Sleep measurement tools, circadian strategies, and dietary factors

pertaining to sleep in athletes and non-athletes.

by

Dean Jason Miller

A Thesis Submitted in Fulfilment of the Requirement for the Degree of

**Doctor of Philosophy**

in the School of Health, Medical and Applied Sciences

Central Queensland University

July 2021

**Table of Contents**

**Summary**

Obtaining sufficient sleep is important for humans to maintain optimal performance. This is especially true for elite athletes striving to improve performance while contending with unique stimuli that may impact their sleep (e.g., training, competition, travel). Measurement tools, countermeasures, and/or strategies that can be used to mitigate factors that negatively affect sleep are of great interest to athletes, coaches, and practitioners. This dissertation addresses (1) the validity of a wearable device to measure sleep; (2) the impact of evening exercise modality on sleep; (3) the effectiveness of strategies to facilitate circadian adaptation following transmeridian travel; and (4) the effectiveness of combined nutritional ingredients on sleep.

The first two studies (Chapters 3 and 4) in this dissertation were validation studies of the commercially available device WHOOP strap 2.0. These studies were conducted in a controlled laboratory environment, using the gold standard of sleep measurement – polysomnography. In both studies, the classification of sleep stage for WHOOP and polysomnographic systems were arranged into 30-second epochs. The only difference between the studies was the way in which the data were collected using the WHOOP strap. In Chapter 3, the start and end clock times for each sleep period were entered into the WHOOP iOS application – allowing the WHOOP system to classify the sleep within the specified period (i.e., manual adjustment). In Chapter 4, no information was entered into the WHOOP iOS application – allowing the device to automatically detect and classify sleep based on proprietary algorithms (i.e., automatic detection). Once automatically detected data were obtained, manual adjustments were made to the same sleep periods to obtain a dataset for direct comparison. Detailed epoch-by-epoch analyses revealed that the WHOOP strap was comparable to research grade actigraphy when classifying two stage sleep (i.e., sleep or wake; 86% agreement). Importantly, this finding was consistent when the WHOOP strap was used under manual adjustment or automatic detection functionalities. Compliance for completing self-report measurements, such as bedtimes and getup times, is a common hurdle when collecting data with athletes. These studies highlight an important practical implication of modern wearable technology, in that athletes may wear these devices and acquire accurate data without actively providing contextual information.

Chapters 5 and 7 address a common factor that can affect the sleep of athletes who are required to travel across time zones for training and competition – jet lag. In both studies, data were collected from young elite athletes following transmeridian travel. Chapter 5 highlights an association between a single question to assess jet lag and the nocturnal levels of urinary melatonin (6-sulphatoxymelatonin). That is, concentration of 6-sulphatoxymelatonin reduced as subjective jet lag increased. In comparison to other jet lag scales, the one used in this study can be conducted more efficiently and is now the first to be validated against an objective circadian biomarker in an athlete population. The aim of the study in Chapter 7 was to examine the effectiveness of a light exposure schedule, implemented via hard copy and electronic reminders, to enhance circadian adaptation following transmeridian travel. Participants adapted to the new time zone within 7 days and no significant differences in 6-sulphatoxymelatonin were found between a light exposure group and a control group. It is possible that the intervention was ineffective or that participants did not adhere to the intervention because compliance with the schedule (i.e., light exposure measurements) was not assessed.

The aim of the study associated with Chapter 6 was to examine the impact of evening exercise modality on the sleep of healthy young males. Traditional sleep medicine recommendations advise against exercise in the evening due to potential disruptions to the biological preparations to initiate sleep. One of the main concerns is that the increase in core body temperature associated with muscular contraction may interrupt the natural decline in core body temperature prior to sleep. However, the main finding of this study was that sleep outcomes did not differ across the three conditions (i.e., no exercise, 30 minutes of aerobic exercise, 30 minutes of resistance exercise). Interestingly, aerobic exercise resulted in a higher core body temperature than the other conditions during exercise; but returned to pre-exercise levels in the 90 minutes between exercise termination and sleep (i.e., 21:30h-23:00h).

The study in Chapter 8 examined the effectiveness of a formulated sleep drink on the sleep quality and quantity of healthy young males. Next day cognitive function was also assessed. The ingredients included in the drink are known to improve sleep when consumed independently but

were examined in combination for this study. Sleep onset latency was significantly reduced following the consumption of an optimal sleep drink compared with consumption of a non-optimal sleep drink and consumption of a placebo. Importantly, the optimal sleep drink did not impair next-day cognitive performance – which is a side effect of some pharmaceutical sleep medications commonly used by athletes. The findings of this study suggest that the optimal sleep drink may be used to promote the onset of sleep without impacting next-day performance.

Several conclusions and practical implications may be drawn from the results in this dissertation. First, wearable technology capable of estimating sleep – in this case the WHOOP strap – appear to be a practical alternative to commonly used research grade actigraphy. Athletes and coaches may utilise the device to automatically detect and estimate sleep and wake, thus eliminating the need to provide contextual inputs (i.e., bed/wake times). For travelling athletes, a single question to assess subjective jet lag appears to be a suitable option for estimating jet lag. For factors impacting sleep, the findings suggest that moderate intensity aerobic or resistance exercise does not impact the sleep of healthy young males; a light exposure protocol does not enhance circadian adaptation following transmeridian travel; and a formulated sleep drink may be an effective non-pharmaceutical option to initiate sleep.

## Acknowledgements

I would like to acknowledge the Traditional Owners of the land on which this dissertation was written and pay respect to the Kaurna Peoples, Elders, past, present, and future. I acknowledge the culture, stories, songlines, sacred sites, and histories that run through this land.

Firstly, I would like to thank my primary supervisor, Associate Professor Charli Sargent for your tireless effort, guidance, and support throughout the past few years. I would also like to thank Professor Greg Roach for affording me the same support, despite not being a formal part of the supervisory team.

To Dr Aaron Scanlan, I would like to thank you for helping me back when I was an undergrad trying to gain entry into the Master's program. Despite not getting the chance to meet face-to-face, your unique insights, lightning speed edits, and guidance were invaluable.

I would also like to acknowledge Associate Professor Shona Halson, Rosie Stanimirovic, and the Australian Institute of Sport for allowing me to collect data through their collaborations with CQUniversity.

I owe a huge amount to Dr Michele Lastella. Without your willingness to take a scrappy 3rd year Psych student under your wing, I would certainly not be in this position today. Your mentorship has been the catalyst to any past or future opportunities I have been granted. And for that I cannot thank you enough. Still a nuffy though.

I would like to thank my family, in particular my parents, who have made significant sacrifices to afford me the opportunities to pursue my goals. Thank you to the ECR writing group at the Appleton Institute, and writing buddy Taylor Swain – our writing sessions really progressed the latter stages of this thesis. To my wife Brooke, from high school maths, all the way through University, you have been a constant pillar of support and motivation. I could not have achieved any of this without you. Special mention to a good boy – Chase Miller – who was also by my side  throughout this journey.

**Declaration**

By submitting this thesis for formal examination at CQUniversity Australia, I declare that it meets all requirements as outlined in the Research Higher Degree Theses Policy and Procedure.

By submitting this thesis for formal examination at CQUniversity Australia, I declare that all of the research and discussion presented in this thesis is original work performed by the author. No content of this thesis has been submitted or considered either in whole or in part, at any tertiary institute or university for a degree or any other category of award. I also declare that any material presented in this thesis performed by another person or institute has been referenced and listed in the reference section.

By submitting this thesis for formal examination at CQUniversity Australia, I acknowledge that thesis may be freely copied and distributed for private use and study; however, no part of this thesis or the information contained therein may be included in or referred to in any publication without prior written permission of the author and/or any reference fully acknowledged.

This research has not been submitted for an award by another research degree candidate (Co-Author), either at CQUniversity or elsewhere.

## Publications arising from thesis

**Published:**

1. **Miller, D. J.,** Lastella, M., Scanlan, A. T., Bellenger, C., Halson, S. L., Roach, G. D., and Sargent, C. (2020). A validation study of the WHOOP strap against polysomnography to assess sleep. *Journal of Sports Sciences*, 38(22), 2631–2636. Candidate contributed to: conceptualisation (70%), methodology (70%), data collection (100%), formal analyses (90%), writing orginal draft (100%), review and editing (70%), visualisation (90%) and project administration (90%). The remaining work was distributed across co-authors.

2. **Miller, D. J**., Sargent, C., Roach, G. D., Scanlan, A. T., Vincent, G. E., and Lastella, M. (2020). Moderate-intensity exercise performed in the evening does not impair sleep in healthy males. *European Journal of Sport Science*, 20(1), 80–89. Candidate contributed work related to: conceptualisation (50%), methodology (80%), data collection (80%), formal analyses (100%), writing orginal draft (100%), review and editing (70%), visualisation (90%) and project administration (90%). The remaining work was distributed across co-authors.

3. Halson, S. L., Shaw, G., Versey, N., **Miller, D. J.,** Sargent, C., Roach, G. D., Nyman, L., Carter, J. M., and Baar, K. (2020). Optimisation and Validation of a Nutritional Intervention to Enhance Sleep Quality and Quantity. *Nutrients*, 12(9), 2579. **For phase 2 of the project (i.e., Chapter 8)**, the candidate contributed to work related to: methodology (50%), data collection (80%), formal analyses (100%), review and editing (70%), and project administration (90%). The remaining work was distributed across Sargent, C., and Roach, G. D.

4. **Miller, D.J.,** Roach G.D., Lastella M., Scanlan A.T., Bellenger C., Halson S.L., Sargent C. (2021). A validation study of a commercial wearable device to automatically detect and estimate sleep. *Biosensors*, 11(6):185. Candidate contributed to: conceptualisation (70%), methodology (70%), data collection (100%), formal analyses (90%), writing orginal draft

(100%), review and editing (70%), visualisation (90%) and project administration (90%). The remaining work was distributed across co-authors.

5. **Miller, D. J.,** Roach, G. D., Lastella, M., Scanlan, A. T., & Sargent, C. (2021). Implementing a Circadian Adaptation Schedule after Eastward Flight in Young Male Athletes. Applied Sciences, 11(21), 9962. Candidate contributed to: conceptualisation (70%), methodology (90%), data collection (100%), formal analyses (100%), writing orginal draft (100%), review and editing (70%), visualisation (90%) and project administration (90%). The remaining work was distributed across co-authors.

**Candidate contribution to un-published chapters:**

Chapter 5: The validity of a single question for assessing subjective jet lag in athletes. Candidate contributed to: conceptualisation (50%), methodology (70%), data collection (60%), formal analyses (100%), writing orginal draft (100%), review and editing (70%), visualisation (90%) and project administration (80%). The remaining work was distributed across co-authors.

**Other publications contributed to during candidature:**

1. **Miller, D.J.,** Sargent, C., Vincent, G. E., Roach, G. D., Halson, S. L., & Lastella, M. (2017). Sleep/Wake Behaviours in Elite Athletes from Three Different Football Codes. *Journal of sports science & medicine*, *16*(4), 604–605.

2. Sargent, C., Lastella, M., Romyn, G., Versey, N., **Miller, D.,** & Roach, G. (2018). How well does a commercially available wearable device measure sleep in young athletes?. *chronobiology international*, 35(6), 754-758.

3. Lastella, M., Roach, G., **Miller, D.,** Versey, N., Romyn, G., & Sargent, C. (2018). Athletes underestimate sleep quantity during daytime nap opportunities. *Chronobiology international*, *35*(6), 869-871.

4. Romyn, G., Lastella, M., Miller, D., Versey, N., Roach, G., & Sargent, C. (2018). Daytime naps can be used to supplement night-time sleep in athletes. *Chronobiology international*, *35*(6), 865-868.

5. Vincent, G., Sargent, C., Roach, G., **Miller, D.,** Kovac, K., & Scanlan, A. et al. (2019). Exercise before bed does not impact sleep inertia in young healthy males. *Journal of sleep research*, *29*(3).

6. Lastella, M., **Miller, D.,** Vincent, G., Elsworthy, N., Scanlan, A., & Duffield, R. (2021). The influence of training and competition on sleep behaviour of soccer referees. *Science and medicine in football*, 1-7.

7. **Miller, D.J.,** Capodilupo J.V., Lastella M., Sargent C., Roach G.D., et al. (2020). Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. *PLOS ONE* 15(12).

8. Bellenger, C. R., Miller, D., Halson, S. L., Roach, G., & Sargent, C. (2021). Wrist-Based Photoplethysmography Assessment of Heart Rate and Heart Rate Variability: Validation of WHOOP. *Sensors*, *21*(10), 3571.

9. Lastella, M., **Miller, D. J.,** Quilelli, M., Roberts, S., Aisbett, B., & Condo, D. (2021). The Impact of Chronotype on the Sleep and Training Responses of Elite Female Australian Footballers. Clocks & Sleep, 3(4), 528–535.

**List of Tables**

## List of Figures

## List of key terminology and abbreviations

Throughout this dissertation there will be references to specific terms related to sleep and/or circadian rhythms. These terms, and where necessary, abbreviations, are detailed in Table i.

**Table i.** Key terminology and abbreviations.

| Term (Abbreviation) | Definition |
|---|---|
| Time in bed (TIB) | Total amount of time in bed from the start of a sleep period to the end of a sleep period. |
| Total sleep time (TST) | Total amount of sleep within the time in bed period. |
| Wake in sleep | Total amount of wake within the time in bed period. |
| Wake after sleep onset (WASO) | Total amount of wake between the initiation of sleep and the end of a sleep period. |
| Sleep onset latency (SOL) | Total amount of time between the start of a sleep period and the initiation of sleep. |
| Sleep efficiency | Percentage of time in bed spent asleep. |
| Rapid eye movement sleep (REM) | The sleep stage associated with dreaming and restoration of mental functions. |
| Slow wave sleep (SWS) | The "deep" sleep stage associated with muscular restoration. |
| Polysomnography (PSG) | The gold standard measurement of sleep. |
| Electro-encephalogram (EEG) | Polysomnographic technique for measuring brain activity. |
| Electro-oculogram (EOG) | Polysomnographic technique for measuring eye movements. |
| Electromyogram (EMG) | Polysomnographic technique for measure muscular activity. |
| 6-suplhatoxymelatonin (aMT6s) | The urinary metabolite for melatonin. |

# 1. Introduction

Elite athletes and their coaches aim to improve athletic performance by managing and modifying factors such as training, diet, psychological wellbeing, and physical recovery (Mujika, Halson et al., 2018). Sleep has been described as the optimal recovery strategy for athletes (Halson, 2008), as well as a potential factor that may impact athletic performance (Mah et al., 2011). As a cohort, it appears that athletes do not obtain sufficient sleep (Leeder, Glaister et al., 2012; Sargent, Halson et al., 2014). Inadequate sleep is a major concern for athletes and coaches given sleep restriction (i.e., <6 h per day) impairs a range of human functions including cognitive capacity (Belenky, Wesensten et al., 2003), glucose metabolism (Spiegel, Leproult et al., 1999; Van Cauter, Spiegel et al., 2008), regulation of appetite (Spiegel, Tasali et al., 2004) and immune function (Vgontzas, Zoumakis et al., 2004), all of which may impact physical performance. There are many potential mechanisms through which sleep may be impaired in athletes; these include training (Sargent, Halson et al., 2014; Sargent, Lastella et al., 2014), competition (Lastella, Lovell et al., 2014; Juliff, Halson et al., 2015; Lastella, Roach et al., 2015b), dietary intake (Halson, 2008), and transmeridian travel (Youngstedt and O'Connor, 1999; Leeder, Gardner et al., 2009; Lastella, Roach et al., 2014).

Due to the negative impact of inadequate sleep on human physiology (Spiegel, Leproult et al., 1999; Belenky, Wesensten et al., 2003; Spiegel, Tasali et al., 2004; Vgontzas, Zoumakis et al., 2004; Van Cauter, Spiegel et al., 2008), strategies to improve sleep and/or minimise the impact of external factors on sleep, are of great interest to professional athletes, practitioners, and coaches. This is reflected by the marked increase in research outputs in the past decade. Specifically, 56 experimental studies examining sleep in athletes had been published in peer-reviewed journal articles prior to 2010, while 257 peer-reviewed journal articles were published between 2010 and 2020 (Lastella, Memon et al., 2020).

To examine the factors that may affect the sleep of athletes, wearable sleep monitoring devices have become common practice in the field of sport science. However, few of these wearable devices have been validated against polysomnography, which is the gold standard for monitoring sleep. If

such devices cannot validly measure sleep, they are of limited use in sports science (Halson, Peake, and Sullivan, 2016; de Zambotti, Cellini et al., 2019).

Independent of sleep measurement, athletes and coaches must also consider the impact of training, competition, and travel on sleep. For an athlete to improve performance, the volume and intensity of their training (i.e., training load) is manipulated periodically (Mujika, Halson et al., 2018). The timing of training is a major factor that can affect the amount of sleep an athlete obtains (Sargent, Halson et al., 2014; Sargent, Lastella et al., 2014). Traditional sleep hygiene recommendations advise against exercising near bedtime due to the potential disruption to sleep (Zarcone, 1994). A potential mechanism that may impact sleep following training or competition at night is elevated core body temperature, which may impact the natural evening decline in core body temperature associated with sleep onset (Driver and Taylor, 2000). However, there is some evidence from both epidemiological and experimental studies to suggest that evening exercise may have no detrimental effects on sleep (Alley et al., 2015; Buman et al., 2014; O'Connor and Youngstedt, 1995). Therefore, it is unclear whether core body temperature; or other potential factors induced by evening exercise may impact subsequent sleep.

In addition to their high workload, athletes may be required to travel across multiple time zones to train or compete. Trans-meridian travel results in a misalignment between the human circadian system and the local destination time, often termed "jet lag" (Eastman et al., 2009). The side-effects of jet lag may inhibit an athlete's ability to train and/or compete at an optimal level (Leatherwood and Dragoo, 2013). A common strategy implemented to facilitate entrainment to a new destination time zone is to utilise light exposure to "shift" to the timing of the human circadian system (i.e., the body clock) (Burgess, et al., 2003; Eastman and Burgess, 2009; Eastman, et al., 2005). The theory underpinning the use of light exposure protocols to facilitate entrainment are well established in the literature (Burgess, et al., 2003; Eastman and Burgess, 2009; Eastman, et al., 2005); however the efficacy of such protocols have not been evaluated with athletes following travel across multiple time zones.

To combat sleep disturbances related to training, competition, and travel, athletes may be prescribed a pharmaceutical means of obtaining sleep. The use of "sleeping pills' among athletes is common practice on the night preceding competition (Juliff, Halson et al., 2015). Sleep medications can be effective in reducing sleep onset latency but are also associated with residual side-effects, which could affect the ability of athletes to train and perform the following day (Reilly, Atkinson, and Budgett, 2001). Therefore, non-pharmaceutical means of increasing sleep duration, especially those without residual side effects, are of great interest to athletes and coaches. Several dietary supplements have been shown to impact the neurological processes associated with sleep (Halson, 2014). Dietary precursors such as tart cherry juice and L-theanine may stimulate neurotransmitters aligned with sleep processes (Halson, 2014). While evidence suggests that dietary supplements may influence sleep, the optimal doses and effectiveness of these supplements as a sleep aid have not been examined. Therefore, research aimed at formulating the optimal combination and dose of dietary precursors may provide athletes with a non-pharmaceutical means of facilitating sleep onset.

Strategies aimed at monitoring and improving the sleep of athletes during training, competition and travel phases are of great interest to athletes and sport scientists. This dissertation aims to examine the measurement tools, countermeasures, and circadian strategies to mitigate compromised sleep in athletes.

**1.1 References**

Alley, J.R., Mazzochi, J.W., Smith, C.J., Morris, D.M; Collier, S.R (2015). Effects of resistance exercise timing on sleep architecture and nocturnal blood pressure. *Journal of Strength and Conditioning Research,* 29(5):1378-1385.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., . . . Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research*, 12(1), 1-12.

Buman, M. P., Phillips, B. A., Youngstedt, S. D., Kline, C. E., and Hirshkowitz, M. (2014). Does nighttime exercise really disturb sleep? Results from the 2013 National Sleep Foundation Sleep in America Poll. *Sleep Medicine*, 15(7), 755-761.

Burgess, H. J., Savic, N., Sletten, T., Roach, G., Gilbert, S. S., and Dawson, D. (2003). The relationship between the dim light melatonin onset and sleep on a regular schedule in young healthy adults. *Behavioral Sleep Medicine*, 1(2), 102-114.

de Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M., and Baker, F. C. (2019). Wearable Sleep Technology in Clinical and Research Settings. *Medicine and Science in Sports and Exercise*, 51(7), 1538-1557.

Driver, H. S., and Taylor, S. R. (1996). Sleep disturbances and exercise. *Sports Medicine*, 21(1), 1-6.

Eastman, C.I., and Burgess, H. J. (2009). How to travel the world without jet lag. *Sleep Medicine Clinics*, 4(2), 241.

Eastman, C.I., Gazda, C.J., Burgess, H. J., Crowley, S. J., and Fogg, L. F. (2005). Advancing circadian rhythms before eastward flight: a strategy to prevent or reduce jet lag. *Sleep*, 28(1), 33.

Halson, S. (2008). Nutrition, sleep and recovery. *European Journal of Sport Science*, 8(2), 119-126.

Halson, S.L. (2014a). Sleep in elite athletes and nutritional interventions to enhance sleep. Sports *Medicine*, 44(1), 13-23.

Halson, S. L., Peake, J. M., and Sullivan, J. P. (2016). Wearable technology for athletes: Information overload and pseudoscience? *The International Journal of Sports Physiology and Performance, 11*(6), pp. 705-706.

Juliff, L.E., Halson, S.L., and Peiffer, J.J. (2015). Understanding sleep disturbance in athletes prior to important competitions. *Journal of Science and Medicine in Sport*, 18(1), 13-18.

Leatherwood, W. E., and Dragoo, J. L. (2013). Effect of airline travel on performance: a review of the literature. *British journal of sports medicine*, 47(9), 561–567.

Leeder, J. D. C., Gardner, A. S., Foley, S., van Someren, K., and Pedlar, C. R. (2009). The effect of jet lag on parameters of sleep in elite divers quantified by actigraphy. *Medicine and Science in Sports and Exerci*se, 41(5), 57.

Leeder, J., Glaister, M., Pizzoferro, K., Dawson, J., and Pedlar, C. (2012). Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *Journal of sports science*, 30(6), 541-545.

Lastella, M., Lovell, G. P., and Sargent, C. (2014). Athletes' precompetitive sleep behaviour and its relationship with subsequent precompetitive mood and performance. *European Journal of Sport Science*, 14 (Suppl 1), S123-130.

Lastella, M., Roach, G. D., Halson, S. L., Gore, C. J., Garvican-Lewis, L. A., and Sargent, C. (2014b). The effects of transmeridian travel and altitude on sleep: preparation for football competition. *Journal of Sports Science and Medicine*, 13(3), 718-720.

Lastella, M., Roach, G. D., Halson, S. L., Martin, D. T., West, N. P., and Sargent, C. (2015b). Sleep/wake behaviour of endurance cyclists before and during competition. *Journal of Sports Science*, 33(3), 293-299.

Lastella, M., Memon, A.R., and Vincent, G.E. (2020). Global research output on sleep research in athletes from 1966 to 2019: a bibliometric analysis. *Clocks and Sleep*, 2(2), 99-119.

Mah, C. D., Mah, K. E., Kezirian, E. J., and Dement, W. C. (2011). The effects of sleep extension on the athletic performance of collegiate basketball players. *Sleep*, 34(7), 943.

Mujika, I., Halson, S., Burke, L. M., Balague, G., and Farrow, D. (2018). An Integrated, Multifactorial Approach to Periodization for Optimal Performance in Individual and Team Sports. *The International Journal of Sports Physiology and Performance*, 13(5), 538-561.

O'Connor, P. B., MJ; Youngstedt, SD. (1998). Exercise-induced increase in core temperature does not disrupt a behavioral measure of sleep. *Physiology and Behaviour*, 64(3), 213-217.

Reilly, T., Atkinson, G., and Budgett, R. (2001). Effect of low-dose temazepam on physiological variables and performance tests following a westerly flight across five time zones. *International Journal of Sports Medicine*, 22(3), 166-174.

Sargent C, Halson S, Roach G.D. (2014). Sleep or swim? Early-morning training severely restricts the amount of sleep obtained by elite swimmers. *European Journal of Sport Science*. 14 Suppl 1:S310-315.

Sargent, C., Lastella, M., Halson, S. L., and Roach, G.D. (2014). The impact of training schedules on the sleep and fatigue of elite athletes. *Chronobiology International*, 31(10), 1160-1168.

Spiegel, K., Leproult, R., and Van Cauter, E. (1999). Impact of sleep debt on metabolic and endocrine function. *Lancet*, 354(9188), 1435-1439.

Spiegel, K., Tasali, E., Penev, P., and Van Cauter, E. (2004). Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. *Annals of internal medicine*, 141(11), 846.

Van Cauter, E., Spiegel, K., Tasali, E., and Leproult, R. (2008). Metabolic consequences of sleep and sleep loss. *Sleep Medicine*, 9(3), S23-S28.

Vgontzas, A., Zoumakis, E., Bixler, E., Lin, H. M., Follett, H., Kales, A., and Chrousos, G. (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *Journal of Clinical Endocrinology and Metabolism*, 89(5), 2119-2126.

Youngstedt S.D., and O'Connor PJ. (1999). The influence of air travel on athletic performance. *Sports Medicine*, 28(3), 197–207.

Zarcone VP. (1994). Sleep hygiene. In. *Principles and Practice of Sleep Medicine*. Philadelphia, PA: W.B. Saunders Co. , pp. 542-546.

## 2. Literature Review

### 2.1 Sleep

Sleep is defined as a reversible state of mind and body resulting in disengagement and unresponsiveness from surroundings (Carskadon and Dement, 2011). Approximately one third of the human lifespan is spent in a state of sleep (Fuller, Gooley, and Saper, 2006). Normal human sleep is comprised of two states; rapid eye movement (REM) sleep and non-rapid eye movement (NREM) sleep, which alternate cyclically throughout a sleep episode (Carskadon and Dement, 2011). Polysomnography (PSG) is considered the gold standard method of objectively measuring human sleep and wakefulness (Kushida et al., 2005).

Sleep begins in NREM sleep (i.e., Stages 1 and 2) and gradually progresses into deeper NREM sleep (i.e., Stage 3) before reaching REM sleep (Carskadon and Dement, 2011; Fuller et al., 2006). The first cycle of sleep begins with Stage 1 sleep, which typically indicates sleep onset. Stage 1 sleep is associated with a low arousal threshold and lasts 1-7 minutes. During this stage, sleep is easily disrupted and acts as a transitional stage between wake and sleep throughout a sleep episode. The cycle then transitions into Stage 2 NREM sleep. Stage 2 sleep usually lasts for 10- 25 minutes, and arousal is less likely during this stage (Carskadon and Dement, 2011). As Stage 2 progresses, the low amplitude slow wave electroencephalographic (EEG) signals transitions into Stage 3 NREM sleep. Stage 3 sleep is categorised by slow-wave amplitude EEG waveform, thus referred to as slow wave sleep (SWS) (Iber, Ancoli-Israel, Chesson, and Quan, 2007; Rechtschaffen and Kales, 1968). SWS is the restorative stage of sleep and is associated with increased parasympathetic tone with decreases in heart rate, blood pressure, core body temperature, and metabolic rate (Davenne, 2009; Savis, 1994). The first episode of REM sleep generally occurs 70-90 minutes after sleep is initiated (Carskadon and Dement, 2011). During REM, the blocking of corticospinal pathways results in total muscle relaxation, thus allowing for muscle recovery via myofibril restoration (Davenne, 2009). REM sleep is also associated with learning and memory consolidation (Tilley and Empson, 1978). The shifting of sleep stages throughout the night facilitates recovery

processes within the body and allows for humans to awake feeling fresh and alert (Davenne, 2009; Santos, Tufik, and De Mello, 2007).

The human sleep/wake cycle is regulated by two mechanisms; (1) a homeostatic sleep drive and; (2) a circadian oscillator that regulates sleep and wake across each 24-hour day (Frank, 2013). The homeostatic sleep drive is the sleep-wake dependant component of sleep regulation (Borbély and Achermann, 1999). This mechanism increases the "sleep pressure" as the duration of wake increases and can only be diminished by initiating sleep (Borbely and Tobler, 1989; Kattler, Dijk et al., 1994; Borbély and Achermann, 1999; 2009).

Many physiological processes exhibit peaks and troughs over a 24-hour period (i.e., circadian rhythm; Dardente, Dardente et al., 2007). The fluctuation of these physiological processes (e.g., core body temperature) are governed by a circadian pacemaker within the suprachiasmatic nucleus of the hypothalamus (Czeisler, Duffy et al., 1999; Dardente, Dardente et al., 2007). Sleep and wake states are influenced by two processes governed by the suprachiasmatic nucleus; the ascending arousal system and the ventrolateral preoptic nucleus (Schwartz and Roth, 2008). The ascending arousal system comprises of monoaminergic cell populations distributed across two branches within the brain. This monoaminergic system discharges in a coordinated manner to promote arousal and sustained wakefulness (Schwartz and Roth, 2008). However, every 24 hours these processes are inhibited by y-aminobutyric acid and galaninergic neurons located in the ventrolateral preoptic nucleus (Schwartz and Roth, 2008). The pathway between the ventrolateral preoptic nucleus and the ascending arousal system is mutually inhibiting, creating an "on-off" switch of wakefulness and sleep (Schwartz and Roth, 2008). The control of the suprachiasmatic nucleus over these processes are sensitive to external cues (i.e., sunlight, darkness) referred to as zeitgebers (time givers). Following the period of day, cycles are entrained by external cues to structure the rhythms over 24-hour periods (Czeisler, Duffy et al., 1999; Czeisler and Gooley, 2007).

*2.1.2 Measuring sleep*

Polysomnography (PSG) is considered the gold standard method of objectively measuring human sleep and wakefulness (Kushida, Littner et al., 2005). However, implementation of PSG is expensive, time consuming, and impractical for use in some field settings (Yi Shin et al., 2006). Traditionally, the most commonly utilised alternative to PSG is actigraphy (de Zambotti, Cellini et al., 2019). Actigraphs are devices worn on the wrist like a watch that contain an accelerometer capable of detecting movement. The devices operate on the principle that movement is correlated with wake and long periods of inactivity are correlated with sleep (Pollak, Tryon et al., 2001). Using this principle, actigraphs then provide a binary classification of sleep or wake in 30-second intervals (i.e., epochs; (Pollak, Tryon et al., 2001; Sargent, Lastella et al., 2018). Despite actigraphy providing a practical alternative to PSG, some devices are limited to binary sleep/wake detection and a certain level of expertise and manual data handling is required to collect accurate measurements (de Zambotti, Cellini et al., 2019).

With advancements in technology, sleep wearables such as armbands, rings, and smartwatches provide an immediate measure of sleep that can be acquired with little expertise. In comparison to actigraphy, modern sleep wearables utilise a combination of accelerometers, thermometers, and photoplethysmographs to provide multi-stage sleep data (i.e., wake, light sleep [stages 1 and 2], slow wave sleep [stage 3] and rapid-eye-movement sleep [REM]; de Zambotti, Rosas et al., 2017; de Zambotti, Cellini et al., 2019). However, few sleep wearables have been validated against gold-standard PSG (Shambroom, Fabregas et al., 2012; de Zambotti, Rosas et al., de Zambotti, Cellini et al., 2019). When assessing the ability of modern sleep wearables to obtain accurate sleep data, it is important to consider the way in which these devices obtain raw data. For example, sleep wearables typically provide data under two functionalities: (1) by automatically detecting the onset and offset of sleep; and (2) manual adjustment of bedtimes and wakeup times by the user. Despite this distinction, most validations of sleep wearables have only analysed the manual adjustment function of devices, particularly when conducting epoch-by-epoch analyses.

Additionally, inconsistencies can be seen in the framework used to validate these devices against PSG.

### *2.1.3 Validating sleep wearables*

Recent recommendations regarding the validation of sleep wearables suggest a multi-level analysis of agreement against PSG (de Zambotti, Cellini et al., 2019). The recommended three-tiered analysis consists of difference testing, Bland-Altman analysis, and epoch-by-epoch analysis.

#### 2.1.3.1 Difference testing

Within-subject difference testing should be conducted to indicate if a device overestimates or underestimates raw sleep parameters in comparison to PSG. For example, if PSG-derived sleep data indicate that an individual slept 480 minutes during a sleep opportunity, and a sleep wearable estimated 460 minutes of sleep, that equates to a 20-minute underestimation. In previous validations of sleep wearables against PSG, a clinically satisfactory range of < 30 min for TST has been utilised (de Zambotti, Cellini et al., 2019). This guideline indicates the difference for TST between a device and PSG should fall within 30 minutes for the device to be considered clinically accurate. However, the rationale behind this range is not known (de Zambotti, Cellini et al., 2019), and there are inconsistencies in the use of this comparison statistic used to describe the "clinically satisfactory range". For example, several studies have utilised the mean difference (bias) between devices as the comparison statistic (Meltzer, Walsh et al., 2012; Meltzer, Hiruma et al., 2015; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017), while others have used the limits of agreement between devices (Werner, Molinari et al., 2008; de Zambotti, Baker et al., 2016; de Zambotti, Rosas et al., 2017). Descriptive statistics and difference testing provide indication of a devices tendency to overestimate or underestimate sleep measures; however there is no clear threshold at which a device can be considered "validated". Therefore, it is necessary to conduct further analyses using a more comprehensive approach.

### 2.1.3.2 Epoch-by-epoch comparisons

The next stage in determining the accuracy of sleep wearables against PSG are epoch-by-epoch comparisons. PSG data are typically scored in 30-s periods (i.e., epochs) – with each epoch classified as either wake, stage 1, stage 2, SWS, or REM (Iber, Ancoli-Israel et al., 2007). For sleep wearables capable of providing 30-s epochs, a robust epoch-by-epoch comparison can be made with PSG. Corresponding records from the sleep wearable are aligned with PSG, with each epoch classified into one of four categories assessing sleep and wake - True Sleep (TS), False Sleep (FS), True Wake (TW), and False Wake (FW) – on the basis of their agreement with PSG. Statistical measures of epoch-by-epoch concordance for each variable are then calculated from the number of epochs in each category:

- Agreement = [(TS + TW)/(TS + FW + TW + FS)]*100 = percentage of all sleep and wake epochs correctly detected by the device;

- Sensitivity = [TS/(TS + FW)]*100 = percentage of sleep epochs correctly detected by the device;

- Specificity = [TW(TW + FS)]*100 = percentage of wake epochs correctly detected by the device.

Conducting epoch-by-epoch comparisons controls for the possibility of false accuracy in the device's ability to detect sleep and wake (i.e., sensitivity and specificity). For example, during a 9-h sleep opportunity, it can be assumed that a healthy young person will be in a state of sleep for most (i.e., ~90%) of the sleep opportunity (Luca, Haba Rubio et al., 2015). If a device estimates that an individual obtained 100% sleep in the 9-h sleep opportunity, and PSG data indicate sleep occurred for 90% of the sleep opportunity, the agreement between devices equates to 90%. This may lead to a false indication of accuracy of a sleep wearable. It is therefore essential to include a measure of sensitivity (e.g., the ability to detect sleep) and specificity (e.g., the ability to detect wake) when assessing the validity of a device for measuring sleep.

