Intraindividual variability in sleep among athletes: A systematic review of definitions, operationalisations, and key correlates

Shana Kemp[1,2], Angela L. Spence[1,3], Brad S. Keller[2], Kagan J. Ducker[1,3], and Daniel F. Gucciardi[1]

[1]*Curtin School of Allied Health, Curtin University*

[2]*Fremantle Football Club*

[3]*Curtin Health Innovation Research Institute, Curtin University*

Author Notes

*Address correspondence to Daniel Gucciardi, Curtin School of Allied Health, Curtin University, GPO Box U1987, Perth, Australia, 6845. Email: daniel.f.gucciardi@gmail.com

Conflicts of interest: we have none to report.
Word count (manuscript) = 4191
Word count (abstract) = 240
Number of tables = 0
Number of figures = 3

Kemp, S., Spence, A.L., Keller, B.S., Ducker, K.J., & Gucciardi, D.F. (in press). Intraindividual variability in sleep among athletes: A systematic review of definitions, operationalisations, and key correlates. *Scandinavian Journal of Medicine and Science in Sport*. doi: 10.1111/sms.14453

Abstract

Via systematic review with narrative synthesis of findings, we aimed to document the ways by which researchers have defined, operationalised, and examined sleep variability among athletes. We identified studies in which scholars examined intra-person variability in sleep among athletes via a search of six databases (Web of Science, Embase, Medline, PsycINFO, CINHAL Plus, and ProQuest Dissertations and Theses Global) using a protocol that included keywords for the target outcome (sleep*), population (athlet* OR sport*), and outcome operationalisation (variability OR variation OR “standard deviation” OR fluctuate OR fluctuation OR stability OR instability OR reactivity OR IIV OR intraindividual). We complemented this primary search with citation searching of eligible articles. Assessments of study quality captured 8 core elements, namely aims/hypotheses, sample size justification, sample representativeness, number of days sleep assessed, measures of sleep and its correlates, missing data, and inferences and conclusions. From a total of 1,209 potentially relevant papers, we identified 16 studies as meeting our eligibility criteria. Concept definitions of variability were notably absent from this work and where available were vague. Quantitative deviations from one’s typical level of target sleep metrics reflected the essence by which all but one of the research teams operationalised sleep variability. We assessed the overall quality of empirical work as moderate in nature. We propose a working definition of sleep variability that can inform knowledge generation on the temporal, day-to-day dynamics of sleep functioning that is required for personalised interventions for optimising sleep health.

Keywords: heterogenous variances; location-scale models; PRISMA

Review registration: We pre-registered the protocol for this systematic review on 6th April 2022 via the Open Science Framework (OSF: https://bit.ly/3MNy6DI)

Disclosure statement: We have no competing interests to declare.

1. Introduction

Good sleep is essential for optimal human health and functioning, particularly elite athletes who experience arduous physical and psychological strain [1]. The American National Sleep Foundation recommends that young adults (18-25 years) and adults (26-64 years) accrue at least 7-9 hours of sleep per 24-hour cycle [2] and experience short time to fall asleep after lights out (0-30 min), few awakenings greater than 5 min in duration (< 2), reduced wake after sleep onset (0-20 min), and sleep efficiency > 85% [3] to reap the full benefits of this health preventative and restorative bodily function. The extent to which these recommendations generalise to elite athletes remains unknown [4] and is likely challenging to pinpoint because of the diverse, often multifactorial physical and psychological conditioning programs required for certain sports. It is generally accepted among the sport science community that many elite athletes accrue insufficient amounts of sleep [4] and the quality of their sleep is suboptimal[5]. Much of the available work prioritises evidence on differences in mean levels of sleep metrics between individuals (inter-individual variability) for interventions or strategies [6–8], with little consideration of intra-individual variability for optimising sleep health. This gap in knowledge limits our potential to generate innovative tactics that might optimise sleep health, that is, multidimensional sleep patterns (e.g., duration, efficiency, quality) contextualised to personal and contextual factors which give rise to positive health and well-being [9]. The development of robust interventions for optimising sleep health requires knowledge on the temporal, day-to-day dynamics of sleep functioning in relation to personal (e.g., training load, psychological factors) and contextual (e.g., air travel, social dynamics) factors. We propose that this knowledge is best acquired via estimates of intra-individual variability (IIV) alongside mean levels in sleep metrics, yet such evidence is fragmented across the literature making it insufficient for theory and practice[1].

[1] We acknowledge there exists numerous ways by which to operationalise sleep consistency or variability, including others which focus on day-to-day variations in sleep episodes (awake/sleep) on consecutive days like the sleep regularity index [10] rather than within-person fluctuations across multiple days.

The scholarly literature on athlete sleep is relatively young yet burgeoning, with approximately 80% of total outputs published since 2011 [1]. Scholars have published several systematic reviews and meta-analyses to summarise what is currently known about athlete sleep and its role with athlete functioning. Regarding the importance of sleep for athlete health and performance, one statistical synthesis (k = 77, 227 effects, N = 959) indicated that sleep loss is detrimental to exercise performance (mean %Δ = − 7.56%, 95% CI − 11.9 to − 3.13), with subgroup analyses clarifying the maladaptive nature of sleep deprivation, sleep restriction (combination of late and early) and late restriction (earlier than normal waking), but not early restriction or delayed sleep onset [11]. Among the general population, meta-analytic evidence (k = 72, N = 8,608) supports a causal relation between sleep and mental health and specific indices including depression, anxiety, rumination, and stress ($\geq$ g = -.49), as well as a dose-response effect, whereby greater improvements in sleep lead to more adaptive experiences of mental health [12]. Regarding factors that promote optimal athlete sleep, one systematic review of sleep interventions (k = 10, N = 218) found that sleep extension provided the most benefit for performance, with mixed results for napping, sleep hygiene, and post-exercise recovery [6]. In another systematic review and meta-analysis (k = 27, N = 617), narrative synthesis supported the benefits of sleep hygiene, assisted sleep, and sleep extension interventions for sleep, performance, and mood; meta-analytic synthesis of randomised controlled trials (k =12) supported the effectiveness of sleep interventions, irrespective of their type, on subjective sleep quality (*g* = 0.62, 95% CI [0.21, 1.02]), reduced sleepiness (*g* = 0.81, 95% CI [0.32, 1.30]) and decreased negative affect (*g* = 0.63, 95% CI [0.27, 0.98]), with no meaningful effects on device-assessed or self-reported sleep and aerobic or anaerobic performance [7]. Napping as a specific sleep strategy is also potentially beneficial for physical and cognitive performance, as well as perceptual and psychological factors (k= 36, N = 3,489) [11]. Collectively, the available evidence supports athletes’ sleep as essential to recovery, training, and performance, making it a cornerstone of holistic intervention approaches for the modern athlete.

