidelines [27–29]. Although subjective assessments obtained via self-report instruments are rightly regarded as fundamental [30–32], the information gathered needs to be specific. In insomnia, this includes clinical interview, prospective assessment of perceived sleep continuity (e.g., sleep diaries) and quality, and structured evaluation of symptom severity [29] with respect to characteristic nighttime manifestations (difficulty initiating sleep, difficulty maintaining sleep, early morning awakenings) and associated daytime impairment (as captured by dedicated scales such as the insomnia severity index [33] and the sleep condition indicator [34]). By contrast, although several studies have proposed the PSQI as a screening tool for insomnia and reported acceptable sensitivity [35–37], and although its usefulness in tracking insomnia treatment effectiveness has been documented [38–40], the instrument does not meet the requirements expected of a clinical measure of insomnia severity. Beyond the items embedded within the sleep disturbance component that assess difficulty initiating and maintaining sleep and early awakenings, multiple sleep-related factors unrelated to insomnia contribute to the total score, and hence the PSQI global score cannot be considered a transparent index of insomnia severity [29]. The same reasoning applies more broadly to the domain of clinical sleep disturbance. Higher PSQI scores may reflect heterogeneous complaints, including subjective dissatisfaction, breathing-related symptoms, nightmares, sleep habits, medication use, or daytime impairment; such scores may indicate clinically relevant impairment, but they do not necessarily correspond to a specific sleep disorder construct.

To illustrate how this ambiguity may emerge in a particularly relevant clinical context, namely that of insomnia assessment and treatment, we conducted a focused audit of recent publications in which the PSQI was explicitly linked to insomnia severity. A PubMed search was performed on February 16, 2026 using the query: ("Pittsburgh Sleep Quality Index"[Title/Abstract] AND "insomnia severity"[Title/Abstract]). In line with the aim of addressing current usages of the instrument, the 100 most recent records were screened. Each study was coded according to whether the PSQI global score was explicitly interpreted as (1) a measure of sleep quality distinct from insomnia or sleep disturbance, (2) a measure of insomnia severity, (3) a measure of sleep disturbance, or (4) unclearly defined. Coding criteria were defined after preliminary inspection of a subset of records and applied to information contained in the full texts, with particular attention to the methods and interpretation sections. Overall, the audit was intended as a time-stamped and targeted illustration of interpretive practices in insomnia-related literature, rather than as an estimate of the prevalence of PSQI misuse across the full literature.

Of the 100 retrieved records (published between January 16, 2020 and December 25, 2025), three were excluded (two due to inaccessible full text and one retracted article), leaving 97 studies for analysis. Nineteen studies (19.6%) did not clearly distinguish PSQI-assessed sleep quality from insomnia and/or sleep disturbance. Specifically, four studies explicitly described the PSQI as a measure of insomnia symptoms or severity, and seven framed it as a measure of sleep disturbance. The remaining eight studies displayed conceptual ambiguity, conflating sleep quality with insomnia, sleep continuity, or clinical screening constructs. Notably, even among studies coded as category 1 (i.e., interpreting the PSQI as a measure of sleep quality distinct from insomnia), seven relied on PSQI cutoffs to classify participants as having insomnia and/or clinical sleep disturbances. The complete results of the audit search are available in Supplementary Table S1.

Clearly, these figures should not be interpreted as prevalence estimates for the broader PSQI literature; rather, they document the presence of interpretive ambiguity within a focused insomnia-related search context. Within this literature and research landscape, despite updated diagnostic criteria and clear guidelines on the assessment of insomnia, the interpretative confusion between sleep quality, insomnia severity, and generic sleep disturbance remains observable in recent insomnia-related literature. A redefinition of sleep quality within a more explicit conceptual framework, as discussed in the constructive section below, may help overcome this longstanding ambiguity.

3.4. The conventional cutoff as a shortcut, overshadowing informative heterogeneity

As noted above, Buysse et al.'s initial study emphasized discrimination between "poor" and "good" sleepers as a primary goal of the PSQI [7] and to this end, the authors sought to identify a total score that best differentiated healthy sleepers from patients with sleep complaints of varying etiologies. A cutoff >5 was found to optimize sensitivity and specificity within their sample, and since then this threshold has been widely adopted to dichotomize PSQI scores across heterogeneous populations [10], with alternative cutoffs occasionally proposed in specific contexts [41].

More generally, the practice of dichotomizing ordinal or continuous scores to derive clinical cutoffs has been extensively debated and criticized, with a central concern being the unjustified reduction of informative variability and the collapse of heterogeneous presentations into non-exhaustive categories derived from specific, and often non-representative, samples [42–46]. Two implicit assumptions underlie dichotomization: first, that there are no meaningful differences within categories (i.e., categories are internally homogeneous); and second, that the proposed cutoff captures the most informative variation in the underlying construct [42,45]. Both these assumptions are difficult to sustain in the case of the PSQI. The global score is the sum of seven components reflecting distinct and multifaceted sleep-related domains, and similar total scores may therefore arise from substantially different configurations of complaints. Moreover, from a numerical standpoint, the range of possible scores above the cutoff is substantially larger than the range below it.

To illustrate this point, we conducted a simple simulation using R (version 4.5.1) [47]. All theoretically possible combinations of the seven PSQI component scores (each ranging from 0 to 3) were exhaustively enumerated, yielding $4^7 = 16{,}384$ distinct component profiles. For each profile, the global PSQI score was computed as the sum of the seven components, allowing examination of the full theoretical distribution of total scores and the structural heterogeneity of profiles exceeding the conventional cutoff (>5). Notably, this simulation does not assume that all component configurations are equally likely to occur in empirical samples. Rather, it illustrates the structural heterogeneity permitted by the scoring system itself.