With many sleep wearables providing measurements of sleep stage, the same epoch-by-epoch analyses must also be applied to individual sleep stages (i.e., four-stage sleep – wake, light sleep, slow wave sleep and REM). Comparisons should be conducted by aligning 30-s epoch data for each sleep stage obtained with a sleep wearable with a corresponding PSG record. The same calculations should be conducted for each sleep stage. For example, to analyse the ability of a sleep wearable to measure REM sleep, the aligned data should be split into True REM (TR), False REM (FR), True non-REM (TNR), and False non-REM (FNR). The following calculations can then be made based on their agreement with PSG:

- Agreement = [(TR + TNR)/(TR + FNR + TNR + FR)]*100 = percentage of all REM and non-REM epochs correctly detected by the device;

- Sensitivity = [TR/(TR + FNR)]*100 = percentage of REM epochs correctly detected by the device;

- Specificity = [TNR/(TNR + FR)]*100 = percentage of non-REM epochs correctly detected by the device.

Given that epoch-by-epoch comparisons for four-stage sleep results in multiple outputs, and that wake is included in the four-stage classification, it is reasonable for validations of sleep wearables to only report the sensitivity of the device for each stage.

### 2.1.3.4 Bland-Altman analysis

The Bland-Altman method of differences is used to visualise the agreement between two different methods of clinical measurement (Altman and Bland, 1983). The method allows for a repeated measures comparison of two devices for measuring a continuous variable by plotting the average and difference between each pair of data. In a Bland-Altman plot, each pair of data is presented as a single point; a solid line represents the mean bias between devices; and dashed lines indicate the 95% confidence interval between the devices. However, these methods have been misused (Ludbrook, 2010) since their initial publication in 1983 (Altman and Bland, 1983), and several improvements have been suggested (Ludbrook, 2010). For example, Ludbrook (2010) notes that

the traditional Bland-Altman method does not account for different combinations of bias (i.e., proportional, or non-proportional) or variance (i.e., homoscedasticity or heteroscedasticity) within the data (Ludbrook, 2010).

In the context of sleep wearable validations, proportional bias is present if differences between a device and PSG change as a function of duration and heteroscedasticity is present if variance of the differences changes as a function of duration. Ludbrook (2010) provides recommendations on how to adjust Bland-Altman plots for data with different bias and variance. For data that has no proportional bias and are homoscedastic, Ludbrook (2010) suggests constructing Bland-Altman plots with a horizontal bias and adjacent 95% confidence limits (Ludbrook, 2010). For data that has proportional bias and are homoscedastic, Ludbrook suggests constructing adjacent 95% confidence limits around the sloped bias line (Ludbrook, 2010). For data that has proportional bias and is heteroscedastic, Ludbrook suggests constructing V-shaped limits of agreement with a horizontal bias. Ludbrook (2010) does not provide specific recommendations for data that do not have proportional bias and are heteroscedastic. For these data, a Bland-Altman plot with a horizontal bias line, with V-shaped confidence limits can be constructed (Figure 2.1; +-1.96*SD). The four possible combinations of bias and variance are visualised in Figure 2.1. A flow chart for identifying which Bland-Altman plot is suitable for certain data is presented in Figure 2.2.

**Figure 2.1** The four possible combinations of bias and variance for validations of sleep wearables against polysomnography.
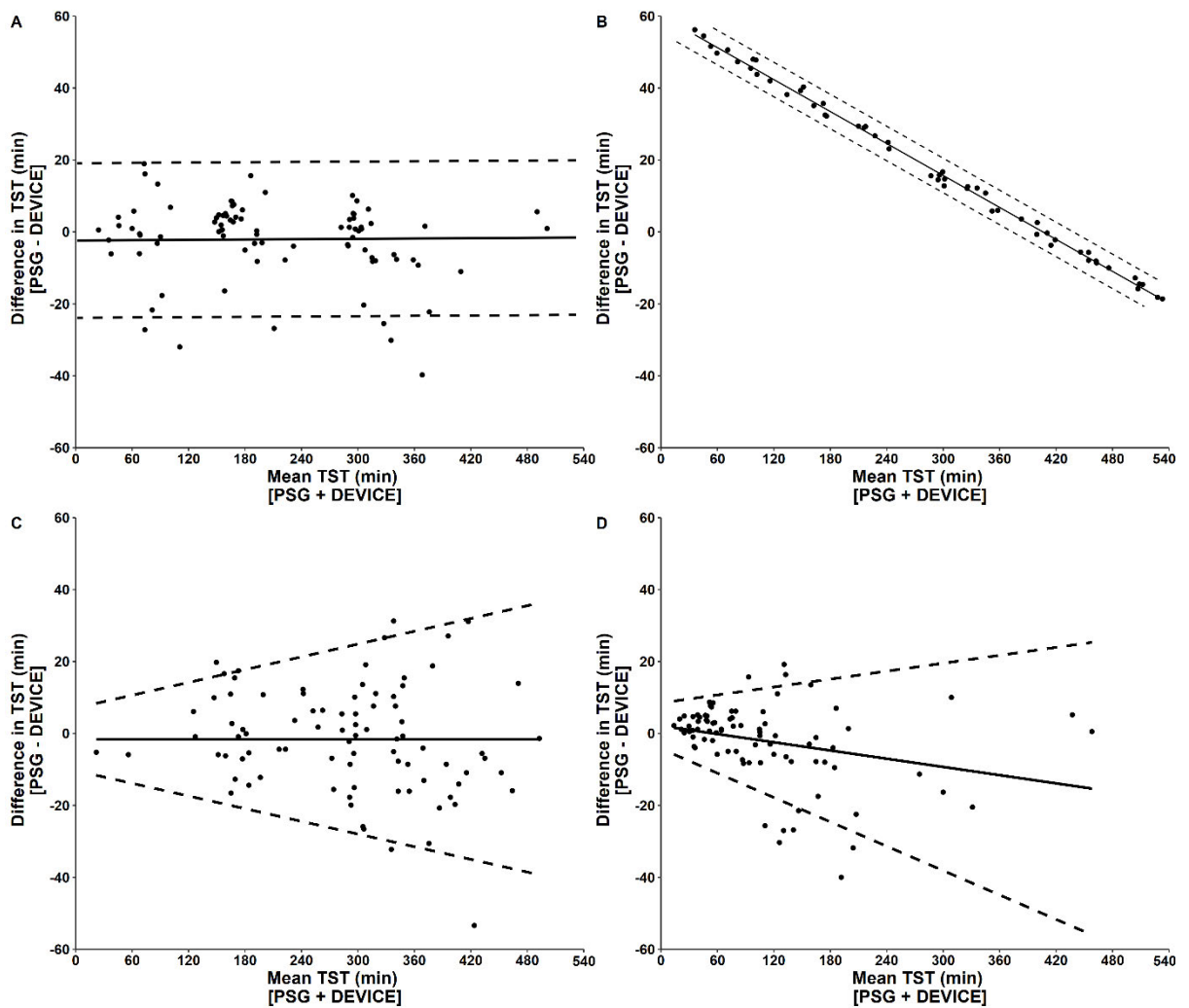


Notes: These are hypothetical data. Panel A presents data that does not have proportional bias and is homoscedastic; Panel B presents data that has proportional bias and is heteroscedastic; Panel C presents data that has no proportional bias and is heteroscedastic; and Panel D presents data that has proportional bias and is heteroscedastic. TST = total sleep time; PSG = polysomnography.

**Figure 2.2** Flow chart for identifying the appropriate application of Bland-Altman plots.
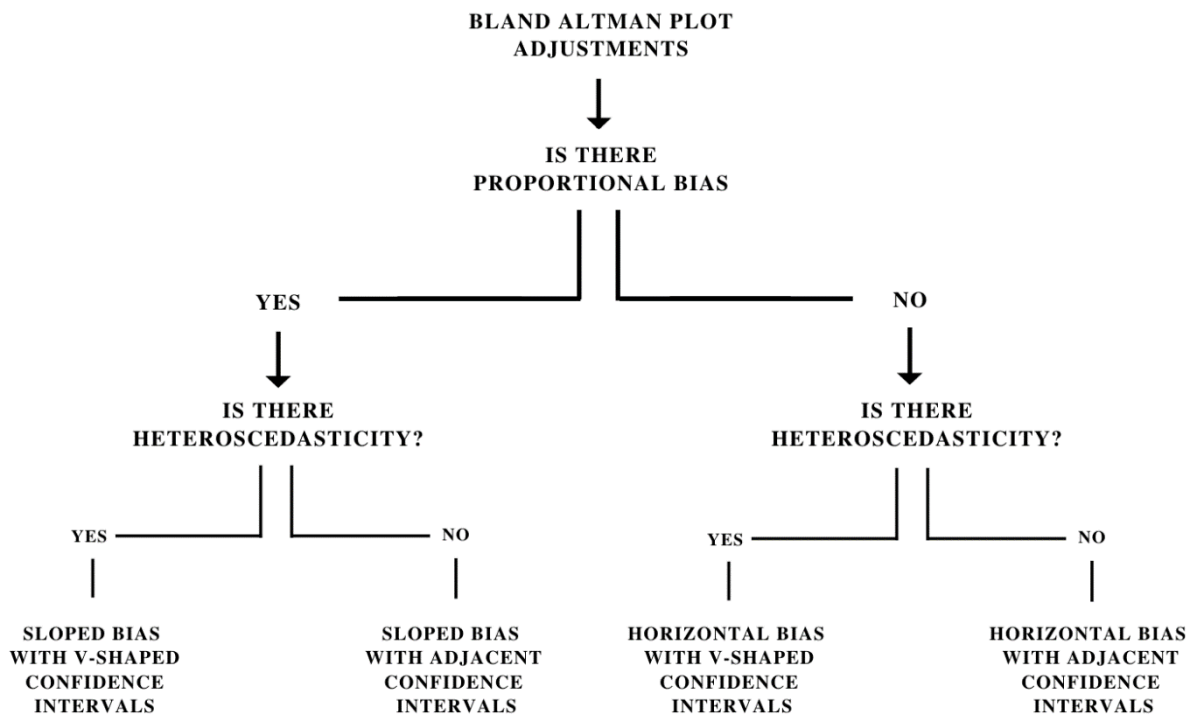


2.1.3.5 Previous sleep wearable validations

While all sleep wearable validations include a comparison between PSG and a wearable device, there are several methodological differences between approaches. For comparison, searches of PubMed, Embase, and PsycInfo were conducted to include only laboratory-based validations of consumer wearable devices to measure sleep. All studies were published in English and were published prior to July 2020. Aggregated performance matrices were collated from a total of 24 validation studies using the following search (Table 2.1-2.3):

"(sleep) AND (polysomnography OR PSG) AND (epoch OR epoch-by-epoch OR epoch-for-epoch) AND (sleep wearable OR wearable device OR consumer sleep technology OR fitbit OR whoop OR actical OR actigraph OR actigraphy OR readiband OR jawbone OR garmin OR apple watch OR withings) AND (validity OR validation OR validate OR comparison OR comparisons OR comparative OR reliability OR accuracy OR positive predictive value OR negative predictive value OR bland altman)".

In terms of sleep wearable functionality, most of the studies either manually adjusted the sleep times (N=11) or did not describe how sleep times were defined (N=11). Only two of the studies utilised the auto-detection from the wearable devices. Out of the 24 studies, 14 included healthy participants and 10 included participants with a range of sleep-related and/or other ailments that may impact sleep. Across all included studies, the wearable devices had a sensitivity for sleep of 91%, specificity for wake of 54%, TST absolute bias of 36.7 minutes, wake absolute bias of 28.1 minutes, light sleep absolute bias of 33.8 minutes, SWS absolute bias of 34.2, minutes and REM absolute bias of 12.4 minutes.

The devices included in the studies were obtained from eight different manufacturers, four of which were included in more than one study. The most common device manufacturer was Fitbit, with 10 studies including one of their devices. On average, Fitbit devices had a sensitivity for sleep of 88%, specificity for wake of 58%, TST absolute bias of 40.1 minutes, wake absolute bias of 32.7 minutes, light sleep absolute bias of 40.1 minutes, SWS absolute bias of 38.7 minutes, and REM absolute bias of 13.0 minutes. The second most common manufacturer was Jawbone, with five studies including one of their devices. On average, Jawbone devices had a sensitivity for sleep of 97%, specificity for wake of 41%, TST absolute bias of 42 minutes and wake absolute bias of 26.5 minutes. Only one study included measures of sleep stage for a Jawbone device (Table 2.3). Four studies included Sensewear devices, which yielded an average sensitivity for sleep of 91%, specificity for wake of 47%, TST absolute bias of 18.1 minutes and wake absolute bias of 12.8 minutes. Only one study included measures of sleep stage for a Sensewear device (Table 2.3).

In terms of sleep wearable functionality, studies that manually adjusted the sleep times had a sensitivity for sleep of 88%, specificity for wake of 51%, TST absolute bias of 38.8 minutes, wake absolute bias of 33.8, minutes light sleep absolute bias of 49.9 minutes, SWS absolute bias of 33.9 minutes, and REM absolute bias of 14.3 minutes. The studies that utilised sleep auto-detection had a sensitivity for sleep of 94%, specificity for wake of 88%, TST absolute bias of 44.4 minutes, wake absolute bias of 31.5, minutes light sleep absolute bias of 11.3 minutes, SWS absolute bias of 35.2 minutes, and REM absolute bias of 4.9 minutes. These findings suggest that auto-detection of sleep,

compared to manual adjustment of sleep, may be a more accurate function to utilise among wearable devices. However, it is important to note that this interpretation is based on reports in two studies. Further cross-device validations are needed to evaluate the optimal functionality for wearable devices to estimate sleep.

There were numerous differences in the characteristics of the participants examined in each of the studies that were reviewed (Table 2.1). For example, participants can be categorised as clinical (i.e., participants with sleep related and-or other ailments) or healthy. Given that wearable devices often have difficulty in classifying wake, it is reasonable to expect that their performance may decrease when used in clinical populations where a higher proportion of wake is present. For healthy participants, the wearable devices had a sensitivity for sleep of 93%, specificity for wake of 56%, TST absolute bias of 40.1 minutes, wake absolute bias of 27.7 minutes, light sleep absolute bias of 21.8 minutes, SWS absolute bias of 25.7 minutes, and REM absolute bias of 5.0 minutes. For clinical participants, the wearable devices had a sensitivity for sleep of 89%, specificity for wake of 49%, TST absolute bias of 48.6 minutes, wake absolute bias of 37.1 minutes, light sleep absolute bias of 54.6 minutes, SWS absolute bias of 44 minutes, and REM absolute bias of 27.6 minutes. These findings suggest that the accuracy of wearable technology for estimating sleep/wake parameters may be lower in clinical populations.

Overall, a wide variation in the study methodology and data analysis limits robust comparisons of wearable devices. Factors such as device functionality (i.e., manual manipulation; auto detection), population (i.e., good vs bad sleepers) and manufacturer may impact results. Further research is needed to compare devices across population cohorts under controlled laboratory settings.

**Table 2.1** Characteristics of studies included in the systematic search

| Paper | Device | Method | Gender | Age | Number of sleep records | Population |
|---|---|---|---|---|---|---|
| Roane (2015) | Sensewear Armband | N/A | 9 M; 11 F | $15.5 \pm 2.0$ | 20 | healthy adolescents |
| Cook (2019) | Fitbit Alta Hr | Manual | 46 F; 3 M | $30.3 \pm 9.8$ | 49 | sleep clinic patients being tested for hypersomnolence |
| O'Driscoll (2013) | Sensewear Armband | N/A | 34 M; 16 F | $45.5 \pm 2.0$ | 50 | adults suspected of sleep disorders |
| Alsaadi (2014) | Sensewear-Pro3 | N/A | 24 F; 26 M | $42.7 \pm 15.2$ | 50 | back pain sufferers |
| Meltzer (2015) | Fitbit Ultra | Manual | 31 M; 32 F | $9.7 \pm 4.6$ | 63 | children scheduled for clinical PSG |
| Lee (2019) | Fitbit Alta Hr | Automatic detection | 30 M; 28 F | $16.6 \pm 0.9$ | 386 | healthy adolescents |
| Toon (2016) | Up | Manual | 51 M; 27 F | $8.4 \pm 4.0$ | 64 | children with suspected obstructive sleep apnoea |
| Roberts (2020) | Apple Watch; Oura Ring | N/A | 5 M; 3 F | $40.8 \pm 4.8$ | 25 (apple watch), 26 (oura) | Healthy adults |
| de Zambotti (2015) | Jawbone UP | Manual | 28 F | $50.1 \pm 3.9$ | 46 | middle-aged women |

**Table 2.1 continued.** Characteristics of studies included in the systematic search

| Paper | Device | Method | Gender | Age | Number of sleep records | Population |
|---|---|---|---|---|---|---|
| de Zambotti (2018) | Fitbit Charge 2 | N/A | 12 M; 23 F (control); 6 M, 3 F (PLM) | 35.0 ± 12.0 (control); 42.0 ± 15.0 (PLM) | 44 | healthy adults |
| Kang (2017) | Fitbit Flex | Manual | 14 M, 19 F (insomnia); 6 M; 11 F (good sleepers) | 38.4 ± 11.2 (insomnia); 22.0 ± 2.6 (good sleepers) | 33 insomnia; 17 good sleepers | healthy adults and adults with insomnia |
| Cook (2017) | Fitbit Flex | Manual | 4 M; 17 F | 26.5 ± 4.6 | 21 | Major Depressive Disorder sufferers |
| Maskevich (2017) | Jawbone UP2, Fitbit One | N/A | 1 M; 6 F | 54.1 ± 6.4 | 7 | Huntington's gene carriers |
| Moreno-Pino (2019) | Fitbit Charge 2 And Fitbit Alta HR | N/A | 42 M; 23 F | 58.8 ± 13.8 | 31 (Charge); 34 (Alta) | adults referred to an outpatient sleep clinic |
| Kahawage (2020) | Fitbit Alta HR | Manual | 19 M; 23 F | 49.1 ± 17.5 | 38 | adults with chronic insomnia |
| Cook (2018) | Jawbone UP3 | Manual | 29 M; 14 F | 33.3 ± 11.0 | 43 | clinical patients referred for PSG |

**Table 2.1 continued.** Characteristics of studies included in the systematic search

| Paper | Device | Method | Gender | Age | Number of sleep records | Population |
|---|---|---|---|---|---|---|
| Montgomery (2012) | Fitbit | Manual | 14 M; 10 F | 26.1 | 24 | healthy adults |
| Kanady (2020) | Basis B1 | N/A | 5 M; 13 F | 26.8 ± 3.4 | 88 | healthy young adults |
| Ameen (2019) | Mi Band | Automatic detection | 13 F | 29.0 ± 13.0 | 21 | healthy adults |
| Pesonen (2018) | Polar Fitness Tracker | N/A | 8 M; 9 F (younger), 9 M; 8 F (older) | 11.0 ± 0.8 (younger); 17.8 ± 1.8 (older) | 34 | healthy children |
| de Zambotti (2015) | Jawbone UP | Manual | 37 M; 28 F | 15.8 ± 2.5 | 65 | healthy adolescents |
| de Zambotti (2016) | Fitbit Charge HR | Manual | 17 M; 15 F | 17.3 ± 2.5 | 32 | healthy adolescents |
| Dunican (2018) | Readiband | NS | 20 M; 30 F | 57.0 ± 5.0 | 50 | middle aged adults |
| Shin (2015) | Sensewear Armband | N/A | 6 M; 3 F | 23.3 ± 4.1 | 81 | healthy adults |

Notes: N/A = not available; M = male; F = female.

**Table 2.2** Aggregated sleep comparisons for included studies.

| Paper | Device | TST Bias | Wake Bias | Light Sleep Bias | SWS Bias | REM Bias |
|---|---|---|---|---|---|---|
| Roane (2015) | Sensewear Armband | -10.1 | 7.7 | | | |
| Cook (2019) | Fitbit Alta Hr | 11.6 | -7.63 | -11.2 | 18.2 | 10.1 |
| O'Driscoll (2013) | Sensewear Armband | 26 | -17.5 | - | - | - |
| Alsaadi (2014) | Sensewear-Pro3 | -3.23 | -4.96 | - | - | - |
| Meltzer (2015) | Fitbit Ultra | 41 (normal setting), -105 (sensitive setting) | -32 (normal setting), 106 (sensitive setting) | - | - | - |
| Lee (2019) | Fitbit Alta Hr | -24.1 (5h TIB); -37.1 (6.5h TIB); -47.0 (9 h TIB) | 21.0 (5h TIB); 29.4 (6.5h TIB); 41.9 (9h TIB) | 9.9 (5h TIB); 3.3 (6.5h TIB); -20.9 (9h TIB) | -37.8 (5h TIB); -46.4 (6.5h TIB); -21.5 (9h TIB) | 3.8 (5h TIB); 6.0 (6.5h TIB); -4.9 (9h TIB) |
| Toon (2016) | Up | 9 | -9 (UP) | - | - | - |
| Roberts (2020) | Oura Ring | 15.6 (oura) | -14.1 (oura) | - | - | - |
| de Zambotti (2015) | Jawbone UP | 26.6 | 31.2 | - | - | - |
| de Zambotti (2018) | Fitbit Charge 2 | -9 (control); -8 (PLMS) | -5 (control); -2 (PLMS) | -34 (control); -35 (PLMS) | 24 (control), 28 (PLMS) | 1 (control); 0 (PLMS) |

**Table 2.2 (continued)** Aggregated sleep comparisons for included studies.

| Paper | Device | TST Bias | Wake Bias | Light Sleep Bias | SWS Bias | REM Bias |
|---|---|---|---|---|---|---|
| Kang (2017) | Fitbit Flex | 32.9 (insomnia) <br> 6.5 (good sleepers) | -30.5 (insomnia) <br> -7.1 (good sleepers) | - | - | - |
| Cook (2017) | Fitbit Flex | 46 (normal setting) <br> -86.3 (sensitive setting) | -44 (normal setting) <br> 74.8 (sensitive setting) | - | - | - |
| Maskevich (2017) | Jawbone UP2, Fitbit One | 78.7 (jawbone) <br> 88.1 (fitbit) | -36 (jawbone) <br> -39 (fitbit) | - | - | - |
| Moreno-Pino (2019) | Fitbit Charge 2 And Fitbit Alta HR | -59.78 (mean of both devices) | 36.1 | -68.8 | 74.2 | 54.2 |
| Kahawage (2020) | Fitbit Alta HR | 53.3 | -48.4 | 138.3 | -59.4 | 18.5 |
| Cook (2018) | Jawbone UP3 | 39.6 | -34.3 | 0.24 | 24.2 | |
| Montgomery (2012) | Fitbit | 67.1 (fitbit) | | | | |
| Kanady (2020) | Basis B1 | 15.11 | -15.11 | 16.1 | -8.6 | 5.2 |
| Ameen (2019) | Mi Band | 69.6 | -33.6 | - | - | - |
| Pesonen (2018) | Polar Fitness Tracker | -28.9 (younger) <br> -20.6 (older) | 24.4 (younger) <br> 12.5 (older) | - | - | - |

**Table 2.2 (continued).** Aggregated sleep comparisons for included studies.

| Paper | Device | TST Bias | Wake Bias | Light Sleep Bias | SWS Bias | REM Bias |
|---|---|---|---|---|---|---|
| de Zambotti (2015) | Jawbone UP | -10.0 | 9.3 | - | - | - |
| de Zambotti (2016) | Fitbit Charge HR | 8.0 | -5.6 | - | - | - |
| Dunican (2018) | Readiband | 58.0 | -70.0 | - | - | - |
| Shin (2015) | Sensewear Armband | -33.0 | 20.9 | - | - | - |

Notes: TIB = time in bed; PLMS = periodic leg movements syndrome.

**Table 2.3** Epoch-by-epoch concordance statistics.

| Paper | Device | Two-stage agreement | Sensitivity for sleep | Specificity for wake | Sensitivity for light sleep |
|---|---|---|---|---|---|
| Roane (2015) | Sensewear Band | 88% | 94% | 39% | |
| Cook (2019) | Fitbit Alta Hr | 90% | 96% | 58% | 73% |
| O'Driscoll (2013) | Sensewear Band | 79.9% | 88.7 | 50% | - |
| Alsaadi (2014) | Sensewear-Pro3 | 85% | 90% | 54% | - |
| Meltzer (2015) | Fitbit Ultra | 84% (normal setting)<br>71% (sensitive setting) | 87% (normal setting)<br>70% (sensitive setting) | 52% (normal)<br>79% (sensitive) | - |
| Lee (2019) | Fitbit Alta Hr | 90% | 90% | 88% (5h TIB)<br>90% (6.5h TIB)<br>88% (9h TIB) | 71% (5h)<br>68% (6.5h)<br>71% (9h) |
| Toon (2016) | Up | | 92% | 66% | - |
| Roberts (2020) | Apple Watch; Oura Ring | 89% (Cole-Kripke)<br>88% (Sadeh)<br>89% (Oura) | 94% (Cole-Kripke)<br>91% (Sadeh)<br>96% (oura) | - | - |

**Table 2.3 (continued).** Epoch-by-epoch concordance statistics.

| Paper | Device | Two-stage agreement | Sensitivity for sleep | Specificity for wake | Sensitivity for light sleep |
|---|---|---|---|---|---|
| de Zambotti (2015) | Jawbone UP | - | 96% | 37% | - |
| de Zambotti (2018) | Fitbit Charge 2 | - | 96% | 61% | 81% |
| Kang (2017) | Fitbit Flex | 87% (normal setting, insomnia) | 97% (normal setting, insomnia) | 35% (normal setting, insomnia) | - |
| | | 68% (sensitive setting, insomnia) | 64% (sensitive setting, insomnia) | 89% (sensitive setting, insomnia) | |
| | | 93% (normal setting, good sleepers) | 97% (normal setting, good sleepers) | 36% (normal setting, good sleepers) | |
| | | 66% (sensitive setting, good sleepers) | 65% (sensitive setting, good sleepers) | 82% (sensitive setting, good sleepers) | |
| Cook (2017) | Fitbit Flex | 88% (normal setting) | 98% (normal setting) | 35% (normal setting) | - |
| | | 78% (sensitive setting) | 78% (sensitive setting) | 80% (sensitive setting) | |

**Table 2.3 (continued)** Epoch-by-epoch concordance statistics.

| Paper | Device | Two-stage agreement | Sensitivity for sleep | Specificity for wake | Sensitivity for light sleep |
|---|---|---|---|---|---|
| Maskevich (2017) | Jawbone UP2, Fitbit One | 83% (Jawbone) <br> 81% (Fitbit) | 99% (Jawbone) <br> 99% (Fitbit) | 34% (Jawbone) <br> 27% (Fitbit) | |
| Moreno-Pino (2019) | Fitbit Charge 2 And Fitbit Alta | | 87% (Charge) <br> 88% (Alta) | 35% (Charge) <br> 52% (Alta) | |
| Kahawage (2020) | Fitbit Alta | 83% | 97% | 39% | 79% |
| Cook (2018) | Jawbone UP3 | 87% | 97% | 39% | 60% |
| Montgomery (2012) | Fitbit | | 98% | 19.8% | |
| Kanady (2020) | Basis B1 | 54% | 99% | | 62% |
| Ameen (2019) | Mi Band | 53% | 99% | | 71% |
| Pesonen (2018) | Polar Fitness Tracker | younger: 91% <br> older: 90% | younger: 93% <br> older: 91% | younger: 77% <br> older: 83% | |

Notes: TIB = time in bed

## 2.2 Sleep in the athlete population

### 2.2.1 Physical training

Athletes and coaches design short and long-term training programs to maximise adaptations and optimise performance prior to competitive periods (i.e., periodisation; Mujika, Halson et al., 2018). Training programs typically consist of short-term periods of intensive training followed by periods of reduced training (i.e., taper) to maximise the physiological effects of the previous training period (Mujika, Halson et al., 2018). Periodisation is typically consistent across most sports; however the structure of training can vary greatly between and within sports and athletes (e.g., timing of training, training load; Halson, 2014b), which may impact sleep/wake behaviour.

Regarding the timing of training, there is a tendency for individual sports (e.g., swimmers) to train in the early morning (Greyson, Kelly et al., 2018), while team-based field sports (i.e., soccer) typically train in the late morning or afternoon (Whitworth-Turner, Di Michele et al., 2018). Thus, athletes from individual sports tend to go to bed earlier, wake up earlier, and obtain less sleep than athletes from team sports (Lastella, Roach et al., 2015). For example, sleep duration is severely restricted on the nights prior to early morning training sessions compared to nights prior to rest days in swimmers. Specifically, swimmers who trained between 05:00h and 06:00h obtained <5h of sleep, while swimmers who trained between 10:00h and 11:00h obtained >7h sleep (Sargent, Halson et al., 2014). Regardless of what time training occurs, sleep duration is shortened prior to training days when compared to rest days (i.e., days on which no training occurs) in both individual and team sport athletes (Sargent, Halson et al., 2014; Sargent, Lastella et al., 2014; Kolling, Steinacker et al., 2016; Caia, Scott et al., 2017), and is severely restricted prior to training scheduled before 07:00h (Sargent, Halson et al., 2014; Kolling, Steinacker et al., 2016; Caia, Scott et al., 2017; Whitworth-Turner, Di Michele et al., 2018; Gudmundsdottir, 2019). While these findings suggest that training commitments impact the sleep/wake behaviours of athletes, it is also important to consider the physiological impact of training on sleep.

Differences in the physiological demands between sports influences the amount and type of training athletes are exposed to, which in turn, may influence the sleep obtained by athletes. Large increases in training load (>25%) reduce sleep duration and sleep efficiency in rugby league players, synchronised swimmers and cyclists (Lastella, Roach et al., 2015a; Schaal, Y et al., 2015; Thornton, Duthie et al., 2017; Roberts, Teo et al., 2019). Differences in training demands can be highlighted in sports that may be considered similar in nature, such as those identified as "football" codes. Australian rules football, rugby union, and soccer are all classified as football codes, yet there are major differences in the physiological demands of each code. Australian rules football is a high-impact sport (i.e., tackling; high velocity collisions) with high aerobic demand, while rugby is characterised by repeated physical collisions combined with short bursts of high-intensity efforts (Varley, Gabbett et al., 2014), and soccer requires frequent changes of direction, vertical leaps, and high aerobic demand (Iaia, Rampinini et al., 2009). In this regard, Australian rules football players experienced more sleep disturbances compared to rugby union and soccer players (Miller, Sargent et al., 2017). In addition to aforementioned factors such as timing of training, it is reasonable to suggest that athletes who are exposed to training with higher physiological demand may be more likely to experience disturbed sleep.

### 2.2.2 Competition

The aim of any training program is to prepare athletes to perform optimally during competition. During competition, the focus shifts to acute performance rather than developing physiological adaptations as during training. Immediately after competition, athletes aim to facilitate adequate recovery for subsequent competition or training bouts (Mujika, Halson et al., 2018). On the night prior to competition, athletes typically obtain sufficient sleep (>7h total sleep time and >85% sleep efficiency (Fowler, Duffield et al., 2014; Fowler, Duffield et al., 2015; Lastella, Roach et al., 2015; Shearer, Jones et al., 2015; Eagles and Lovell, 2016; Fullagar, Duffield et al., 2016; Sargent and Roach, 2016; Caia, Scott et al., 2017; Lalor, Halson et al., 2018). In comparison, athletes typically obtain less sleep on the night of competition compared to surrounding nights (Richmond, Dawson et al., 2004; Fowler, Duffield et al., 2014; Fowler, Duffield et al., 2015; Shearer, Jones et al., 2015; Eagles and Lovell, 2016; Fullagar, Duffield et al., 2016; Caia, Scott et al., 2017; Dunican, Higgins et al., 2018; Lalor, Halson et al., 2018;

O'Donnell, Beaven et al., 2018; Roberts, Teo et al., 2019). The time of day that competition occurs can also affect sleep duration. Reductions in sleep duration are greater (~80 minutes) after competition starting later than 18:00h compared to competition starting during the day (~20 minutes; Roberts, Teo et al., 2019). A reduction in sleep duration or sleep efficiency on the night of competition may be due to various factors such as delayed bedtime (Roberts, Teo et al., 2019), consumption of caffeine before or during competition (Dunican, Higgins et al., 2018), post-game alcohol consumption (Barnes, 2014), and/or exercise-induced elevations in core body temperature (Sargent and Roach, 2016). However, further research is needed to isolate mechanisms that may have a significant impact on athlete sleep.

### 2.2.3 Transmeridian travel

Athletes often travel across multiple time zones to take part in training camps and competition (i.e., transmeridian travel). Following transmeridian travel, it is likely athletes may experience travel fatigue and jet lag. Travel fatigue is caused by sub-optimal sleep environments on the plane and/or in airports, extended periods of inactivity, and the dehydrating effects of recycled air during long-haul flights (Reilly, Atkinson et al., 1997; Lastella et al., 2019). Collectively, these factors have been shown to compromise the sleep of athletes during travel (Lastella; Fowler, Duffield et al., 2015b; Fullagar, Duffield et al., 2016; Thornton, Miller et al., 2018). While travel fatigue directly impacts the homeostatic mechanism of sleep/wake regulation, the second mechanism of sleep/wake regulation, known as the human circadian oscillator, is affected when an individual rapidly crosses time zones. Following transmeridian travel, the circadian system is aligned with the environmental cues (i.e., sunlight) at the point of departure rather than the new destination (i.e., jet lag; Burgess, Crowley et al., 2003; Lastella, Roach et al., 2014). For example, if an athlete travels from London to Sydney, a habitual bedtime of 23:00h in London will correspond to 12:00h in Sydney. The circadian system is primed for sleep at an inappropriate time for the new destination and environmental cues. Symptoms of jet lag include fatigue, insomnia, daytime sleepiness, and gastrointestinal discomfort (Boulos, Campbell et al., 1995) which gradually subside as the circadian system aligns with environmental time cues (Eastman and Burgess, 2009). Similarly, if an athlete is required to train or compete in the days immediately following arrival at a new destination following transmeridian travel, they may be doing so at a time

that their body is primed for rest. Therefore, strategies to facilitate body clock adaption are of interest to athletes and coaches seeking to reduce the time spent in circadian misalignment following travel.

*2.2.4 Exercise timing, sleep and core body temperature*

The human circadian rhythms of core body temperature and sleep are closely linked (Lack and Lushington, 1996). Within the brain, the anterior hypothalamus plays a major role in both sleep and thermoregulation (McGinty, 1990). The onset of sleep is associated with peripheral heat loss in the body, which is then followed by a gradual decrease in core body temperature during sleep (Glotzbach and Heller, 1976). Core body temperature is heightened during periods of wakefulness (Van Dongen and Dinges, 2000) and rises significantly during exercise as a by-product of muscular contraction (Waterhouse, Drust et al., 2005; Sawka, 2011). Elevated core body temperature following exercise can negatively affect subsequent sleep (Horne and Staff, 1983; Horne and Moore, 1985; Waterhouse, Drust et al., 2005). In particular, evening exercise is thought to result in disturbed sleep due to increased arousal and increased core body temperature (Davies, 1979; Browman, 1980; Driver and Taylor, 1996; Waterhouse, Drust et al., 2005).