The modest outcomes of existing sleep strategies or interventions [6–8] suggests that our understanding of optimal approaches to maximising athlete sleep health is underdeveloped. Among the limitations (e.g., small sample sizes, underrepresentation of females, precision in reporting of interventions) of past work on athlete sleep, we contend that the prioritisation on mean levels for the operationalisation of sleep is one area that requires immediate attention because such data provides limited insight into strategies that might optimise sleep health. The focus on mean levels of sleep duration within scientific research is unsurprising given the prioritisation of this metric by the American National Sleep Foundation in their guidelines [2]. Strategies informed by mean levels of sleep indices are often based on the presumption that barriers to optimal sleep (e.g., travel across time zones, unfamiliar sleeping environments, well-being) are static and enduring, and therefore a 'one size fits all' approach to mitigating such problems is ideal. The one size fits all approach erroneously assumes that specific strategies will be effective for all people and all types of barriers; this assumption is inadequate for athletes who experience numerous and diverse stressors across the various ecologies of their occupational context (e.g., training, competition, organisational). For example, sleep hygiene tactics (e.g., regular sleep-wake cycle, optimal sleeping environment) might be possible when athletes remain in the same geographical location for their training and competition schedule, yet challenging or impossible when regular travel is characteristic of their sport (e.g., altering time zones, unusual sleeping environments, competition scheduling). An alternative approach to sleep intervention is one that embraces complexity and therefore encompasses a personalised repertoire of tactics that can be activated reactively to unanticipated stressors, or proactively to known challenges. Resolving these gaps in knowledge and inadequacies with past work is important because sleep health interventions represent unrealised potential for optimising athlete performance and health.

Intra-individual variability in sleep metrics provides rich information about the in/stability of person-situation dynamics over time and across contexts that is unavailable from mean levels alone. For example, two athletes might accrue roughly equivalent mean levels of total sleep duration (see Figure 1) and other core metrics (e.g., onset), yet differ meaningfully regarding the quantitative deviations from their typical level across a finite period (e.g., low versus high variability). Sleep variability provides unique information on physical and mental functioning beyond that which is explained by habitual estimates of sleep health [14]. Depending on the context, IIV might suggest resistance or maladaptive responses to barriers to optimal sleep, which might be missed by a myopic focus on mean levels alone. Within the context of elite military forces, for example, reductions in IIV in sleep duration and efficiency across the first 7-days following an intensive 3-week selection course provides insight into emergent resilience [15]. Typically, high sleep variability increases the risk of numerous health and behavioural issues, including physical health conditions, body mass, psychopathology, and stress [14]. Knowledge of sleep IIV and its determinants among athletes represents an untapped source of evidence to inform a new generation of individualised sleep interventions that address dynamic, intra-individual networks of barriers and enablers to optimal sleep health.

Sleep patterns are characterised by several metrics across dimensions of continuity (e.g., latency, efficiency), architecture (e.g., rapid eye movement, slow wave), and naps (e.g., duration, frequency per 24 h) [3]. Guidelines and recommendations almost exclusively focus on mean levels of these sleep metrics [1,2], despite evidence suggesting that IIV is salient for health and functioning [14]. Thus, via a systematic review, we aimed to document the ways by which researchers have defined and operationalised sleep variability among athletes. In so doing, we lay the foundations for an empirical and practical shift to one that considers sleep IIV among the conversation on sleep health for athletes.

2. Methods

We pre-registered the protocol for this systematic review on 6th April 2022 via the Open Science Framework (OSF: https://bit.ly/3MNy6DI) and report the results according to the 2020 version of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [16].

2.1. Literature Search

DG searched databases (Web of Science, Embase, Medline, PsycINFO, CINHAL Plus, and ProQuest Dissertations and Theses Global) from inception to 6th April 2022 using a protocol that included keywords for the target outcome (sleep*), population (athlete* OR sport*), and outcome operationalisation (variability OR variation OR "standard deviation" OR fluctuate OR fluctuation OR stability OR instability OR reactivity OR IIV OR intraindividual). We also completed a backward (reference lists) and forward (citations) search of articles identified as eligible via the primary database search on 5th July 2022.

2.2. Selection Criteria

We considered studies eligible for inclusion if they (i) sampled athletes, defined as individuals who are behaviourally engaged in sport, that is, "involving physical exertion and skill as the primary focus of the activity, with elements of competition where rules and patterns of behaviour governing the activity exist formally through organisations and is generally recognised as a sport" [17]; and (ii) assessed sleep metrics (e.g., duration, quality, efficiency) daily for at least 3 nights utilising either self-report or wearable devices. We excluded studies when (i) they assessed sleep using polysomnography only because such studies typically reflect controlled experimental environments rather than the complexities of everyday life; (ii) assessed sleep repeatedly other than daily measures (e.g., weekly, monthly); the (iii) article was written in any language other than English; (iv) full-text was unavailable via our university library subscriptions or directly from the corresponding author (i.e., 2 email requests/reminders, separated by 2 weeks); (v) results were published as an abstract rather than a full-text (e.g., dissertation); or (vi) article presented no new primary data on sleep variability among athletes (e.g., narrative or systematic review, commentary).