Across the complete combinatorial space, global scores range from 0 to 21 and follow a symmetric distribution centered around values of 10–11. Of the 16,384 possible profiles, 15,648 (95.5%) yield a global score greater than 5, whereas 728 profiles correspond to a score of exactly 6, the minimum value above the cutoff. Although this distribution reflects the mathematical structure of the scoring system rather than empirical prevalence, it highlights that the conventional threshold encompasses a very large region of the possible score space. Furthermore, profiles exceeding the cutoff display marked structural heterogeneity. Among configurations with global scores >5, some involve only two non-zero components ($n = 21$), whereas others reflect high scores across all seven components ($n = 2187$). The majority of profiles above the threshold include at least one maximally scored component (component score = 3; $n = 14{,}001$), while others reach the same classification status through more diffuse moderate increases across multiple domains. Even among profiles yielding the minimal value above the threshold (global score = 6; $n = 728$), configurations range from two highly scored components to more evenly distributed moderate scores.

Taken together, these numerical results illustrate that, at the level of the scoring system, the >5 cutoff does not correspond to a single phenomenological pattern, but rather aggregates a wide spectrum of structurally distinct profiles under the common label of "poor sleep". Clinically, this translates into similar global scores reflecting markedly different individual situations. As illustrated in Table 1, individuals with diverse sleep-related profiles (e.g., circadian delay, pain-related sleep fragmentation, or excessive daytime sleepiness) may obtain identical or comparable global scores, while meaningful heterogeneity is observed both below and above the conventional threshold.

Although the cutoff may still be useful in large-scale epidemiological studies [8–10], these considerations call into question its use for cohort splitting and, even more so, for individual-level classification without inspection of the underlying component profile. Moreover, as recalled above, alternative cutoffs have also been proposed across samples and clinical contexts, including a >8 threshold for the Korean version of the PSQI in the general population [41] and a higher cut point of 9 for breast cancer survivors [48]. Once again, this variability does not necessarily indicate that any single threshold is "wrong", but rather highlights that optimal cutoffs are sample-, purpose-, and criterion-dependent. A threshold chosen to maximize sensitivity in one context may not yield

Table 1
Illustrative Pittsburgh sleep quality index (PSQI) component configurations, including cases yielding similar global scores despite reflecting distinct underlying patterns. These examples highlight the interpretive limitations of relying on the conventional PSQI >5 cut-off to classify “poor sleepers”, showing how comparable global scores can arise from heterogeneous constellations of sleep-related complaints, such as circadian delay, restricted sleep opportunity, pain-related fragmentation, medication-supported sleep, early-morning awakening, excessive daytime sleepiness, or insomnia-like initiation and continuity difficulties. Profiles are intended to exemplify plausible clinical presentations rather than represent empirical prevalence.

Profile	Subjective quality	Latency	Duration	Efficiency	Disturbance	Medication	Daytime dysfunction	Global score	Pattern description
A. Short sleeper with restricted opportunity	1	0	2	0	1	0	0	4	Short time in bed with high sleep efficiency
B. Circadian delay (late chronotype)	1	2	1	1	1	0	0	6	Long sleep latency due to delayed phase, with otherwise restorative sleep
C. Pain-related nocturnal awakenings	2	0	1	1	2	0	0	6	Frequent awakenings from pain, normal sleep onset
D. Medication-supported sleep with residual issues	2	0	0	0	1	2	0	5	Hypnotic use stabilizes sleep initiation but residual disturbances persist.
E. Early-morning awakening pattern	2	0	2	1	1	0	1	7	Wakes too early with reduced sleep duration; mild daytime fatigue
F. High sleepiness but intact nocturnal sleep	1	0	0	0	1	0	3	5	Excessive daytime sleepiness despite normal sleep continuity
G. Sleep initiation + sleep continuity complaints	2	2	1	1	1	0	1	8	Difficulty initiating sleep; fragmented nights; mild daytime impairment

the same interpretive value when used for epidemiological stratification, treatment response, or individual-level classification.

3.5. *Retrospective sleep timing and continuity metrics are redundant and vulnerable to recall bias*

As a final point, the sleep timing and continuity metrics derived from the PSQI warrant specific consideration. Among its items, the PSQI requires respondents to retrospectively estimate habitual sleep latency, duration, and timing (bedtime and wake-up time). Based on this information, three components are scored: sleep latency, sleep duration, and sleep efficiency. These may appear as the most “objective” PSQI components, insofar as they map estimated quantities (minutes of latency, hours of duration, and percentages of efficiency) onto categorical 0–3 scores. Yet, several assumptions underlying this PSQI-derived triad deserve attention.

First, given the natural interdependence among latency, duration, and efficiency, the scoring of these components is intrinsically redundant. Two partially overlapping items contribute to the latency component, and both duration and latency mechanically constrain estimates of sleep efficiency. Moreover, while respondents directly report sleep duration (often rounded to full hours), time in bed is inferred from retrospective bedtime and wake-up times, and inconsistent combinations of reported sleep duration and time in bed may therefore yield implausible outputs (e.g., reported sleep duration exceeding inferred time in bed), unless responses are carefully screened and corrected.

Second, retrospective estimation is inherently vulnerable to recall bias [49], and sleep behaviors are similarly affected [50,51], with irregular sleep patterns such as those characterizing insomnia and other common sleep complaints being particularly prone to biased recall [4,51,52]. Accordingly, even within the domain of subjective measures, sleep questionnaires may produce estimates that diverge from prospective monitoring and may also be skewed by cognitively simplified responding or socially desirable patterns when compared with sleep diaries [51,52]. This risk may be even more pronounced for the PSQI, which asks respondents to summarize a broad and potentially variable period from the preceding month rather than reporting night-by-night variability.