While traditional recommendations of sleep hygiene have generally advised against evening exercise (Zarcone, 1994), epidemiological evidence indicates that there may be a positive impact of evening exercise on sleep (Vuori, Urponen et al., 1988). For example, Vuori et al. (1988) examined self-reported sleep measures in a random sample of adults (N=1,600) who exercised in the afternoon (i.e., 16:00h-20:00h) or in the evening (i.e., after 20:00h). A negative effect of evening exercise was uncommon, with participants reporting reduced sleep onset latency, improved sleep quality and waking up feeling refreshed following evening exercise compared to days when they did not exercise (Vuori, Urponen et al., 1988). In addition, the American National Sleep Foundation conducted a poll among the general population assessing the relationship between sleep and the timing of exercise. Sleep outcomes (e.g., subjective sleep quality) were similar in people who exercised in the evening compared to people who exercised in morning and afternoon (Buman et al., 2013). Despite evidence suggesting that evening exercise may not be harmful to sleep (Buman et al., 2013), and in some cases may even improve sleep (Vuori, Urponen et al., 1988), these data only highlight an association between the time

of day of exercise and sleep. Such associations do not necessarily confirm that a variable (e.g., exercise timing) is directly responsible for changes in another variable (e.g., sleep quality). For example, there may be differences in other daily behaviours of people who exercise in the evening compared to those who exercise in the morning or afternoon that could account for the differences in sleep outcomes. For a robust comparison, an experimental protocol can be used to isolate the effects of the time of day of exercise on sleep and to reduce or eliminate the effect of potential confounds on the variables of interest. Thus, experimental studies are needed to manipulate the timing of exercise and measure the subsequent impact on sleep.

Several laboratory studies have reported evening exercise does not negatively impact sleep (O'Connor and Youngstedt, 1995; Dworak, Wiater et al., 2008; Myllymaki, Kyrolainen et al., 2011; Brand, Kalak et al., 2014; Alley et al., 2015). In a study examining the effects of evening exercise on the sleep of healthy young males, Flausino et al. (2012) found increased sleep efficiency and increased REM onset latency following evening aerobic exercise in comparison to a no exercise baseline night. Dworak et al. (2008) compared the impact of moderate-intensity aerobic exercise and high-intensity aerobic exercise performed 3-4 h before bedtime on sleep efficiency, sleep duration, sleep onset latency, and sleep architecture (i.e., the time spent in different stages of sleep) and observed no differences in sleep variables following moderate-intensity exercise compared to the no exercise baseline night, but increases in SWS following high-intensity exercise (Dworak, Wiater et al., 2008). Similarly, O'Connor et al. (1998) examined the effect of evening low-intensity aerobic cycling exercise and evening moderate-intensity aerobic cycling exercise performed 30-90 minutes before bedtime on the sleep of healthy young adults; there was no difference in sleep onset, sleep duration, or sleep efficiency between conditions (O'Connor, 1998). In one of the few studies incorporating resistance exercise, Alley et al. (2015) compared the effects of the timing of resistance exercise (i.e., sessions at 07:00h, 13:00h and 19:00h) with major muscular groups on sleep against a control condition (i.e., no resistance exercise). During the exercise sessions, participants performed 3 sets of 10 repetitions of each exercise at 65% of their 10-repitition maximum. Participants experienced significantly less wake episodes in all resistance exercise conditions compared to the control condition and took longer to fall asleep after completing resistance exercise at 13:00h and 19:00h compared to 07:00h. These results suggest that while evening

resistance exercise may reduce the amount of wake in sleep, it may also increase the time taken to fall asleep. Overall, there are conflicting findings regarding the effects of evening exercise on sleep, with limited data available exploring resistance exercise. Some of the differences in results between studies can be explained by the type of exercise examined (e.g., aerobic vs. resistance), the duration of the exercise session (e.g., short or long), and the intensity of the exercise performed (e.g., low, moderate, high).

Several methodological differences can be highlighted in the literature regarding evening exercise and sleep. Exercise modality differs in most protocols. Generally, studies have either not reported exercise modality (Brand, Kalak et al., 2014), or have utilised one exercise modality (O'Connor, 1998; Dworak, Wiater et al., 2008; Myllymaki, Kyrolainen et al., 2011; Brand, Kalak et al., 2014; Alley, 2015). For instance, Brand et al. (2014) examined the impact of evening self-perceived exercise exertion on sleep without controlling for exercise modality, intensity, timing, or duration. Therefore, it is difficult to interpret the results when the exercise stimulus has not been controlled. In comparison, Dworak et al. (2008), O'Connor et al. (1998), and Alley et al. (2015) implemented session-specific exercise protocols related to exercise mode, exercise intensity, and exercise timing. Despite these studies implementing robust methodologies regarding exercise in the evening, there are several combinations of exercise methodology and timing that have not yet been investigated. This limits the degree to which meaningful inferences regarding the overall effect of evening exercise on sleep can be made across studies.

**2.3 Interventions to minimise compromised sleep**

*2.3.1. Jet lag interventions*

Effective jet lag countermeasures are based upon manipulation of the human circadian system. By using established chronobiological principles, it is possible to estimate the timing of the human circadian system and to apply an appropriate intervention in situations when the circadian system becomes misaligned with the external environment (i.e., following transmeridian travel).

The circadian rhythms of core body temperature and melatonin are strongly linked with the sleep/wake cycle (Cagnacci, Soldani et al., 1996; Lewy, Cutler et al., 1999). Endogenous melatonin secretion typically begins 2 h prior to habitual bedtime (Burgess, Savic et al., 2003), while the daily minimum core body temperature (CBTmin) typically occurs ~7 h after melatonin onset, which coincides with the low point of the circadian cycle (Cagnacci, Soldani et al., 1996). Based on the rhythms of these two variables, it is then possible to make certain assumptions regarding their relationship with an individual's sleep/wake behaviour. For example, melatonin onset will occur at 21:00h, CBTmin will occur at 04:00h, and CBTmax will occur at 16:00h in an individual who normally sleeps from 23:00h to 07:00h (Roach and Sargent, 2019).

Stimuli such as exercise, bright light, and exogenous melatonin are known to shift the timing of the circadian system (Roach and Sargent, 2019). Shifting of the circadian system is most sensitive to retinal light exposure in a "critical period" in the hours either side of CBTmin (Wever, Polášek et al., 1983; Czeisler, Allan et al., 1986; Khalsa, Jewett et al., 2003). This critical period has been identified using the human phase response curve to light, whereby light exposure directly before CBTmin delays (i.e., shifts earlier in time) the circadian system and light exposure after the CBTmin advances (i.e., shifts later in time) the circadian system (Khalsa, Jewett et al., 2003). Using estimations of CBTmin and application of appropriately timed light exposure (Roach and Sargent, 2019) is an effective strategy to enhance circadian adaptation to destination time zones following transmeridian travel (Burgess, Crowley et al., 2003; Eastman, Gazda et al., 2005; Eastman and Burgess, 2009). For example, if an athlete travels from London to Perth, their CBTmin (04:00h London time) would occur at 12:00h Perth

time. Subsequently, this athlete will feel the need for sleep at a time that is not appropriate for the destination. However, if the athlete avoids light in the three hours prior to CBTmin (i.e., 09:00-12:00h Perth time) and seeks light exposure in the three hours following CBTmin (i.e., 12:00-15:00h Perth time), they should experience a phase advance and enhance their entrainment to the destination time zone. Interventions to minimise jet lag recommend an adaptation period of approximately 1 day per hour of time zone change (Roach and Sargent, 2019).

Light exposure protocols to prevent or reduce jet lag have been conducted in laboratory conditions (Czeisler, Kronauer et al., 1989; Burgess, Crowley et al., 2003; Eastman, Gazda et al., 2005); however very few studies have examined the effectiveness of a light exposure protocol on circadian adaptation in the field with athletes who have rapidly crossed time-zones. In theory, light exposure schedules should facilitate body clock adaptation following transmeridian travel. However, their efficacy in the field has not been established because of the difficulties associated with measuring the timing of the circadian system. For this reason, people use subjective measures, but these can be affected by other symptoms (fatigue; Thompson, Batterham et al., 2013). Indeed, experimental studies have utilised circadian biomarkers such as salivary dim light melatonin onset to measure jet lag (Boulos, Macchi et al., 2002). So although the measurement of dim light melatonin onset provide an accurate marker of circadian phase, the data collection method can be time-consuming and may be impractical for field studies (Burgess, Wyatt et al., 2015). While the circadian principles of bright light interventions are well understood, the practicality of implementing them successfully, and acquiring an objective measure of circadian phase in field settings is not. Given the difficulty in measuring the timing of the circadian system, subjective jet lag measures (e.g., Liverpool Jet Lag Scale) have been used to track the alleviation of jet lag symptoms (Waterhouse, et al 2009). However, such scales have not been validated against objective biomarkers of circadian phase. If a subjective measure could be used to measure jet lag in the field with some degree of confidence, it would allow for convenient estimation of adaptation following transmeridian travel.

*2.3.2 Sleeping aids*

To combat factors impacting sleep (e.g., transmeridian travel, training, competition), athletes may resort to alternative means of initiating sleep (e.g., pharmaceutical sleep medication; Waterhouse, Reilly, and Edwards, 2004). Medications such as benzodiazepines are effective for inducing sleep; however they do not necessarily result in a prolonged, quality sleep period and have been associated with adverse side–effects (Longo and Johnson, 2000; Thomas Reilly and Edwards, 2007). Prescription of benzodiazepines is associated with significant risk of dependence syndrome, psychomotor impairments, daytime sleepiness, and withdrawal symptoms following long-term use (Hajak, Muller, Wittchen, Pittrow, and Kirch, 2003; Vgontzas, Kales, and Bixler, 1995). For this reason, prescription of non-benzodiazepines, such as zolpidem and zopiclone, have been favoured by practitioners (Hajak et al., 2003). While both classes of medications possess similar sedative effects, non-benzodiazepines act upon a different molecular pathway to benzodiazepines. Subsequently, properties linked to the residual effects of benzodiazepines are not activated by non-benzodiazepines (Hajak et al., 2003). Side-effects related to daytime functioning may be detrimental to athletes utilising benzodiazepines to obtain sleep prior to training or competition. While non-benzodiazepines medications are preferred by practitioners, usage of zolpidem and zopiclone may result in dependence syndrome and withdrawal symptoms if use is ceased (Hajak et al., 2003). Therefore, a non-pharmaceutical means of enhancing sleep (i.e., dietary intake) is of great interest to elite athletes and sport scientists.

It is well established that athletes should maintain a healthy balanced diet and obtain an adequate amount sleep to maintain optimal athletic performance (Mujika, Halson et al., 2018). However, it is important to consider the effect of dietary supplementation on sleep. For example, diets high in carbohydrate may result in shorter sleep onset latency, diets high in protein may result in improved sleep quality, and diets high in fat may negatively influence sleep duration (Lindseth et al., 2013). The mechanism behind the effects of diet on sleep are likely associated with hormonal release within the brain (Halson, 2008). For example, melatonin is a hormone released from the pineal gland, which influences the human sleep/wake cycle (Van Cauter and Tasali, 2011). The neurotransmitters within the brain associated with the release of melatonin include serotonin, gamma-aminobutyric acid,

orexin, melanin-concentrating hormone, cholinergic, galanin, noradrenaline, and histamine (Halson, 2008; Halson, 2014). Dietary factors such as carbohydrate, tart cherry juice, valerian, L-theanine, and nucleotides may be used to stimulate some of these neurotransmitters to increase melatonin release (e.g., serotonin; Halson, 2014).

The release of melatonin via the synthesis of serotonin is subject to the availability of the amino acid tryptophan. Tryptophan is capable of crossing the blood-brain barrier and is required for synthesising serotonin, which synthesises melatonin (Halson, 2014). Melatonin can be increased by increasing the uptake of tryptophan or reducing the concentration of large neutral chain amino acids (Figure 2.3; Halson, 2014). Increasing the uptake of tryptophan can be facilitated by: (1) ingestion of a high fat meal increasing free fatty acids and increased tryptophan; and (2) exercise-induced increases in free fatty acids. The main mechanism for reducing the concentration of amino acids is decreasing its competition with tryptophan to cross the blood brain barrier, which can be facilitated by: (1) a diet high in protein containing more tryptophan than large neutral amino acids; and (2) ingestion of carbohydrate resulting in an increased tryptophan to branch chain amino acid ratio.

Despite knowledge of these molecular pathways, no studies have examined the optimal doses of macronutrient ingredients (i.e., combinations of carbohydrate, proteins, and fats) to stimulate the release of melatonin and subsequently improve sleep. Such a supplement would be of particular interest to athletes, with natural ingredients unlikely to result in the side-effects associated with pharmaceutical sleep medications.

**Figure 2.3** Effects of diet on tryptophan uptake and the central nervous system (Halson, 2014).
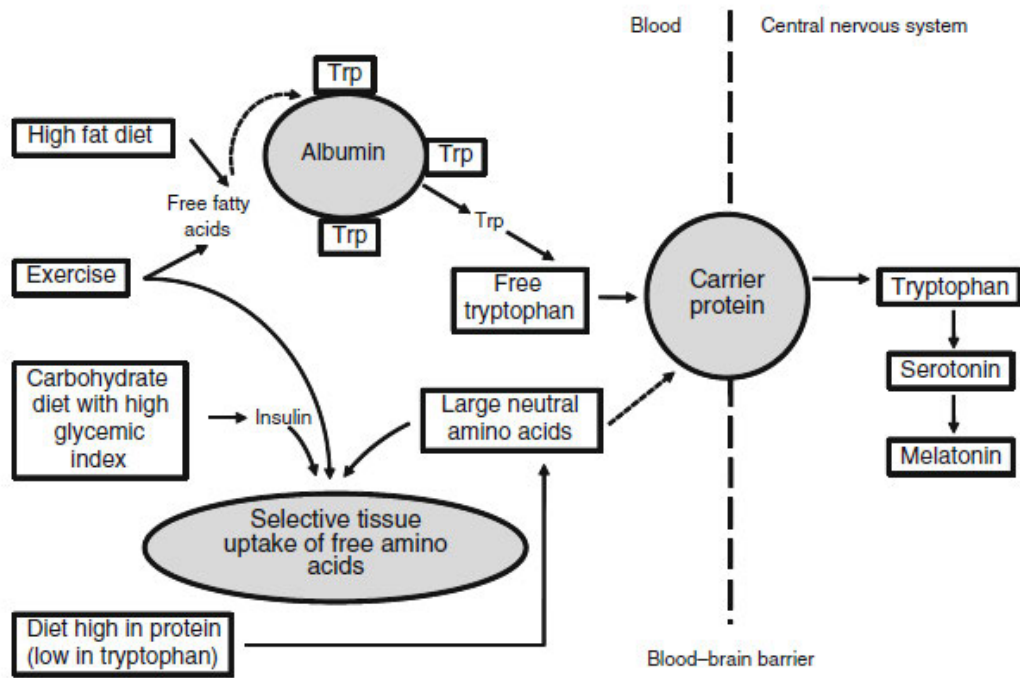
## 2.4 Rationale for research studies

The proposed studies aim to examine the measurement tools, countermeasures, and circadian strategies to mitigate compromised sleep in athletes.

*Chapter 3: A validation of the WHOOP strap to measure sleep duration and stages of sleep.*

Wearable sleep technology is a common tool used to monitor the sleep of elite athletes. However, few devices have been validated against the gold standard measurement of sleep – PSG. The aim of this Chapter was to conduct an epoch-by-epoch validation of the WHOOP strap to measure two-stage (i.e., sleep, wake) and four-stage sleep (i.e., wake, light sleep, SWS, REM) when accurate bedtimes are manually entered into the WHOOP smartphone application.

*Chapter 4: A validation of the WHOOP strap to detect and measure sleep and sleep staging*

Validations of sleep wearables have often neglected to state whether sleep was manually adjusted (i.e., accurate bed and wake times manually added for each sleep period), or whether sleep onset and offset was automatically detected by the device. This omission has consequences for athletes seeking clarification on how to best use sleep wearable devices. Therefore, the two aims of this Chapter were to: (1) compare the ability of actigraphy (i.e., ACTICAL), WHOOP-AUTO (i.e., when sleep is automatically detected by WHOOP) and WHOOP-MANUAL (i.e., when accurate bed times are manually entered into the WHOOP smartphone application) to measure 2-stage sleep (i.e., sleep or wake) against polysomnography; and (2) compare the ability of WHOOP-AUTO and WHOOP-MANUAL to measure 4-stage sleep (i.e., wake, light sleep, SWS, REM) against PSG.

*Chapter 5 The validity of a single question for assessing subjective jet lag in athletes.*

Commonly utilised measures of jet lag are primarily based upon subjective perceptions of jet lag and not circadian biomarkers. In previous research investigating strategies to cope with jet lag, urinary concentrations of a melatonin metabolite known as 6-suplhatoxymelatonin (aMT6s) has been used as a marker of circadian phase. Currently, there are no subjective measures of jet lag that have been validated

against circadian biomarkers. The aim of this Chapter was to validate the use of a single question to assess subjective jet lag in athletes against night-time urinary melatonin production.

*Chapter 6: The impact of moderate-intensity aerobic exercise and moderate-intensity resistance exercise performed in the evening on the sleep of healthy young males.*

To maintain a healthy lifestyle, it is recommended that humans obtain regular exercise and sufficient nightly sleep. However, sleep and exercise compete against each other in a time poor society. Individuals may choose to sacrifice sleep for exercise, or vice versa. In some cases, stringent training and competition schedules may require athletes to perform exercise (i.e., training or competition) in the evening. Traditional sleep recommendations advise against evening exercise due to potential detrimental effects to sleep. However, recent evidence suggests that evening exercise may not be detrimental to sleep and that evening exercise may be a viable option for athletes, or time poor individuals seeking to obtain the recommended amounts of sleep and exercise. Therefore, the aim of this Chapter was to examine the effect of a single bout of moderate-intensity aerobic exercise and a single bout of moderate-intensity resistance exercise completed 90 min before bedtime on subsequent night-time sleep in well-trained individuals.

*Chapter 7: Implementing a circadian adaptation schedule after eastward flight in young male athletes.*

Elite athletes who travel over multiple time zones (i.e., transmeridian travel) to train or compete are likely to experience jet lag. Jet lag is caused by the misalignment between the human circadian system and the destination time zone. Symptoms such as daytime sleepiness, gastrointestinal discomfort, and difficulty initiating night-time sleep gradually subside as the circadian system aligns to destination time cues (i.e., zeitgebers). Appropriately scheduled light exposure can be an effective strategy to enhance circadian adaptation to destination time zones following transmeridian travel. Due to complexities in implementing data collection protocols in the field, few studies have assessed the effectiveness a circadian adaptation using an appropriate phase marker in athletes. Therefore, the aim of this Chapter was to examine the effectiveness of light exposure/avoidance schedule to facilitate circadian adaptation

in young male cricket players using the night-time urinary melatonin production as a measure of circadian adaptation.

*Chapter 8: The effectiveness of a formulated sleep drink on sleep quality and quantity.*

Research has highlighted that many athletes do not obtain sufficient sleep. To combat potential sleep loss, athletes may turn to a pharmaceutical means of obtaining sleep. Medications such as benzodiazepines, zolpidem and zaleplon may be used to assist in falling asleep; however, they are subject to residual next-day side-effects such as drowsiness and decreased psychomotor performance. Research has highlighted certain macronutrients that stimulate neurotransmitters in the brain that can also improve sleep. However, the combined impact and optimal dosage of ingredients have not been empirically tested. Therefore, the aim of this Chapter was to examine the effectiveness of a formulated sleep drink on sleep quality and quantity.

**2.5 References**

Alley, J.R., Mazzochi, J.W., Smith, C.J., Morris, D.M; Collier, S.R (2015). Effects of resistance exercise timing on sleep architecture and nocturnal blood pressure. *Journal of Strength and Conditioning Research* 29(5):1378-1385.

Altman, D.G., and Bland, J.M. (1983). Measurement in Medicine: The Analysis of Method Comparison Studies. *The Statistician*, 32(3), 307.

Barnes, M.J. (2014). Alcohol: impact on sports performance and recovery in male athletes. Sports medicine, 44(7), 909-919.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., . . . Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of sleep research*, 12(1), 1-12.

Borbély, A.A., and Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. Journal of Biological Rhythms, 14(6), 559-570.

Borbély, A.A., and Achermann, P. (2009). Concepts and models of sleep regulation: an overview. Journal of Sleep Research, 1(2), 63-79. Borbely AA, Tobler I. (1989). Endogenous sleep-promoting substances and sleep regulation. *Physiological Reviews*. 69:605-670.

Boulos, Z., Campbell, S. S., Lewy, A. J., Terman, M., Dijk, D. J., and Eastman, C. I. (1995). Light treatment for sleep disorders: consensus report VII. Jet lag. *Journal of biological rhythms*, 10(2), 167-176.

Boulos, Z., Macchi, M. M., Stürchler, M. P., Stewart, K. T., Brainard, G. C., Suhner, A., . . . Steffen, R. (2002). Light visor treatment for jet lag after westward travel across six time zones. *Aviation, space, and environmental medicine*, 73(10), 953-963.

Brand, S., Kalak, N., Gerber, M., Kirov, R., Puhse, U., and Holsboer-Trachsler, E. (2014). High self-perceived exercise exertion before bedtime is associated with greater objectively assessed sleep efficiency. *Sleep medicine*, 15(9), 1031-1036.

Browman, C.P. (1980). Sleep following sustained exercise. *Psychophysiology*, 17(6), 577-580.

Buman, M. P., Phillips, B. A., Youngstedt, S. D., Kline, C. E., and Hirshkowitz, M. (2014). Does nighttime exercise really disturb sleep? Results from the 2013 National Sleep Foundation Sleep in America Poll. *Sleep Medicine*, 15(7), 755-761.

Burgess, H. J., Crowley, S. J., Gazda, C. J., Fogg, L. F., and Eastman, C. I. (2003). Preflight adjustment to eastward travel: 3 days of advancing sleep with and without morning bright light. *Journal of biological rhythms*, 18(4), 318-328.

Burgess, H. J., Savic, N., Sletten, T., Roach, G., Gilbert, S. S., and Dawson, D. (2003). The relationship between the dim light melatonin onset and sleep on a regular schedule in young healthy adults. B*ehavioral sleep medicine*, 1(2), 102-114.

Burgess, H. J., Wyatt, J. K., Park, M., and Fogg, L. F. (2015). Home circadian phase assessments with measures of compliance yield accurate dim light melatonin onsets. *Sleep*, 38(6), 889-897.

Cagnacci, A., Soldani, R., Laughlin, G. A., and Yen, S. S. (1996). Modification of circadian body temperature rhythm during the luteal menstrual phase: role of melatonin. *Journal of applied physiology*, 80(1), 25-29.

Caia, J., Scott, T. J., Halson, S. L., and Kelly, V. G. (2017). Do players and staff sleep more during the pre- or competitive season of elite rugby league? *European journal of sport science*, 17(8), 964-972.

Carskadon, M. A., and Dement, W. C. (2011). Normal human sleep: an overview. *Principles and practice of sleep medicine*, 5, 16-26.

Czeisler, C., Allan, J., Strogatz, S., Ronda, J., Sanchez, R., Rios, C., . . . Kronauer, R. (1986). Bright light resets the human circadian pacemaker independent of the timing of the sleep-wake cycle. *Science*, 233(4764), 667-671.

Czeisler, C. A., Duffy, J. F., Shanahan, T. L., Brown, E. N., Mitchell, J. F., Rimmer, D. W., . . . Emens, J. S. (1999). Stability, precision, and near-24-hour period of the human circadian pacemaker. *Science*, 284(5423), 2177-2181.

Czeisler, C. A., and Gooley, J. J. (2007). *Sleep and circadian rhythms in humans,* 72, 579-597.

Czeisler, C. A., Kronauer, R. E., Allan, J. S., Duffy, J. F., Jewett, M. E., Brown, E. N., and Ronda, J. M. (1989). Bright light induction of strong (type 0) resetting of the human circadian pacemaker. *Science*, 244(4910), 1328-1333.

Dardente, H., Dardente, H., and Cermakian, N. (2007). Molecular circadian rhythms in central and peripheral clocks in mammals. *Chronobiology international*, 24(2), 195-213.

Davenne, D. (2009). Sleep of athletes – problems and possible solutions. Biological Rhythm Research, 40(1), 45-52.

Davies, C. T. (1979). Thermoregulation during exercise in relation to sex and age. *European journal of applied physiology*, 42(2), 71-79.

De Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M., and Baker, F. C. (2019). Wearable Sleep Technology in Clinical and Research Settings. *Medicine and science in sports and exercise*, 51(7), 1538-1557.

De Zambotti, M., Baker, F. C., Willoughby, A. R., Godino, J. G., Wing, D., Patrick, K., and Colrain, I. M. (2016). Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiology and behavior*, 158, 143-149.

De Zambotti, M., Rosas, L., Colrain, I. M., and Baker, F.C.(2017). The Sleep of the Ring: Comparison of the OURA Sleep Tracker Against Polysomnography. *Behavourial sleep med*icine, 1-15.

Driver, H. S., and Taylor, S. R. (1996). Sleep disturbances and exercise. *Sports medicine*, 21(1), 1-6.

Duffy, J., Kronauer, R., and Czeisler, C. (1996). Phase-shifting human circadian rhythms: influence of sleep timing, social contact and light exposure. *The journal of physiology*, 495(Pt 1), 289-297.

Dunican, I. C., Higgins, C. C., Jones, M. J., Clarke, M. W., Murray, K., Dawson, B., . . . Eastwood, P. R. (2018). Caffeine use in a Super Rugby game and its relationship to post-game sleep. *European journal of sport science*, 18(4), 513-523.

Dworak, M., Wiater, A., Alfer, D., Stephan, E., Hollmann, W., and Strüder, H. K. (2008). Increased slow wave sleep and reduced stage 2 sleep in children depending on exercise intensity. *Sleep medicine*, 9(3), 266-272.

Eagles, A. N., and Lovell, D. I. (2016). Changes in sleep quantity and efficiency in professional rugby union players during home-based training and match play. *The journal of sports medicine and physical fitness*, 56(5), 565-571.

Eastman, C.I., and Burgess, H. J. (2009). How to travel the world without jet lag. *Sleep medicine clinics*, 4(2), 241.

Eastman, C.I., Gazda, C.J., Burgess, H. J., Crowley, S. J., and Fogg, L. F. (2005). Advancing circadian rhythms before eastward flight: a strategy to prevent or reduce jet lag. *Sleep*, 28(1), 33.

Flausino, N.H., Da Silva Prado, J.M., de Queiroz, S.S., Tufik, S., and de Mello, M.T. (2012). Physical exercise performed before bedtime improves the sleep pattern of healthy young good sleepers. *Psychophysiology*, 49(2), 186-192.

Fowler, P., Duffield, R., Howle, K., Waterson, A., and Vaile, J. (2015). Effects of northbound long-haul international air travel on sleep quantity and subjective jet lag and wellness in professional Australian soccer players. *The international journal of sports physiology and performance*, 10(5), 648-654.

Fowler, P., Duffield, R., and Vaile, J. (2014). Effects of Domestic Air Travel on Technical and Tactical Performance and Recovery in Soccer. *International journal of sports physiology and performance,* 9(3), 378.

Fowler, P., Duffield, R., and Vaile, J. (2015). Effects of simulated domestic and international air travel on sleep, performance, and recovery for team sports. *Scandinavian journal of medicine and science in sports*, 25(3), 441-451.

Frank, M. G. (2013). Astroglial regulation of sleep homeostasis. *Current opinion in neurobiology*, 23(5), 812-818.

Fullagar, H. H., Duffield, R., Skorski, S., White, D., Bloomfield, J., Kolling, S., and Meyer, T. (2016). Sleep, Travel, and Recovery Responses of National Footballers During and After Long-Haul International Air Travel. *The international journal of sports physiology*, 11(1), 86-95.

Fuller, P. M., Gooley, J. J., and Saper, C. B. (2006). Neurobiology of the sleep-wake cycle: sleep architecture, circadian regulation, and regulatory feedback. *Journal of biological rhythms*, 21(6), 482-493.

Glotzbach, S. F., and Heller, H. C. (1976). Central nervous regulation of body temperature during sleep. *Science*, 194(4264), 537-539.

Greyson, I., Kelly, S., Peyrebrune, M., and Furniss, B. (2018). Interpreting and implementing the LTAD model: English Swimming Coaches' views on the Swimming LTAD in practice. *International journal of sports science and coaching*, 5, 403.

Gudmundsdottir, S. L. (2019). Training Schedule and Sleep in Adolescent Swimmers. *Pediatric exercise science*, 1-7.

Guo, H., Brewer, J. M., Lehman, M. N., and Bittman, E. L. (2006). Suprachiasmatic regulation of circadian rhythms of gene expression in hamster peripheral organs: effects of transplanting the pacemaker. *The journal of neuroscience*, 26(24), 6406-6412.

Hajak, G., Muller, W. E., Wittchen, H. U., Pittrow, D., and Kirch, W. (2003). Abuse and dependence potential for the non-benzodiazepine hypnotics zolpidem and zopiclone: a review of case reports and epidemiological data. *Addiction*, 98(10), 1371-1378.

Halson, S. (2008). Nutrition, sleep and recovery. *European journal of sport science*, 8(2), 119-126.

Halson, S.L. (2014a). Sleep in elite athletes and nutritional interventions to enhance sleep. Sports *Medicine*, 44(1), 13-23.

Halson S.L. (2014). Monitoring training load to understand fatigue in athletes. *Sports medicine,* 44:139-147.

Horne, J.A., and Moore, V.J. (1985). Sleep EEG effects of exercise with and without additional body cooling. *Electroencephalography and clinical neurophysiology*, 60(1), 33-38.

Horne, J. A., and Staff, L. H. (1983). Exercise and sleep: Body-heating effects. Sleep: *Journal of sleep research and sleep medicine,* 6(1), 36-46.

Iaia, F. M., Rampinini, E., and Bangsbo, J. (2009). High-intensity training in football. *The international journal of sports physiology and performance*, 4(3), 291-306.

Iber, C., Ancoli-Israel, S., Chesson, A., and Quan, S. (2007). The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. 2007. American Academy of Sleep Medicine, Westchester, IL.

Juliff, L.E., Halson, S.L., and Peiffer, J.J. (2015). Understanding sleep disturbance in athletes prior to important competitions. *Journal of science and medicine in sport*, 18(1), 13-18.

Kang, S. G., Kang, J. M., Ko, K. P., Park, S. C., Mariani, S., and Weng, J. (2017). Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of psychosomatic research.*, 97, 38-44.

Kattler, H., Dijk, D.J., and Borbely, A.A. (1994). Effect of unilateral somatosensory stimulation prior to sleep on the sleep EEG in humans. *Journal of sleep research*, 3(3), 159-164.

Khalsa, S.B., Jewett, M.E., Cajochen, C., and Czeisler, C.A. (2003). A phase response curve to single bright light pulses in human subjects. *Journal of physiology*, 549(Pt 3), 945-952.

Kolling, S., Steinacker, J.M., Endler, S., Ferrauti, A., Meyer, T., and Kellmann, M. (2016). The longer the better: Sleep-wake patterns during preparation of the World Rowing Junior Championships. *Chronobiology international*, 33(1), 73-84.

Kushida, C. A., Littner, M. R., Morgenthaler, T., Alessi, C. A., Bailey, D., Coleman Jr, J., . . . Kramer, M. (2005). Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*, 28(4), 499-521.

Lack, L., and Lushington, K. (1996). The rhythms of human sleep propensity and core body temperature. *Journal of sleep research*, 5(1), 1.

Lalor, B. J., Halson, S. L., Tran, J., Kemp, J. G., and Cormack, S. J. (2018). No Compromise of Competition Sleep Compared With Habitual Sleep in Elite Australian Footballers. *International journal of sports physiology and performance,* 13(1), 29.

Lastella, M., Lovell, G. P., and Sargent, C. (2014). Athletes' precompetitive sleep behaviour and its relationship with subsequent precompetitive mood and performance. *European journal of sport science*, 14 Suppl 1, S123-130.

Lastella, M., Memon, A.R., and Vincent, G.E. (2020). Global Research Output on Sleep Research in Athletes from 1966 to 2019: A Bibliometric Analysis. *Clocks and sleep*, 2(2), 99-119.

Lastella, M., Roach, G. D., Halson, S. L., Gore, C. J., Garvican-Lewis, L. A., and Sargent, C. (2014b). The effects of transmeridian travel and altitude on sleep: preparation for football competition. *Journal of sports science and medicine*, 13(3), 718-720.

Lastella, M., Roach, G. D., Halson, S. L., Martin, D. T., West, N. P., and Sargent, C. (2015). The impact of a simulated grand tour on sleep, mood, and well-being of competitive cyclists. *Journal of sports medicine and physical fitness*, 55(12), 1555-1564.

Lastella, M., Roach, G. D., Halson, S. L., Martin, D. T., West, N. P., and Sargent, C. (2015b). Sleep/wake behaviour of endurance cyclists before and during competition. *Journal of sports science*, 33(3), 293-299.

Lastella, M., Roach, G. D., Halson, S. L., and Sargent, C. (2015). Sleep/wake behaviours of elite athletes from individual and team sports. *European journal of sport science*, 15(2), 94-100.

Lastella, M., Roach, G. D., and Sargent, C. (2019). Travel fatigue and sleep/wake behaviors of professional soccer players during international competition. *Sleep health*, 5(2), 141-147.

Leeder, J., Glaister, M., Pizzoferro, K., Dawson, J., and Pedlar, C. (2012). Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *Journal of sports science*, 30(6), 541-545.

Leeder, J., Glaister, M., Pizzoferro, K., Dawson, J., and Pedlar, C. (2012). Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *Journal of sports science*, 30(6), 541-545.

Lewy, A. J., Cutler, N. L., and Sack, R. L. (1999). The Endogenous Melatonin Profile as a Marker for Circadian Phase Position. *Journal of biological rhythms,* 14(3), 227-236

Lindseth, G., Lindseth, P., and Thompson, M. (2013). Nutritional effects on sleep. *Western journal of nursing research.*, 35(4), 497-513.

Longo, L.P., and Johnson, B. (2000). Addiction: Part I. Benzodiazepines–side effects, abuse risk and alternatives. *American family physician*, 61(7), 2121-2128.

Luca, G., Haba Rubio, J., Andries, D., Tobback, N., Vollenweider, P., Waeber, G., . . . Tafti, M. (2015). Age and gender variations of sleep in subjects without sleep disorders. *Annalls of medicine*, 47(6), 482-491.

Ludbrook, J. (2010). Confidence in Altman-Bland plots: a critical review of the method of differences. *Clinical and experimental pharmacology and physiology*, 37(2), 143-149.

Maskevich, S., Jumabhoy, R., Dao, P. D. M., Stout, J. C., and Drummond, S. P. A. (2017). Pilot Validation of Ambulatory Activity Monitors for Sleep Measurement in Huntington's Disease Gene Carriers. *Journal of huntingtons disease*, 6(3), 249-253.

McGinty, D., Szymusiak, R. (1990). Keeping cool, a hypothesis about the mechanisms and functions of slow-wave sleep. *Trends in neurosciences,* 13(12), 480–487.

Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. (2015). Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. *Sleep*. 38:1323-1330.

Meltzer, L. J., Walsh, C. M., Traylor, J., and Westin, A. M. (2012). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep*, 35(1), 159-166.

Miller, D. J., Sargent, C., Vincent, G. E., Roach, G. D., Halson, S. L., and Lastella, M. (2017). Sleep/Wake Behaviours in Elite Athletes from Three Different Football Codes. *Journal of sports science and medicine*, 16(4), 604-605.

Mujika, I., Halson, S., Burke, L. M., Balague, G., and Farrow, D. (2018). An Integrated, Multifactorial Approach to Periodization for Optimal Performance in Individual and Team Sports. *The international journal of sports physiology and performance*, 13(5), 538-561.