2.3. Screening Approach

SK and DG collaboratively reviewed eligible articles (titles and abstracts) using a web-application (Research Screener [https://researchscreener.com]) that enables assessors to screen all research abstracts from scientific databases using machine learning to optimise the review process

[18]. Research Screener ranks the abstracts in order of relevance based on relevant existing articles known to the team for inclusion based on the screening criteria, and continuously updates the learning algorithm every 50 abstracts screened based on what is deemed in/eligible by the reviewer. We used three seed articles known to our team that met the eligibility criteria outlined above to initiate the screening process in Research Screener [19–21]. Typically, Research Screener is used to optimise the review process, that is, review no more than 50% of eligible articles [e.g., 22,23]; however, we used this application solely to screen articles because of the user-friendly interface.

2.4. Data Extraction

SK extracted data items from primary studies using a pre-determined form, or requested such information from the corresponding author of eligible studies where that information was unavailable in the full text. DG assessed a random sample of 30% of data extraction forms to check accuracy and consistency. We captured information on the publication details (e.g., publication year, study location), nature of the scientific work (e.g., number of nights assessed), sampled participants (e.g., age, sport/athletic pursuit) including application of the tiered Participant Classification Framework [24], outcome assessments (e.g., definition of sleep variability, method used to assess sleep), and study quality. Using an 8-item tool, we assessed study quality as good, fair, or poor regarding indicators relevant to examinations of sleep variability, namely aims/hypotheses, sample size justification, sample representativeness, number of days sleep assessed, measures of sleep and its correlates, missing data, and inferences and conclusions [14]. The complete data extraction form is available on the OSF project page (https://bit.ly/3NE1jB3).

2.5. Protocol Deviations

We deviated from our registered protocol in one way. We considered studies eligible for inclusion when authors explicitly stated, *or it could be inferred from their narrative*, that they were directly interested in sleep variability.

3. Results

3.1. Characteristics of Eligible Studies

An overview of the search and selection process is depicted in Figure 2, courtesy of a Shiny app for generating PRISMA 2020-compliant diagrams [25]. In total, we identified 16 studies as meeting our eligibility criteria. Authors published their research in peer-reviewed journals (n = 16) between 2015 and 2022; conducted the research in the UK (n = 5), Australia (n = 4), Portugal (n = 4), USA (n = 1), Iceland (n = 1), and the Netherlands (n = 1); and sampled 937 athletes aged approximately 15-36 years from single sports only (e.g., swimming, rugby league, soccer; n = 13), or a mix of multiple sports (n = 3), who were highly trained/national level (Tier 3, n = 4), elite/international level (Tier 4, n = 6), or mixed levels (n = 6); implemented observational studies (n = 16) across 7 days to 15 months, during competition (n = 7), out of competition (n = 5), or unspecified (n = 4). Representation of athlete sex indicated studies in males only (n = 4), females only (n = 4), mixed-sex cohort (n = 4), male versus female sub-analysis (n = 2) and participant sex not explicitly mentioned (n = 2). Full details of each study are available in the supplementary material (see Table S1; https://bit.ly/3NE1jB3).

3.2. Overview of Methodological Quality

An overview of study quality assessments is presented in Figure 3. Key strengths of this work include the quality of the key correlate (100% good ++, excluding 2 not applicable) and sleep measure (62.5% good ++), and inferences and conclusions (87.5% good ++). There were both strengths and limitations to other study features, including aims and hypotheses (50% good ++, 50% fair +), sample representativeness (100% fair +), and number of days (nights of sleep) assessed (62.5% fair [≥ 7 & < 14], 25% good [> 14 days], and 12.5% [poor < 7 days]). Key weaknesses of this work included the absence of sample size justifications (100% poor) and description of missing data (62.5% poor [not reported], 25% fair [≤ 20% & >10%; none reported > 20%], and 12.5% good [≤10%]).

3.3. Definitions and Operationalisations of Sleep Variability

Authors reported an explicit definition of sleep variability in only 4 of 16 eligible studies. The most common features of these definitions included “differences within individuals time” (n =

2) or variation across days (n = 1) or nights (n =1). The most common operationalisations of sleep variability included the intra-individual standard deviation (n = 8) and coefficient of variation (n = 6); the other representation included within-individual z scores, calculated as ([individual player's score – individual player's average]/individual player's SD). One paper excluded any explicit information on how the authors operationalised sleep statistically. Researchers assessed sleep primarily via devices (n = 4; e.g., Actigraph), self-report (n = 5), or a combination of both approaches (n = 7).

3.4. Key Correlates of Sleep Variability

Regarding key correlates of sleep variability, researchers have examined demographic (e.g., type of athlete), contextual (e.g., time of competition or training session), biological (e.g., nocturnal cardiac activity), physical (e.g., load), psychological (e.g., well-being, perceived effort) factors as correlates of sleep variability. Among the 4 studies in which the authors explicitly defined sleep variability, 2 reported intra-individual variability estimates descriptively (9-22% for sleep duration and 2-11% for sleep efficiency), with no consideration of determinants or outcomes of variability. Two other studies examined differences in sleep variability between playing level and athletes categorised as ir/regular sleepers. Among rugby league players, elite juniors demonstrated greater variability in sleep onset time, time in bed, and sleep duration than both elite seniors and sub-elite seniors, as well as greater variability in sleep efficiency and subjective sleep quality than elite seniors[19]. Regarding differences among elite team sport athletes, regular sleeper displayed less variability in total sleep time, sleep efficiency, and sleep onset and offset[5].

4. Discussion

Via a systematic review, we narratively synthesised the literature on sleep variability among athletes, particularly regarding considerations of definition and operationalisation. We found that this body of evidence is small relative to work on mean levels of sleep metrics among athletes where there exists expert consensus recommendations [1], with all but one of these outputs generated since 2017. Concept definitions of variability were notably absent from this work; where available, definitions were vague and therefore insufficient for guiding robust operationalisations of sleep variability as well as conceptual development and integration of findings. Quantitative deviations from one's typical level of some sleep metric reflected the essence by which all but one of the research teams operationalised sleep variability; this feature is also characteristic of definitions of intra-individual sleep variability with non-athlete populations [e.g., 'quantifies daily variation around the mean', 14(p108)]. Finally, the overall quality of empirical work is moderate in nature. Current knowledge on key determinants or outcomes of sleep variability is best described as being in its infancy.