Consistent with these concerns, a seminal study of PSQI criterion validity in 112 healthy volunteers examined associations between the PSQI global score and component scores and sleep parameters derived from weekly sleep diaries and actigraphic monitoring [53]. While the PSQI global score and the sleep latency, duration, and efficiency components showed significant, moderate-to-strong correlations with corresponding sleep diary outcomes (sleep disturbances and medication use were not correlated with sleep diaries, and daytime dysfunction showed only a small association with sleep diary-derived sleep latency), no significant association emerged between PSQI variables and actigraphy-derived sleep parameters. Notably, the PSQI global score was significantly associated with self-reported depressive symptoms, suggesting that part of what the questionnaire captures may reflect broader psychological status and processes (e.g., negativity bias or pessimistic thinking) rather than sleep-related experience alone [1,53].

In a subsequent study focusing specifically on self-reported sleep duration, PSQI-derived duration and sleep diary duration were compared against objective estimates from polysomnography (PSG) and actigraphy [54]. PSQI sleep duration showed statistically significant but small correlations with both PSG- and actigraphy-derived duration, whereas its association with sleep diary-reported duration was strong. Moreover, correlations between sleep diary estimates and objective sleep duration measured on corresponding nights were substantially higher than those observed between PSQI month-level estimates and objective measures, further emphasizing the distinctive nature of broad retrospective judgments.

Taken together, these considerations suggest that the temptation to interpret PSQI-derived quantitative sleep metrics as straightforward proxies of their objective counterparts should be tempered. Rather, these metrics should be regarded as distinct self-reported outcomes that likely reflect a complex interaction between habitual sleep patterns, their subjective perception, and individual trait- and state-dependent factors (e.g., extreme sleep or circadian patterns, mood states, vulnerability to recall distortions, or sleep-related attentional bias [55]).

4. Reframing the PSQI within sleep health: a path forward

4.1. *From sleep quality to multidimensional sleep health*

When considering the future of the PSQI, a major conceptual shift in sleep research plays a pivotal role: the introduction of multidimensional sleep health more than a decade ago. In 2014, Daniel Buysse, the first author of the original PSQI, proposed moving the focus from sleep disturbances to sleep health and its determinants, explicitly framing sleep in positive terms [56]. Drawing on broader conceptualizations of health centered on multifaceted well-being rather than the mere absence of disease [57], Buysse acknowledged the complexity of sleep and sought to identify those measurable dimensions most consistently associated with well-being.

Five core dimensions were eventually identified: sleep satisfaction/quality (subjective appraisal of "good" versus "poor" sleep), alertness/sleepiness (capacity to maintain active wakefulness), timing (placement of sleep within the 24-h cycle), sleep continuity or efficiency (ease of initiating and maintaining sleep), and sleep duration (amount of sleep obtained per day). These dimensions were operationalized in the SATED scale (an acronym reflecting the five domains) [56] and later expanded to include regularity (Ru-SATED), referring to the stability of sleep timing over time [58–60]. Subsequently, the multidimensional sleep health framework rapidly gained recognition within public health research [61] and has been associated with a range of physical and mental health outcomes [62–64], as well as overall mortality [65,66].

Notably, the growing relevance of sleep health has also prompted reflection on how outcomes from the multitude of existing self-report instruments assessing sleep-related factors [67] may be interpreted within this paradigm [68,69]. In this context, the large body of PSQI data (collected in the past, currently being collected, and likely to be collected in the future) may benefit from reconsideration under a multidimensional sleep health framework. Specifically, rather than treating the PSQI as a monolithic index of "sleep quality", its items and components may be mapped onto SATED dimensions, aligning the instrument's outputs with contemporary conceptual models.

While the intrinsic limitations discussed thus far remain relevant (e.g., reliance on retrospective reporting and the absence of repeated or period-specific assessments, which restricts evaluation of sleep variability and regularity over time), substantial correspondence can nevertheless be identified. As illustrated in Table 2, the PSQI global sleep quality rating readily maps onto the satisfaction dimension; daytime sleepiness (component 7, question 8) captures alertness; habitual bedtime and wake-up time allow derivation of timing-related indices; and sleep duration and efficiency are directly assessed through dedicated components. In contrast, the sleep disturbance and medication use components do not map cleanly onto SATED domains and may instead reflect complementary or disorder-specific information, whereas the increasingly recognized regularity dimension included in Ru-SATED is not directly assessed by PSQI items. However, even though it remains fundamentally heuristic and does not provide a validated reconstruction of the SATED/Ru-SATED scale, this mapping scheme could be explored as an alternative scoring approach, or as a harmonization template in multicenter studies employing heterogeneous sleep measures [68,70]. In this way, the PSQI can be repositioned not as an outdated or flawed instrument, but rather as a legacy tool whose content can be reinterpreted within the contemporary, multidimensional framework of sleep health.

Table 2
Proposed mapping of the Pittsburgh sleep quality index (PSQI) items and components onto the five sleep health dimensions (SATED). The original version of the Ru-SATED scale can be accessed on the website of the University of Pittsburgh Center for Sleep and Circadian Science (https://www.sleep.pitt.edu/research/measures-and-study-instruments; accessed on February 26, 2026). *Abbreviations:* Ru-SATED = regularity, satisfaction, alertness, timing, efficiency, duration.