Myllymaki, T., Kyrolainen, H., Savolainen, K., Hokka, L., Jakonen, R., Juuti, T., . . . Rusko, H. (2011). Effects of vigorous late-night exercise on sleep quality and cardiac autonomic activity. *Journal of sleep research*, 20(1 Pt 2), 146-153.

O'Connor, P. B., MJ; Youngstedt, SD. (1998). Exercise-induced increase in core temperature does not disrupt a behavioral measure of sleep. *Physiology and behaviour*, 64(3), 213-217.

O'Connor, P. J., and Youngstedt, S. D. (1995). Influence of exercise on human sleep. *Exercise and sport sciences reviews*, 23, 105.

O'Donnell, S., Beaven, C. M., and Driller, M. J. S. S. f. H. (2018). Sleep/wake behavior prior to and following competition in elite female netball athletes. *Nature and science of sleep*, 14(2), 289-295.

Pollak, C.P., Tryon, W.W., Nagaraja, H., and Dzwonczyk, R. (2001). How accurately does wrist actigraphy identify the states of sleep and wakefulness? *Sleep*, 24(8), 957-965.

Rechtschaffen, A., and Kales, A. (1968). A manual of standardized terminology, technique and scoring system for sleep stages of human sleep. Los Angeles Brain Information Service. Brain Information Institute, UCLA.

Reilly, T., Atkinson, G., and Waterhouse, J. (1997). Travel fatigue and jet-lag. *Journal of sports sciences*, 15(3), 365-369.

Reilly, T., and Edwards, B. (2007). Altered sleep–wake cycles and physical performance in athletes. *Physiology and behavior*, 90(2/3), 274-284.

Richmond, L., Dawson, B., Hillman, D., and Eastwood, P. (2004). The effect of interstate travel on sleep patterns of elite Australian Rules footballers. *Journal of science and medicine in sport*, 7(2), 186-196.

Roach, G. D., and Sargent, C. (2019). Interventions to Minimize Jet Lag After Westward and Eastward Flight. *Frontiers in physiology,* 10(927).

Roberts, S. S. H., Teo, W.-P., and Warmington, S. A. (2019). Effects of training and competition on the sleep of elite athletes: a systematic review and meta-analysis. *British Journal of sports medicine,* 53(8), 513-522.

Santos, R.V., Tufik, S., and De Mello, M.T. (2007). Exercise, sleep and cytokines: is there a relation? *Sleep medicine reviews*, 11(3), 231-239.

Sargent C, Halson S, Roach G.D. (2014). Sleep or swim? Early-morning training severely restricts the amount of sleep obtained by elite swimmers. *European journal of sport science*. 14 Suppl 1:S310-315.

Sargent, C., Lastella, M., Halson, S. L., and Roach, G.D. (2014). The impact of training schedules on the sleep and fatigue of elite athletes. *Chronobiology international*, 31(10), 1160-1168.

Sargent, C., Lastella, M., Romyn, G., Versey, N., Miller, D. J., and Roach, G. D. (2018). How well does a commercially available wearable device measure sleep in young athletes? *Chronobiology international*, 35(6), 754-758.

Sargent, C., and Roach, G. D. (2016). Sleep duration is reduced in elite athletes following night-time competition. Chronobiology international, 33(6), 667–670.

Savis, J. C. (1994). Sleep and athletic performance: Overview and implications for sport psychology. *The sport psychologist*, 8(2), 111-125.

Sawka, M.N., Wenger, C.B. and Pandolf, K.B. (2011). Thermoregulatory responses to acute exercise-heat stress and heat acclimation. In. *Comprehensive physiology*.

Schaal, K., Le Meur, Y., Louis, J., Filliard, J. R., Hellard, P., Casazza, G., and Hausswirth, C. (2015). Whole-Body Cryostimulation Limits Overreaching in Elite Synchronized Swimmers. *Medicine and science in sports and exercise*, 47(7), 1416-1425.

Schwartz, J. R., and Roth, T. (2008). Neurophysiology of sleep and wakefulness: basic science and clinical implications. *Current neuropharmacology*, 6(4), 367-378.

Shambroom, J.R., Fabregas, S.E., and Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of sleep research*, 21(2), 221-230.

Shearer, D. A., Jones, R. M., Kilduff, L. P., and Cook, C. J. (2015). Effects of competition on the sleep patterns of elite rugby union players. *European journal of sport science*, 15(8), 681-686.

Spiegel, K., Leproult, R., and Van Cauter, E. (1999). Impact of sleep debt on metabolic and endocrine function. *Lancet*, 354(9188), 1435-1439.

Spiegel, K., Tasali, E., Penev, P., and Van Cauter, E. (2004). Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. *Annals of internal medicine*, 141(11), 846.

Thompson, A., Batterham, A. M., Jones, H., Gregson, W., Scott, D., and Atkinson, G. (2013). The practicality and effectiveness of supplementary bright light for reducing jet-lag in elite female athletes. *International journal of sports medicine*, 34(07), 582-589.

Thornton, H. R., Duthie, G. M., Pitchford, N. W., Delaney, J. A., Benton, D. T., and Dascombe, B. J. (2017). Effects of a 2-week high-intensity training camp on sleep activity of professional rugby league athletes. *International journal of sports physiology and performance*, 12(7), 928.

Thornton, H. R., Miller, J., Taylor, L., Sargent, C., Lastella, M., and Fowler, P. M. (2018). Impact of short- compared to long-haul international travel on the sleep and wellbeing of national wheelchair basketball athletes. *Journal of sports science*, 36(13), 1476-1484.

Tilley, A. J., and Empson, J. A. (1978). REM sleep and memory consolidation. *Biological psychology*, 6(4), 293-300.

Van Cauter, E., Spiegel, K., Tasali, E., and Leproult, R. (2008). Metabolic consequences of sleep and sleep loss. *Sleep medicine*, 9(3), S23-S28.

Van Dongen, H. P. A., and Dinges, D. F. (2000). Circadian rhythms in fatigue, alertness, and performance. *Principles and practice of sleep medicine (3$^{rd}$ ed.),* 20, 391-399.

Varley, M. C., Gabbett, T., and Aughey, R. J. (2014). Activity profiles of professional soccer, rugby league and Australian football match play. *Journal of sports science*, 32(20), 1858-1866.

Vgontzas, A., Zoumakis, E., Bixler, E., Lin, H. M., Follett, H., Kales, A., and Chrousos, G. (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *Journal of clinical endocrinology and metabolism*, 89(5), 2119-2126.

Vgontzas, A.N., Kales, A., and Bixler, E.O. (1995). Benzodiazepine side effects: role of pharmacokinetics and pharmacodynamics. *Pharmacology*, 51(4), 205-223.

Vuori, I., Urponen, H., Hasan, J., and Partinen, M. (1988). Epidemiology of exercise effects on sleep. *Acta physiologica Scandinavica*. Supplementum, 574, 3–7.

Waterhouse, J., Drust, B., Weinert, D., Edwards, B., Gregson, W., Atkinson, G., . . . Reilly, T. (2005). The circadian rhythm of core temperature: origin and some implications for exercise performance. *Chronobiology international*, 22(2), 207-225.

Waterhouse, J., Reilly, T., and Edwards, B. (2004). The stress of travel. *Journal of sports sciences*, 22(10), 946-966.

Werner, H., Molinari, L., Guyer, C., and Jenni, O. G. (2008). Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. *The Archives of pediatrics and adolescent medicine*, 162(4), 350-358.

Wever, R. A., Polášek, J., and Wildgruber, C. M. (1983). Bright light affects human circadian rhythms. *Pflügers archiv european journal of physiology*, 396(1), 85-87.

Whitworth-Turner, C., Di Michele, R., Muir, I., Gregson, W., and Drust, B. (2018). A comparison of sleep patterns in youth soccer players and non-athletes. *Science and medicine in football*, 2(1), 3-8.

Yi, H., Shin, K., Shin, C. (2006). Development of the sleep quality scale. *Journal of sleep research,* 15(3), 309–316.

Youngstedt S.D., and O'Connor PJ. (1999). The Influence of Air Travel on Athletic Performance. *Sports medicine*, 28(3), 197–207.

Zarcone VP. (1994). Sleep hygiene. In. *Principles and Practice of Sleep Medicine*. Philadelphia: W.B. Saunders Co. , pp. 542-546.

**3. A validation study of the WHOOP strap against polysomnography to assess sleep.**

Published peer-reviewed publications associated with this Chapter (Appendix D):

Miller, D. J., Lastella, M., Scanlan, A. T., Bellenger, C., Halson, S. L., Roach, G. D., and Sargent, C. (2020). A validation study of the WHOOP strap against polysomnography to assess sleep. *Journal of sports sciences,* 38(22), 2631–2636.

**Abstract**

The aim of the study was to compare the WHOOP strap – a wearable device that estimates sleep based on measures of movement and heart rate derived from actigraphy and photoplethysmography, respectively. Twelve healthy adults (6 female, 6 male, aged 22.9±3.4 years) participated in a 10-day, laboratory-based protocol. A total of 86 sleeps were independently assessed in 30-second epochs using polysomnography and WHOOP. For WHOOP, bed times were entered by researchers and sleeps were scored by the company based on proprietary algorithms. WHOOP overestimated total sleep time by 8.2±32.9 minutes compared to polysomnography, but this difference was non-significant. WHOOP was compared to polysomnography for 2-stage (i.e., wake, sleep) and 4-stage categorisation (i.e., wake, light sleep [N1 or N2], slow wave sleep [N3], REM) of sleep periods. For 2-stage categorisation, the agreement, sensitivity to sleep, specificity for wake, and Cohen's kappa were 89%, 95%, 51%, and 0.49, respectively. For 4-stage categorisation, the agreement, sensitivity to light sleep, SWS, REM, and wake, and Cohen's kappa were 64%, 62%, 68%, 70%, 51%, and 0.47, respectively. In situations where polysomnography is impractical (e.g., field settings), WHOOP is a reasonable method for estimating sleep, particularly for 2-stage categorisation, if accurate bedtimes are manually entered.

**3.1 Introduction**

Technology capable of measuring the quality and quantity of sleep is of great interest to researchers and practitioners in sport science (Halson, Peake, and Sullivan, 2016). An objective measure of sleep is valuable for monitoring the impact of training and competition on the sleep/wake behaviours of athletes (Lastella et al., 2014; Miller et al., 2017; Sargent, Halson, and Roach, 2014; Sargent and Roach, 2016). Due to the negative impacts of insufficient sleep (e.g., cognitive impairment; Dinges et al., 1997; Edwards and Waterhouse, 2009; Van Dongen, Maislin, Mullington, and Dinges, 2003; Vgontzas et al., 2004), and the importance of restorative processes that occur during sleep (e.g., muscle regeneration; Dattilo et al., 2011; Davenne, 2009), devices that can accurately measure sleep may provide a valuable objective indication of cognitive and physiological recovery in the demanding environments of elite sport.

Laboratory-based polysomnography (PSG) is an objective approach that is considered the gold standard of sleep measurement. PSG provides useful information regarding the amount and quality of sleep, but it is expensive and can be impractical for use in field settings (Ancoli-Israel et al., 2003). The main validated alternative to PSG is actigraphy (de Zambotti., et al., 2019). Actigraphs are small devices typically worn on the wrist that utilise accelerometer-based technology to estimate sleep and wake behaviour (Roach et al., 2013). These estimates provide a two-stage detection (i.e., sleep or wake) by assuming that activity is substantially lower during sleep compared to wake (Roach et al., 2013; Sargent et al., 2018). Compared to PSG, actigraphy may be a more practical method of measuring of sleep, however a certain level of expertise and manual data handling is required to obtain accurate data (de Zambotti., et al., 2019). For example, inconsistencies between objective data and self-report data must be identified prior to formal analyses (de Zambotti., et al., 2019).

With advancements in consumer sleep technology, wearable devices such as armbands, rings and smartwatches provide an objective measurement of sleep that can be acquired with little expertise. In comparison to actigraphy, some sleep wearables also provide a four-stage detection of sleep (i.e., wake, light sleep, slow wave sleep [SWS] and rapid-eye-movement sleep [REM]) based on measures including movement detection, heart rate, heart rate variability, skin conductance and/or skin

temperature (de Zambotti et al ., 2017; de Zambotti., et al., 2019). Due to ease of use and provision of detailed daily feedback, the use of sleep wearables has become commonplace in sports science practice (Halson, et al., 2016). However, few devices capable of providing four-stage sleep data have been validated against PSG using recommended, best-practice comparisons (de Zambotti, et al., 2017; de Zambotti., et al., 2019; Shambroom, Fabregas, and Johnstone, 2012).

Previous validations of sleep wearables capable of providing four-stage sleep data have yielded high sensitivity to sleep, low to moderate sensitivity to wake, and varying sensitivity to different stages of sleep (de Zambotti, et al., 2017; Shambroom, et al., 2012). The WHOOP strap (CB Rank, Greater Boston, New England, USA) is a wrist-worn wearable device that provides measures of heart rate, heart rate variability and sleep staging. In conjunction with the WHOOP smart phone application, athletes and coaches can monitor daily averages of heart rate, heart rate variability and sleep. However, the WHOOP strap (in conjunction with the WHOOP smart phone application) has yet to be validated against a gold standard approach for measuring sleep (i.e., polysomnography). The aim of this study was to conduct an epoch-by-epoch validation of the WHOOP strap to measure two-stage (i.e., sleep, wake) and four-stage sleep (i.e., wake, light sleep, SWS, REM).

## 3.2 Methods

### 3.2.1 Participants

Twelve healthy, young adults (male: n = 6, female: n = 6; age: mean ± SD, 22.9 ± 3.4yr) participated in this study. Participants were excluded if they reported any existing medical conditions or sleep disorders, or if they had a recent history of undertaking shift work and/or transmeridian travel. The study was approved by the Central Queensland University Human Research Ethics Committee (HREC:H16/06-168).

### 3.2.2 Laboratory setting

The study was conducted in a purpose-built accommodation suite at the Appleton Institute of Behavioural Science, Central Queensland University. The suite is sound attenuated, free from external environmental cues, and simultaneously houses six participants with private bedrooms and bathrooms.

### 3.2.3 Design

Data were collected as part of a larger experimental study. Participants lived in a sleep laboratory for ten consecutive nights/days and were given sleep opportunities of varying durations. All participants were given a 9-h sleep opportunity on the first night of the study (23:00h to 08:00h), followed by a delayed 9-h hour sleep opportunity on the following night (03:00h-12:00h). Six of the participants were then given daytime sleep opportunities of 2h (08:30h-10:30h) and 5h (16:30h-21:30) for six consecutive days. The remaining six participants were given a daytime sleep consisting of 7-h sleep opportunity (14:30h-21:30h) for six consecutive days. All participants were then given a 7-h sleep opportunity (08:30h-15:30h) on the final experimental day. During all sleep periods (excluding the 2-h sleep opportunities due to logistical reasons), participants wore the WHOOP strap on their non-dominant wrist and had electrodes attached to their face and scalp for measuring sleep using PSG. During wake periods, participants performed sedentary tasks such as reading or watching movies. Data were collected between the 30th October 2018 and the 6th of December 2018.

### 3.2.4 Measures and procedures

Sleep was measured simultaneously using the WHOOP strap (Generation 2.0 hardware, Generation 3 algorithm, CB Rank, Greater Boston, New England) and PSG. Prior to the study, clock time was manually synchronised on all devices (i.e., laboratory computers, mobile devices running the WHOOP smart phone application). To acquire WHOOP strap sleep data, the start and end times of each sleep opportunity were manually entered into the WHOOP smart phone application by the researchers. The manufacturer then provided data in 30-s epochs for wake, light sleep, SWS and REM for comparison with PSG. To measure sleep using PSG, a standard montage of electrodes were attached to the face and scalp of participants including three electroencephalography electrodes (i.e., C4-M1, F4-M1, O2-M1), two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram. PSG data were recorded directly to data acquisition, storage, and analysis systems (Grael, Compumedics; Victoria, Australia). PSG records were manually scored in 30-s epochs by an experienced registered polysomnographic technician in compliance with standard criteria (Iber, Ancoli-Israel, Chesson, and

Quan, 2007). The following sleep variables were extracted from the PSG records and the WHOOP strap data:

- Total sleep time (TST): the sum of minutes spent in any stage of sleep (stages N1, N2, N3, REM);

- Wake: the sum of minutes spent awake;

- Light sleep: the sum of minutes spent in stage N1 or N2 sleep.

- Slow wave sleep (SWS): the sum of minutes spent in stage N3 sleep.

- Rapid-eye-movement sleep: (REM): the sum of minutes spent in stage R.

To ensure that the WHOOP and PSG data were properly aligned for each individual sleep, two-stage agreement (see below) was calculated for offset adjustments of $\pm$ 3 minutes in 30-second increments (Sargent, Lastella, Halson, and Roach, 2016). In all cases, agreement was not improved by applying offset adjustments (mean $\pm$ SD, 88.6 $\pm$ 0.2%; range = 88.15 – 88.90%) when compared to unadjusted data (agreement = 88.94%). Therefore, subsequent analyses were based on unadjusted data.

*3.2.5 Data analysis*

Differences in sleep variables (i.e., TST, wake, light sleep, SWS and REM) between measures (i.e., PSG and the WHOOP strap) were analysed using separate linear mixed models with the R package lme4 (R Core Team, 2016). A random intercept for participants was included in each model to account for intraindividual dependencies and interindividual heterogeneity.

In order to conduct epoch-by-epoch comparisons for two-stage categorisation of sleep (i.e., sleep, wake), the data from the WHOOP strap was arranged in 30-s epochs and aligned with the corresponding PSG record. The following measures were then calculated:

- sensitivity: the percentage of PSG-determined sleep epochs correctly identified as sleep by the device;

- specificity: the percentage of PSG-determined wake epochs correctly identified as wake by the device;

- agreement: the percentage of PSG-determined sleep and wake epochs correctly identified as sleep or wake by the WHOOP strap.

In order to conduct epoch-by-epoch comparisons for four-stage categorisation of sleep, data from the WHOOP strap was then arranged into corresponding 30-s PSG epochs indicating wake, light sleep, SWS or REM. The following measures were calculated:

- sensitivity for wake: the percentage of PSG-determined wake epochs correctly identified as wake the WHOOP strap;

- sensitivity for light sleep: the percentage of PSG-determined N1 and N2 epochs correctly identified as light sleep by the WHOOP strap;

- sensitivity for SWS: the percentage of PSG-determined N3 epochs correctly identified as SWS the WHOOP strap;

- sensitivity for REM: the percentage of PSG-determined REM epochs correctly identified as REM the WHOOP strap;

- agreement: the percentage of PSG-determined N1, N2, REM, and wake epochs correctly identified as light sleep, SWS, REM, or wake by the WHOOP strap.

Cohen's kappa (κ) was calculated to evaluate the agreement of the WHOOP strap with PSG beyond what could be expected by chance (Sim and Wright, 2005). Agreement was interpreted against recommended guidelines – 0.00–0.20 indicates *slight agreement*, 0.21–0.40 is *fair*, 0.41–0.60 is *moderate*, 0.61–0.80 is *substantial*, and 0.81–1.0 is *almost perfect* (Landis and Koch, 1977). Agreement between PSG and the WHOOP strap was also examined using the Bland-Altman plots with 95% limits of agreement for repeated measurements (Bland and Altman, 2007). For each sleep variable, the difference between the WHOOP strap and PSG (i.e., bias) and the 95% limits of agreement (bias ± 1.96*SD) were plotted. Each plot was tested for heteroscedasticity (i.e., whether variance changes as a function of duration) and proportional bias (i.e., whether the differences between WHOOP and PSG change as a function of duration) using the Breusch-Pagan test and least ordinary squares regression, respectively. If proportional bias or heteroscedasticity were present, the bias and 95% limits of agreement were adjusted (Ludbrook, 2010).

**3.3 Results**

WHOOP strap data for 22 out of a possible 108 (i.e., 9 sleep opportunities for 12 participants) sleeps were lost or compromised. Participant withdrawal resulted in loss of data (n=8), while interference from nearby devices compromised some data during download due to operator error (n=6). WHOOP strap data for an additional 8 sleep opportunities were lost. It is unclear whether the loss of data was due to the device, the software, the wearer, or the operator. Thus, 86 manually adjusted sets of WHOOP strap sleep data (5-h sleep opportunity = 32 sleeps, 7-h sleep opportunity = 33 sleeps, 9-h sleep opportunity =21 sleeps) were included in the analyses for comparison to PSG. PSG data acquired during the study reinforced that the sample were healthy sleepers (Table 3.1).

**Table 3.1** Sleep variables determined by polysomnography.

| Sleep opportunity | TST (min) | WASO (min) | SE (%) | SOL (min) |
| --- | --- | --- | --- | --- |
| 9 hours | 480.1 ± 33.6 | 60.2 ± 33.7 | 88.4 ± 6.8 | 13.6 ± 18.4 |
| 7 hours | 380.6 ± 35.8 | 39.2 ± 35 8 | 90.7 ± 9.3 | 2.7 ± 2.0 |
| 5 hours | 245.4 ± 51.6 | 54.6 ± 51.6 | 78.6 ± 18.4 | 11.9 ± 20.2 |

Notes: TST; total sleep time, WASO; wake after sleep onset, SOL; sleep onset latency. Data are mean ± SD.

For two-stage sleep, there were no significant differences between the WHOOP strap and PSG for TST and wake (Table 3.2). Two-stage epoch-by-epoch data showed high sensitivity (i.e., ability to detect sleep), but low specificity (i.e., ability to detect wake; Table 3). Cohen's kappa coefficient for two-stage sleep ($\kappa = 0.49$) indicated moderate agreement between the WHOOP strap and PSG (Landis and Koch, 1977).

For four-stage sleep, there were no significant differences between the WHOOP strap and PSG for wake, light sleep and SWS, however the WHOOP strap significantly overestimated REM compared with PSG (Table 3.2). Four-stage epoch-by-epoch data had the highest sensitivity for REM and lowest

sensitivity for wake (Table 3.3). Cohen's kappa coefficient for four-stage sleep ($\kappa = 0.47$) indicated moderate agreement between the WHOOP strap and PSG (Landis and Koch, 1977).

Bland-Altman plots comparing the WHOOP strap to PSG for each sleep variable are depicted in Figure 3.1. Proportional bias and heteroscedasticity were present for wake and SWS, but not for TST and REM. For light sleep, there was no proportional bias, but the data were heteroskedastic.

A four-stage error matrix comparing the WHOOP strap and PSG is presented in Table 3.4. When the WHOOP strap misclassifies wake, it is because it is classifying it as either light sleep or REM. When the WHOOP strap misclassifies light sleep, it is because it is classifying it as either SWS sleep or REM. When the WHOOP strap misclassifies SWS, it is because it is classifying it as light sleep. When the WHOOP strap misclassifies REM, it is because it is classifying it as either wake or light sleep.

**Table 3.2** Comparison of sleep variables determined by polysomnography and the WHOOP strap.

| Variable (min) | Measure | | WHOOP vs. PSG | | | |
|---|---|---|---|---|---|---|
| | WHOOP | PSG | MAE | Bias | F (df) | p |
| TST | $358.7 \pm 98.5$ | $350.4 \pm 105.2$ | 22.5 | $8.2 \pm 32.9$ | 0.4 (1,159) | 0.54 |
| Wake | $46.0 \pm 37.7$ | $54.2 \pm 46.9$ | 22.5 | $-8.2 \pm 32.9$ | 2.3 (1,160) | 0.12 |
| Light sleep | $174.4 \pm 52.5$ | $178.1 \pm 58.7$ | 34.6 | $-3.7 \pm 44.4$ | 0.2 (1,159) | 0.62 |
| SWS | $80.0 \pm 27.1$ | $83.7 \pm 32.1$ | 20.1 | $-3.7 \pm 26.4$ | 0.9 (1,160) | 0.33 |
| REM | $104.2 \pm 50.2$ | $88.6 \pm 33.8$ | 34.1 | $15.6 \pm 39.7$ | 7.2 (1,159) | 0.01 |

Notes: Negative bias indicates an underestimation of the sleep variable by the WHOOP strap. TST; total sleep time, SWS; slow wave sleep, REM; rapid eye movement sleep; MAE; mean absolute error. Data are mean ± SD.

**Table 3.3** Two-stage and four-stage epoch-by-epoch concordance statistics for the WHOOP strap and polysomnography.

| Measure | Value (%) |
|---|---|
| **Two-stage comparison** | |
| Sensitivity for sleep | 95 |
| Specificity for wake | 51 |
| Overall agreement | 89 |
| **Four-stage comparison** | |
| Sensitivity for wake | 51 |
| Sensitivity for light sleep | 62 |
| Sensitivity for SWS | 68 |
| Sensitivity for REM | 70 |
| Overall agreement | 64 |

Notes: SWS; slow wave sleep, REM; rapid eye movement sleep.

**Table 3.4** Four-stage error matrix for the WHOOP strap and PSG.

| | Stage | WHOOP | | | |
|---|---|---|---|---|---|
| | | Wake | Light sleep | SWS | REM |
| PSG | Wake | 51% | 29% | 2% | 18% |
| | Light sleep | 8% | 62% | 12% | 18% |
| | SWS | <1% | 31% | 68% | <1% |
| | REM | 3% | 25% | 2% | 70% |

Notes: This matrix presents the percentage of each sleep stage that the WHOOP strap has correctly classified or misclassified when compared to PSG. Shaded cells indicate correctly classified sleep. SWS; slow wave sleep, REM; rapid eye movement sleep.
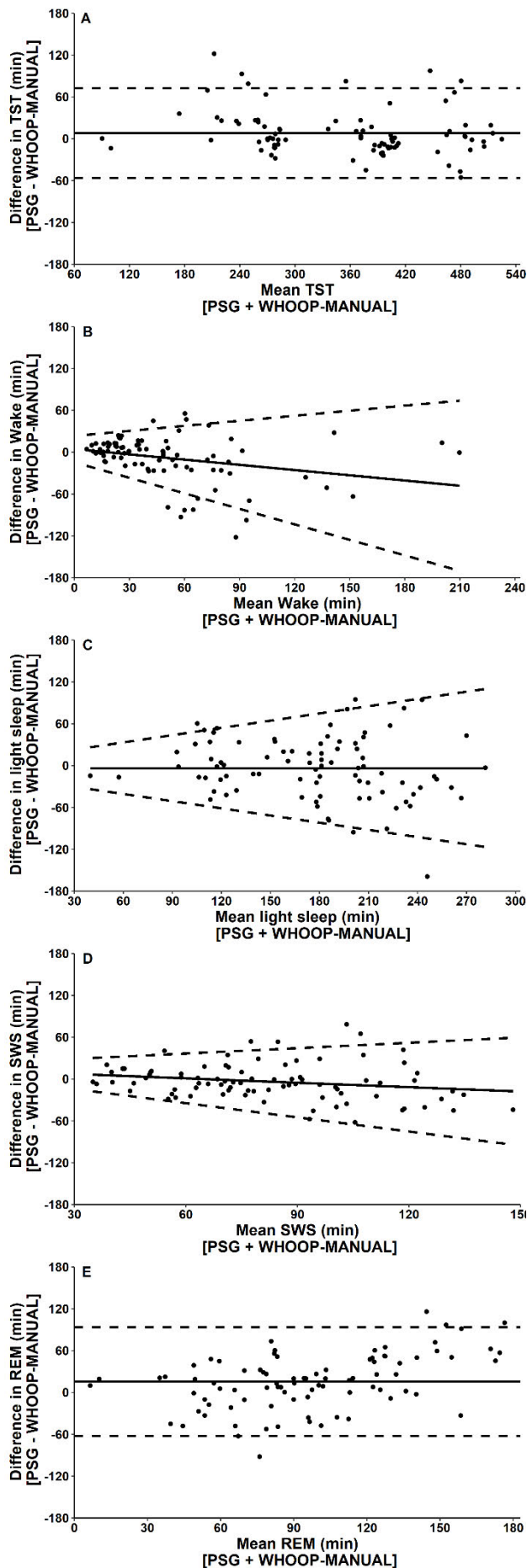
**Figure 3.1** Bland-Altman plots for WHOOP and PSG: total sleep time (A; TST), wake (B), light sleep (C), SWS (D; slow wave sleep) and REM (E; rapid eye movement sleep). The x-axes represent the mean of the values obtained from the WHOOP strap and PSG; and the y-axes represent the difference between the values, such that positive values indicate that the WHOOP strap overestimated relative to PSG and negative values indicate that the WHOOP strap underestimated relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement (±1.96SDs; Ludbrook, 2010).

**3.4 Discussion**

In this study, the validity of the WHOOP strap as a measure of sleep was assessed in comparison to the gold standard – polysomnography – using difference testing, Bland-Altman analyses, and epoch-for-epoch comparisons. Most research-grade sleep wearables use a two-stage framework to categorise sleep periods, i.e., each epoch is scored as sleep or wake. In contrast, most commercial-grade sleep wearables use a four-stage framework to categorise sleep periods, i.e., each epoch is scored as light sleep (non-REM stage 1 and non-REM stage 2), SWS (non-REM stage 3), REM sleep, or wake. Therefore, for these analyses, the validity of the WHOOP strap for both two-stage and four-stage categorisation of sleep was assessed.

For the two-stage categorisation of sleep, the WHOOP strap overestimated sleep by 8.2 minutes and underestimated wake by 8.2 minutes; the 95% confidence intervals around the sleep bias were $\pm$ 64.5 minutes; sensitivity to sleep was 95%, specificity for wake was 51%, overall agreement was 89%; and chance-corrected agreement was moderate, i.e., Cohen's kappa was 0.49. These outcomes are similar to those previously reported for research-grade sleep wearables. In particular, research-grade sleep wearables overestimate sleep by 7-67 minutes (Cook, et al., 2019; Cook, Prairie, and Plante, 2018; de Zambotti, Baker, and Colrain, 2015; de Zambotti et al., 2016; de Zambotti et al., 2015; Kang et al., 2017; Maskevich et al., 2017); they have sensitivity to sleep of 81-97%, specificity for wake of 38–82%, and overall agreement of 81-97% (de Zambotti, et al., 2018; Kosmadopoulos et al., 2014; Paquet, et al., 2007; Signal, et al., 2005); and they have chance-corrected agreement of 0.31–0.65 (Kang et al., 2017; Kosmadopoulos et al., 2014; Maskevich et al., 2017; Shambroom, et al., 2012). As a guide to the maximum level of agreement that could be reasonably expected between any two methods of assessing sleep, note that the chance-corrected agreement between expert sleep scorers for independently scoring a common set of PSG records using a three-stage categorisation of sleep (i.e., non-REM, REM, wake) is 0.78 (Magalang et al., 2013), i.e., substantial rather than almost perfect (Landis and Koch, 1977). Taken together, these comparisons indicate that the WHOOP strap is comparable to research-grade sleep wearables. Therefore, in situations where polysomnography is

impractical (e.g., field settings), the WHOOP strap is reasonable method for estimating 2-stage sleep – provided that accurate bed times are manually entered.

For the four-stage categorisation of sleep, the WHOOP strap overestimated wake by 8.2 minutes; underestimated light sleep by 3.7 minutes; underestimated SWS by 3.7 minutes; and overestimated REM by 15.6 minutes. Sensitivity to wake was 51%, sensitivity for light sleep was 62%, sensitivity for SWS was 68%, sensitivity for REM was 70%, and overall agreement was 64%. Chance-corrected agreement was moderate, i.e., Cohen's kappa was 0.47. In comparison, research-grade sleep wearables have a sensitivity for light sleep of 60-81%, a sensitivity to SWS of 36-67%, and a sensitivity to REM of 30-74% (Cook, et al., 2019; Cook, Prairie, and Plante, 2018; de Zambotti, et al., 2017; de Zambotti, et al., 2018). Note that the chance-corrected agreement between expert sleep scorers for independently scoring a common set of PSG records using a five-stage categorisation of sleep (i.e., wake, N1, N2, N3, REM) is 0.63 (Magalang et al., 2013), i.e., substantial rather than almost perfect (Landis and Koch, 1977). Taken together, these comparisons indicate that the WHOOP strap, in situations where polysomnography is not available, may be as reasonable method for estimating four-stage sleep.

### 3.4.1 Limitations, boundary conditions, and future research

It should be made clear that this validation study was conducted only on the WHOOP strap two-stage and four-stage sleep capabilities and not for heart rate, heart rate variability, sleep need, or other WHOOP metrics. The findings of this study should only be applied to settings where accurate sleep times are manually entered for each sleep opportunity. Additionally, these findings should only be applied to sleep opportunities between 5-9h in duration; it is not clear how the WHOOP strap may perform for shorter sleep opportunities. The participants in this study were healthy, young, good sleepers exposed to ideal sleeping conditions. It is unclear how the WHOOP strap may perform when individuals are exposed to less ideal sleeping environments or when individuals experience a high proportion of wake during a sleep episode. It should be acknowledged that the algorithms used by WHOOP to score sleep are proprietary, and that epoch-by-epoch data is not accessible through the WHOOP smart phone application. As a result, using the WHOOP strap in conjunction with the

WHOOP smart phone application may be limited in situations where epoch-by-epoch analyses are required (i.e., clinical and/or research settings). Future investigations should examine the accuracy of the WHOOP strap to detect and estimate two-stage sleep and four-stage sleep without manual adjustment of sleep opportunity in the WHOOP smart phone application.

**3.5 References**

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., and Pollak, C. (2003). The role of actigraphy in the study of sleep and circadian rhythms. American Academy of Sleep Medicine Review Paper. *Sleep, 26*(3), pp. 342-392.

Bland, J. M., and Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics, 17*(4), pp. 571-582.

Buysse, D. J., Monk, T. H., Carrier, J., and Begley, A. (2005). Circadian patterns of sleep, sleepiness, and performance in older and younger adults. *Sleep, 28*(11), pp. 1365-1376.

Cook, J. D., Eftekari, S. C., Dallmann, E., Sippy, M., and Plante, D. T. (2019). Ability of the Fitbit Alta HR to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: A comparison against polysomnography. *Journal of Sleep Research, 28*(4), p e12789.

Cook, J. D., Prairie, M. L., and Plante, D. T. (2018). Ability of the multisensory jawbone up3 to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: a comparison against polysomnography and actigraphy. *Journal of Clinical Sleep Medicine: JCSM : official publication of the American Academy of Sleep Medicine, 14*(5), pp. 841-848.

Dattilo, M., Antunes, H., Medeiros, A., Mônico Neto, M., Souza, H., Tufik, S., and de Mello, M. (2011). Sleep and muscle recovery: endocrinological and molecular basis for a new and promising hypothesis. *Medical Hypotheses, 77*(2), pp. 220-222.

Davenne, D. (2009). Sleep of athletes – problems and possible solutions. *Biological Rhythm Research, 40*(1), pp. 45-52.

De Zambotti, M., Baker, F. C., Willoughby, A. R., Godino, J. G., Wing, D., Patrick, K., and Colrain, I. M. (2016). Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiology and Behavior, 158*, pp. 143-149. M, D. E.

de Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M., and Baker, F. C. (2019). Wearable sleep technology in clinical and research settings. *Medicine and Science in Sports and Exercise*, *51*(7), pp. 1538-1557.

De Zambotti, M., Claudatos, S., Inkelis, S., Colrain, I. M., and Baker, F. C. (2015). Evaluation of a consumer fitness-tracking device to assess sleep in adults. *Chronobioly International, 32*(7), pp. 1024-1028.

De Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M., and Baker, F. C. (2018). A validation study of Fitbit Charge 2 compared with polysomnography in adults. *Chronobiology International, 35*(4), pp. 465-476.