Precise and unambiguous definitions of concepts and their operationalisation via methodological procedures are fundamental to evidence quality and accumulation, and the translation of knowledge into practice and policy [26,27]. The absence of an explicit concept definition of sleep IIV within much of the existing research is a major weakness yet, given the infancy of this body of work, represents an opportunity to set the foundations for conceptually robust work in the future. Quality data depends on quality concept definitions, which in turn is a prerequisite for robust theory, measurement, and application. The reasons for the absence of a high-quality definition of sleep variability in existing scientific work on sleep variability are unclear (e.g., regularly spoken in everyday life, scientific infancy of the field). Irrespective of such reasons, the field is in urgent need of a scientific definition that provides a precise, clear, and cohesive understanding of the meaning and defining features of sleep variability [27]. Thus, we propose a working definition of sleep variability as a quantitative approximation of the magnitude of temporally heterogeneous deviations for each 24-hour sleep cycle from one's typical level of indicators of sleep across a finite period. This definition encompasses several requirements for high-quality concept definitions [27]. We clarify the nature of the phenomenon (magnitude of temporally heterogeneous deviations for each sleep session) and event to which this property applies (some indicator of people's sleep, e.g., temporal elements like duration or perceptual elements like quality). We also clarify the conceptual theme that summarises the nature of the necessary and sufficient conditions for IIV in sleep (deviations from one's typical level across a finite period). In this way, IIV reflects within-person fluctuations in metrics of sleep functioning for each sleep episode (typically each night within a 24-hour period including daytime napping) around some measure of central tendency, across a defined temporal period.

Our working definition incorporates key features among all eligible studies identified via our systematic review regarding how sport science researchers have operationalised sleep variability in terms of processes, tests, and measurements. The first consideration is the temporal period across which one assesses sleep functioning and for which they can make inferences regarding IIV of specific metrics (e.g., duration, efficiency). We found no clear consensus regarding the minimum or optimal temporal window for which to assess athlete sleep functioning when interested in intra-individual variability. Others have recommended at least 7-days [28] are required to estimate sleep variability reliably, though this recommendation is derived solely from an empirical analysis of 166 older adults aged 60 years and over across a 14-day period. Conversely, others found that guidance for minimum number of nights depends on the sleep metric (e.g., duration, variability), measurement window of interest (e.g., weekly or monthly), and desired reliability threshold of interest[29]. In the absence of robust empirical evidence (e.g., Monte Carlo simulations), we suggest that researchers justify their selection of temporal window according to the research questions and contextual factors. For example, Gucciardi et al. (2021) sampled the 7-day 'recovery' period between a 3-week selection course for entry into elite special forces and the start of a subsequent 15-month training cycle because it permitted inferences regarding emergent resilience via reductions in within-person sleep variability during this window. Alternatively, one may need to assess sleep across several nights to capture sufficient information 'on' (e.g., training, competition) and 'off' days when interested in the phenomenon of social jet lag (e.g.,[30]).

The second consideration concerns the ways by which researchers quantify and statistically model IIV in sleep functioning. We found that researchers relied on quantifications that characterise the amplitude or amount of fluctuation, namely the intra-individual standard deviation or coefficient of variation, which are subsequently employed as an aggregate index of variability in statistical models that either ignore (e.g., general linear models) or incorporate (e.g., mixed effects models) dependency inherent within repeated measurements of sleep. Despite the simplicity and practical intuitiveness of the intra-individual standard deviation or coefficient of variation as quantifications of the amplitude or amount of fluctuation, they are characterised by several disadvantages (e.g., sensitive to systematic within-person change, correlated with overall mean) that limit their usefulness for operationalising sleep variability, as reviewed elsewhere [27]. Monte Carlo simulations also indicate that indexes of IIV such as intra-individual standard deviation or coefficient of variation often have poor reliability [32]. Relatedly, utilising intra-individual standard deviations or coefficient of variations as an aggregate index of variability in statistical models is fraught with danger because it excludes uncertainty in the variability estimate and therefore inflates Type I error [33,34]. Aggregate indices of variability also and prevents analysts' from incorporating predictors that vary across time, such as daily indices of physical workload and psychosocial stress, and correlations among random effects (e.g., variability and mean levels) that are likely of substantive interest [33,34]. Thus, the two-step approach commonly employed within the sport science literature limits the congruence between concept and statistical modelling, and the repertoire of substantive questions regarding sleep variability that can be addressed. Mixed-effects location-scale models represent an alternative approach to alleviate these shortcomings of existing approaches and maximise congruence between concept and design and test combinations[34].

Our assessment of study quality identified that the weaknesses (e.g., absence of sample size justification or missing data) outweighed the strengths (e.g., device-assessed sleep, quality of key correlates) evident among existing research on sleep variability among athletes. Given the reliance on inferential statistics to test key research questions among this body of work, the absence of sample size justifications means readers are unable to judge the informativeness of the data, given the design and test combination. Sample size justifications are rarely reported in the sport science literature [35]. Within the eligible work summarised here, the median sample size of 36.5 provides 80% power ($\alpha = 0.5$) to detect moderate-to-large effects (e.g., ~$r = .44$; between-group differences,

~d = .66) for statistical models that assume normal distributions, homogeneity of variances, and independence in the data. Accounting for non-independence and/or relaxing the assumption of homogeneity of variances complicates statistical modelling because it involves several fixed and random effects across multiple levels, which is challenging to estimate in the absence of prior work to guide plausible population effects [36,37]. Of course, under certain circumstances (e.g., equal cluster sizes), mixed effects modelling produces approximately identical results to summary-based-statistics such as *t*-tests and linear regression [38] which can simplify sample size justifications. We also observed an absence of information regarding missing data – whether present or not and, if so, to what degree – on key variables of interest within the eligible body of work. Particularly in applied settings, longitudinal monitoring of factors that occur and likely vary daily are inherently plagued by missing data [e.g., training load, 39]. Handling of missing data is potentially problematic, rather than the presence of it per se, because what we do to deal with incomplete data can introduce bias into statistical models and undermine statistical power [40]. Reporting clear sample size justifications [41], pre-registering method and data analysis protocols [42], and maximising transparency in research reporting represent key opportunities for an area of research in its infancy [43].