Sleep health dimension	SATED definition	PSQI item(s)/component(s)	Mapping type	Notes
Satisfaction	Subjective appraisal of sleep quality	Q6 (component 1: subjective sleep quality)	Direct	Retrospective monthly evaluation
Alertness	Ability to maintain daytime wakefulness	Q8 (part of component 7: daytime dysfunction)	Direct	Conflated with enthusiasm/motivation (Q9) in original component 7
Timing	Placement of sleep within 24h cycle	Q1 (bedtime), Q3 (wake-up time)	Derived	Sleep midpoint must be inferred from reported bedtime and wake-up time; no information on regularity or variability
Efficiency	Ease of sleep initiation and maintenance	Component 2 (sleep latency), component 4 (sleep efficiency), and night awakenings item from component 5	Partial	Mixed scoring logic; retrospective
Duration	Total sleep obtained	Q4 (component 3: sleep duration)	Direct	Retrospective; monthly summary

4.2. *Beyond global score and cutoff: profiles, domains, and item-level outputs*

From a practical perspective, reconsidering PSQI outputs offers an opportunity to address the previously discussed concerns regarding reliance on the global score or isolated component scores for decision making in research and clinical contexts. This approach starts from the assumption that the instrument itself contains structurally rich information at the item level, and that this signal is not fully replaceable. Aggregated scores intended to reflect latent or superordinate constructs may satisfy discrimination criteria, yet they inevitably reduce information that may carry interpretable and potentially predictive value [13,71]. Nonetheless, in several straightforward ways, the informational potential inherent in the PSQI can be more fully leveraged.

First, rather than reporting only the global score or listing individual component scores, presenting the full component configuration in a synthetic visual format may facilitate recognition of domain-specific patterns (e.g., relatively higher scores across components). Such profile-based representations make explicit how similar global or component scores may correspond to structurally distinct sleep presentations. In this way, individual configurations can be more easily compared with data-driven profiles identified in clustering or latent profile analyses [72,73]. To illustrate this point, Fig. 2 displays graphical component profiles corresponding to the seven PSQI score configurations reported in Table 1.

While such profile-based representations should not be interpreted as a validated alternative scoring system, they illustrate a possible direction for future work: empirically derived profiles or domains could be tested in large samples, examined for stability across populations, and evaluated for their ability to predict clinical outcomes or treatment response. Until such validation is available, profile visualization may be best understood as an interpretive aid complementing, rather than replacing, conventional scoring.

Second, in research settings and clinical trials targeting specific sleep-related outcomes, aggregation may be replaced or complemented by selecting the most appropriate unit of analysis within the PSQI item set, thus encompassing the full range of information from individual items to global scores [13]. As noted earlier, several clinical trials of insomnia interventions have adopted the PSQI global score as a primary endpoint [17,19], yet more focused outcomes (such as sleep initiation difficulties, sleep efficiency, or specific nocturnal disturbances) can be examined directly using relevant components or items without the need for additional instruments. Furthermore, consistent with the mapping approach discussed in the previous section, proxies for sleep-related constructs not explicitly targeted by the original PSQI may be derived from theoretically grounded item combinations. For example, indices

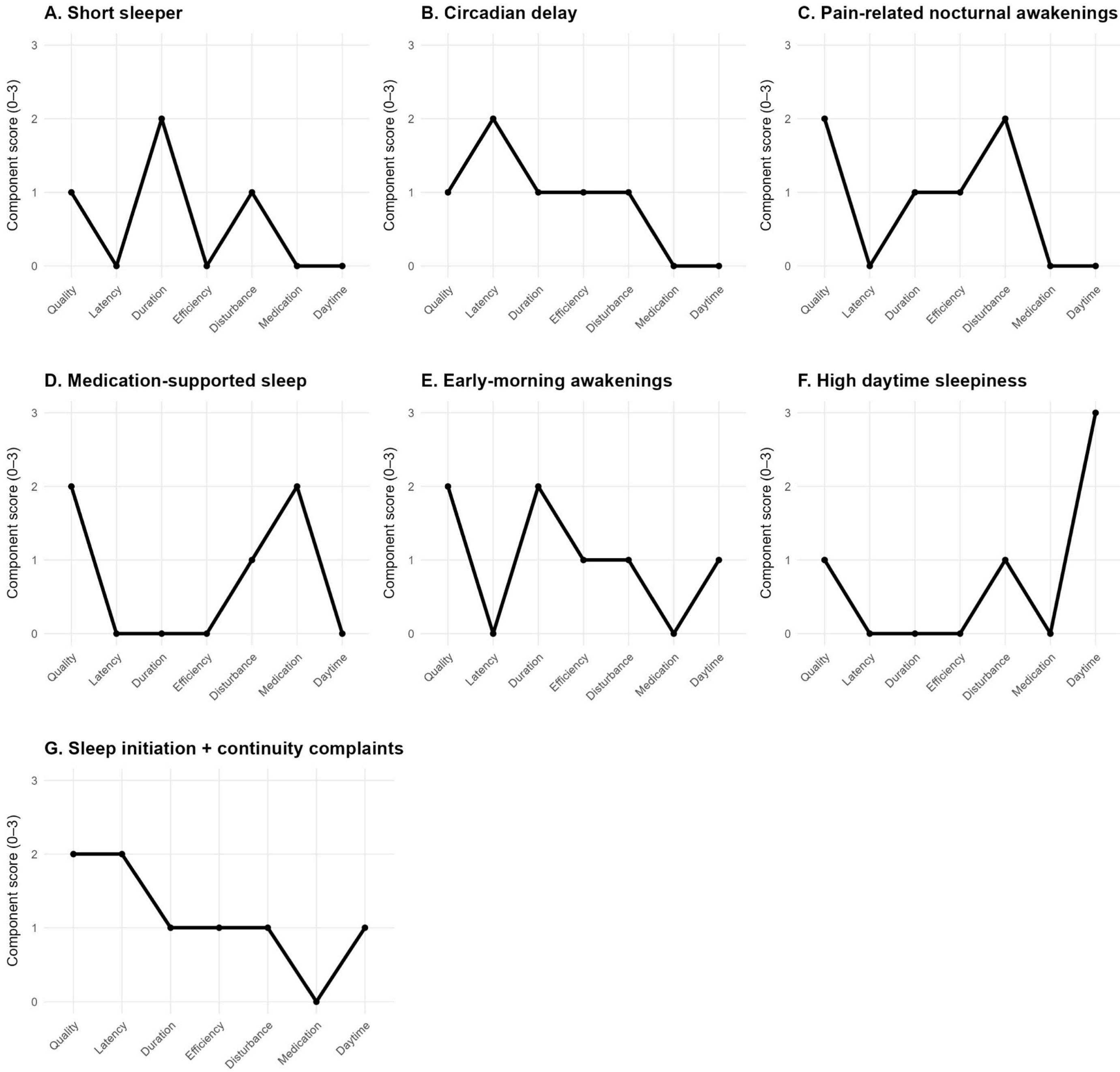

Fig. 2. Graphical representation of seven illustrative Pittsburgh sleep quality index (PSQI) component configurations corresponding to the score profiles presented in Table 1. Each panel shows the relative pattern of component scores (0–3) for a given profile. The figure illustrates how comparable global or component scores may correspond to structurally distinct self-reported sleep patterns.