De Zambotti, M., Rosas, L., Colrain, I. M., and Baker, F. C. (2017). The sleep of the ring: comparison of the oura sleep tracker against polysomnography. *Behavioral Sleep Medicine*, pp. 1-15.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., . . . Pack, A. I. (1997). Cumulative sleepiness, mood disturbance and psychomotor vigilance performance decrements during aweek of sleep restricted to 4-5 hours per night. *Sleep, 20*(4), pp. 267-277.

Edwards, B. J., and Waterhouse, J. (2009). Effects of one night of partial sleep deprivation upon diurnal rhythms of accuracy and consistency in throwing darts. *Chronobiology International, 26*(4), pp. 756-768.

Halson, S. L., Peake, J. M., and Sullivan, J. P. (2016). Wearable technology for athletes: Information overload and pseudoscience? *The International Journal of Sports Physiology and Performance, 11*(6), pp. 705-706.

Iber, C., Ancoli-Israel, S., Chesson, A., and Quan, S. (2007). The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. 2007. *American Academy of Sleep Medicine, Westchester, IL.*

Kang, S. G., Kang, J.M., Ko, K.P., Park, S.C., Mariani, S., and Weng, J. (2017). Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of Psychosomatic Research, 97*, pp. 38-44.

Kosmadopoulos, A., Sargent, C., Darwent, D., Zhou, X., and Roach, G. D. (2014). Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavior Research Methods, 46*(4), pp. 1032-1041.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), pp. 159-174.

Lastella, M., Roach, G. D., Halson, S. L., Gore, C. J., Garvican-Lewis, L. A., and Sargent, C. (2014). The effects of transmeridian travel and altitude on sleep: preparation for football competition. *Journal of Sports Science and Medicine, 13*(3), pp. 718-720.

Ludbrook, J. (2010). Confidence in Altman-Bland plots: a critical review of the method of differences. *Clinical and Experimental Pharmacology and Physiology, 37*(2), pp. 143-149.

Magalang, U. J., Chen, N.-H., Cistulli, P. A., Fedson, A. C., Gíslason, T., Hillman, D., . . . Investigators, f. t. S. (2013). Agreement in the Scoring of Respiratory Events and Sleep Among International Sleep Centers. *Sleep*, 36(4), pp. 591-596

Maskevich, S., Jumabhoy, R., Dao, P. D. M., Stout, J. C., and Drummond, S. P. A. (2017). Pilot validation of ambulatory activity monitors for sleep measurement in huntington's disease gene carriers. *Journal of Huntington's Disease, 6*(3), pp. 249-253.

Meltzer, L. J., Hiruma, L. S., Avis, K., Montgomery-Downs, H., and Valentin, J. (2015). Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. *Sleep, 38*(8), pp. 1323-1330.

Meltzer, L. J., Walsh, C. M., Traylor, J., and Westin, A. M. (2012). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep, 35*(1), pp. 159-166.

Miller, D. J., Sargent, C., Vincent, G. E., Roach, G. D., Halson, S. L., and Lastella, M. (2017). Sleep/wake behaviours in elite athletes from three different football codes. *Journal of Sports Science and Medicine, 16*(4), pp. 604-605.

Paquet, J., Kawinska, A., and Carrier, J. (2007). Wake detection capacity of actigraphy during sleep. *Sleep, 30*(10), pp. 1362-1369.

Roach, G. D., Schmidt, W. F., Aughey, R. J., Bourdon, P. C., Soria, R., Claros, J. C., . . . Sargent, C. (2013). The sleep of elite athletes at sea level and high altitude: a comparison of sea-level natives and high-altitude natives (ISA3600). *British Journal of Sports Medicine, 47 Suppl 1*, pp. i114-120.

Sargent, C., Halson, S., and Roach, G. D. (2014). Sleep or swim? Early-morning training severely restricts the amount of sleep obtained by elite swimmers. *European Journal of Sport Science, 14 Suppl 1*, pp. S310-315.

Sargent, C., Lastella, M., Halson, S. L., and Roach, G. D. (2016). The validity of activity monitors for measuring sleep in elite athletes. *The Journal of Science and Medicine in Sport, 19*(10), pp. 848-853.

Sargent, C., Lastella, M., Romyn, G., Versey, N., Miller, D. J., and Roach, G. D. (2018). How well does a commercially available wearable device measure sleep in young athletes? *Chronobiology International, 35*(6), pp. 754-758.

Sargent, C., and Roach, G. D. (2016). Sleep duration is reduced in elite athletes following night-time competition. *Chronobiology International*, pp. 1-4.

Shambroom, J. R., Fabregas, S. E., and Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research, 21*(2), pp. 221-230.

Signal, T. L., Gale, J., and Gander, P. H. (2005). Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation, Space, and Environmental Medicine, 76*(11), pp. 1058-1063.

Sim, J., and Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85(3), pp. 257-268.

Toon, E., Davey, M. J., Hollis, S. L., Nixon, G. M., Horne, R. S., and Biggs, S. N. (2016). Comparison of commercial wrist-based and smartphone accelerometers, actigraphy, and PSG in a clinical cohort of children and adolescents. *Journal of Clinical Sleep Medicine, 12*(3), pp. 343-350.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M., and Dinges, D. F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26*(2), pp. 117-126.

Vgontzas, A., Zoumakis, E., Bixler, E., Lin, H. M., Follett, H., Kales, A., and Chrousos, G. (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *Journal of Clinical Endocrinology and Metabolism, 89*(5), pp. 2119-2126.

Werner, H., Molinari, L., Guyer, C., and Jenni, O. G. (2008). Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. *The Archives of Pediatrics and Adolescent Medicine, 162*(4), pp. 350-358.

**4. A validation of the WHOOP strap to detect and measure sleep and sleep staging.**

Published peer-reviewed publications associated with this Chapter (Appendix D):

**Miller, D.J.,** Roach G.D., Lastella M., Scanlan A.T., Bellenger C., Halson S.L., Sargent C. (2021). A validation study of a commercial wearable device to automatically detect and estimate sleep. *Biosensors*, 11(6):185.

**Abstract**

The aims of this study were to: (1) compare actigraphy (ACTICAL) and a commercially available sleep wearable (i.e., WHOOP) under two functionalities (i.e., sleep autodetection [WHOOP-AUTO]; and manual adjustment of sleep [WHOOP-MANUAL]) for 2-stage categorisation of sleep (sleep or wake) against polysomnography, and; (2) compare WHOOP-AUTO and WHOOP-MANUAL for 4-stage categorisation of sleep (wake, light sleep, slow wave sleep [SWS], or rapid eye movement sleep [REM]) against polysomnography. Six healthy adults (male: n = 3; female: n = 3; age: 23.0 ± 2.2 yr) participated in the 9-night protocol. Fifty-four sleeps assessed by ACTICAL, WHOOP-AUTO and WHOOP-MANUAL were compared to polysomnography using difference testing, Bland-Altman comparisons, and epoch-by-epoch comparisons. Compared to polysomnography, ACTICAL overestimated total sleep time (37.6 min) and underestimated wake (-37.6 min); WHOOP-AUTO underestimated SWS (-15.5 min); and WHOOP-MANUAL underestimated wake (-16.7 min). For ACTICAL, sensitivity for sleep, specificity for wake and overall agreement were 98%, 60% and 89%, respectively. For WHOOP-AUTO, sensitivity for sleep, wake, light sleep, SWS and REM, and agreement for 2-stage and 4-stage categorisation of sleep were 90%, 60%, 61%, 63%, 66%, 86% and 63%, respectively. For WHOOP-MANUAL, sensitivity for sleep, wake, light sleep, SWS and REM, and agreement for 2-stage and 4-stage categorisation of sleep were 97%, 45%, 67%, 61%, 66%, 90% and 62%, respectively. WHOOP-AUTO and WHOOP-MANUAL have a similar sensitivity and specificity to actigraphy for 2-stage categorisation of sleep and can be used as a practical alternative to polysomnography for 2-stage categorisation of sleep and 4-stage categorisation of sleep.

**4.1 Introduction**

Polysomnography (PSG) is the gold standard method of objectively assessing sleep (Kushida, Littner et al., 2005). However, PSG is expensive, time consuming and impractical in some field settings (Yi, Shin et al., 2006a; Yi, Shin et al., 2006b). The most commonly accepted alternative to PSG is research-grade actigraphy (de Zambotti, Cellini et al., 2019). Actigraphy uses algorithms based on the association of movement and wakefulness, allowing for an objective measurement of sleep and wake (Pollak, Tryon et al., 2001). However, acquiring accurate actigraphy data can be cumbersome and requires certain resources and expertise (e.g., proprietary software, reliance of self-reported sleep times, retrospective data extraction; (de Zambotti, Cellini et al., 2019; Miller, Lastella et al., 2020). Modern commercial wearable technology provides a user-friendly, accessible alternative to PSG and actigraphy that provides an easily accessible aggregated sleep data (de Zambotti, Cellini et al., 2019; Miller, Lastella et al., 2020). Unlike actigraphy, which relies solely on accelerometer-based movement detection to measure sleep and wake, sleep wearable technology utilises accelerometers and heart rate tracking technology (photoplethysmography) to provide 2-stage categorisation and 4-stage categorisation of sleep.(de Zambotti, Cellini et al., 2019). Photoplethysmography provides convenient and accurate indication of autonomic nervous system status by measuring heart rate variability (Stein and Pu, 2012).

While providing a convenient alternative to PSG for measuring sleep, use of actigraphy relies on self-report data from the user (i.e., bedtime, get up time). For retrospective manual adjustment of bed and wake times by the researcher (de Zambotti, Cellini et al., 2019). In comparison, most commercial sleep wearables estimate sleep under two functions; auto-detection (i.e., an algorithm is used to automatically detect sleep onset and sleep offset) or manual adjustment (i.e., the user manually adjusts bed and wake times after a sleep period is recorded). The distinction between auto-detected and manually adjusted sleep records is important as they are two different methods of measuring sleep. However, most validation studies for sleep wearables, have either analysed manually adjusted data (de Zambotti, Baker et al., 2016; de Zambotti, Rosas et al., 2017; Miller, Lastella et al., 2020) or have not reported their methods of data acquisition (i.e., auto-detection versus manual; Meltzer, Hiruma et al., 2015; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017). For manually adjusted data

validations, the adjustments are performed by researchers in a controlled laboratory setting (Sargent, Lastella et al., 2018; Miller, Lastella et al., 2020). Therefore, the accuracy of sleep wearables in situations where manual adjustment of sleep times is performed by the user may vary. In this context, actigraphy and sleep wearables utilising a manual adjustment function are subject to compliance of wearing the device and accurately reporting bed and wake times. Compliance for self-report measures is a common obstacle in acquiring accurate data in clinical settings (Stone, 2002) and elite sport (Saw, Main et al., 2015). Therefore, sleep wearables that are capable of accurately auto-detecting sleep, and therefore eliminating non-compliance of users reporting bed and wake times, would provide an attractive alternative for measuring sleep in non-laboratory settings.

The WHOOP strap is a sleep wearable capable of estimating sleep (Miller, Lastella et al., 2020). When using manually adjusted sleep records (i.e., WHOOP-MANUAL), the WHOOP strap has been validated as an alternative for 2-stage categorisation of sleep (i.e., sleep and wake) and 4-stage categorisation of sleep (i.e., wake, light sleep, slow wave sleep [SWS], rapid eye movement sleep [REM]) when PSG is impractical (Miller, Lastella et al., 2020). However, the ability of the WHOOP strap to automatically detect (i.e., WHOOP-AUTO) and categorise 2-stage sleep and 4-stage sleep has not been examined. Therefore, the two aims of this study were to; (1) compare the ability of WHOOP-AUTO, WHOOP-MANUAL and research grade actigraphy (ACTICAL) for 2-stage categorisation of sleep against PSG, and; (2) compare the ability of WHOOP-AUTO and WHOOP-MANUAL for 4-stage categorisation of sleep against PSG.

**4.2 Methods**

*4.2.1 Participants*

Six healthy, young adults (male: n = 3; female: n = 3; age: $23.0 \pm 2.2$ yr; height: $170.5 \pm 7.2$ cm; weight: $65.8 \pm 3.6$ kg) participated in this study. In order to maintain a healthy sample, participants were excluded if they reported any existing medical conditions or sleep disorders or had a recent history (in the prior 3 months) of shift work and/or transmeridian travel. The study was approved by the Central Queensland University Human Research Ethics Committee (HREC:H16/06-168).

*4.2.2 Laboratory setting*

The study was conducted in a purpose-built accommodation suite at the Appleton Institute for Behavioural Science, Central Queensland University, Adelaide, Australia. The suite is sound-attenuated, free from external environmental cues and simultaneously houses six participants with private bedrooms and bathrooms.

*4.2.3 Design*

Data were collected as part of a larger experimental study. Participants lived in a sleep laboratory for ten consecutive nights/days and were given sleep opportunities of varying durations. Participants were given 9h sleep opportunities on nights 1 (23:00–08:00h) and 2 (03:00–12:00h); 7h sleep opportunities on days 3–8 (14:30–21:30h) and day 9 (08:30–15:30h; Figure 4.1). Participants completed sedentary simulated work shifts on days 3-9 and performed sedentary tasks during free time (i.e., reading, watching movies; Figure 4.1). Data were collected between the 16th and the 25th of July 2019.
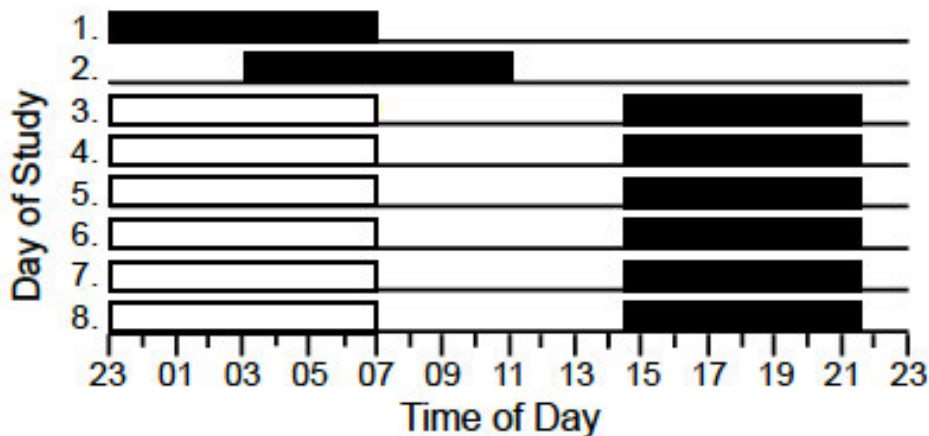
**Figure 4.1.** Illustration of the study design. Black horizontal bars indicate sleep opportunities. White horizontal bars indicate simulated work periods.

### 4.2.4 Measures and procedure

Sleep was measured using PSG. A standard montage of electrodes was attached to the face and scalp of participants (i.e., C4-M1, F4-M1, O2-M1), including two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram (Miller, Sargent et al., 2019). PSG data were recorded directly to data acquisition, storage, and analysis systems (Grael, Compumedics; Victoria, Australia). PSG records were manually scored in 30-s epochs by an experienced registered polysomnographic technician in compliance with standard criteria (Iber, Ancoli-Israel et al., 2007). The commercially available wearable device used in this study was the WHOOP strap (Generation 2.0, CB Rank, Greater Boston, New England). The research grade activity monitor used in this study was the Actical Z-series (ACTICAL; Mini-Mitter Philips Respironics, Inc.; Kosmadopoulos, Sargent et al., 2014). Participants wore the WHOOP strap and ACTICAL on their non-dominant wrist, with the WHOOP strap placed proximally. Prior to the study, clock time was manually synchronised on all devices (i.e., laboratory computers, ACTICAL, mobile devices running the WHOOP iOS application).

Data for automatically detected sleep (WHOOP-AUTO) and for manually-adjusted sleep (WHOOP-MANUAL) were provided by the manufacturer for comparison to PSG. WHOOP-AUTO data were provided first to ensure that the manufacturer was blind to sleep times. Once WHOOP-AUTO

data were received, the start and end times of each sleep opportunity were manually entered by a researcher into the WHOOP iOS application and the manually adjusted data (WHOOP-MANUAL) was then provided by the manufacturer. Epoch-by-epoch ACTICAL data were obtained using accompanying software (30-s epochs, medium sensitivity threshold; Actiware version 3.4; Mini-Mitter Philips Respironics, Inc.; Kosmadopoulos, Sargent et al., 2014)

The following sleep variables were collected during the study:

- Total sleep time (TST): the sum of minutes spent in any stage of sleep (N1, N2, N3, REM).

- Wake: the sum of minutes spent awake during the sleep opportunity.

- Light sleep: the sum of minutes spent in stage N1 or N2 sleep.

- Slow wave sleep (SWS): the sum of minutes spent in stage N3 sleep.

- Rapid-eye-movement sleep: (REM): the sum of minutes spent in stage REM.

- Sleep onset latency (SOL): the duration of time from lights out to the first epoch of any stage of sleep.

PSG and WHOOP-MANUAL provided records of all the above variables. WHOOP-AUTO provided records of all the above variables, except for sleep onset latency. ACTICAL provided records of total sleep time and wake only.

To ensure that the WHOOP-AUTO, WHOOP-MANUAL and ACTICAL data were properly aligned to PSG data for each sleep record, agreement was calculated for offset adjustments of ±3 min in 30-s increments (Sargent, Lastella et al., 2016). In all cases, agreement was not substantially improved by applying an offset, so all subsequent analyses were based on unadjusted data.

*4.2.5 Data Analysis*

Differences in TST, wake, light sleep, SWS and REM between PSG, WHOOP-AUTO and WHOOP-MANUAL were tested using separate General Linear Mixed Models (R package lme4; R Core Team, 2016). Differences in TST and wake time between PSG and ACTICAL were analysed using separate General Linear Mixed Models (R package lme4; R Core Team, 2016). A random intercept for participants was included in each model to account for intraindividual dependencies and interindividual heterogeneity.

Agreement between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL was tested using the Bland-Altman limits of agreement method for repeated measurements.(Bland and Altman, 2007) For each sleep variable, the difference between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL (i.e., bias) and the 95% limits of agreement (i.e., bias ± 1.96*SD) were plotted. Each plot was examined for heteroscedasticity and proportional bias using the Breusch-Pagan test and least ordinary squares regression, respectively. If proportional bias or heteroscedasticity was present, the bias and 95% limits of agreement were adjusted accordingly (Ludbrook, 2010).

To conduct epoch-by-epoch comparisons for 2-stage categorisation of sleep, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL data were arranged in 30-s epochs and aligned with the corresponding PSG record. The following measures were then calculated for WHOOP-AUTO, WHOOP-MANUAL and ACTICAL:

- sensitivity: the percentage of PSG-determined sleep epochs correctly identified as sleep by each method;
- specificity: the percentage of PSG-determined wake epochs correctly identified as wake by each method;
- agreement: the percentage of PSG-determined sleep and wake epochs correctly identified as sleep or wake by each method.

To conduct epoch-by-epoch comparisons for 4-stage categorisation of sleep, WHOOP-AUTO and WHOOP-MANUAL data were arranged in 30-s epochs and aligned with the corresponding PSG record. The following measures were then calculated for WHOOP-AUTO and WHOOP-MANUAL:

- sensitivity for wake: the percentage of PSG-determined wake epochs correctly identified as wake by each method;

- sensitivity for light sleep: the percentage of PSG-determined N1 and N2 epochs correctly identified as light sleep by each method;

- sensitivity for SWS: the percentage of PSG-determined N3 epochs correctly identified as SWS by each method;

- sensitivity for REM: the percentage of PSG-determined REM epochs correctly identified as REM by each method;

- agreement: the percentage of PSG-determined N1, N2, REM, and wake epochs correctly identified as light sleep, deep sleep, REM, or wake by each method.

Cohen's kappa (κ) was calculated to evaluate agreement between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL beyond what could be expected by chance (Sim and Wright, 2005). Agreement was interpreted against recommended guidelines as: *slight agreement* = 0-0.20; *fair* agreement = 0.21-0.40; *moderate* agreement = 0.41-0.60; *substantial* agreement = 0.61-0.80; *almost perfect* agreement = 0.81-0.99; and *perfect* agreement = 1 (Landis and Koch, 1977).

Aggregated data were collated from previous studies to compare WHOOP-AUTO, WHOOP-MANUAL and ACTICAL, respectively, against previous validations of sleep wearables (Figure 4.5; Meltzer, Hiruma et al., 2015; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017; Cook, Prairie et al., 2018; de Zambotti, Goldstone et al., 2018; Pesonen and Kuula, 2018; Sargent, Lastella et al., 2018; Cook, Eftekari et al., 2019; Walch, Huang et al., 2019; Miller, Lastella et al., 2020).

**4.3 Results**

Data acquired using WHOOP-AUTO (n=54), WHOOP-MANUAL (n=54) and ACTICAL (n=54) were included in the analyses for comparison to PSG. No data were lost and WHOOP-AUTO correctly identified 100% of the 54 sleep opportunities.

For 2-stage categorisation of sleep, there was no difference between WHOOP-AUTO and PSG for TST or wake time (Table 4.1). Epoch-by-epoch data showed high sensitivity and moderate specificity for WHOOP-AUTO against PSG (Table 4.2). Cohen's kappa coefficient indicated *moderate* agreement ($\kappa = 0.44$) between the WHOOP-AUTO and PSG for 2-stage categorisation of sleep (Landis and Koch, 1977).

For 2-stage categorisation of sleep, there was no difference between WHOOP-MANUAL and PSG for TST, but WHOOP-MANUAL significantly underestimated wake time compared to PSG (Table 4.1). Epoch-by-epoch data showed high sensitivity, but low specificity compared to PSG (Table 2). Cohen's kappa coefficient indicated *moderate* agreement ($\kappa = 0.48$) between the WHOOP-MANUAL and PSG for 2-stage categorisation of sleep.

ACTICAL significantly overestimated TST and underestimated wake when compared to PSG (Table 4.1). For two-stage categorisation of sleep, ACTICAL had high sensitivity (i.e., ability to detect sleep) and moderate specificity (i.e., ability to detect wake; Table 4.2). Cohen's kappa coefficient for 2-stage categorisation of sleep ($\kappa = 0.23$) indicated *fair* agreement between ACTICAL and PSG.

For 4-stage categorisation of sleep, there was no significant difference between WHOOP-AUTO and PSG for TST, wake time, light sleep or REM. WHOOP-AUTO significantly underestimated SWS and overestimated sleep onset latency (Table 4.1). There was moderate overall agreement for 4-stage categorisation of sleep between PSG and WHOOP-AUTO and moderate sensitivity for wake, light sleep, SWS and REM (Table 4.2). Cohen's kappa coefficient indicated moderate agreement ($\kappa = 0.47$) between WHOOP-AUTO and PSG for 4-stage categorisation of sleep (Landis and Koch, 1977). A 4-stage error matrix comparing WHOOP-AUTO and PSG is presented in Table 4.3. When WHOOP-AUTO misclassifies wake, it classifies it as light sleep. When WHOOP-AUTO misclassifies light sleep,

it classifies it as either wake or REM. When WHOOP-AUTO misclassifies SWS, it classifies it as light sleep. When WHOOP-AUTO misclassifies REM, it classifies it as light sleep. Bland-Altman plots comparing WHOOP-AUTO to PSG for each sleep variable are depicted in Figure 4.2. Proportional bias (i.e., whether the differences between a device and PSG change as a function of duration) and heteroscedasticity (i.e., whether variance changes as a function of duration) were present for TST, wake time, and sleep onset latency, but not for light sleep, SWS or REM.

For 4-stage categorisation of sleep, there was no difference between WHOOP-MANUAL and PSG for TST, light sleep, SWS, REM or sleep onset latency. WHOOP-MANUAL significantly underestimated wake time compared to PSG (Table 4.1). There was *moderate* overall agreement for 4-stage categorisation of sleep, *moderate* sensitivity for light sleep, SWS and REM, and *low* sensitivity for wake time between PSG and WHOOP-MANUAL (Table 4.2). Cohen's kappa coefficient indicated *moderate* agreement ($\kappa = 0.49$) between WHOOP-MANUAL and PSG for 4-stage categorisation of sleep (Landis and Koch, 1977). A 4-stage error matrix comparing the WHOOP-MANUAL and PSG is presented in Table 4.4. When WHOOP-MANUAL misclassifies wake, it classifies it as light sleep. When WHOOP-MANUAL misclassifies light sleep, it classifies it as REM. When WHOOP-MANUAL misclassifies SWS, it classifies it as light sleep. When WHOOP-MANUAL misclassifies REM, it classifies it as light sleep. Bland-Altman plots comparing WHOOP-MANUAL to PSG for each sleep variable are depicted in Figure 4.2. Proportional bias and heteroscedasticity were present for wake time and sleep onset latency, but not for TST, light sleep, SWS or REM.

**Table 4.1** Comparison of sleep variables determined by PSG, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL.

| Variable (min) | PSG | PSG vs. WHOOP-AUTO | | | PSG vs. WHOOP-MANUAL | | | PSG vs. ACTICAL | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Bias | AE | F | Bias | AE | F | Bias | AE | F |
| TST | 392.8 (*60.7*) | -17.8 (*61.1*) | 40.0 | 1.7 | 16.7 (*35.6*) | 25.4 | 2.4 | 37.6* (*85.6*) | 38.1 | 12.2 |
| Wake | 53.9 (*45.7*) | 17.8 (*61.1*) | 40.0 | 2.8 | -16.7* (*35.6*) | 25.4 | 6.3 | -37.6* (*85.6*) | 38.1 | 35.1 |
| Light | 197.1 (*50.8*) | -8.9* (*55.9*) | 43.8 | 0.8 | 13.9 (*59.9*) | 47.0 | 2.0 | N/A | N/A | N/A |
| SWS | 101.4 (*21.6*) | -15.5** (*30.1*) | 24.7 | 13.1 | -6.1 (*25.4*) | 20.7 | 2.8 | N/A | N/A | N/A |
| REM | 94.3 (*28.9*) | 6.5 (*39.5*) | 33.0 | 0.9 | 8.8 (*42.0*) | 33.0 | 1.9 | N/A | N/A | N/A |
| SOL | 5.3 (*5.9*) | | | | -0.2 (*4.8*) | 2.8 | .01 | N/A | N/A | N/A |

Notes: PSG; polysomnography, AE; absolute error (minutes), *F; F-statistic,* TST; total sleep time, Wake; wake time; Light; light sleep; SWS; slow wave sleep, REM; rapid eye movement sleep; SOL; sleep onset latency. Negative bias indicates an underestimation of the sleep variable by WHOOP-AUTO, WHOOP-MANUAL and ACTICAL when compared to PSG. * indicates significant difference to PSG with p<0.05; ** indicates significant difference to PSG with p<0.001. Data are mean (*SD*).

**Table 4.2** Epoch-by-epoch concordance statistics for WHOOP-AUTO (2-stage and 4-stage categorisation of sleep), WHOOP-MANUAL (2-stage and 4-stage categorisation of sleep) and ACTICAL (2-stage categorisation of sleep) against PSG.

| Measure | Value (%) |
| --- | --- |
| **2-stage comparison** | |
| WHOOP-AUTO | |
| Sensitivity for sleep | 90 |
| Specificity for wake | 60 |
| Overall agreement | 86 |
| WHOOP-MANUAL | |
| Sensitivity for sleep | 97 |
| Specificity for wake | 45 |
| Overall agreement | 90 |
| ACTICAL | |
| Sensitivity for sleep | 98 |
| Specificity for wake | 60 |
| Overall agreement | 89 |
| **4-stage comparison** | |
| WHOOP-AUTO | |
| Sensitivity for wake | 60 |
| Sensitivity for light sleep | 61 |
| Sensitivity for SWS | 63 |
| Sensitivity for REM | 66 |
| Overall agreement | 63 |
| WHOOP-MANUAL | |
| Sensitivity for wake | 45 |
| Sensitivity for light sleep | 67 |
| Sensitivity for SWS | 61 |
| Sensitivity for REM | 66 |
| Overall agreement | 62 |

Notes: SWS; slow wave sleep, REM; rapid eye movement sleep.

**Table 4.3** 4-stage error matrix for WHOOP-AUTO and PSG.

| | | WHOOP-AUTO | | | |
|---|---|---|---|---|---|
| | **Stage** | Wake | Light sleep | SWS | REM |
| **PSG** | Wake | 60% | 26% | <1% | 12% |
| | Light sleep | 14% | 61% | 10% | 15% |
| | SWS | 6% | 28% | 64% | 2% |
| | REM | 6% | 27% | 1% | 66% |

Notes: This matrix presents the percentage of each sleep stage that WHOOP-AUTO has correctly or incorrectly classified compared to PSG. Shaded cells indicate correctly classified sleep. SWS; slow wave sleep, REM; rapid eye movement sleep.

**Table 4.4** 4-stage error matrix for WHOOP-MANUAL and PSG.

| | | WHOOP-MANUAL | | | |
|---|---|---|---|---|---|
| | **Stage** | Wake | Light sleep | SWS | REM |
| **PSG** | Wake | 45% | 37% | <1% | 18% |
| | Light sleep | 7% | 67% | 11% | 15% |
| | SWS | <1% | 38% | 61% | <1% |
| | REM | 1% | 31% | 2% | 66% |

Notes: This matrix presents the percentage of each sleep stage that the WHOOP-MANUAL has correctly or incorrectly classified compared to PSG. Shaded cells indicate correctly classified sleep. SWS; slow wave sleep, REM; rapid eye movement sleep.
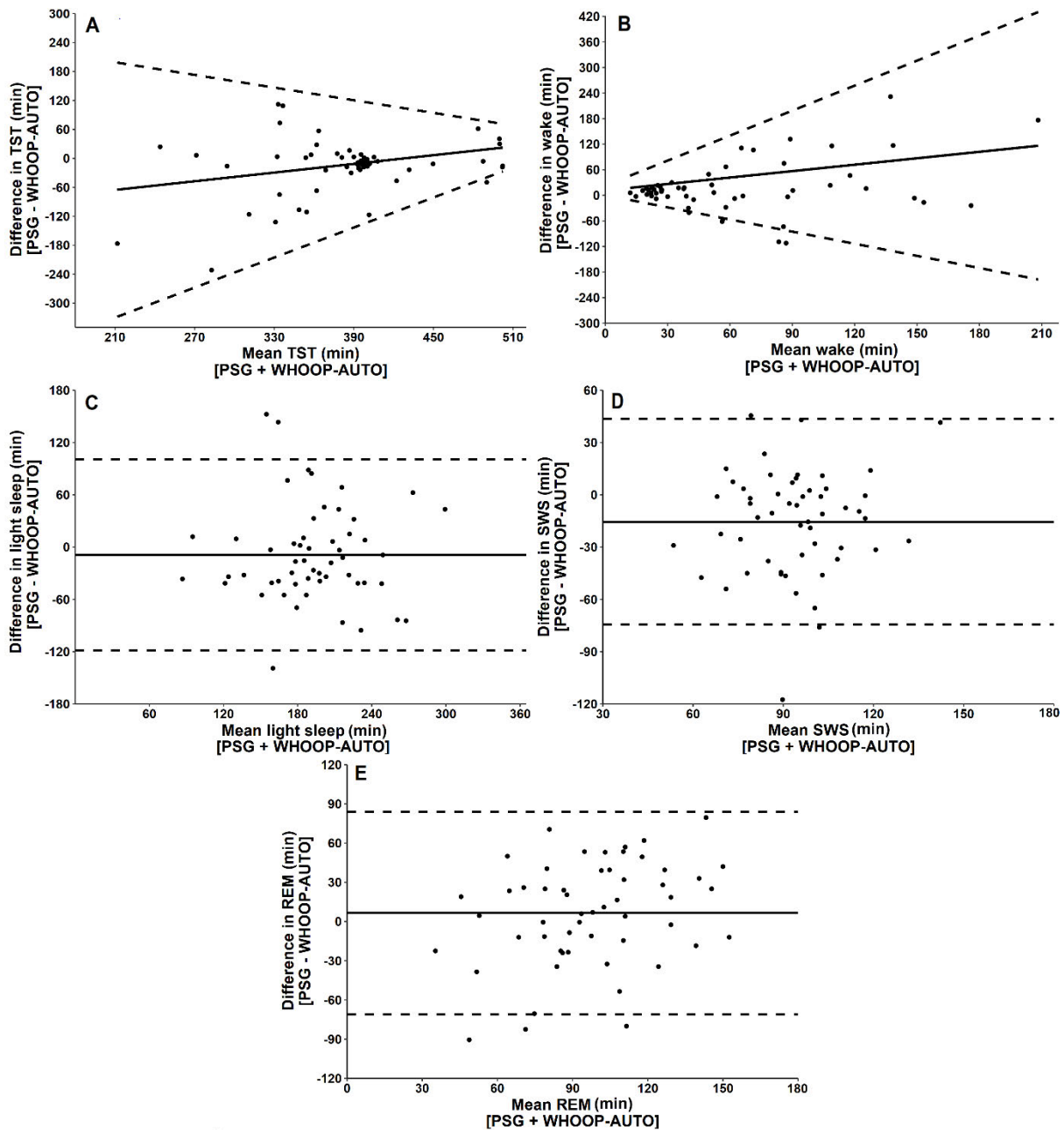
**Figure 4.2** Bland-Altman plots for WHOOP-AUTO and PSG derived measures of (A) total sleep time (TST), (B) wake time, (C) light sleep, (D) slow wave sleep (SWS), (E) rapid eye movement sleep (REM). Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from WHOOP-AUTO and PSG. The y-axes represent the difference between the values, such that positive values indicate that WHOOP-AUTO overestimates relative to PSG and negative values indicate that WHOOP-AUTO underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement (±1.96 standard deviations; Ludbrook, 2010).