Strengths of our study include a pre-registered protocol, open data, transparency regarding deviations from our intentions, and search of peer-reviewed literature and dissertations. Nevertheless, the findings reported here are best interpreted within the context of the limitations, including a traditional search approach only (e.g., no direct contact with individual researchers for unpublished work, such as pre-prints), capture of articles written in English only, and application of a bespoke methodological quality assessment tool [14].

5. Conclusion

Most research on athlete sleep has prioritised the generation of data and knowledge on mean levels of indicators of sleep, which has largely overshadowed research on IIV across a finite period. We argue that this oversight and oversimplification of the essence of athlete sleep has serious implications for knowledge development (theory) and translation into evidence-based strategies (application). Essentially, differences between athletes in mean level estimates of sleep alone are informative only when IIV is small or trivial; as deviations from one's typical level of indicators of sleep across a finite period increase, which represent systematic rather than random error, mean level estimates alone likely offer flawed data for making inferences about differences between people and the effectiveness of strategies to optimise sleep health. Thus, the next frontier of strategies to optimise athlete sleep health demands knowledge of both mean levels and IIV of metrics of sleep functioning.

6. Perspective

Optimised sleep enables biological and psychological restoration to cost inflicted by everyday activities[12,44,45]. Acute and chronic inadequate sleep quantity and quality can trigger negative effects on bodily functioning. For example, healthy individuals who obtain less than six hours of sleep per night may be at risk of disturbances to bodily functions such as glucose metabolism, immune processes, and cognitive capacity[46]. Thus, identifying ways to intervene before acute sleep issues develop into chronic considerations is imperative, especially for populations (e.g., elite athletes) and contexts (e.g., shift work) where achieving sleep health might be thwarted. Currently, the development and execution of personalised sleep health interventions typically relies primarily on mean-levels of sleep indicators rather than a complementary view that also considers within-person variability in these estimates. As concept definitions underpin high-quality research and practice, we systematically reviewed the literature to examine how researchers have defined sleep variability to date, and leverage this knowledge to propose a new working definition that conforms to guidelines for high-quality concept definitions [27].

References

1. Walsh N, Halson SL, Sargent C, Roach G, Nedelec M, Gupta L, Leeder J, Fullagar H, Coutts A, Edwards B, Pullinger S, Robertson C, Burniston J, Lastella M, Le Meur Y, Hausswirth C, Bender A, Grandner M, Samuels C. Sleep and the athlete: narrative review and 2021 expert consensus recommendations. Br J Sports Med 2021;55:356–368.

2. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, Hazen N, Herman J, Adams Hillard PJ, Katz ES, Kheirandish-Gozal L, Neubauer DN, O'Donnell AE, Ohayon M, Peever J, Rawding R, Sachdeva RC, Setters B, Vitiello MV, Ware JC. National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015;1:233–243.

3. Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, Dauvilliers Y, Ferri R, Fung C, Gozal D, Hazen N, Krystal A, Lichstein K, Mallampalli M, Plazzi G, Rawding R, Scheer FA, Somers V, Vitiello MV. National Sleep Foundation's sleep quality recommendations: first report. Sleep Health 2017;3:6–19.

4. Sargent C, Lastella M, Halson SL, Roach GD. How Much Sleep Does an Elite Athlete Need? Int J Sports Physiol Perform 2021;16:1746–1757.

5. Halson SL, Johnston RD, Piromalli L, Lalor BJ, Cormack S, Roach GD, Sargent C. Sleep Regularity and Predictors of Sleep Efficiency and Sleep Duration in Elite Team Sport Athletes. Sports Med - Open 2022;8:79.

6. Bonnar D, Bartel K, Kakoschke N, Lang C. Sleep Interventions Designed to Improve Athletic Performance and Recovery: A Systematic Review of Current Approaches. Sports Med 2018;48:683–703.

7. Gwyther K, Rice S, Purcell R, Pilkington V, Santesteban-Echarri O, Bailey A, Walton CC. Sleep interventions for performance, mood and sleep outcomes in athletes: A systematic review and meta-analysis. Psychol Sport Exerc 2022;58:102094.

8. Silva A, Silva A, Edwards B, Tod D, Amaral A, Borba D, Grade I, de Mello M. Sleep extension in athletes: what we know so far - A systematic review. Sleep Med 2021;77:128–135.

9. Buysse DJ. Sleep Health: Can We Define It? Does It Matter? Sleep 2014;37:9–17.

10. Fischer D, Klerman EB, Phillips AJK. Measuring sleep regularity: theoretical properties and practical usage of existing metrics. Sleep 2021;44:zsab103.

11. Craven J, McCartney D, Desbrow B, Sabapathy S, Bellinger P, Roberts L, Irwin C. Effects of Acute Sleep Loss on Physical Performance: A Systematic and Meta-Analytical Review. Sports Med 2022;

12. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: A meta-analysis of randomised controlled trials. Sleep Med Rev 2021;60:101556.

13. Lastella M, Halson SL, Vitale JA, Memon AR, Vincent GE. To Nap or Not to Nap? A Systematic Review Evaluating Napping Behavior in Athletes and the Impact on Various Measures of Athletic Performance. Nat Sci Sleep 2021;13:841–862.

14. Bei B, Wiley JF, Trinder J, Manber R. Beyond the mean: A systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. Sleep Med Rev 2016;28:108–124.

15. Gucciardi DF, Lang JWB, Lines RLJ, Chapman MT, Ducker KJ, Peeling P, Crane M, Ntoumanis N, Parker SK, Thøgersen-Ntoumani C, Quested E, Temby P. The emergence of resilience: Recovery trajectories in sleep functioning after a major stressor. Sport Exerc Perform Psychol 2021;10:571–589.

16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;n71.

17. May C. What is Sport? Sport Aust 2021;

18. Chai KEK, Lines RLJ, Gucciardi DF, Ng L. Research Screener: a machine learning tool to semi-automate abstract screening for systematic reviews. Syst Rev 2021;10:93.