reflecting specific disturbances such as breathing difficulties or periodic limb movements may be extracted from items contributing to the sleep disturbance component [68], and circadian-related proxies can be obtained from habitual bedtime and wake-up time information [74–76]. Although circadian assessment clearly lies outside the original scope of the PSQI, both habitual bedtime [75] and habitual sleep midpoint [76] derived from PSQI items have been shown to provide information that is independent of sleep duration.

Third, within comprehensive assessment batteries [67], the full PSQI protocol can serve as a structured informational resource for clinicians. Beyond recording aggregate scores for reporting purposes, item-level responses cover a wide range of sleep-related experiences and may help identify domains requiring further exploration during the clinical encounter. In this sense, the questionnaire can function not merely as a scoring tool but as a guide to deepen understanding of the patient's presentation, helping identify sleep-related domains requiring further exploration (e.g., sleep timing, nocturnal disturbances, daytime dysfunction) and potentially integrating quantitative responses with qualitative inquiry and supporting case conceptualization [77].

5. Conclusions

In this review, we critically examined the PSQI, arguably the most widely used self-report sleep instrument in research and clinical settings worldwide [10,67]. As the instrument approaches its fourth decade of use, it is timely to reassess its conceptual foundations, methodological assumptions, and common patterns of application within contemporary sleep research. Our aim was not to dismiss the PSQI, but to evaluate its limitations and potential within the evolving landscape of sleep science, integrating a clinical perspective with methodological and psychometric reflection.

In the critical part of our discussion, we emphasized structural and conceptual ambiguities around the construct of "sleep quality", highlighting the risk of conflating sleep quality, sleep disturbance, and insomnia severity. Furthermore, we discussed the implications of relying heavily on the global score and conventional cutoff, and we illustrated how such practices may overshadow the informative structural heterogeneity inherent in the PSQI's item composition.

In the constructive part, we proposed that these limitations can be interpreted as opportunities for methodological refinement. In particular, reframing the PSQI within the multidimensional sleep health framework [56,68] allows its content to be aligned with contemporary conceptual models. Moreover, leveraging item-level information and domain-specific configurations may enhance both research precision and clinical utility, supporting applications ranging from sleep epidemiology [78] to personalized sleep medicine [79,80].

Ultimately, critical appraisal of established instruments is part of the normal evolution of scientific measurement. In line with the notion of epistemic iteration [11,81], progress in sleep scale development and application depends on a reciprocal refinement of theory and measurement. Re-examining the PSQI through this lens may encourage novel psychometric investigation and empirical testing of alternative scoring approaches in large cohorts and well-characterized clinical samples, and such efforts can help the PSQI remain a valuable and adaptable instrument within contemporary sleep research rather than a legacy measure used uncritically.

Ethical approval and informed consent

No new data were generated or analyzed for this study and, therefore, ethical approval was not sought.

Author contributions

Matteo Carpi: Conceptualization; Methodology; Investigation; Data curation; Visualization; Writing - Original Draft; Writing - Review & Editing; Sofia Fontoura Dias: Investigation; Data curation; Visualization; Writing - Review & Editing; Daniel Ruivo Marques: Methodology; Visualization; Writing - Review & Editing; Claudio Liguori: Supervision; Visualization; Writing - Review & Editing.

Data availability statement

All data supporting the findings of this study are available within the article and its supplementary materials.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors used OpenAI's ChatGPT to refine the manuscript's grammar and language. All suggestions were critically reviewed, edited, and verified by the authors to ensure consistency with the intended meaning and scientific accuracy.

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.smrv.2026.102326.