**Figure 4.3** Bland-Altman plots for WHOOP-MANUAL and PSG derived measures of (A) total sleep time (TST), (B) wake time, (C) light sleep, (D) slow wave sleep (SWS), (E) rapid eye movement sleep (REM) and (F) sleep onset latency (SOL). Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from WHOOP-MANUAL and PSG. The y-axes represent the difference between the values, such that positive values indicate WHOOP-MANUAL overestimates relative to PSG and negative values indicate WHOOP-MANUAL underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement (±1.96 standard deviations; Ludbrook, 2010).
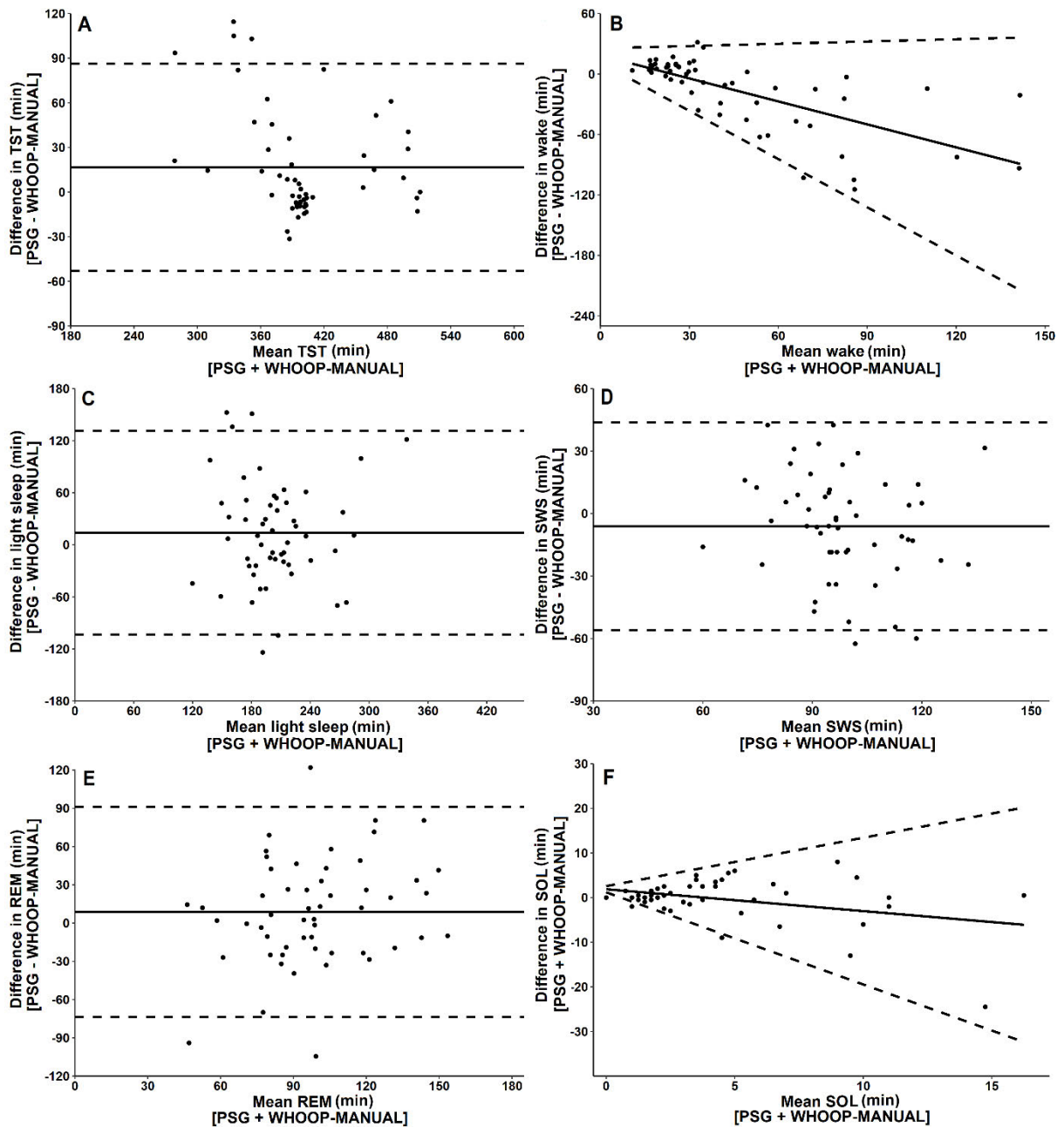
**Figure 4.4** Bland-Altman plots for ACTICAL and PSG derived measures of (A) total sleep time (TST) and (B) wake time. Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from ACTICAL and PSG. The y-axes represent the difference between the values, such that positive values indicate that ACTICAL overestimates relative to PSG and negative values indicate that ACTICAL underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement (±1.96 standard deviations; Ludbrook, 2010).
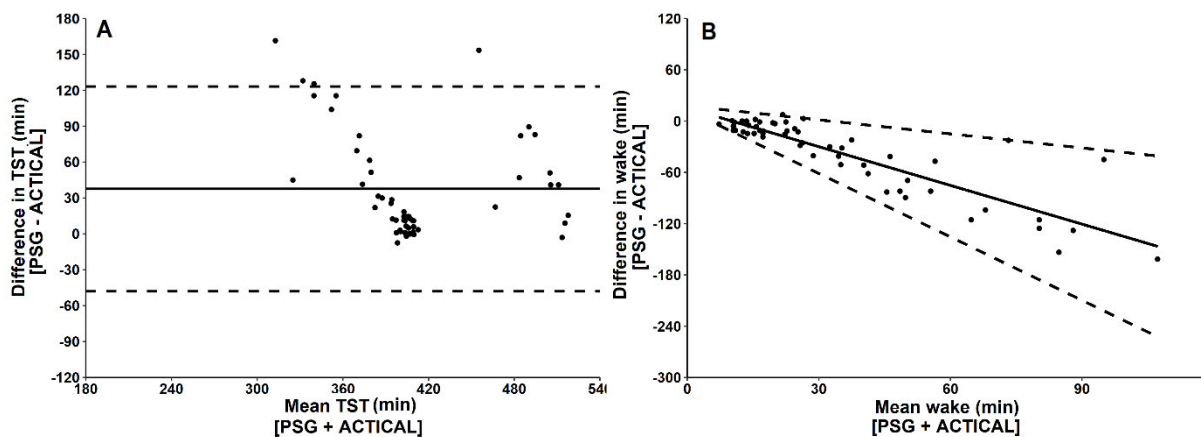
## 4.5 Discussion

The two aims of this study were to; (i) compare WHOOP-AUTO, WHOOP-MANUAL and research grade actigraphy (ACTICAL) for 2-stage categorisation of sleep against PSG, and; (ii) compare WHOOP-AUTO and WHOOP-MANUAL for 4-stage categorisation of sleep against PSG.

### 4.5.1 Two-stage categorisation of sleep

Actigraphy is commonly utilised as an objective measure of sleep and wake by practitioners (e.g., sleep scientists, sports scientists; Zambotti, Cellini et al., 2019; Halson, 2019) However, the process of acquiring sleep data using actigraphy requires certain expertise and is usually a retrospective analysis of an extended data collection period – rather than the immediate day-by-day data that is provided by modern sleep wearables. The accuracy of actigraphy and sleep wearables utilising a manual adjustment function are subject to compliance of the user wearing the device and accurately reporting bed and wake

times. In this context, it is important to compare the performance of actigraphy (i.e., ACTICAL) to modern sleep wearables that can automatically detect sleep and provide easily accessible data (e.g., the WHOOP strap).

Regarding the 2-stage detection of sleep, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL had high sensitivity for sleep (97, 90 and 98% respectively) but WHOOP-MANUAL had lower specificity for wake (45%) than WHOOP-AUTO (60%) and ACTICAL (60%). Chance-corrected agreement was *fair* for ACTICAL (κ = 0.23) and *moderate* for WHOOP-AUTO (κ = 0.44) and WHOOP-MANUAL (κ = 0.48). It should be noted that a previous validation study conducted in the same laboratory (i.e., Chapter 3) found WHOOP-MANUAL to have a 51% specificity for wake when compared to PSG (Miller, Lastella et al., 2020). These findings support a previous validation of WHOOP-MANUAL 2-stage categorisation of sleep (Miller, Lastella et al., 2020) and provide novel support for WHOOP-AUTO as practical alternative for 2-stage categorisation of sleep in absence of PSG.

### 4.5.2 Four-stage categorisation of sleep

For 4-stage categorisation of sleep, WHOOP-AUTO and WHOOP-MANUAL had similar overall agreement (63% and 62% respectively) and sensitivity to light sleep (61% and 67% respectively), SWS (63% and 61% respectively) and REM (66% and 66% respectively). Chance-corrected agreement for 4-stage categorisation of sleep was *moderate* for WHOOP-AUTO (κ = 0.47) and WHOOP-MANUAL (κ = 0.49). This supports a previous validation of WHOOP-MANUAL to measure 4-stage and provides the first validation of WHOOP-AUTO as a practical alternative for 4-stage categorisation of sleep in the absence of PSG. The main disparity between WHOOP-AUTO and WHOOP-MANUAL for 4-stage categorisation of sleep compared to PSG was that WHOOP-AUTO exhibited 16% higher sensitivity for wake compared to WHOOP-MANUAL. However, WHOOP-MANUAL is able to provide an accurate measure of onset latency (Table 4.1). Depending on the variable of interest, practitioners seeking to utilise the WHOOP strap to measure sleep can selectively utilise WHOOP-AUTO or WHOOP-MANUAL functions. For example, in situations where the WHOOP strap is utilised for 2-stage or 4-stage categorisation of sleep for sleep opportunities between 7-9 h, WHOOP-AUTO appears to be the

more practical, better performing function. However, given that WHOOP-MANUAL utilises a reference point for when an individual begins to attempt sleep, it should be used in situations where sleep onset latency is the variable of interest.

The difference for estimating wake between WHOOP-AUTO and WHOOP-MANUAL in this study highlight the need for future validation research to report the ability of sleep wearables to measure a range of sleep measures under auto-detection and manual function. Previous validation studies for consumer sleep wearables do not explicitly report whether data were acquired using automatic detection of sleep or manual entering of sleep times (Meltzer, Hiruma et al., 2015; de Zambotti, Baker et al., 2016; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017) Thus, limiting practitioners' ability to best utilise sleep wearables to measure specific sleep variables. Overall, the findings of this study suggest that WHOOP-AUTO and WHOOP-MANUAL may be used as a practical alternative for 2-stage categorisation of sleep and 4-stage categorisation of sleep when PSG is not available.

### 4.5.3 Comparison to other sleep wearables

Due to an increase in consumer devices providing measures of sleep, it is important to conduct cross-device comparisons. Ideally, within-study comparisons like in the present study should be made to provide meaningful comparison. However, interpretations of cross-study comparisons can be made with consideration to differences in study methodologies (i.e., sleep opportunity, sample, sleep environment). A comparison of the performance of WHOOP-AUTO, WHOOP-MANUAL and ACTICAL respectively, to previous sleep wearable validations can be seen in Figure 4.5 (Meltzer, Hiruma et al., 2015; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017; Cook, Prairie et al., 2018; de Zambotti, Goldstone et al., 2018; Pesonen and Kuula, 2018; Sargent, Lastella et al., 2018; Cook, Eftekari et al., 2019; Walch, Huang et al., 2019; Miller, Lastella et al., 2020).

The WHOOP strap, in both automatic and manual functions, fell within the standard deviation for TST bias, sensitivity for sleep, specificity for wake, sensitivity for light sleep, sensitivity for SWS,

and sensitivity for REM compared to previous validations (Figure 4.5). Previous validations of sleep wearables have shown that there is an apparent "trade-off" between sensitivity and specificity (de Zambotti, Cellini et al., 2019) Such that higher sensitivity may result in decreased specificity, and vice versa. For example, a validation study conducted with the Fitbit One had high sensitivity but had low specificity compared to PSG (Figure 4.5; Meltzer, Hiruma et al., 2015). Compared to the WHOOP strap, other sleep wearables have shown higher sensitivity to individual sleep stages (Figure 4.5). However, both WHOOP-AUTO and WHOOP-MANUAL appear to be consistent across all 4 sleep stages and do not seem to exhibit a large "trade-off" between sensitivities for all sleep stages.

According to the methodologies of previous studies, WHOOP-AUTO provides the only epoch-by-epoch comparison to PSG using sleep auto detection (Meltzer, Hiruma et al., 2015; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017; Cook, Prairie et al., 2018; de Zambotti, Goldstone et al., 2018; Pesonen and Kuula, 2018; Sargent, Lastella et al., 2018; Cook, Eftekari et al., 2019; Walch, Huang et al., 2019). From a practical perspective, WHOOP-AUTO provides a measure of sleep comparable to manually adjusted data and eliminates the risk of non-compliance for entering bed times. Overall, the findings of this validation study suggest that the WHOOP strap, under both automatic and manual detection of sleep, performs well in comparison to other commercially available sleep wearables.
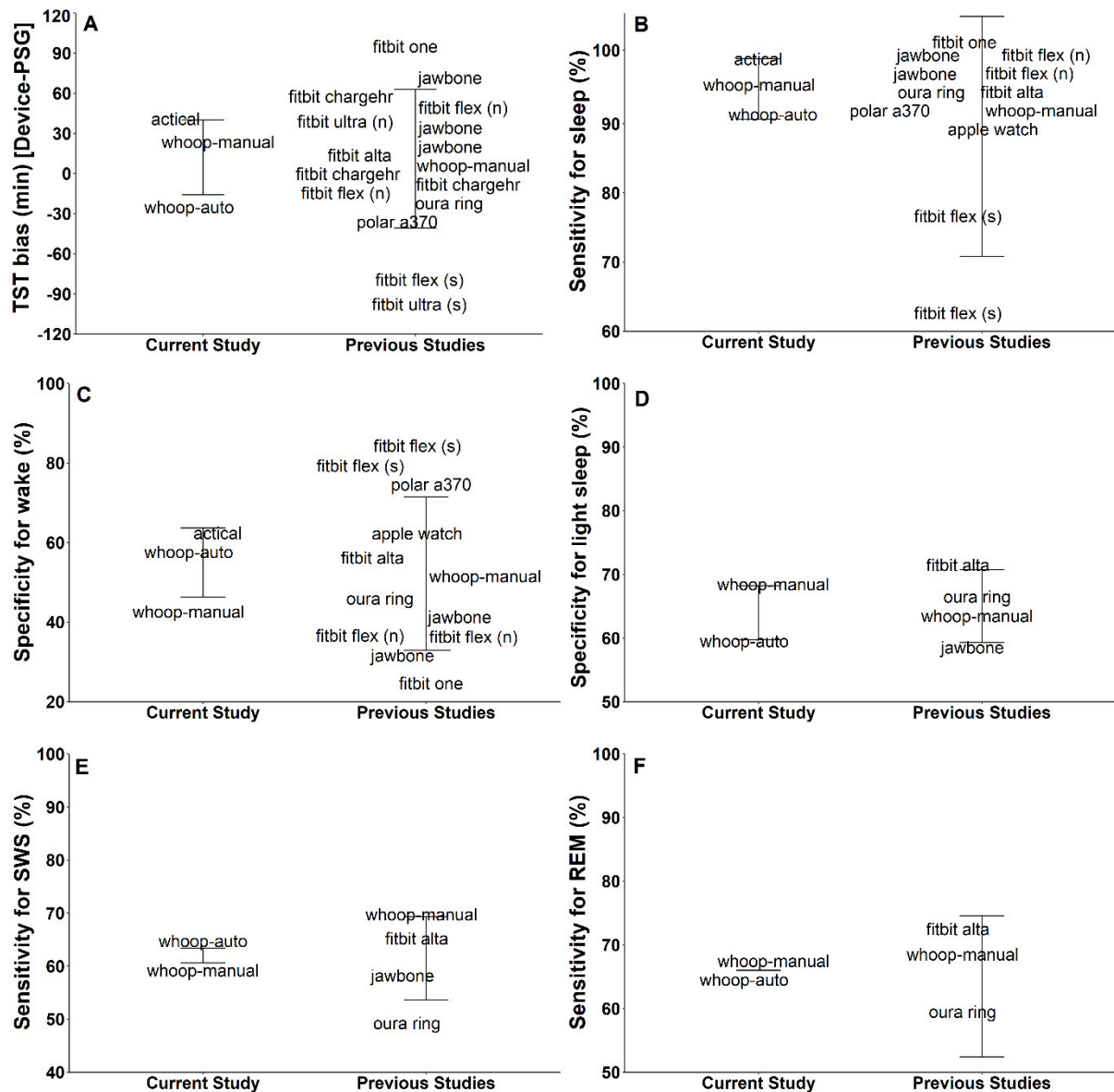
Figure 4.5 Performance of ACTICAL, WHOOP-AUTO, WHOOP-MANUAL and other sleep wearables for (A) total sleep time bias (TST), (B) sensitivity for sleep, (C) specificity for wake, (D) sensitivity for light sleep (E) sensitivity for slow wave sleep (SWS) and (F) sensitivity for rapid eye movement sleep (REM). Error bars represent standard deviation. Fitbit Flex (N); Fitbit Flex with normal sensitivity, Fitbit Flex (S); Fitbit Flex with high sensitivity.Superscript numbers represent respective validation studies (de Zambotti, Baker et al., 2015; Meltzer, Hiruma et al., 2015; Cook, Prairie et al., 2017; de Zambotti, Rosas et al., 2017; Kang, Kang et al., 2017; Maskevich, Jumabhoy et al., 2017; Cook, Prairie et al., 2018; de Zambotti, Goldstone et al., 2018; Pesonen and Kuula, 2018; Sargent, Lastella et al., 2018; Cook, Eftekari et al., 2019; Walch, Huang et al., 2019; Miller, Lastella et al., 2020).

*4.5.4 Boundary conditions and future research*

This validation study was conducted on the WHOOP-AUTO and WHOOP-MANUAL functions of the WHOOP strap. Validation of other WHOOP metrics (i.e., heart rate, heart rate variability) were outside of the scope of this project. The algorithms used by WHOOP to score sleep are proprietary, and epoch-by-epoch data are not accessible through the WHOOP smart phone application. Findings should also be interpreted within the boundary conditions of the sleep environment (laboratory), time in bed opportunities (7-9 h) and sample (healthy young adults). Future investigations should validate the WHOOP strap and other sleep wearables with reference to all available functionalities (i.e., auto-detection and manual adjustment) and across a wider range of conditions (sleep opportunities of different lengths, disturbed sleep periods, unhealthy and/or older populations).

**4.6 References**

Bland, J. M., and Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of biopharmaceutical statistics*, 17(4), 571-582.

Cook, J. D., Eftekari, S. C., Dallmann, E., Sippy, M., and Plante, D. T. (2019). Ability of the Fitbit Alta HR to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: A comparison against polysomnography. *Journal of sleep research*, 28(4),

Cook, J. D., Prairie, M. L., and Plante, D. T. (2017). Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *Journal of affective disorders*, 217, 299–305.

Cook, J. D., Prairie, M. L., and Plante, D. T. (2018). Ability of the Multisensory Jawbone UP3 to Quantify and Classify Sleep in Patients With Suspected Central Disorders of Hypersomnolence: A Comparison Against Polysomnography and Actigraphy. Journal of clinical sleep medicine : *JCSM : official publication of the American Academy of Sleep Medicine*, 14(5), 841–848.

de Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M., and Baker, F. C. (2019). Wearable Sleep Technology in Clinical and Research Settings. *Medicine and science in sports and exercise*, 51(7), 1538-1557.

de Zambotti, M., Baker, F. C., and Colrain, I. M. (2015). Validation of sleep-tracking technology compared with polysomnography in adolescents. *Sleep*, 38(9), 1461–1468.

de Zambotti, M., Baker, F. C., Willoughby, A. R., Godino, J. G., Wing, D., Patrick, K., and Colrain, I. M. (2016). Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiology and behavior*, 158, 143-149.

de Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M., and Baker, F. C. (2018). A validation study of Fitbit Charge 2™ compared with polysomnography in adults. *Chronobiology international*, 35(4), 465–476.

de Zambotti, M., Rosas, L., Colrain, I. M., and Baker, F.C.(2017). The Sleep of the Ring: Comparison of the OURA Sleep Tracker Against Polysomnography. *Behavourial sleep med*icine, 1-15.

Halson S. L. (2019). Sleep Monitoring in Athletes: Motivation, Methods, Miscalculations and Why it Matters. *Sports medicine*, 49(10), 1487–1497.

Iber, C., Ancoli-Israel, S., Chesson, A., and Quan, S. (2007). The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. 2007. American Academy of Sleep Medicine, Westchester, IL.

Kang, S. G., Kang, J. M., Ko, K. P., Park, S. C., Mariani, S., and Weng, J. (2017). Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of psychosomatic research*., 97, 38-44.

Kosmadopoulos, A., Sargent, C., Darwent, D., Zhou, X., and Roach, G. D. (2014). Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavioural research methods*, 46(4), 1032–1041.

Kushida, C. A., Littner, M. R., Morgenthaler, T., Alessi, C. A., Bailey, D., Coleman, J., Jr, Friedman, L., Hirshkowitz, M., Kapen, S., Kramer, M., Lee-Chiong, T., Loube, D. L., Owens, J., Pancer, J. P., and Wise, M. (2005). Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*, 28(4), 499–521.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.

Ludbrook J. (2010). Confidence in Altman-Bland plots: a critical review of the method of differences. *Clinical and experimental pharmacology and physiology*, 37(2), 143–149.

Maskevich, S., Jumabhoy, R., Dao, P. D. M., Stout, J. C., and Drummond, S. P. A. (2017). Pilot Validation of Ambulatory Activity Monitors for Sleep Measurement in Huntington's Disease Gene Carriers. *Journal of huntingtons disease*, 6(3), 249-253.

Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. (2015). Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. *Sleep*. 38:1323-1330.

Miller, D. J., Lastella, M., Scanlan, A. T., Bellenger, C., Halson, S. L., Roach, G. D., and Sargent, C. (2020). A validation study of the WHOOP strap against polysomnography to assess sleep. Journal of sports sciences, 38(22), 2631–2636.

Miller, D. J., Sargent, C., Roach, G. D., Scanlan, A. T., Vincent, G. E., & Lastella, M. (2020). Moderate-intensity exercise performed in the evening does not impair sleep in healthy males. *European journal of sport science*, 20(1), 80–89.

Pesonen AK, Kuula L. The Validity of a New Consumer-Targeted Wrist Device in Sleep Measurement: An Overnight Comparison Against Polysomnography in Children and Adolescents. *Journal of Clinical Sleep Medicine*. 2018 Apr 15;14(4):585-591.

Pollak, C.P., Tryon, W.W., Nagaraja, H., and Dzwonczyk, R. (2001). How accurately does wrist actigraphy identify the states of sleep and wakefulness? Sleep, 24(8), 957-965.Sargent C, Lastella M, Halson SL, Roach GD. (2016). The validity of activity monitors for measuring sleep in elite athletes. Journal of science and medicine in sport. 19:848-853.

Sargent, C., Lastella, M., Romyn, G., Versey, N., Miller, D. J., and Roach, G. D. (2018). How well does a commercially available wearable device measure sleep in young athletes? *Chronobiology international*, 35(6), 754-758.

Saw, A. E., Main, L. C., and Gastin, P. B. (2015). Monitoring athletes through self-report: factors influencing implementation. *Journal of sports science and medicine*, 14(1), 137–146.

Sim, J., and Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*, 85(3), 257–268.

Stein, P. K., and Pu, Y. (2012). Heart rate variability, sleep and sleep disorders. *Sleep medicine reviews*, 16(1), 47–66.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., and Hufford, M. R. (2002). Patient non-compliance with paper diaries. *British medical journal*, 324(7347), 1193–1194.

Walch, O., Huang, Y., Forger, D., and Goldstein, C. (2019). Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. Sleep, 42(12), Yi H, Shin K, Shin C. (2006). Development of the sleep quality scale. *Journal of sleep research*. 15:309-316.

Yi, H., Shin, K., Shin, C. (2006). Development of the sleep quality scale. *Journal of sleep research*, 15(3), 309–316.

**5. The validity of a single question for assessing subjective jet lag in athletes.**

**Abstract**

The aim of this study was to validate a single question for assessing jet lag in athletes. To do this, responses to the question were compared to urinary 6-sulphatoxymelatonin (aMT6s) – the urinary metabolite of melatonin. Concentrations of aMT6s are naturally elevated during nocturnal habitual sleep – therefore low levels of aMT6s during sleep indicate misalignment between the circadian system (e.g., jet lag). Twenty-seven athletes (age: $17.6 \pm 1.3$ yr) that were traveling over multiple time zones participated in the study. On each experimental night, participants responded to a 7-point scale rating the severity of the jet lag they experienced that day (0 = "none"; 7 = "extremely high"). Participants then collected all urine during sleep into a collection container. Samples were frozen and assayed for the concentration of aMT6s. To acquire a secretion rate, the concentration of aMT6s was divided by the volume of the sample and multiplied by the duration of the collection period for the corresponding night. Kendall's tau-b correlation tests with repeated measures were conducted to assess the relationship between the urinary concentration of aMT6s and the subjective jet lag responses. There was a negative correlation between the nightly secretion rate of aMT6s and the subjective jet lag scale, such that levels of aMT6s decreased as subjective jet lag increased ($\tau b = -0.23$, moderate, $p = <.001$). The results suggest that a single jet lag question may be used as a moderate predictor of urinary concentrations of aMT6s – an accurate biomarker. In absence of biomarkers of circadian phase, the scale used in this study may be used as an indicator of jet lag among athletes.

**5.1 Introduction**

Jet lag is caused by misalignment between the body's circadian clock and the environmental time cues of a new destination after transmeridian travel (Eastman, Gazda et al., 2005). Athletes are often required to travel over multiple time zones, and strategies to enhance the entrainment of the circadian clock to the new destination time zone would be beneficial. In a recent publication, detailed instructions on how to minimise jet lag were described for athletes travelling to the Tokyo Olympic games (Roach and Sargent, 2019). These instructions are based upon principles related to the entrainment of the human circadian cycle to environmental time cues via the avoidance/exposure of light, exercise and melatonin (Roach and Sargent, 2019). In principle, these recommendations should be effective, however there are few convenient measures of circadian phase available to athletes.

Commonly utilised measures of jet lag are primarily based on subjective perceptions of jet lag, rather than objective circadian biomarkers (Waterhouse, Edwards et al., 2000). For example, the Liverpool Jet Lag Questionnaire consists of 15 visual analogue scales that assess the severity of jet lag symptoms (e.g., sleep, alertness, fatigue, hunger, bowel movements) and it is commonly used to assess jet lag symptoms in athletes (Ledger, Bin et al., 2020). While such scales are considered beneficial, it is difficult to know how well they reflect adaption to a new time zone because they have not been compared against a biomarker of circadian phase. Gold standard measurements of circadian phase include biomarkers such as melatonin or core body temperature (Khalsa, Jewett et al., 2003). However, it is impractical for athletes and support staff to repeatedly collect, analyse and interpret biomarkers within the window of circadian misalignment (i.e., jet lag) following travel. Therefore, a simple and practical subjective approach to measure jet lag would be of benefit to athletes and support staff.

Previous research investigating jet lag strategies among athletes have measured urinary concentrations of the melatonin metabolite 6-suplhatoxymelatonin (aMT6s) as a biomarker of circadian phase (Cardinali, Bortman et al., 2002). Given that melatonin is excreted nocturnally, low levels of urinary melatonin secreted during daytime hours are expected during circadian misalignment following transmeridian travel. As the circadian clock entrains to the new destination, gradual increases in nocturnal melatonin secretion should occur. Therefore, urinary concentration of nocturnal aMT6s can

be used as an objective measure indicating entrainment of the circadian body clock to a new time zone (i.e., jet lag). Currently, there are no subjective measures of jet lag that have been validated against circadian biomarkers. With this in mind, and with considerations for athlete compliance, the aim of this Chapter was to validate the use of a single question to assess subjective jet lag in athletes against night-time urinary melatonin production..

## 5.2 Methods

Twenty-seven young elite athletes (10 cricket players and 17 basketball players; mean ± standard deviation (SD); age: $17.6 \pm 1.3$ yr) volunteered to participate in the study. Participants, as well as parents and/or guardians where applicable, were provided with an information sheet detailing the benefits and risks of participation as well as their right to withdraw from the study at any stage. Written informed consent was obtained from participants or parents/guardians and ethical approval was obtained from the Central Queensland University Human Research Ethics Committee (HREC:21129).

Data were collected for 8 days following;

- travel from the United Kingdom to Adelaide, Australia (8.5-h eastward time change; n= 9);
- travel from Venice, Italy to Canberra, Australia (10-h eastward time change; n=17).
- travel from Canberra, Australia to Venice, Italy (10-h westward time change; n=17);

### 5.2.1 Measures and Procedures

*Subjective jet lag*

Prior to going to bed each evening, participants provided a subjecting rating of jet lag using a self-report diary. Participants responded on a 7-point scale to the statement "Please rate the level of jet lag you have experienced today", where 0 = "none", 1 = "extremely low", 2 = "very low", 3 = "low", 4 = "moderate", 5 = "high", 6 = "very high" and 7 = "extremely high".

*Urinary aMT6s*

The production rate of melatonin was inferred from concentrations of the metabolite aMT6s. Participants voided urine (i.e., pre-sleep void) immediately prior to each main sleep period. Participants then collected all subsequent urine passed during the sleep period, including the first void upon waking in the morning (post-sleep void) into a 2-L collection container. Participants recorded the clock times for the pre- and post-sleep voids using labels on the collection container. A researcher noted the final volume of the urine and obtained a 1-mL aliquot from each container in 1.5-mL tubes prepared with 10 mg of boric acid as a preservative (concentration = 250 g of boric acid per litre of water). Samples were frozen and assayed for the concentration of aMT6s (aMT6s human radioimmunoassay, Stockgrand, Guildford, UK). The concentration of aMT6s was divided by the volume of the sample (ml) and multiplied by the duration (hours) of the void period for the corresponding night to acquire a secretion rate (ng/hour).

Urine samples were collected during the days immediately following travel. For athletes travelling from Canberra to Venice (and travelling from Venice to Canberra), baseline urine samples were collected prior to travel. Baseline measures could not be collected prior to athletes travelling from the United Kingdom to Adelaide, thus baseline measures of overnight urine were collected 30 days after arrival in these cases (n=9).  Samples were excluded if an error was reported by a participant (i.e., accidental void of urine; n=64 samples) or if urine production and/or aMT6s concentration were outside normative values (n=28 samples; Rosenberg, Fiserova-Bergerova et al., 1989; Mahlberg, Tilmann et al., 2006).

### 5.2.2 Data Analysis

A total of 302 samples were included in the analyses. Kendall's tau-b correlation tests with repeated measures were conducted to assess the relationship between the urinary concentration of aMT6s (continuous variable) and the subjective jet lag measure (ordinal variable; Khamis, 2008). Correlations were interepreted as: <0.10 = very weak; 0.10 to 0.19 = weak; 0.20 to 0.29 = moderate; >0.30 = strong.

All data were analysed with RStudio. An α-level of 0.05 was used to indicate statistical significance. Data are presented as mean ± standard error.

**5.3 Results**

A total of 197 and 105 samples following Eastward and Westward travel were analysed, respectively. On average, participants produced 567.5ml of urine during nocturnal sleep periods. There was a significant negative correlation between aMT6s concentration and subjective jet lag (moderate strength; $\tau b = -0.23$; $z = -5.56$; $p < 0.001$), such that the concentration of urinary aMT6s reduced as subjective jet lag increased (Figure 5.1).
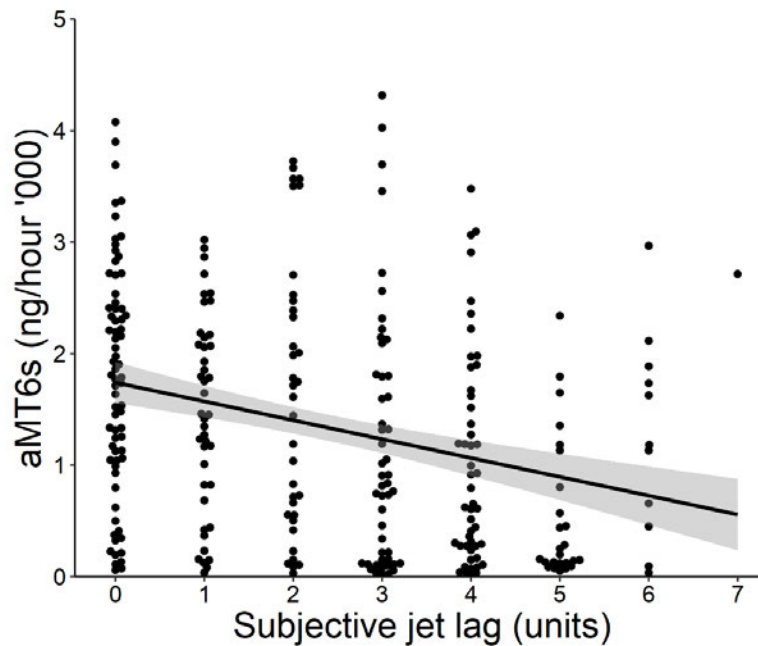


**Figure 5.1** Urinary aMT6s concentration as a function of subjective jet lag. The solid black line represents smoothed conditional means (R package: ggplot2). The grey shaded area represents standard error.

**5.4 Discussion**

The aim of this study was to validate a single question jet lag scale for athletes. The main finding was a significant negative correlation between subjective jet lag and aMT6s concentration (Figure 5.1). This suggests that the single jet lag question may be used as a moderate predictor jet lag.

This is the first study to examine the correlation between a single-question jet lag scale and an established circadian biomarker (i.e., urinary melatonin) in athletes. Due to the difficulty in collecting bio-specimens in the field, previous validations of subjective jet lag scales have relied on correlations between subjective jet lag and associated factors such as fatigue (Ledger, Bin et al., 2020). A potential limitation for this type of validation (i.e., subjective to subjective) is that uncontrolled factors such as fatigue may change for other reasons not associated with jet lag. The novel data collection methods utilised in this study allowed for comparison against an objective circadian biomarker. Indeed, the data reflect the expected relationship between subjective jet lag and the concentration of nocturnal urinary melatonin. Specifically, subjective jet lag is likely to be high when there the concentration of nocturnal urinary melatonin is a low.

Current tools for assessing jet lag typically include a number of questions related to different symptoms associated with jet lag (e.g., Liverpool jet lag questionnaire). These tools are useful, but one disadvantage is that they can take 5-10 min to complete. In the present study, a single question was used to assess subjective jet lag. In situations where time is critical and/or compliance in completing a long questionnaire is an issue, the single question might be a suitable alternative. Compliance for reporting subjective data is a common hurdle when collecting data in the athlete population (Stone, 2002). For this reason, jet lag scales with multiple questions (e.g., Liverpool Jet Lag Questionnaire) may impact the motivation for athletes to self-report. For these reasons, the scale used in this study may provide a practical alternative to other questionnaires.

The findings of this study should be interpreted under the boundary conditions of the direction of travel undertaken (i.e., 8.5-h east, 10-h east and 10-h west) and the population examined (i.e., young, elite, team-sport athletes). With consideration to these boundary conditions, athletes and support staff

may use a single question jet lag scale as an estimation of jet lag. The analysis in this study did not control for potential confounding variables related to jet lag such as fatigue or exercise performed while jet lagged. To make robust comparisons between jet lag and aMT6s, future investigations should compare several subjective jet lag scales against circadian biomarkers using a repeated measures design.

**5.5 References**

Cardinali, D. P., Bortman, G. P., Liotta, G., Pérez Lloret, S., Albornoz, L. E., Cutrera, R. A., Batista, J., and Ortega Gallo, P. (2002). A multifactorial approach employing melatonin to accelerate resynchronization of sleep-wake cycle after a 12 time-zone westerly transmeridian flight in elite soccer athletes. *Journal of pineal research*, 32(1), 41–46.

Eastman, C.I., Gazda, C.J., Burgess, H. J., Crowley, S. J., and Fogg, L. F. (2005). Advancing circadian rhythms before eastward flight: a strategy to prevent or reduce jet lag. *Sleep*, 28(1), 33.

Khalsa, S. B., Jewett, M. E., Cajochen, C., and Czeisler, C. A. (2003). A phase response curve to single bright light pulses in human subjects. *The journal of physiology*, 549 (Pt 3), 945–952.

Khamis, H. (2008). Measures of Association: How to Choose? *Journal of diagnostic medical sonography*, 24(3), 155–162.

Ledger S, Bin YS, Nour M, Cistulli P, Bauman A, Allman-Farinelli M, Naismith SL, Stamatakis E, Caillaud C, De Chazal P, Simpson SJ. (2020). Internal consistency and convergent and divergent validity of the Liverpool jet lag questionnaire. *Chronobiology international*. 37:218-226.

Mahlberg, R., Tilmann, A., Salewski, L., and Kunz, D. (2006). Normative data on the daily profile of urinary 6-sulfatoxymelatonin in healthy subjects between the ages of 20 and 84. Psychoneuroendocrinology, 31(5), 634–641.

Robert, B. (2011). Significance and Measures of Association. Retrieved from http://polisci.usca.edu

Roach, G. D., and Sargent, C. (2019). Interventions to Minimize Jet Lag After Westward and Eastward Flight. *Frontiers in Physiology*, 10(927).