19. Caia J, Halson SL, Scott T, Kelly V. Intra-individual variability in the sleep of senior and junior rugby league athletes during the competitive season. Chronobiol Int 2017;34:1239–1247.

20. Costa J, Figueiredo P, Nakamura F, Rago V, Rebelo A, Brito J. Intra-individual variability of sleep and nocturnal cardiac autonomic activity in elite female soccer players during an international tournament. PLOS ONE 2019;14:.

21. Whitworth-Turner C, Di Michele R, Muir I, Gregson W, Drust B. A comparison of sleep patterns in youth soccer players and non-athletes. Sci Med Footb 2018;2:3–8.

22. Lines RLJ, Hoggan BL, Nahleen S, Temby P, Crane M, Gucciardi DF. Enhancing shared mental models: A systematic review and meta-analysis of randomized controlled trials. Sport Exerc Perform Psychol 2022;

23. Murdoch EM, Lines RLJ, Crane MF, Ntoumanis N, Brade C, Quested E, Ayers J, Gucciardi DF. The effectiveness of stress regulation interventions with athletes: A systematic review and multilevel meta-analysis of randomised controlled trials. Int Rev Sport Exerc Psychol 2021;1–37.

24. McKay AKA, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, Sheppard J, Burke LM. Defining Training and Performance Caliber: A Participant Classification Framework. Int J Sports Physiol Perform 2021;17:317–331.

25. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. Campbell Syst Rev 2022;18:.

26. Impellizzeri FM, Jeffries AC, Weisman A, Coutts AJ, McCall A, McLaren SJ, Kalkhoven J. The ‘training load’ construct: Why it is appropriate and scientific. J Sci Med Sport 2022;25:445–448.

27. Podsakoff PM, MacKenzie SB, Podsakoff NP. Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. Organ Res Methods 2016;

28. Rowe M, McCrae C, Campbell J, Horne C, Tiegs T, Lehman B, Cheng J. Actigraphy in older adults: comparison of means and variability of three different aggregates of measurement. Behav Sleep Med 2008;6:127–145.

29. Lau T, Ong JL, Ng BKL, Chan LF, Koek D, Tan CS, Müller-Riemenschneider F, Cheong K, Massar SAA, Chee MWL. Minimum number of nights for reliable estimation of habitual sleep using a consumer sleep tracker. SLEEP Adv 2022;3:zpac026.

30. Henderson SEM, Brady EM, Robertson N. Associations between social jetlag and mental health in young people: A systematic review. Chronobiol Int 2019;36:1316–1333.

31. Wang L (Peggy), Hamaker E, Bergeman CS. Investigating inter-individual differences in short-term intra-individual variability. Psychol Methods 2012;17:567–581.

32. Estabrook R, Grimm KJ, Bowles RP. A Monte Carlo simulation study of the reliability of intraindividual variability. Psychol Aging 2012;27:560–576.

33. Lester HF, Cullen-Lester KL, Walters RW. From Nuisance to Novel Research Questions: Using Multilevel Models to Predict Heterogeneous Variances. Organ Res Methods 2021;24:342–388.

34. McNeish D. Specifying location-scale models for heterogeneous variances as multilevel SEMs. Organ Res Methods 2021;24:630–653.

35. Sainani K, Chamari K. Wish List for Improving the Quality of Statistics in Sport Science. Int J Sports Physiol Perform 2022;17:673–674.

36. Arend MG, Schäfer T. Statistical power in two-level models: A tutorial based on Monte Carlo simulation. Psychol Methods 2019;24:1–19.

37. Lang JWB, Bliese PD, Runge JM. Detecting consensus emergence in organizational multilevel data: Power simulations. Organ Res Methods 2019;1094428119873950.

38. Murayama K, Usami S, Sakaki M. Summary-statistics-based power analysis: A new and practical method to determine sample size for mixed-effects modeling. Psychol Methods 2022;

39. Windt J, Ardern CL, Gabbett TJ, Khan KM, Cook CE, Sporer BC, Zumbo BD. Getting the most out of intensive longitudinal data: a methodological review of workload–injury studies. BMJ Open 2018;8:e022626.

40. Graham JW. Missing Data Analysis: Making It Work in the Real World. Annu Rev Psychol 2009;60:549–576.

41. Lakens D. Sample Size Justification. Collabra Psychol 2022;8:33267.

42. Abt G, Boreham C, Davison G, Jackson R, Wallace E, Williams AM. Registered Reports in the Journal of Sports Sciences. J Sports Sci 2021;39:1789–1790.

43. Caldwell AR, Vigotsky AD, Tenan MS, Radel R, Mellor DT, Kreutzer A, Lahart IM, Mills JP, Boisgontier MP, Boardley I, Bouza B, Cheval B, Chow ZR, Contreras B, Dieter B, Halperin I, Haun C, Knudson D, Lahti J, Miller M, Morin JB, Naughton M, Neva J, Nuckols G, Peters S, Roberts B, Rosa-Caldwell M, Schmidt J, Schoenfeld BJ, Severin R, Skarabot J, Steele J, Twomey R, Zenko Z, Lohse KR, Nunan D, Consortium for Transparency in Exercise Science

(COTES) Collaborators. Moving Sport and Exercise Science Forward: A Call for the Adoption of More Transparent Research Practices. Sports Med 2020;50:449–459.

44. Morrison M, Halson SL, Weakley J, Hawley JA. Sleep, circadian biology and skeletal muscle interactions: Implications for metabolic health. Sleep Med Rev 2022;66:101700.

45. Zhu B, Wang Y, Yuan J, Mu Y, Chen P, Srimoragot M, Li Y, Park CG, Reutrakul S. Associations between sleep variability and cardiometabolic health: A systematic review. Sleep Med Rev 2022;66:101688.