References

[1] Krystal AD, Edinger JD. Measuring sleep quality. Sleep Med 2008;9:S10–7. https://doi.org/10.1016/S1389-9457(08)70011-X.
[2] Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, et al. National sleep Foundation's sleep quality recommendations: first report. Sleep Health 2017;3:6–19. https://doi.org/10.1016/j.sleh.2016.11.006.
[3] Stephan AM, Siclari F. Reconsidering sleep perception in insomnia: from misperception to mismeasurement. J Sleep Res 2023;32:e14028. https://doi.org/10.1111/jsr.14028.
[4] Harvey AG, Tang NKY. (Mis)perception of sleep in insomnia: a puzzle and a resolution. Psychol Bull 2012;138:77–101. https://doi.org/10.1037/a0025730.
[5] Willroth EC, Atherton OE, Sutin AR, Robins RW. Wellbeing trajectories and dementia risk among mexican-origin adults living in the U.S. Health Psychol 2026. https://doi.org/10.1037/hea0001570.
[6] Simpson DB, Gay CL, Berg S, English C, Ihle-Hansen H, Berg Helland G, et al. Sleep quality, not sleep apnoea, is associated with post-stroke fatigue in acute ischemic stroke: a cross-sectional Norwegian study (NORFAST). Disabil Rehabil 2026:1–11. https://doi.org/10.1080/09638288.2026.2620503.
[7] Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213. https://doi.org/10.1016/0165-1781(89)90047-4.
[8] Hinz A, Glaesmer H, Brähler E, Löffler M, Engel C, Enzenbach C, et al. Sleep quality in the general population: psychometric properties of the Pittsburgh sleep quality index, derived from a German community sample of 9284 people. Sleep Med 2017;30:57–63. https://doi.org/10.1016/j.sleep.2016.03.008.
[9] Lee SY, Ju YJ, Lee JE, Kim YT, Hong SC, Choi YJ, et al. Factors associated with poor sleep quality in the Korean general population: providing information from the Korean version of the Pittsburgh sleep quality Index. J Affect Disord 2020;271:49–58. https://doi.org/10.1016/j.jad.2020.03.069.
[10] Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: a systematic review and meta-analysis. Sleep Med Rev 2016;25:52–73. https://doi.org/10.1016/j.smrv.2015.01.009.
[11] Fried EI, Flake JK, Robinaugh DJ. Revisiting the theoretical and methodological foundations of depression measurement. Nat Rev Psychol 2022;1:358–68. https://doi.org/10.1038/s44159-022-00050-2.
[12] Sancho-Domingo C, Carballo JL, Coloma-Carmona A, Buysse DJ. Brief version of the Pittsburgh Sleep Quality Index (B-PSQI) and measurement invariance across gender and age in a population-based sample. Psychol Assess 2021;33:111–21. https://doi.org/10.1037/pas0000959.
[13] Carpi M, Marques DR. Shall we rush to be brief? Methodological and clinical costs of shortening insomnia scales. Sleep Breath 2026;30(59). https://doi.org/10.1007/s11325-026-03620-3. s11325-026-03620–3.
[14] Fried EI. What are psychological constructs? On the nature and statistical modelling of emotions, intelligence, personality traits and mental disorders. Health Psychol Rev 2017;11:130–4. https://doi.org/10.1080/17437199.2017.1306718.
[15] Borsboom D, Mellenbergh GJ, Van Heerden J. The theoretical status of latent variables. Psychol Rev 2003;110:203–19. https://doi.org/10.1037/0033-295X.110.2.203.
[16] Monti MM. The subcortical correlates of self-reported sleep quality. Sleep 2025;48. https://doi.org/10.1093/sleep/zsaf115. zsaf115.
[17] Wang S, Lan Y, Liu Z, Xu S, Wu X. Effects of different interventions on insomnia in adults: systematic review and network meta-analysis. J Psychiatr Res 2023;165:140–9. https://doi.org/10.1016/j.jpsychires.2023.07.004.
[18] Fatemeh G, Sajjad M, Niloufar R, Neda S, Leila S, Khadijeh M. Effect of melatonin supplementation on sleep quality: a systematic review and meta-analysis of randomized controlled trials. J Neurol 2022;269:205–16. https://doi.org/10.1007/s00415-020-10381-w.
[19] Zhang S, Zhao Y, Qin Z, Han Y, He J, Zhao B, et al. Transcutaneous auricular vagus nerve stimulation for chronic insomnia disorder: a randomized clinical trial. JAMA Netw Open 2024;7:e2451217. https://doi.org/10.1001/jamanetworkopen.2024.51217.
[20] Manzar MD, BaHammam AS, Hameed UA, Spence DW, Pandi-Perumal SR, Moscovitch A, et al. Dimensionality of the Pittsburgh sleep quality index: a systematic review. Health Qual Life Outcomes 2018;16:89. https://doi.org/10.1186/s12955-018-0915-x.
[21] Cole JC, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-Factor scoring model for the Pittsburgh sleep quality index in older adults. Sleep 2006;29:112–6. https://doi.org/10.1093/sleep/29.1.112.
[22] De La Vega R, Tomé-Pires C, Solé E, Racine M, Castarlenas E, Jensen MP, et al. The Pittsburgh sleep quality index: validity and factor structure in young people. Psychol Assess 2015;27:e22–7. https://doi.org/10.1037/pas0000128.