Rosenberg J, Fiserova-Bergerova V, Lowry LK. (1989). Horizons: Biological Monitoring IV: Measurements in Urine. *Applied Industrial Hygiene*. 4:F-16-F-21.

Waterhouse, J., Edwards, B., Nevill, A., Atkinson, G., Reilly, T., Davies, P., and Godfrey, R. (2000). Do subjective symptoms predict our perception of jet-lag?. *Ergonomics*, 43(10), 1514–1527.

**6. Moderate-intensity exercise performed in the evening does not impair sleep in healthy males.**

Published peer-reviewed publications associated with this Chapter (Appendix D):

Miller, D. J., Sargent, C., Roach, G. D., Scanlan, A. T., Vincent, G. E., and Lastella, M. (2020). Moderate-intensity exercise performed in the evening does not impair sleep in healthy males. European journal of sport science, 20(1), 80–89.

**Abstract**

The aim of this study was to examine the effect of single bouts of moderate-intensity aerobic exercise and moderate-intensity resistance exercise performed in the evening on the sleep of healthy young males. The study employed a repeated-measures, counterbalanced, crossover design with three conditions (control, evening aerobic exercise, evening resistance exercise). Twelve male participants (mean ± SD; age:  27.3 ± 3.4 yr) attended the laboratory on three occasions separated by one day between each visit. Between 20:45h and 21:30h, participants completed either no exercise, 30 minutes of aerobic exercise at 75%HRmax, or 30 minutes of resistance exercise corresponding to 75% of 10-repetition maximum. A 9-h sleep opportunity was provided between 23:00h and 08:00h. Core body temperature was measured using ingestible temperature capsules and sleep was measured using polysomnography. Core body temperature was higher during the aerobic exercise and resistance exercise compared to control (p=0.001). There was no difference in core body temperature at bedtime between the conditions. Sleep onset latency, total sleep time, slow-wave sleep duration, REM sleep duration, wake after sleep onset and sleep efficiency were similar in each condition (p>0.05). Single bouts of moderate-intensity aerobic exercise or moderate-intensity resistance exercise performed in the evening did not impact subsequent night-time sleep. Core body temperature increased during both forms of exercise, but returned to pre-exercise levels in the 90 minutes prior to bedtime. Healthy young males can engage in a single bout of moderate-intensity aerobic exercise or moderate-intensity resistance exercise ceasing 90 minutes before bed without compromising their subsequent sleep.

## 6.1 Introduction

To maintain a healthy lifestyle, humans should perform at least 150 to 300 minutes of moderate-intensity exercise each week (Piercy and Troiano, 2018) and obtain 7 to 9 hours of sleep each night (Haskell et al., 2007; Hirshkowitz et al., 2015). In most cases, physical activity and sleep are complementary; active individuals are healthy sleepers, and good sleepers tend to be physically active (Hasan, Urponen, Vuori, and Partinen, 1988; Vuori, Urponen, Hasan, and Partinen, 1988). However, in a time-poor society, exercise and sleep compete with each other for time on a daily basis (Vincent et al., 2017; Yao, 2018). Lack of time is reported as a common perceived barrier to completing daily physical activity (Cerin, Leslie, Sugiyama, and Owen, 2010). For individuals that may not have time during the day, the advent of 24-hour gyms accommodates for evening exercise.

While the chronic effects of exercise on sleep may be beneficial, the acute effects of exercise may be detrimental to sleep. According to traditional sleep hygiene recommendations (Zarcone, 1994), those who exercise in the evening may be at risk of compromising subsequent sleep. Acutely, increased core body temperature, increased energy expenditure and skeletal muscle damage are potential factors that may influence sleep (Driver and Taylor, 2000). The human circadian rhythms of core body temperature and sleep are closely linked (Lack and Lushington, 1996). During periods of high core body temperature, the human body is awake and active; during periods of low core body temperature the human body is inactive and asleep (Van Dongen and Dinges, 2000). The onset of sleep is associated with a natural evening decline in core body temperature (Driver and Taylor, 2000; Murphy and Campbell, 1997). The thermogenic effect of exercise may interfere with this decline in core body temperature if conducted in the evening (Murphy and Campbell, 1997; Sawka, 2011). Even in the absence of exercise, elevated core body temperature can impact subsequent sleep (Horne and Reid, 1985; Shapiro, Allan, Driver, and Mitchell, 1989). It is reasonable to suggest that elevated core body temperature is one mechanism through which evening exercise may impair night-time sleep. Thus, traditional sleep hygiene recommendations advise against exercising near bedtime (Zarcone, 1994).

Recent epidemiological data has challenged traditional sleep hygiene recommendations, highlighting a positive association between evening exercise and self-reported sleep quality and

quantity (Buman, Phillips, Youngstedt, Kline, and Hirshkowitz, 2014). In a cross-sectional study examining sleep and exercise, evening exercise was not associated with disturbed sleep (Buman et al. 2014). Furthermore, experimental data suggests that evening exercise may result in increased slow wave sleep (Dworak et al., 2008), increased sleep efficiency and reduced REM onset latency (Flausino, Da Silva Prado, de Queiroz, Tufik, and de Mello, 2012) and have no impact on subjective sleep quality (Myllymaki et al., 2011) or quantity (Alley, 2015; O'Connor, Breus, and Youngstedt, 1998). When interpreting such data, it is important to consider differences in exercise modality. Both aerobic exercise and resistance exercise are recommended for health benefits (e.g., decreased risk of hypertension, diabetes and depression; Warburton, Nicol, and Bredin, 2006), but the effect of these modalities on sleep when performed close to bedtime is not well understood. Acutely, aerobic exercise results in significantly increased heart rate and oxygen consumption compared to resistance exercise (Pontifex, Hillman, Fernhall, Thompson, and Valentini, 2009). Due to physiological differences between aerobic exercise and resistance exercise, it is possible that they may influence sleep differently when performed in the evening.

To date, few experimental studies have investigated the effect of evening exercise on sleep, and no study has examined the effect of evening exercise modality (e.g., aerobic exercise and resistance exercise) on sleep. If elevated core body temperature does impact sleep, then post-exercise sleep should not be impaired if core body temperature has returned to pre-exercise levels prior to bedtime. Therefore, the aim of this study was to examine the effect of a single bout of moderate-intensity aerobic exercise and single bout of moderate-intensity resistance exercise completed 90 minutes before bedtime on subsequent night-time sleep in well-trained individuals.

**6.2 Methods**

This study employed a repeated measures, counterbalanced, crossover design with three conditions (control, evening aerobic exercise, evening resistance exercise). Participants attended the laboratory on three separate occasions, separated by one day between each visit. During each visit, participants either remained sedentary, completed 30 minutes of aerobic exercise or 30 minutes of resistance exercise between 20:45h and 21:30h (Figure 6.1). Participants were given a 9-h sleep opportunity between 23:00h and 08:00h.

> *6.2.1 Participants*

Twelve healthy young males participated in the study (mean ± SD; age: 27.3 ± 3.4 yr; height: 187.6 ± 5.9 cm; body mass: 82.8 ± 10.9 kg). Potential participants completed a detailed general health questionnaire and underwent a screening interview with a member of the research team. Participation was subject to meeting the following criteria: 18-30 yr; engaging in aerobic (i.e., running, cycling) and/or resistance exercise at least three days per week (inclusive of both exercise modalities); non-smoker for a minimum of six months; no pre-existing psychiatric, neurological or sleep-related conditions; and not taking medication or drugs affecting the central nervous system. Additionally, participation was subject to completion of an aerobic exercise test and resistance exercise test. Participants were supervised by an accredited strength and conditioning coach during all exercise sessions.

A verbal and written description of the study was provided to participants, and written informed consent was obtained prior to participation. Participants were compensated financially for their time. Ethical approval was obtained from the Central Queensland University Human Research Ethics Committee (HREC: 20612).
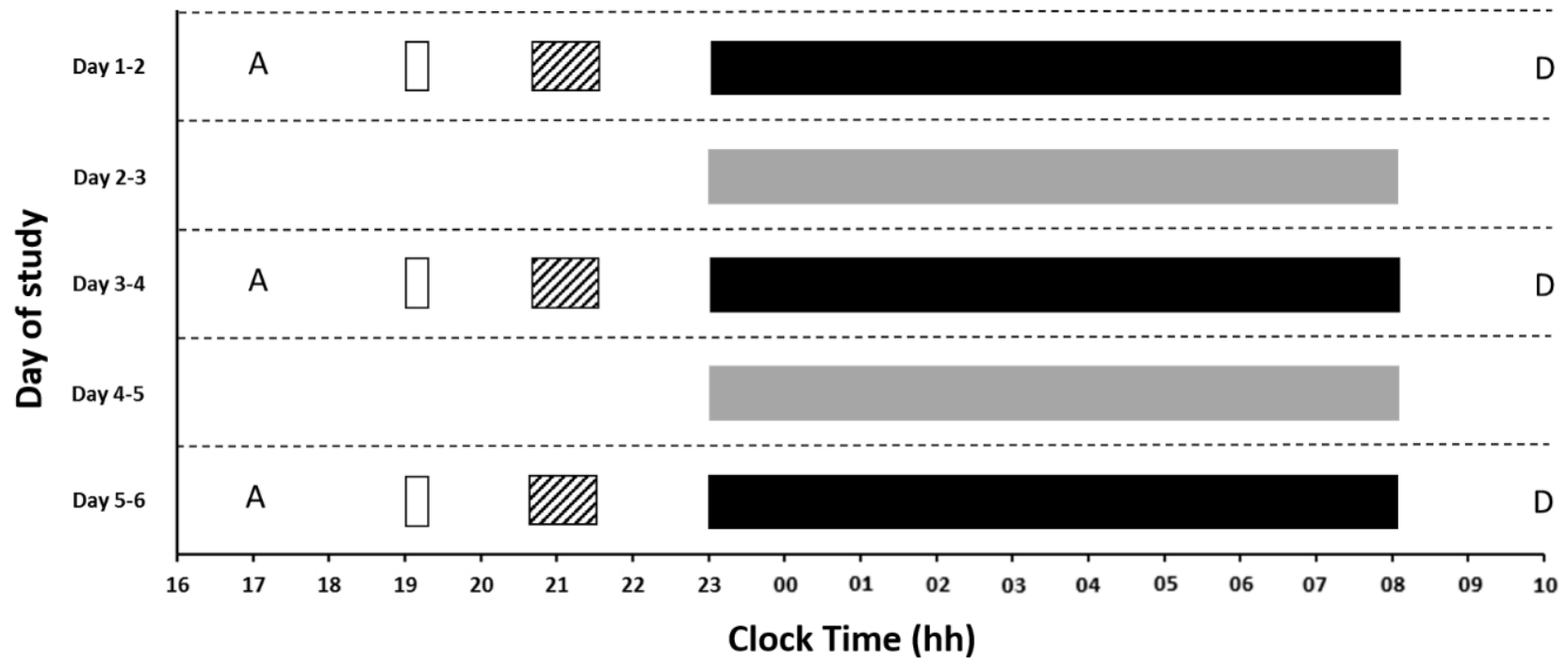
**Figure 6.1** Illustration of the study design. "A": participant arrival, white bars: allocated dinner time, lined bars: experimental conditions (i.e. control, aerobic exercise, resistance exercise), black bars: experimental sleep opportunity in the laboratory, "D": participant departure, grey bars: target sleep opportunity for wash out nights at home. Y-axis: "Day of Study". X-axis: "Clock Time (hh)

### 6.2.2 Laboratory setting

The study was conducted in a purpose-built accommodation suite at the Appleton Institute for Behavioural Science. The suite is sound attenuated, free from external environmental cues and houses six participants at a time. Participants were provided with a private bedroom, living room and bathroom facilities. Meals were served to participants in a common dining room area. During all wake periods, ambient light intensity was maintained at 350 lux (6-foot angle of gaze) and all lights were extinguished during sleep opportunities (<.03 lux). The laboratory was maintained at target temperature between 21-23°C (Kingma et al., 2014). Exercise procedures were conducted in a separate air-conditioned room (20.8 ± 0.8°C; 16.3 ± 0.8% relative humidity).

### 6.3.3 Procedures

*Pre-experimental*

On two separate occasions prior to the study, participants attended the laboratory to perform two exercise tests – a submaximal graded cycling test and a 10-repetition maximum test. In the week prior to the study, participants were instructed to maintain their habitual sleep/wake behaviour. This was measured using wrist activity monitors (Actical Z-series, Mini-Mitter Philips Respironics, Inc. Bend, Oregon; Kosmadopoulos, Sargent, Darwent, Zhou, and Roach, 2014) and self-report sleep diaries (Total sleep time = 7.3 ± 0.9h). In addition, participants were asked to refrain from caffeine and alcohol, and to avoid any moderate to vigorous exercise, in the 48h preceding the first experimental day until the completion of the study.

*Experimental*

The 12 participants completed the protocol over a 6-day period (Figure 6.1). Participants completed the three conditions in six randomised orders (i.e., control, aerobic, resistance; aerobic, resistance, control etc.). On the first experimental day, participants were familiarised with the laboratory and the experimental protocol. On days 1, 3 and 5, participants entered the laboratory at 17:00h. The aerobic exercise session and resistance exercise session began at 20:45h and were completed by

21:30h. Following completion of exercise sessions, participants rated their perceived exertion and whole-body fatigue before returning to their designated living areas to prepare for bed. In the control condition, participants prepared for bed at 20:45h and performed sedentary tasks until bedtime. In each condition, participants were permitted to wash their face and clean their teeth prior to bedtime, but were not allowed to shower.

In the control condition, a research assistant applied electrodes necessary for sleep measurement between 20:45h and 21:30h. In the aerobic exercise and resistance exercise conditions, sleep electrodes were applied immediately following exercise. At 22:30h, participants rated their subjective sleepiness, pre-sleep alertness and pre-sleep tension. In all conditions, participants remained in bed with the lights turned off between 23:00h and 08:00h. Upon awakening participants at 08:00h, research assistants removed electrodes and participants rated their subjective sleep quantity and subjective sleep quality. Participants rated their muscle soreness at 09:00h.

Upon arrival at the laboratory on days 1, 3 and 5, participants ingested a core body temperature capsule at ~17:00h. This allowed time for the capsule to pass into the small intestine before performing exercise and to reduce the risk of inaccurate readings due to ingested fluid or food in the stomach (Livingstone, Grayson, Frim, Allen, and Limmer, 1983). Prior to ingesting the capsule on days 3 and 5, a research assistant confirmed that the previous capsule had been passed as per normal digestive transit time (Roach, 2010). On days 1, 3 and 5 in the laboratory, participants were provided with a meal at 18:00h. The macronutrient profile of the meal was similar to a standard healthy Western diet (Trumbo et al., 2002). Participants were required to consume 100% of each meal and bottled water was available ad libitum (1.6 ± 0.6 litres per experimental night). Between 18:30h and 20:45h, participants remained in their rooms performing sedentary activities (e.g., reading, watching movies).

Following each night in the laboratory, participants returned home for a 'wash-out' day and night (Figure 6.1). During wash-out periods, participants were instructed to maintain a target sleep opportunity (23:00h – 08:00h). The entire experimental protocol was completed by 10:00h on Day 6.

*6.3.4 Measures*

*Pre-experimental submaximal graded exercise test*

The submaximal graded exercise test was performed on a stationary cycle ergometer (Wattbike Trainer, Wattbike Ltd; Nothingham, UK). The test began with a 5-min warm up at 55 W at 60 rev·min$^{-1}$. Thereafter, power output was increased 15 W every min by increasing cadence. During the test, participants rated their level of perceived exertion each min using the Borg Scale (Borg, 1982). The test ended once a participant reached an RPE of 15 (i.e., 'hard') or greater. The values for power output and heart rate obtained during the test were plotted for each participant and the predicted power output corresponding to 75% of maximum heart rate was extrapolated (Tanaka, Monahan, and Seals, 2001).

*Pre-experimental resistance exercise test*

The 10-repetition maximum (10-RM) was established for exercises involving the major muscle groups (i.e., chest press, bicep curl, seated triceps press, lateral shoulder raise, and rear leg elevated lunge). Determination of 10-RM began with a warm-up set of 10 repetitions at a self-selected load, followed by progressively heavier loads separated by 2 minutes of rest. The 10-RM was accepted when a participant could perform 10 but not more than 10 repetitions of the specified exercise (Alley, 2015). Participants also completed an inverted row with body mass providing resistance. For this exercise, participants performed one set to exhaustion.

*Experimental aerobic exercise session*

The aerobic exercise session began with a 5-min warm-up, followed by a 30-min cycle at 75% heart rate maximum (HRmax), and a 5-min warm-down. During the warm-up and the warm-down, participants cycled at an intensity corresponding to 40% of peak power output. During the 30-min cycle, participants maintained a cadence of 80 rev·min$^{-1}$ and an average power output of 149.4 ± 28.6 W. Heart rate was monitored continuously to ensure participants maintained a power output corresponding to 75% HRmax. If HR was below or above the target valve, the air resistance setting was adjusted accordingly. The aerobic exercise session was structured to replicate an aerobic

training session representative of the exercise habits of a healthy young male (Buckworth and Nigg, 2004).

*Experimental resistance exercise session*

The resistance exercise session began with a 5-min warm up, followed by three sets of 10 repetitions of each exercise (i.e., chest press, bicep curl, seated triceps press, side lateral shoulder raise, and rear leg elevated lunge). Each exercise was performed at a weight corresponding to 75% of the pre-determined 10-RM, with a 1-min rest period between each set, and a 2-min rest period between each exercise. For the inverted row, participants performed the number of repetitions corresponding to 75% of exhaustion. Heart rate was monitored continuously during the exercise session. The resistance exercise session was structured to replicate a typical resistance training session representative of the exercise habits of a healthy young male (Buckworth and Nigg, 2004).

*Heart rate*

During all exercise sessions, heart rate was measured continuously using Polar M400 heart rate monitors (M400, Polar Electro; Kempele, Finland).

*Core body temperature*

Core body temperature was measured using the VitalSense™ telemetric physiological monitoring system (Phillips Respironics; Bend, Oregon; Darwent, Zhou, van den Heuvel, Sargent, and Roach, 2011). The system consists of an ingestible capsule thermometer and a wireless transmitter. The ingestible capsule (mass = 1.6 g, length = 22 mm, diameter = 8.6 mm) has a precision of $\pm 0.1$°C. Once activated, the capsule transmits temperature data in 1-min epochs to an external monitor. During wake periods, researchers checked core body temperature monitors every 30-min to ensure connection with the capsule was not lost.

*Sleep*

On experimental nights in the laboratory, sleep was measured using polysomnography (PSG). Prior to each sleep opportunity, a standardised montage of electrodes was attached to the face and scalp of participants. The montage included three electroencephalography electrodes (i.e., C4-M1, F4-M1, O2-M1), two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram. PSG data were recorded directly to data acquisition, storage, and analysis systems (Grael, Compumedics; Victoria, Australia). Sleep records were manually scored in 30-s epochs according to standard criteria by a registered polysomnographic technician who was blind to condition (Iber, Ancoli-Israel, Chesson, and Quan, 2007).

The sleep variables calculated from each PSG recording included:

- Sleep onset latency (min): the duration of time from lights out to the first epoch of any stage of sleep;

- Total sleep time (min): the sum of minutes spent in any stage of sleep (Stages 1, 2, 3, REM);

- Sleep architecture: time in minutes spent in stages 1, 2, 3, REM;

- Wake after sleep onset (min): the duration of time spent awake between sleep onset and lights on;

- Sleep efficiency (%): total sleep time divided by time in bed, multiplied by 100.

*Subjective ratings*

Subjective sleepiness was assessed using the Karolinska Sleepiness Scale, where 1 = "extremely alert" and 9 = "very sleepy, great effort to keep alert, fighting sleep" (Akerstedt and Gillberg, 1990). Subjective sleep quality was assessed using a 7-point scale, where 1 = "extremely poor" and 7 = "extremely good". Subjective sleep quantity and subjective sleep onset latency was assessed by verbally asking participants "how much sleep do you think you got?" and "how long did it take you to fall asleep?". Pre-sleep alertness and pre-sleep tension were measured using a 5-

point Likert scale where 0 = "not at all alert" and 5 = "very alert" and 0 = "not at all tense" and 5 = "very tense".

Muscle soreness was measured using a visual analogue scale with the anchor points "not at all sore" and "very sore". Participants responded by placing a mark through a 100-mm horizontal line between two opposing statements, providing a score out of 100 (Wewers and Lowe, 1990). Whole-body fatigue was measured using a 10-point Likert scale where 0 = "no fatigue at all" and 10 = "extremely high". Perceived exertion was measured using Borg's 6-20 scale where 6 = "no exertion" and 20 = "maximal exertion" (Borg, 1982).

## 6.4 Data analysis

For variables related to aerobic exercise and resistance exercise (i.e., heart rate, RPE), separate repeated measures analysis of variance (ANOVA) were conducted with one within subjects factor (condition: aerobic exercise, resistance exercise). For whole-body fatigue and muscle soreness, separate repeated measures ANOVAs were conducted with one within subjects factor (condition: control, aerobic exercise, resistance exercise). For core body temperature, a repeated measures ANOVA was conducted with two within-groups factors (condition: control, aerobic exercise, resistance exercise; time: between 19:00h and 20:45h, between 20:45h and 21:30h, and lights out time at 23:00h). For the dependent variables related to sleep, separate repeated measures ANOVA were conducted with one within-subjects factor (condition: control, aerobic exercise, resistance exercise). Where a significant main effect was observed, pairwise post-hoc comparisons were conducted and a Bonferroni correction applied. All data were analysed using IBM SPSS Statistics (v24.0; IBM Corp., Armonk, NY, USA). An α-level of 0.05 was used to indicate significance for statistical comparisons. Data are presented mean ± SD.

**6.5 Results**

*6.5.1 Aerobic exercise session and resistance exercise session*

There was a main effect of condition on mean absolute heart rate and mean relative heart rate during exercise (Table 6.1). Mean absolute heart rate and mean relative heart rate were significantly higher during aerobic exercise than during resistance exercise. There was a main effect of condition on perceived exertion and whole-body fatigue (Table 6.1). Perceived exertion was higher following resistance exercise compared to aerobic exercise, and subjective whole-body fatigue was higher after resistance exercise compared to control.

*6.5.2 Core body temperature*

There was a significant main effect of condition and time on core body temperature, and a significant time × condition interaction. Core body temperature was higher during exercise in the aerobic exercise condition when compared to resistance exercise and control conditions. Core body temperature was significantly higher between 20:45h and 21:30h when compared to 19:00h-20:45h and lights out (23:00h). Core body temperature was higher between 20:45h and 21:30h in the aerobic exercise condition when compared to both the resistance exercise condition and the control condition.

*6.5.3 Sleep*

There was no effect of condition on total sleep time, sleep architecture, sleep onset latency, wake after sleep onset, number of awakenings and sleep efficiency (Table 6.2). There was no difference in subjective sleep measures between the control, aerobic exercise or resistance exercise conditions (Table 6.2).

**Table 6.1** Exercise and core body temperature variables as a function of condition.

| Variable | Conditions | | | Statistical outcomes | |
|---|---|---|---|---|---|
| | Control | Aerobic | Resistance | F (df) | P-value |
| **Exercise** | | | | | |
| Morning muscle soreness (units) | $10.7 \pm 14.4$ | $11.7 \pm 12.9$ | $16.5 \pm 15.4$ | 3.285 (2,10) | .080 |
| WBF (units) | $2.5 \pm 1.0$ | $3.3 \pm 1.0$ | $4.3 \pm 1.3^{ab}$ | 10.716 (2,10) | .003 |
| Mean heart rate during exercise (beats·min$^{-1}$) | N/A | $139.3 \pm 2.7$ | $127.3 \pm 16.0^{b}$ | 7.215 (1,11) | .021 |
| Absolute heart rate during exercise (%HR$_{max}$) | N/A | $73.0 \pm 1.3$ | $64.9 \pm 7.5^{b}$ | 15.776 (1,11) | .002 |
| Post-exercise RPE (units) | N/A | $13.0 \pm 1.4$ | $13.7 \pm 1.6$ | 1.615 (1,11) | .230 |
| **Core body temperature (°C)** | | | | | |
| Mean CBT 19:00h – 20:45h | $37.1 \pm 0.1$ | $37.2 \pm 0.2$ | $37.1 \pm 0.2$ | 0.884 (2,10) | .443 |
| Mean CBT 20:45h – 21:45h | $37.3 \pm 0.2$ | $37.7 \pm 0.3^{a}$ | $37.3 \pm 0.2^{b}$ | 16.846 (2,10) | .001 |
| CBT at 23:00h | $37.0 \pm 0.3$ | $37.2 \pm 0.2$ | $37.1 \pm 0.2$ | 1.200 (2,10) | .341 |
| Mean CBT 23:00h – 08:00h | $36.4 \pm 0.1$ | $36.5 \pm 0.2$ | $36.4 \pm 0.1$ | 1.513 (2,10) | .267 |

Data are mean ± SD. WBF, whole-body fatigue; RPE, rating of perceived exertion; beats·min$^{-1}$, heartbeats per minute; %HR$_{max}$, percentage of heart rate maximum; CBT, core body temperature. [a] indicates significant difference from control ($p < 0.05$); [b] indicates significant difference from aerobic exercise ($p < 0.05$).

**Table 6.2** Sleep variables as a function of condition.

| Variable | Conditions | | | Statistical outcomes | |
|---|---|---|---|---|---|
| | Control | Aerobic | Resistance | F (df) | P |
| TST (min) | 488.7 ± 24.3 | 484.7 ± 28.8 | 498.3 ± 19.5 | 3.484 (2,10) | .071 |
| N1 (min) | 24.8 ± 12.7 | 22.2 ± 7.2 | 22.0 ± 7.2 | 2.275 (2,10) | .153 |
| N2 (min) | 219.7 ± 24.8 | 225.4 ± 37.4 | 232.2 ± 37.5 | 1.217 (2,10) | .336 |
| N3 (min) | 143.5 ± 40.7 | 142.5 ± 43.9 | 141.2 ± 28.2 | 0.041 (2,10) | .960 |
| REM (min) | 109.0 ± 23.3 | 104.0 ± 19.2 | 111.2 ± 24.4 | 0.555 (2,10) | .591 |
| SOL (min) | 19.0 ± 13.9 | 26.4 ± 18.6 | 19.8 ± 13.8 | 1.765 (2,10) | .221 |
| WASO (min) | 32.3 ± 20.2 | 26.3 ± 18.6 | 21.9 ± 10.3 | 2.651 (2,10) | .119 |
| Awakenings (count) | 24.4 ± 7.9 | 22.5 ± 6.6 | 23.5 ± 5.7 | 0.927 (2,10) | .427 |
| SE (%) | 90.5 ± 4.5 | 89.8 ± 5.3 | 92.3 ± 3.6 | 3.478 (2,10) | .071 |
| Subjective sleep quantity (min) | 432.0 ± 66.0 | 390.0 ± 90.0 | 414 ± 90.0 | 1.203 (2,10) | .340 |
| Subjective sleep quality (units) | 4.3 ± 0.9 | 4.4 ± 1.1 | 4.3 ± 1.0 | 0.219 (2,10) | .807 |
| Subjective sleepiness (units) | 5.8 ± 1.3 | 5.6 ± 1.4 | 5.3 ± 1.6 | 0.448 (2,10) | .651 |
| Pre-sleep alertness (units) | 2.3 ± 0.9 | 2.9 ± 0.8 | 2.8 ± 1.0 | 2.518 (2,10) | .130 |
| Pre-sleep tension (units) | 2.1 ± 0.7 | 2.3 ± 0.9 | 2.3 ± 0.6 | 0.323 (2,10) | .732 |

Data are mean ± SD. TST = total sleep time; N1 = stage 1 sleep; N2 = stage 2 sleep; N3 = stage 3 sleep; REM = rapid eye movement sleep; SOL = sleep onset latency; WASO = wake after sleep onset; SE = sleep efficiency.
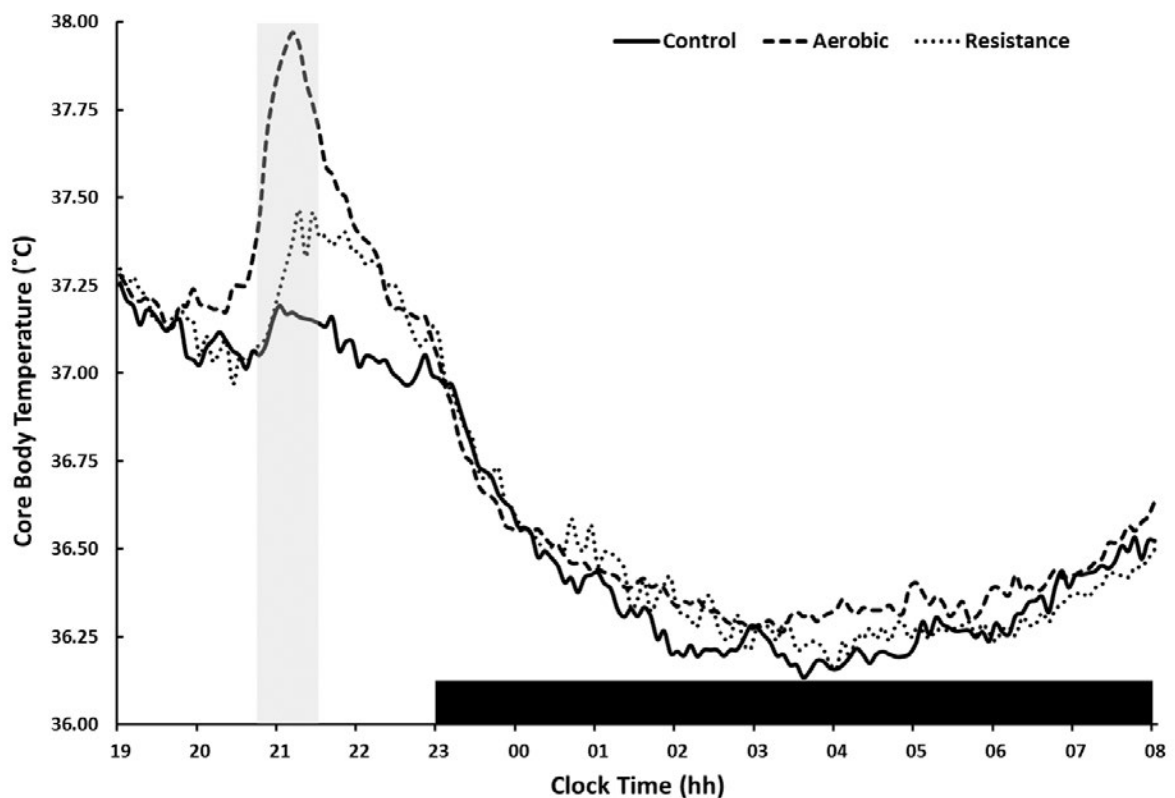
Figure 6.2 Core body temperature. Grey shading: 20:45h – 21:30h (exercise sessions for aerobic and resistance conditions), black horizontal bar : 23:00h – 08:00h (sleep opportunity).

## 6.6 Discussion

This is the first study to examine the effect of evening exercise modality on core body temperature and sleep. A repeated-measures, counterbalanced, crossover design was utilised to examine the impact of three conditions (control, evening aerobic exercise, evening resistance exercise) on subsequent night time sleep. The main finding was that moderate-intensity aerobic exercise or moderate-intensity resistance exercise completed 90 minutes before bedtime does not impair the sleep of healthy young males. Exercise did not affect pre-sleep alertness, pre-sleep physical tension, or subjective sleep quality.

These findings are consistent with previous data indicating that evening exercise may not be detrimental to subsequent sleep (Alley, 2015; Dworak et al., 2008; Myllymaki et al., 2011; O'Connor et al., 1998). For example, O'Connor et al. (1998) found no difference in sleep onset,

total sleep time or sleep efficiency in healthy young adults following a 60-min bout of either low intensity or moderate-intensity intensity cycling completed 90 minutes before bedtime. Alley et al. (2015) reported that healthy young adults experience shorter sleep onset latency and less wake after sleep onset following an evening bout of moderate-intensity resistance exercise performed at 19:00h when compared to no exercise. While these findings suggest that evening exercise may either not influence, or potentially improve, subsequent sleep, the mechanism through which evening exercise may affect sleep remains unclear. Increased core body temperature, increased energy expenditure, and skeletal muscle damage are potential mechanisms whereby exercise may influence sleep (Driver and Taylor, 2000). In the present study, subjective whole-body fatigue was significantly higher during resistance exercise compared to control and aerobic exercise conditions. Additionally, no differences were found for next morning muscle soreness and core body temperature at lights out (23:00h) across conditions. Muscle soreness following exercise is typically experienced 24-48h post exercise; therefore, it is unlikely that muscle soreness influenced sleep in the present study. Given that there was a significantly higher whole-body fatigue in the resistance exercise condition with no impact on sleep, it can be suggested that whole-body fatigue was not the mechanism through which sleep was unchanged.

There is consensus that deviations from homeostatic core body temperature rhythms may impact the sleep of humans (Horne and Moore, 1985; Shapiro et al., 1989). A recent systematic review of the effect of evening exercise on sleep found core body temperature at bedtime was associated with a 3.76min increase in wake after sleep onset for every 0.1∘C increase in core body temperature (Stutz, Eiholzer, and Spengler, 2019). In the present study, core body temperature was significantly elevated during aerobic exercise and resistance exercise compared to the control condition. Furthermore, the increase in core body temperature during exercise was significantly higher in the aerobic condition compared to the resistance condition. Despite the elevated core body temperature during exercise, the 90-min period between the end of exercise and bedtime at 23:00h was sufficient for core body temperature to return to pre-exercise levels following both the aerobic and resistance exercise sessions (Figure 6.2). This decline in core body temperature is consistent with previous data showing no difference in core body temperature between evening exercise and

control conditions 90 min after exercise (Robey et al. 2013). As in the present study, Robey et al. (2013) found no differences in total sleep time, sleep architecture, sleep onset latency, wake after sleep onset and sleep efficiency between the exercise and control conditions. Therefore, a potential explanation for the similar sleep outcomes between conditions in the present study may be the return of core body temperature to pre-exercise levels in the 90 min prior to bedtime. It is unclear how sleep may be influenced if core body temperature remains elevated when attempting to initiate sleep. Therefore, future investigations should establish if the relationship between core body temperature and sleep could be influenced by exercised-induced elevation in core body temperature at bedtime.

Our data, with support of existing evidence (Alley, 2015; Dworak et al., 2008; Myllymaki et al., 2011; O'Connor et al., 1998; Robey et al., 2013), highlight the need to revise sleep hygiene recommendations regarding evening exercise (Zarcone, 1994). The present findings suggest that healthy young males can perform 30 min of either moderate-intensity aerobic exercise or moderate-intensity resistance exercise in the evening without compromising sleep, provided the exercise concludes 90 min before bedtime. Lack of time has been reported as a common barrier for completing daily exercise requirements (Cerin, Leslie, Sugiyama, and Owen, 2010; Yao, 2018) and with the advent of 24-hour gymnasiums, the opportunity to exercise in the evening is now readily available. Our findings may encourage individuals that do not have time during the day to exercise, and ordinarily choose not to exercise at night, to fulfil their daily exercise requirements.