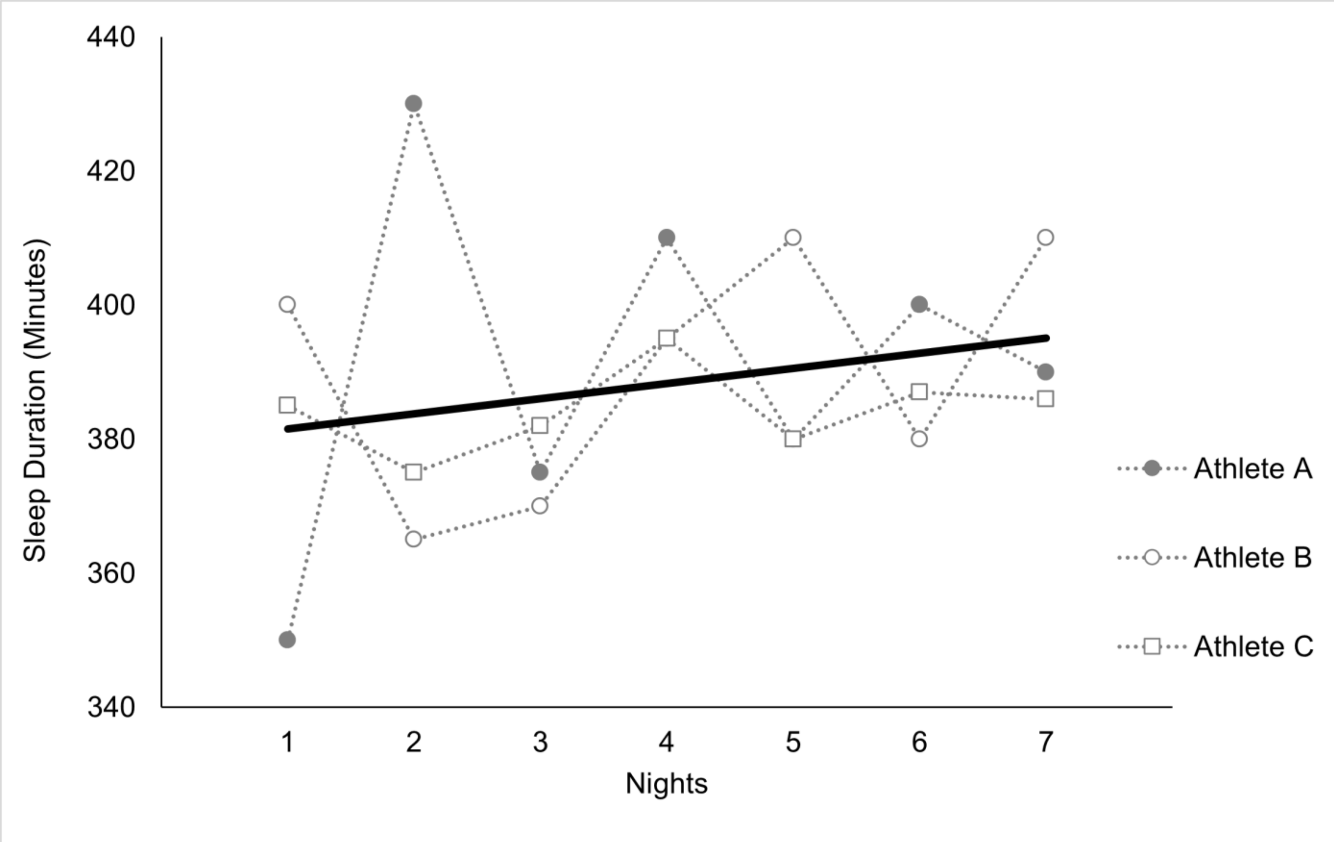
46. Grant M, Glen J. An investigation into sleep patterns and the effect of time of day on performance in youth swimmers. Biol Rhythm Res 2018;49:830–843.

Figure Captions

Figure 1. Hypothetical example of meaningful differences in intra-individual variability in sleep duration (minutes) across a 7-night period for three individual athletes who share the same mean.

Figure 2. PRISMA flow diagram.

Figure 3. Visual depiction of study quality assessments.



440
420
400
380
360
340
Sleep Duration (Minutes)
1
2
3
4
5
6
7
Nights
Athlete A
Athlete B
Athlete C

Figure 1.