[23] Otte JL, Rand KL, Carpenter JS, Russell KM, Champion VL. Factor analysis of the Pittsburgh sleep quality index in breast cancer survivors. J Pain Symptom Manag 2013;45:620–7. https://doi.org/10.1016/j.jpainsymman.2012.03.008.
[24] Raniti MB, Waloszek JM, Schwartz O, Allen NB, Trinder J. Factor structure and psychometric properties of the Pittsburgh sleep quality index in community-based adolescents. Sleep 2018;41. https://doi.org/10.1093/sleep/zsy066.
[25] San L, Arranz B. The night and day challenge of sleep disorders and insomnia: a narrative review. Actas Esp Psiquiatr 2024;52:45–56.
[26] Matsumoto T, Chin K. Prevalence of sleep disturbances: sleep disordered breathing, short sleep duration, and non-restorative sleep. Respir Investig 2019;57:227–37. https://doi.org/10.1016/j.resinv.2019.01.008.
[27] American Academy of Sleep Medicine. International Classification of Sleep Disorders. Text revision. third ed. Darien, IL: American Academy of Sleep Medicine; 2023.
[28] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. American Psychiatric Association Publishing; 2022. DSM-5-TR.
[29] Riemann D, Espie CA, Altena E, Arnardottir ES, Baglioni C, Bassetti CLA, et al. The European insomnia guideline: an update on the diagnosis and treatment of insomnia 2023. J Sleep Res 2023;32:e14035. https://doi.org/10.1111/jsr.14035.
[30] Perlis ML, Grandner M, Posner D, Spiegelhalder K, Riemann D. Sleep diaries and other subjective measures are essential for the assessment of insomnia. J Sleep Res 2024:e14313. https://doi.org/10.1111/jsr.14313.
[31] Dikeos D, Wichniak A, Ktonas PY, Mikoteit T, Crönlein T, Eckert A, et al. The potential of biomarkers for diagnosing insomnia: consensus statement of the WFSBP task force on sleep disorders. World J Biol Psychiatr 2023;24:614–42. https://doi.org/10.1080/15622975.2023.2171479.
[32] Spiegelhalder K, Baglioni C, Morin CM, Nissen C, Palagini L, Perlis ML, et al. Insomnia disorder. Nat Rev Dis Primers 2026;12:17. https://doi.org/10.1038/s41572-026-00693-y.
[33] Bastien CH, Vallières A, Morin CM. Validation of the Insomnia severity index as an outcome measure for insomnia research. Sleep Med 2001;2:297–307. https://doi.org/10.1016/s1389-9457(00)00065-4.
[34] Espie CA, Kyle SD, Hames P, Gardani M, Fleming L, Cape J. The sleep condition indicator: a clinical screening tool to evaluate insomnia disorder. BMJ Open 2014;4:e004183. https://doi.org/10.1136/bmjopen-2013-004183.
[35] Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. Sleep 2006;29:1155–73. https://doi.org/10.1093/sleep/29.9.1155.
[36] Dietch JR, Taylor DJ, Sethi K, Kelly K, Bramoweth AD, Roane BM. Psychometric evaluation of the PSQI in U.S. college students. J Clin Sleep Med 2016;12:1121–9. https://doi.org/10.5664/jcsm.6050.
[37] Chiu H-Y, Chang L-Y, Hsieh Y-J, Tsai P-S. A meta-analysis of diagnostic accuracy of three screening tools for insomnia. J Psychosom Res 2016;87:85–92. https://doi.org/10.1016/j.jpsychores.2016.06.010.
[38] Morin CM. Measuring outcomes in randomized clinical trials of insomnia treatments. Sleep Med Rev 2003;7:263–79. https://doi.org/10.1053/smrv.2002.0274.
[39] Edinger JD, Buysse DJ, Deriy L, Germain A, Lewin DS, Ong JC, et al. Quality measures for the care of patients with insomnia. J Clin Sleep Med 2015;11:311–34. https://doi.org/10.5664/jcsm.4552.
[40] Okajima I, Komada Y, Inoue Y. A meta-analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia: CBT for insomnia: a meta-analysis. Sleep Biol Rhythm 2011;9:24–34. https://doi.org/10.1111/j.1479-8425.2010.00481.x.
[41] Sohn SI, Kim DH, Lee MY, Cho YW. The reliability and validity of the Korean version of the Pittsburgh sleep quality index. Sleep Breath 2012;16:803–12. https://doi.org/10.1007/s11325-011-0579-9.
[42] Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.1. https://doi.org/10.1136/bmj.332.7549.1080.
[43] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127–41. https://doi.org/10.1002/sim.2331.
[44] Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. BMC Med Res Methodol 2012;12:21. https://doi.org/10.1186/1471-2288-12-21.
[45] Binney ZO, Mansournia MA. Methods matter: (mostly) avoid categorising continuous data – a practical guide. Br J Sports Med 2024;58:241–3. https://doi.org/10.1136/bjsports-2023-107599.
[46] MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. Psychol Methods 2002;7:19–40. https://doi.org/10.1037/1082-989x.7.1.19.
[47] R Core Team. R: a language and environment for statistical computing. 2024.
[48] Reinsel RA, Starr TD, O'Sullivan B, Passik SD, Kavey NB. Polysomnographic study of sleep in survivors of breast cancer. J Clin Sleep Med 2015;11:1361–70. https://doi.org/10.5664/jcsm.5264.
[49] Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. Expert Rev Pharmacoecon Outcomes Res 2004;4:159–63. https://doi.org/10.1586/14737167.4.2.159.
[50] Babkoff H, Weller A, Lavidor M. A comparison of prospective and retrospective assessments of sleep. J Clin Epidemiol 1996;49:455–60. https://doi.org/10.1016/0895-4356(95)00529-3.
[51] Dietch JR, Sethi K, Slavish DC, Taylor DJ. Validity of two retrospective questionnaire versions of the consensus sleep diary: the whole week and split week self-assessment of sleep surveys. Sleep Med 2019;63:127–36. https://doi.org/10.1016/j.sleep.2019.05.015.
[52] Mallinson DC, Kamenetsky ME, Hagen EW, Peppard PE. Subjective sleep measurement: comparing sleep diary to questionnaire. NSS 2019;11:197–206. https://doi.org/10.2147/NSS.S217867.
[53] Grandner MA, Kripke DF, Yoon I-Y, Youngstedt SD. Criterion validity of the Pittsburgh sleep quality index: investigation in a non-clinical sample. Sleep Biol Rhythm 2006;4:129–36. https://doi.org/10.1111/j.1479-8425.2006.00207.x.