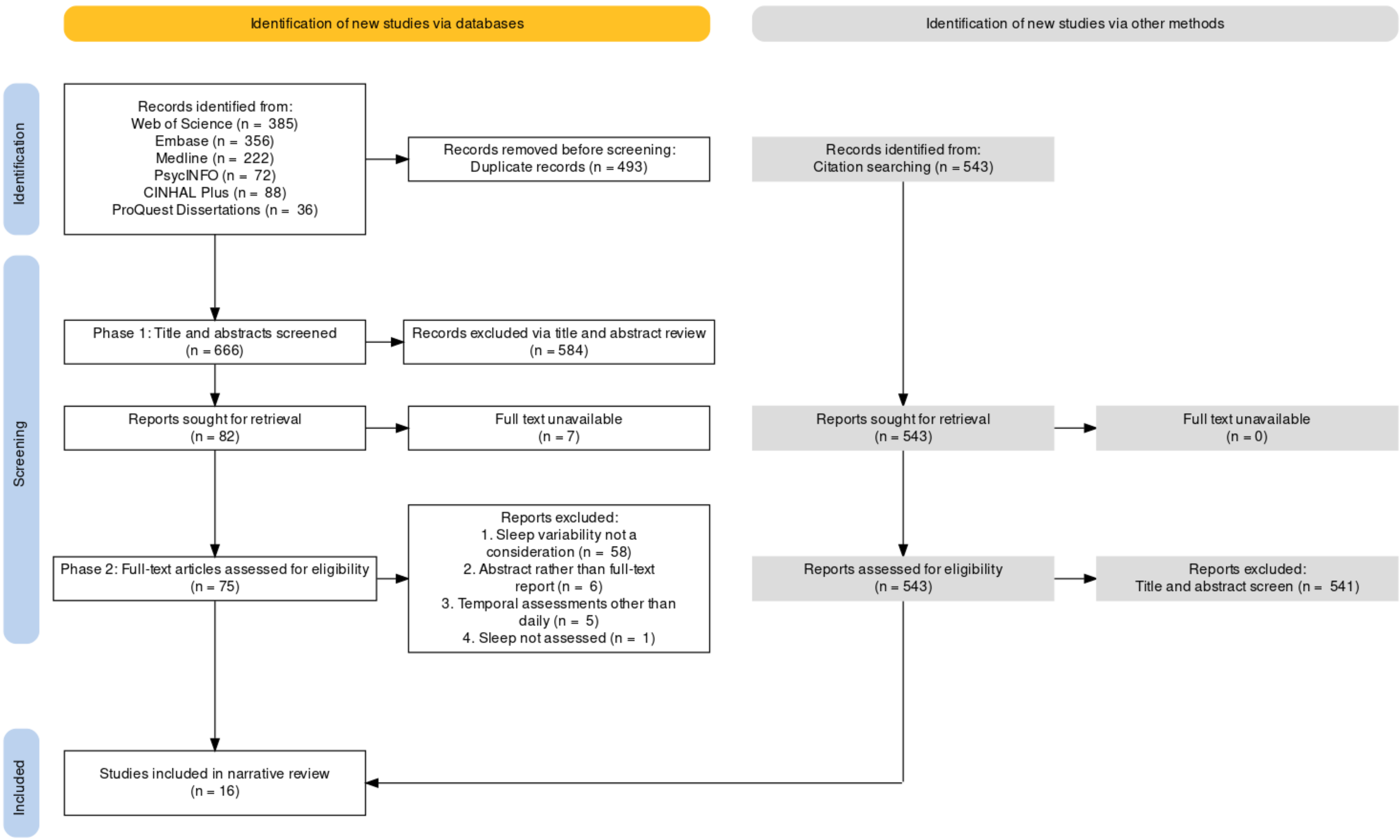

Identification of new studies via databases
Identification of new studies via other methods
Identification
Screening
Included
Records identified from:
Web of Science (n = 385)
Embase (n = 356)
Medline (n = 222)
PsycINFO (n = 72)
CINHAL Plus (n = 88)
ProQuest Dissertations (n = 36)
Records removed before screening:
Duplicate records (n = 493)
Records identified from:
Citation searching (n = 543)
Phase 1: Title and abstracts screened
(n = 666)
Records excluded via title and abstract review
(n = 584)
Reports sought for retrieval
(n = 82)
Full text unavailable
(n = 7)
Reports sought for retrieval
(n = 543)
Full text unavailable
(n = 0)
Reports excluded:
1. Sleep variability not a consideration (n = 58)
2. Abstract rather than full-text report (n = 6)
3. Temporal assessments other than daily (n = 5)
4. Sleep not assessed (n = 1)
Phase 2: Full-text articles assessed for eligibility
(n = 75)
Reports assessed for eligibility
(n = 543)
Reports excluded:
Title and abstract screen (n = 541)
Studies included in narrative review
(n = 16)

Figure 2.

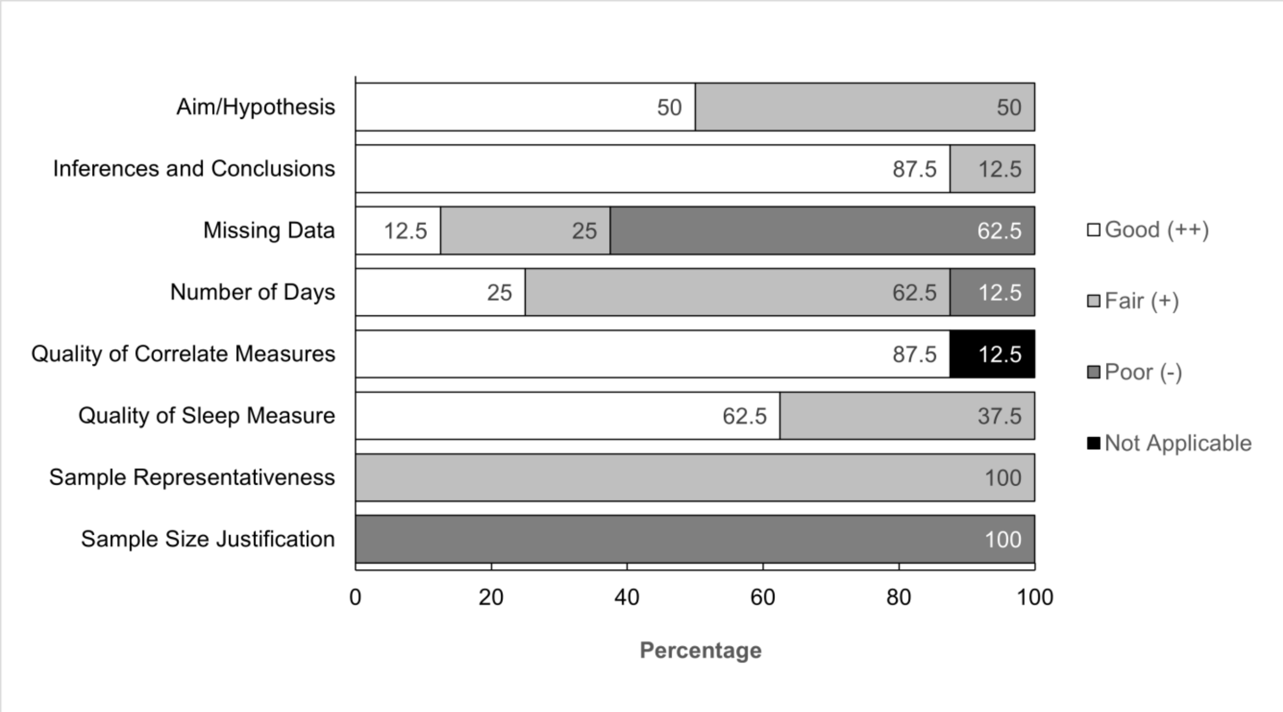

Aim/Hypothesis
Inferences and Conclusions
Missing Data
Number of Days
Quality of Correlate Measures
Quality of Sleep Measure
Sample Representativeness
Sample Size Justification
50
50
87.5
12.5
12.5
25
62.5
25
62.5
12.5
87.5
12.5
62.5
37.5
100
100
0
20
40
60
80
100
Percentage
Good (++)
Fair (+)
Poor (-)
Not Applicable

Figure 3.