[54] Matthews KA, Patel SR, Pantesco EJ, Buysse DJ, Kamarck TW, Lee L, et al. Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, diary, and self-reported habitual sleep in a community sample. Sleep Health 2018;4:96–103. https://doi.org/10.1016/j.sleh.2017.10.011.
[55] Harris K, Spiegelhalder K, Espie CA, MacMahon KMA, Woods HC, Kyle SD. Sleep-related attentional bias in insomnia: a state-of-the-science review. Clin Psychol Rev 2015;42:16–27. https://doi.org/10.1016/j.cpr.2015.08.001.
[56] Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014;37:9–17. https://doi.org/10.5665/sleep.3298.
[57] Larson JS. The conceptualization of health. Med Care Res Rev 1999;56:123–36. https://doi.org/10.1177/107755879905600201.
[58] Wallace ML, Stone K, Smagula SF, Hall MH, Simsek B, Kado DM, et al. Which sleep health characteristics predict all-cause mortality in older men? An application of flexible multivariable approaches. Sleep 2018;41. https://doi.org/10.1093/sleep/zsx189. zsx189.
[59] Ravyts SG, Dzierzewski JM, Perez E, Donovan EK, Dautovich ND. Sleep health as measured by RU SATED: a psychometric evaluation. Behav Sleep Med 2021;19:48–56. https://doi.org/10.1080/15402002.2019.1701474.
[60] Kalkanis A, Lenkens D, Steiropoulos P, Testelmans D. Sleep regularity as an important component of sleep hygiene: a systematic review. Sleep Med Rev 2025;84:102203. https://doi.org/10.1016/j.smrv.2025.102203.
[61] Hale L, Troxel W, Buysse DJ. Sleep health: an opportunity for public health to address health equity. Annu Rev Public Health 2020;41:81–99. https://doi.org/10.1146/annurev-publhealth-040119-094412.
[62] Dalmases M, Benítez ID, Mas A, Garcia-Codina O, Medina-Bustos A, Escarrabill J, et al. Assessing sleep health in a European population: results of the Catalan health survey 2015. PLoS One 2018;13:e0194495. https://doi.org/10.1371/journal.pone.0194495.
[63] Appleton SL, Melaku YA, Reynolds AC, Gill TK, de Batlle J, Adams RJ. Multidimensional sleep health is associated with mental well-being in Australian adults. J Sleep Res 2022;31:e13477. https://doi.org/10.1111/jsr.13477.
[64] Buxton OM, Lee S, Marino M, Beverly C, Almeida DM, Berkman L. Sleep health and predicted cardiometabolic risk scores in employed adults from two industries. J Clin Sleep Med 2018;14:371–83. https://doi.org/10.5664/jcsm.6980.
[65] Wallace ML, Buysse DJ, Redline S, Stone KL, Ensrud K, Leng Y, et al. Multidimensional sleep and mortality in older adults: a machine-learning comparison with other risk factors. J Gerontol A Biol Sci Med Sci 2019;74:1903–9. https://doi.org/10.1093/gerona/glz044.
[66] Xing M, Zhang L, Li J, Li Z, Yu Q, Li W. Development and validation of a novel sleep health score in the sleep heart health study. Eur J Intern Med 2024;127:112–8. https://doi.org/10.1016/j.ejim.2024.05.002.
[67] Shahid A, Wilkinson K, Marcu S, Shapiro CM. STOP, THAT and one hundred other sleep scales. New York, NY: Springer New York; 2012.
[68] Wallace ML, Redline S, Oryshkewych N, Hoepel SJW, Luik AI, Stone KL, et al. Pioneering a multi-phase framework to harmonize self-reported sleep data across cohorts. Sleep 2024;47. https://doi.org/10.1093/sleep/zsae115. zsae115.
[69] Way JAH, Mathieu E, Chau J, Cistulli PA, Bin YS. Assessment tools for population sleep health surveillance in adults: a systematic review of validity and reliability. Sleep Med Rev 2025;84:102196. https://doi.org/10.1016/j.smrv.2025.102196.
[70] Appleton S, Theorell-Haglöw J. Why harmonizing cohorts in sleep is a good idea and the labor of doing so? Sleep 2024;47. https://doi.org/10.1093/sleep/zsae129. zsae129.
[71] Revelle W. The seductive beauty of latent variable models: or why I don't believe in the easter bunny. Pers Indiv Differ 2024;221:112552. https://doi.org/10.1016/j.paid.2024.112552.
[72] Yildirim A, Boysan M. Heterogeneity of sleep quality based on the Pittsburgh sleep quality index in a community sample: a latent class analysis. Sleep Biol Rhythm 2017;15:197–205. https://doi.org/10.1007/s41105-017-0097-7.
[73] Carpi M, Marques DR, Milanese A, Vestri A. Sleep quality and insomnia severity among Italian university students: a latent profile analysis. J Clin Med 2022;11:4069. https://doi.org/10.3390/jcm11144069.
[74] Reiter AM, Sargent C, Roach GD. Finding DLMO: estimating dim light melatonin onset from sleep markers derived from questionnaires, diaries and actigraphy. Chronobiol Int 2020;37:1412–24. https://doi.org/10.1080/07420528.2020.1809443.
[75] Carpi M, Marques DR. The complex association between sleep duration and mental and physical health: are we missing circadian preferences? Sleep Breath 2025;29:331. https://doi.org/10.1007/s11325-025-03520-y.
[76] Carpi M. When you don't have chronotype data: sleep questionnaires as a circadian window. Sleep Sci 2025;s-0045–1809926. https://doi.org/10.1055/s-0045-1809926.
[77] Adamson J, Gooberman-Hill R, Woolhead G, Donovan J. "questerviews": using questionnaires in qualitative interviews as a method of integrating qualitative and quantitative health services research. J Health Serv Res Policy 2004;9:139–45. https://doi.org/10.1258/1355819041403268.
[78] Ferrie JE, Kumari M, Salo P, Singh-Manoux A, Kivimaki M. Sleep epidemiology–a rapidly growing field. Int J Epidemiol 2011;40:1431–7. https://doi.org/10.1093/ije/dyr203.

[79] Garbarino S, Bragazzi NL. Revolutionizing sleep health: the emergence and impact of personalized sleep medicine. J Pers Med 2024;14:598. https://doi.org/10.3390/jpm14060598.
[80] Pevernagie DA, Arnardottir ES, Bruni O, Hartley S, Lammers GJ, Paunio T, et al. Sleep medicine-what's in a name? J Sleep Res 2025;34:e70150. https://doi.org/10.1111/jsr.70150.
[81] Kendler KS. Epistemic iteration as a historical model for psychiatric nosology: promises and limitations. In: Kendler KS, Parnas J, editors. Philosophical issues in psychiatry II. Oxford University Press; 2012. p. 305–22. https://doi.org/10.1093/med/9780199642205.003.0044.