The present study did not obtain energy expenditure data for the aerobic exercise session and the resistance exercise session. Additionally, RPE was not obtained in the control condition. These omissions could be seen as potential limitations due to a lack of objective data comparing the demands of condition. Generalisations of these data should not be applied to settings in which additional stimuli may impair sleep independently of exercise. For instance, the sleep of individuals training or participating in competitive sport in the evening may be subject to additional stressors (Lastella, Lovell, and Sargent, 2014; Miller et al., 2017; Sargent and Roach, 2016). It is also unknown how evening exercise modality may impact core body temperature and sleep when

performed at a higher exercise intensity, a longer exercise duration, or in different populations (i.e., untrained individuals, females, elite athletes, different environmental conditions).

Our data indicate that moderate intensity aerobic exercise and resistance exercise ceasing 90 min prior to bedtime does not impair the sleep of healthy young males. The elevation in core body temperature during aerobic exercise and resistance exercise had subsided by bedtime, such that any potential interference with sleep due to increased core body temperature was avoided. In future, it will be useful to examine the effect of evening exercise of different intensities and/or durations on core body temperature and sleep in a range of populations. This direction of research could be used to inform changes to current sleep hygiene recommendations regarding evening exercise.

## 6.7 References

Akerstedt, T., and Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *The International journal of neuroscience, 52*(1-2), 29-37.

Alley, J. M., Mazzochi, J.W., Smith, CJ., Morris, DM., and Collier, SR. (2015). Effects of Resistance Exercise Timing on Sleep Architecture and Nocturnal Blood Pressure. *Journal of strength and conditioning research 29*(5), 1378-1385.

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise, 14*(5), 377-381.

Buckworth, J., and Nigg, C. (2004). Physical activity, exercise, and sedentary behavior in college students. *The journal of american college health*, 53(1), 28-34.

Buman, M. P., Phillips, B. A., Youngstedt, S. D., Kline, C. E., and Hirshkowitz, M. (2014). Does nighttime exercise really disturb sleep? Results from the 2013 National Sleep Foundation Sleep in America Poll. *Sleep medicine, 15*(7), 755-761.

Cerin, E., Leslie, E., Sugiyama, T., and Owen, N. (2010). Perceived barriers to leisure-time physical activity in adults: an ecological perspective. *Journal of physical activity and health, 7*(4), 451-459.

Darwent, D., Zhou, X., van den Heuvel, C., Sargent, C., and Roach, G. D. (2011). The validity of temperature-sensitive ingestible capsules for measuring core body temperature in laboratory protocols. *Chronobiology international, 28*(8), 719-726.

Driver, H. S., and Taylor, S. R. (2000). Exercise and sleep. *Sleep medicine review, 4*(4), 387-402.

Dworak, M., Wiater, A., Alfer, D., Stephan, E., Hollmann, W., and Strüder, H. K. (2008). Increased slow wave sleep and reduced stage 2 sleep in children depending on exercise intensity. *Sleep medicine, 9*(3), 266-272.

Flausino, N. H., Da Silva Prado, J. M., de Queiroz, S. S., Tufik, S., and de Mello, M. T. (2012). Physical exercise performed before bedtime improves the sleep pattern of healthy young good sleepers. *Psychophysiology,* 49(2), 186-192.

Hasan, J., Urponen, H., Vuori, I., and Partinen, M. (1988). Exercise habits and sleep in a middle-aged Finnish population. *Acta physiologica scandinavica. supplementum, 574*, 33-35.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., . . . Bauman, A. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Medicine and science in sports and exercise, 39*(8), 1423-1434.

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., . . . Adams Hillard, P. J. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep health, 1*(1), 40-43.

Horne, J. A., and Moore, V. J. (1985). Sleep EEG effects of exercise with and without additional body cooling. *Electroencephalography and clinical neurophysiology, 60*(1), 33-38.

Horne, J. A., and Reid, A. J. (1985). Night-time sleep EEG changes following body heating in a warm bath. *Electroencephalography and clinical neurophysiology, 60*(2), 154-157.

Iber, C., Ancoli-Israel, S., Chesson, A., and Quan, S. (2007). The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. 2007. *American Academy of Sleep Medicine, Westchester, IL.*

Kingma, B. R., Frijns, A. J., Schellen, L., and van Marken Lichtenbelt, W. D. (2014). Beyond the classic thermoneutral zone: Including thermal comfort. *Temperature,* 1(2), 142-149.

Kosmadopoulos, A., Sargent, C., Darwent, D., Zhou, X., and Roach, G. D. (2014). Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavior research methods, 46*(4), 1032-1041.

Lack, L., and Lushington, K. (1996). The rhythms of human sleep propensity and core body temperature. *Journal of sleep research*, 5(1), 1-11.

Lastella, M., Lovell, G. P., and Sargent, C. (2014). Athletes' precompetitive sleep behaviour and its relationship with subsequent precompetitive mood and performance. *European journal of sport science, 14(Suppl 1)*, S123-130.

Livingstone, S. D., Grayson, J., Frim, J., Allen, C. L., and Limmer, R. E. (1983). Effect of cold exposure on various sites of core temperature measurements. *Journal of applied physiology : respiratory, environmental and exercise physiology, 54*(4), 1025-1031.

Miller, D. J., Sargent, C., Vincent, G. E., Roach, G. D., Halson, S. L., and Lastella, M. (2017). Sleep/Wake Behaviours in Elite Athletes from Three Different Football Codes. The journal of science and medicine in sport, 16(4), 604-605.

Murphy, P. J., and Campbell, S. S. (1997). Nighttime drop in body temperature: a physiological trigger for sleep onset? *Sleep, 20*(7), 505-511.

Myllymaki, T., Kyrolainen, H., Savolainen, K., Hokka, L., Jakonen, R., Juuti, T., . . . Rusko, H. (2011). Effects of vigorous late-night exercise on sleep quality and cardiac autonomic activity. *Journal of sleep research, 20*(1 Pt 2), 146-153.

O'Connor, P. J., Breus, M. J., and Youngstedt, S. D. (1998). Exercise-induced increase in core temperature does not disrupt a behavioral measure of sleep. *Physiology and behavior, 64*(3), 213-217.

Piercy, K. L., and Troiano, R. P. (2018). Physical Activity Guidelines for Americans From the US Department of Health and Human Services. *Circulation: cardiovascular quality and outcomes*, 11(11).

Pontifex, M. B., Hillman, C. H., Fernhall, B., Thompson, K. M., and Valentini, T. A. (2009). The effect of acute aerobic and resistance exercise on working memory. *Medicine and science in sports and exercise*, 41(4).

Roach, G. D. S., C.; Darwent, D.; Kennaway, D.J.; Ferguson, S.A. (2010). Lost in transit: The journey of ingestible temperature sensors through the human digestive tract. *Ergonomia IJE and HF 32*(1), 49-61.

Robey, E., Dawson, B., Halson, S., Gregson, W., King, S., Goodman, C., and Eastwood, P. (2013). Effect of evening postexercise cold water immersion on subsequent sleep. medicine and science in sports and exercise, 45(7), 1394-1402.

Sargent, C., and Roach, G. D. (2016). Sleep duration is reduced in elite athletes following night-time competition. *Chronobiology international*, 33(6):667-70.

Sawka, M. N. W., C.B; Pandolf, K.B. (2011). Thermoregulatory responses to acute exercise-heat stress and heat acclimation *Comprehensive Physiology*: American Physiological Society.

Shapiro, C. M., Allan, M., Driver, H., and Mitchell, D. (1989). Thermal load alters sleep. *Biological psychiatry, 26*(7), 736-740.

Stutz, J., Eiholzer, R., and Spengler, C. M. (2019). Effects of Evening Exercise on Sleep in Healthy Participants: A Systematic Review and Meta-Analysis. S*ports medicine*, 49(2), 269-287.

Tanaka, H., Monahan, K. D., and Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the american college of cardiology, 37*(1), 153-156.

Trumbo, P., Schlicker, S., Yates, A. A., Poos, M., Food, and Nutrition Board of the Institute of Medicine, T. N. A. (2002). Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids. *The Journal of the american dietetic association*, 102(11), 1621-1630.

Van Dongen, H. P. A., and Dinges, D. F. (2000). Circadian rhythms in fatigue, alertness, and performance. *Principles and practice of sleep medicine*, 20, 391-399.

Vincent, G. E., Jay, S. M., Sargent, C., Vandelanotte, C., Ridgers, N. D., and Ferguson, S. A. (2017). Improving Cardiometabolic Health with Diet, Physical Activity, and Breaking Up Sitting: What about Sleep?, *Frontiers in physiology*, 8(865).

Vuori, I., Urponen, H., Hasan, J., and Partinen, M. (1988). Epidemiology of exercise effects on sleep. *Acta physiologica scandinavica. supplementum, 574*, 3-7.

Warburton, D. E., Nicol, C. W., and Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Canadian medical association journal, 174*(6), 801-809.

Wewers, M. E., and Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in nursing and health, 13*(4), 227-236.

Yao, C. J. Bassner, M. (2018). Healthy behaviors competing for time: associations of sleep and exercise in working Americans. *Sleep health, 5(1):23-30.*

Zarcone, V. P. (1994). Sleep hygiene *Principles and Practice of Sleep Medicine* (Vol. 2, pp. 542-546). Philadelphia: W.B. Saunders Co.

**7. Implementing a circadian adaptation schedule after eastward flight in young male athletes.**

**Abstract**

This study examined the effectiveness of a circadian adaptation schedule in male cricketers after an 8.5-h eastward time zone change. Ten participants (aged 18.7 yr) were randomly assigned to a control group or an intervention group. Participants in the intervention group followed a light exposure schedule in which they were instructed to seek light in the three hours preceding, and avoid light in the three hours following their estimated core body temperature minimum. The rate of adaptation was assessed using the nightly excretion rate of urinary 6-sulphatoxymelatonin (aMT6s). General linear mixed models were conducted to assess the effect of condition (i.e., control and light intervention) on nocturnal secretion of aMT6s. As reflected by the increase in nocturnal melatonin excretion, all participants gradually adapted to the destination time zone. On average, it took 7 days for all participants to return to baseline levels following transmeridian travel. Similarly, it took 7 days for subjective jet lag to alleviate. In the initial 4 days of the protocol, the intervention group registered non-significantly higher levels of nocturnal urinary melatonin, however there was no significant differences in the rate of adaptation between the groups. It is possible that participants did not adhere to the intervention or that they followed the intervention but it was ineffective.

**7.1 Introduction**

Jet lag is a by-product of transmeridian travel and is characterised by difficulty maintaining night-time sleep and feelings of daytime sleepiness and fatigue (Waterhouse et al. 2002). The primary mechanism responsible for jet lag is the misalignment between the body's endogenous circadian system and the local destination time. The effects of jet lag gradually subside as the circadian system aligns with environmental time cues (e.g., sunlight; Eastman and Burgess 2005).

Appropriately scheduled light exposure can be an effective strategy to enhance circadian adaptation (via phase delay or phase advance) to destination time zones following travel (Roach and Sargent 2019). The human circadian rhythms of melatonin and core body temperature are strongly linked with the sleep/wake cycle (Cagnacci et al. 1996; Lewy et al. 1999). Endogenous melatonin secretion typically begins 2 h prior to habitual bedtime (Burgess et al. 2003), and the daily minimum of core body temperature (CBTmin) coincides with the low point of the circadian cycle (Cagnacci et al. 1996). The resetting of the endogenous circadian clock is most sensitive to retinal light exposure in the hours before and after CBTmin (Wever et al. 1983; Khalsa et al. 2003). Light exposure before CBTmin results in a phase delay (i.e., shift later in time) and light exposure after CBTmin results in a phase advance (i.e., shift earlier in time; Khalsa et al. 2003). The implementations of light exposure protocols to prevent or reduce jet lag are well established under laboratory conditions (Burgess et al., 2003; Czeisler et al., 1989; Eastman et al., 2005), but the effectiveness of these protocols in the field—and in particular with elite athletes—have not been thoroughly examined.

Athletes that compete and train internationally and are required to cross multiple time zones, which makes them susceptible to stressors such as jet lag (Fowler et al., 2015; Leatherwood and Dragoo, 2013; Waterhouse, Reilly, and Atkinson, 2000). In addition, athletes may be required compete and/or train very soon after arrival - potentially before they are fully adapted to the new time zone. In such situations, strategies that facilitate adaptation would be useful. Therefore, the aim of this study was to examine the effectiveness of light exposure/avoidance schedule to facilitate circadian

adaptation in young male cricket players using the night-time urinary melatonin production as a measure of circadian adaptation.

**7.2 Methods**

This study employed a randomised, between-groups design. All participants travelled from the United Kingdom to Australia (8.5-h eastward time zone change) and were randomly assigned to a control group or an intervention group. Participants were provided with a self-report diary to record sleep/wake behaviour (i.e., bedtime and getup time) and subjective ratings of sleepiness, sleep quality and jet lag. The production rate of melatonin was inferred from concentrations of the metabolite aMT6s for eight consecutive nights following arrival.

*7.2.1 Participants*

Ten young male cricket players (mean ± SD; age: 18.7 ± 0.9 yr) participated in the study. Participants were provided with an information sheet detailing the benefits and risks of participation as well as their right to withdraw from the study at any stage. Written informed consent was obtained from participants and ethical approval was obtained from the Central Queensland University Human Research Ethics Committee (HREC: 21129).

*7.2.2 Measures*

*Subjective variables*

Self-report diaries were used to collect the following variables:

- Bedtime and get up time: the self-reported clock time a participant went to bed to attempt to sleep and the clock time at which a participant stopped attempting to sleep.
- Subjective sleepiness: assessed using the Karolinska Sleepiness Scale, where 1 = "extremely alert" and 9 = "very sleepy, great effort to keep alert, fighting sleep" (Akerstedt and Gillberg 1990).
- Subjective sleep quality: assessed using a 5-point scale, where 1 = "very poor", 2 = "poor",3 = "average", 4 = "good" and 5 = "very good".

- Subjective jet lag: assessed using 7-point scale, where 0 = "none", 1 = "extremely low", 2 = "very low", 3 = "low", 4 = "moderate", 5 = "high", 6 = "very high" and 7 = "extremely high".

*Urinary aMT6s*

Production of melatonin during all sleep periods was inferred from urinary 6-sulphatoxymelatonin (aMT6s) concentrations.  Urine samples were frozen and subsequently assayed for the concentration of aMT6s (aMT6s human radioimmunoassay, Stockgrand, Guildford, UK). The concentration of aMT6s was divided by the volume of the overnight sample and multiplied by the length of the sleep period for the corresponding night to acquire an excretion rate (i.e., ng/hour). The baseline excretion rate for each participant was then subtracted from the excretion rate of each sample to acquire an excretion rate relative to baseline.

*7.2.3 Procedures*

Upon arrival to Australia, participants were randomly assigned to a control group or an intervention group. A researcher met with each participant and provided them with study materials (i.e., urine collection container and self-report diaries; Appendix C). Immediately prior to each main sleep period, participants voided their urine (i.e., pre-sleep void) and provided ratings of subjective jet lag and sleepiness. Participants were housed in hotel rooms or private service apartments for the duration of the study. Participants then collected all subsequent urine passed during the sleep period in a 2-L collection container, including a final void upon waking (i.e., post-sleep void). Participants recorded the clock times for the pre- and post-sleep void using labels on the urine container. Each morning, participants recorded their bedtime, get up time and rated their subjective sleep quality. A researcher collected the urine containers from participants each morning and transferred a 1-ml aliquot of urine into a 1.5-ml tube (prepared with 10mg of boric acid as a preservative; concentration = 250g of boric acid per litre of water).

Participants in the intervention group were given a paper schedule of when to seek or avoid light during the eight days following arrival. After an 8.5-h eastward time zone change, the human

circadian system can adapt via phase advance or phase delay (Roach and Sargent 2019). A phase delay schedule was chosen as it was more conducive to the timing of daylight hours for the destination time zone. The estimation of CBTmin was formulated under the assumption that participants' nocturnal sleep period was typically from 23:00h to 07:00h and therefore CBTmin will occur at approximately 04:00h (Roach and Sargent 2019). The light exposure schedule advised participants to seek light in the three hours preceding their estimated CBTmin (arrival CBTmin = 04:00h GMT; 14:30h ACST) and to avoid light in the three hours following their estimated CBTmin. Light blocking sunglasses were provided for periods allocated to avoiding light. The estimated CBTmin for each participant was adjusted 1.5h later each experimental day to accommodate for a phase delay (i.e., day 1 CBTmin = 14:30h ACST; day 2 CBTmin = 16:00h ACST; day 3 CBTmin = 18:30h; etc…; Figure 7.1). At 12:00h on each day, electronic reminders (via SMS) of when to seek/avoid light were sent to participants in the intervention group. Baseline measures of overnight urine, bedtime, get up time and subjective ratings could not be collected prior to travel. Instead, baseline measures were collected approximately 30 days after arrival over a 3-day period.

### 7.2.4 Data Analysis

Data were analysed using a General Linear Mixed Model with the R package lme4 (R Core Team 2016). Separate models were tested for each dependant variable (i.e., aMT6s relative to baseline, subjective jet lag, subjective sleepiness, bedtime and get up time). A random intercept for participants was included to account for intraindividual dependencies and interindividual heterogeneity. An α-level of 0.05 was used to indicate statistical significance.
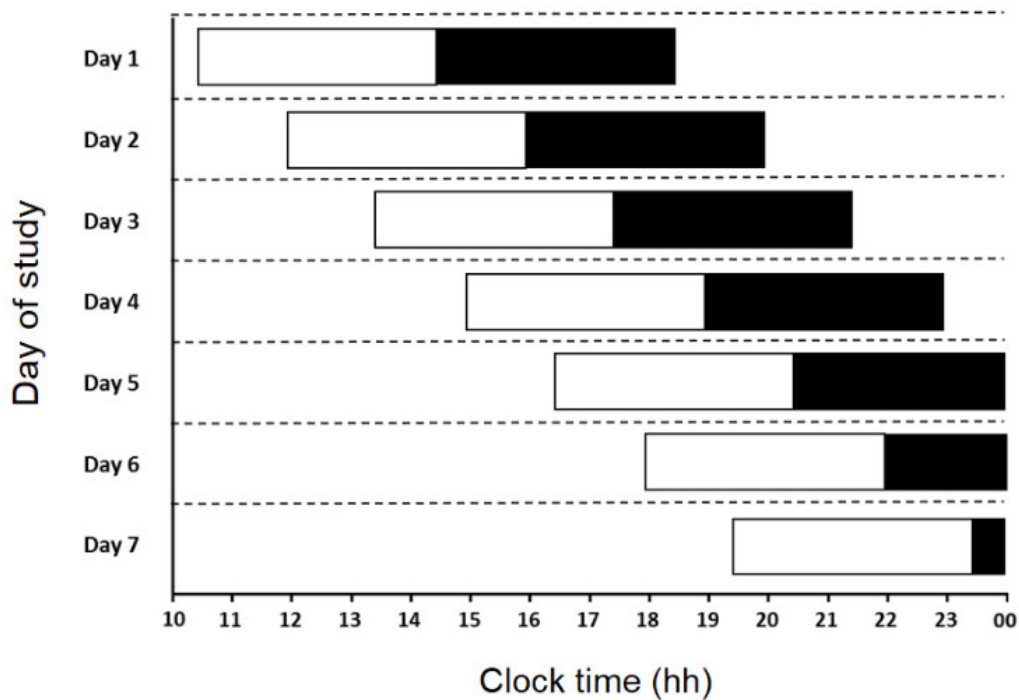
**Figure 7.1** Light exposure protocol. White bars = period to seek bright light, black bars = period to avoid light.

## 7.3 Results

To ensure that all urine was collected during main sleep periods, samples were weighed and compared to average urine production in healthy subjects (Rosenberg, Fiserova-Bergerova et al., 1989). Samples were excluded from the analysis if they were considered insufficient or if an error was reported by a participant (i.e., accidental void of urine). Out of 80 samples, 21 were excluded from analyses (control = 11, intervention = 10). Thus, 59 samples were included in analyses. There was a main effect of day on aMT6S (decreased by day; $F_{(7,35)}=10.4$, $p<0.001$) subjective jet lag (decreased by day; $F_{(7,54)}=22.9$, $p<0.001$), bedtime (later by day; $F_{(7,54)}=3.1$, $p=0.007$) and get up time (earlier by day; $F_{(7,35)}=5.4$, $p<0.001$). There was no main effect of condition or an interaction between condition and day for any variables (Figure 7.2).
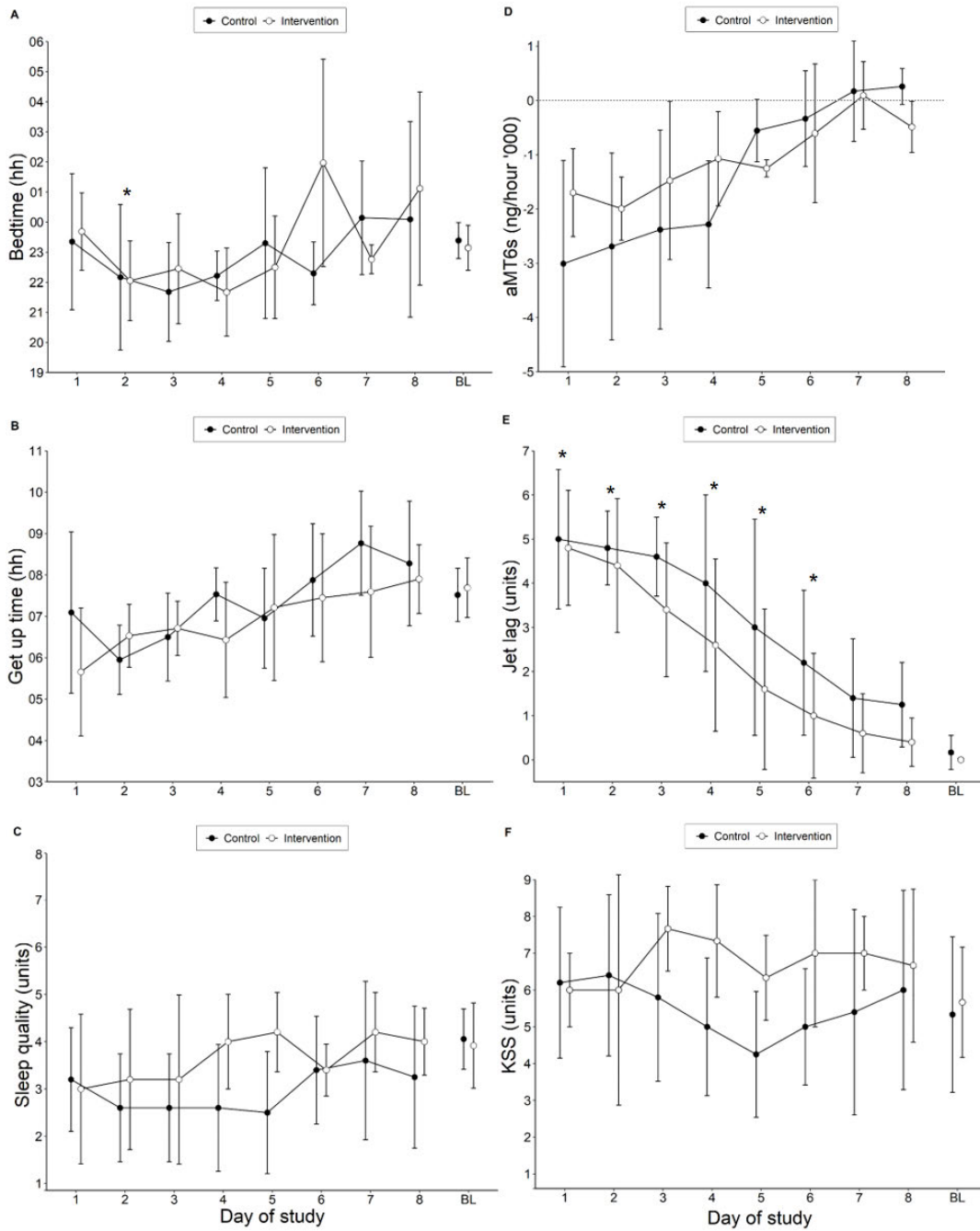
**Figure 7.2** Subjective and objective variables as a function of day: self-reported bedtime (A); self-reported get up time (B); subjective sleep quality (C); urinary aMT6s relative to baseline (D); subjective jet lag (E); subjective sleepiness (F). Data are mean ± SD and have been offset to aid interpretation. Filled circles represent the control condition. Open circles represent the intervention condition. Asterisks indicate significant differences to baseline across both conditions. Dotted line in figure D represents baseline aMT6s. KSS; Karolinska Sleepiness Scale, BL; Baseline.

**7.4 Discussion**

The aim of this study was to examine the effectiveness of a light exposure/avoidance schedule to facilitate circadian adaptation following eastward travel. A randomised, between groups design was utilised to examine the difference between two conditions (control, light exposure). The main finding was that there was no difference in urinary melatonin excretion between conditions. It took participants 7 days to adapt to the 8.5-h eastward time zone change, as demonstrated by the gradual reduction of subjective jet lag and the gradual increase in the excretion rate of aMT6s. However, the light exposure schedule did not facilitate faster adaptation in the intervention group.

A number of review articles have provided detailed guidelines regarding strategies to reduce or potentially eliminate jet lag using light or other phase shifting interventions (i.e., exogenous melatonin) when travelling (Eastman and Burgess, 2009; Roach and Sargent, 2019). Indeed, experimental studies have attempted to implement bright light interventions in the field but have relied on subjective responses or have used salivary dim light melatonin onset (DLMO) as a measure of jet lag (Boulos et al., 2002; Thompson et al., 2013). DLMO can provide an accurate biomarker for circadian phase, however, implementation of the protocol can be time consuming and may be impractical for field studies (Burgess, Wyatt, Park, and Fogg, 2015). While the circadian principles of bright light interventions are well understood, the practicality of implementing them successfully, and acquiring an objective measure of circadian phase in field settings is not. In isolation of the experimental protocol, this is the first study to utilise urinary melatonin as a daily bio marker of the circadian phase of athletes in a field setting. Demonstrated by the return to baseline of night-time melatonin in both conditions, the measures utilised in this study can be used as a valid circadian biomarker for future field studies involving athletes.

For many sports teams or athletes, it is impractical to travel with a staff member solely responsible for formulating and enforcing light exposure schedules to facilitate circadian adaptation. A potential solution is to formulate a pre-planned schedule around training and other activities that will optimise adaptation to the destination time zone. This study indicates that

providing a hard copy adaptation schedule and daily electronic reminders did not enhance adaptation to an 8.5-h eastward time zone change in young cricket players.

There are several potential explanations for why the light exposure schedule did not enhance circadian adaptation. It is possible that participants in the intervention group did not comply with the instructions and reminders to seek/avoid light. Assuming that the instructions and reminders elicited the recommended behaviour in the intervention group, it is also possible that incidental light exposure resulted in a similar rate of adaptation in the control group. On average, the intervention group was closer to baseline measures of aMT6s when compared to the control group between days 1-4 of the protocol (Figure 7.2). However, from days 5-8, the control group was, on average, closer to baseline measures of aMT6s when compared to the intervention group.

The findings of this study should be interpreted under the boundary conditions of the experimental design. This includes the degree of time zone change (i.e., 8.5-h eastward), the direction of adaptation (i.e., phase delay), sample size and characteristics of the participants (i.e., group of 10 young male cricket players), and the method of implementing the adaptation schedule (i.e., generalised hard copy schedule and electronic reminders). Additionally, it is possible that the sample in this study may be underpowered. The present study did not acquire a measure of compliance for the light exposure schedule and did not acquire athlete performance outcomes. The athletes in this study were not given information regarding the expected symptoms of jet lag or the principles underlying circadian misalignment. This was done to reduce the likelihood of the intervention group influencing the behaviour of the control group, given their proximity during the data collection period. Therefore, it is unclear whether the subjective jet lag scale may be more, or less effective among athletes that receive jet lag education. It is reasonable to suggest, that if the intervention group was aware of the chronobiological principle used to allocate times for light exposure, the light exposure protocol may have been more successful. Future research should address these factors and examine alternative strategies for implementing light exposure/avoidance schedules with consideration to the direction of travel and degree of time-zone change and athlete chronotype.

## 7.5 References

Akerstedt, T., and Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *The international journal of neuroscience, 52*(1-2), 29-37.

Burgess, H. J., Crowley, S. J., Gazda, C. J., Fogg, L. F., and Eastman, C. I. (2003). Preflight adjustment to eastward travel: 3 days of advancing sleep with and without morning bright light. *Journal of biological rhythms*, 18(4), 318-328.

Cagnacci, A., Soldani, R., Laughlin, G. A., and Yen, S. S. (1996). Modification of circadian body temperature rhythm during the luteal menstrual phase: role of melatonin. *Journal of applied physiology*, 80(1), 25-29.

Eastman, C.I., Gazda, C.J., Burgess, H. J., Crowley, S. J., and Fogg, L. F. (2005). Advancing circadian rhythms before eastward flight: a strategy to prevent or reduce jet lag. Sleep, 28(1), 33.

Khalsa, S. B., Jewett, M. E., Cajochen, C., and Czeisler, C. A. (2003). A phase response curve to single bright light pulses in human subjects. The journal of physiology, 549(Pt 3), 945–952.

Leatherwood, W. E., and Dragoo, J. L. (2013). Effect of airline travel on performance: a review of the literature. British journal of sports medicine, 47(9), 561–567.

Lewy, A. J., Cutler, N. L., and Sack, R. L. (1999). The Endogenous Melatonin Profile as a Marker for Circadian Phase Position. *Journal of biological rhythms,* 14(3), 227-236

Roach, G. D., and Sargent, C. (2019). Interventions to Minimize Jet Lag After Westward and Eastward Flight. *Frontiers in physiology*, 10(927).

Rosenberg J, Fiserova-Bergerova V, Lowry LK. (1989). Horizons: Biological Monitoring IV: Measurements in Urine. *Applied industrial hygiene*. 4:F-16-F-21.

Waterhouse, J., Edwards, B., Nevill, A., Carvalho, S., Atkinson, G., Buckley, P., Reilly, T., Godfrey, R., and Ramsay, R. (2002). Identifying some determinants of "jet lag" and its symptoms: a study of athletes and other travellers. *British journal of sports medicine*, 36(1), 54–60.

Wever, R. A., Polášek, J., and Wildgruber, C. M. (1983). Bright light affects human circadian rhythms. *Pflügers archiv european journal of physiology*, 396(1), 85-87.

# 8. The effectiveness of a formulated sleep drink on sleep quality and quantity

Published peer-reviewed publications associated with this Chapter (Appendix D):

Halson, S. L., Shaw, G., Versey, N., Miller, D. J., Sargent, C., Roach, G. D., Nyman, L., Carter, J. M., and Baar, K. (2020). Optimisation and Validation of a Nutritional Intervention to Enhance Sleep Quality and Quantity. *Nutrients*, 12(9), 2579.

**Abstract**

Dietary supplements such as carbohydrate, α-lactalbumin and tart cherry juice are known stimulate the neurotransmitters associated with sleep. For this reason, dietary supplementation may be an effective alternative to sleep medication for improving sleep in athletes - but the combined impact and optimal dose of ingredients has not been empirically tested. The aim of this study was to examine the effectiveness of a formulated sleep drink on sleep quality and quantity. Eighteen healthy males participated in the study (age: 25.1 ± 6.2 years). A double-blinded, repeated measures, experimental design was used to compare the effect of three conditions (i.e., most optimal sleep drink, least optimal sleep drink and placebo) on sleep. Participants spent four consecutive nights in CQUniversity's Appleton institute sleep laboratory. Sleep was measured using gold-standard polysomnography and participants were given 9-hour sleep opportunities (23:00h-08:00h) on each experimental night. On each morning, a cognitive test battery was performed to assess subjective and objective measures of performance (e.g., reaction time). There was a main effect of condition on sleep onset latency (p=0.02). Sleep onset latency was lower in the optimal condition (9.9 min) compared to the placebo (19.6 min) and least optimal conditions (26.1 min), but there was no difference in total sleep time or sleep stage duration. There were no differences in next day cognitive performance between conditions. The findings of this study suggest that the concentration of ingredients in the optimal sleep drink may be used to promote the onset of sleep without impacting next-day performance.

**8.1 Introduction**

In recent years, there have been consistent findings suggesting that athletes do not obtain the recommended amount of sleep (Leeder, Glaister et al., 2012; Sargent, Halson et al., 2014). There are several potential mechanisms through which athletes' sleep may be negatively affected (e.g., travel, training commitments, competition, etc). To combat potential sleep loss, athletes may turn to pharmaceutical aids to obtain adequate sleep (Waterhouse, Reilly et al., 2004). Medications such as benzodiazepines, zolpidem and zaleplon reduce the time it takes to fall asleep; however, they are subject to next-day residual side-effects such as drowsiness and decreased psychomotor performance (Waterhouse, Reilly et al., 2004). The residual side-effects of sleep medications are particularly pertinent given that 13% of 176 elite athletes use "sleeping pills" as a strategy to sleep well across the night preceding competition (Juliff, Halson et al., 2015). While there are no studies showing a link between the use of sleep medication and impairments in next-day physical performance, sleep scientists and physicians typically discourage use of sleep medications among athletes.

There is some evidence to indicate that dietary supplements, such as carbohydrate (CHO), α-lactalbumin, tart cherry juice, valerian, L-theanine and nucleotides, can improve sleep outcomes (e.g., reduced sleep onset latency; Halson, 2014; St-Onge, Mikic et al., 2016; Kim, Lee et al., 2018; Schneider, Mutungi et al., 2018; Doherty, Madigan et al., 2019). Consumption of these dietary supplements can influence neurotransmitters that are known to impact sleep (Halson, 2014). For example, the release of melatonin via the synthesis of serotonin is subject to the availability of the amino acid tryptophan. Ingestion of certain foods (e.g., carbohydrate) can result in increased uptake of tryptophan to the brain– therefore increasing melatonin release (Halson, 2014). While there is evidence that some dietary supplements promote sleep, the combined impact and optimal dose of such ingredients have not been empirically tested. Therefore, the aim of this study is to examine the effect of a formulated sleep drink on sleep quality and quantity in healthy young males.

**8.2 Methods**

This study was the second phase of a larger pre-existing project attempting to formulate and optimise a nutritional dietary supplement to enhance sleep. In the first phase of the study, the optimal concentrations of nutritional ingredients were determined using a six-factor Box–Behnken model (Halson et al., 2020). The model was used to predict the optimal and least optimal concentrations for inducing physiological predictors of sleepiness (Halson et al., 2020).

*8.2.1 Participants*

Eighteen healthy males participated in the study (age: 25.1 ± 6.2 years). These participants participated in phase 2 of the project exclusively. None of the participants had a history of a sleep disorder and none were taking sleep medication during testing. Participants were provided with an information sheet detailing the benefits and risks of participation as well as their right to withdraw from the study at any stage. Written informed consent was obtained directly from participants and the projects was approved by the Central Queensland University Human Research Ethics Committee (HREC:H17/09-167).

*8.2.2 Design*

A double-blinded, repeated measures, experimental design was employed to compare the effect of three conditions (i.e., placebo drink, least optimal sleep drink, optimal sleep drink) on sleep and next-day cognitive performance. The protocol was conducted over a 4-day period in a randomised, counterbalanced order (Figure 8.1).

*8.2.3 Laboratory setting*

The study was conducted in a purpose-built accommodation suite at the Appleton Institute of Behavioural Science. The suite accommodates six people at a time, each with their own bedroom, bathroom and living room. Meals were served in a common dining room area. During all wake periods, ambient light intensity was maintained at 350 lux (6-foot angle of gaze) and was

extinguished during sleep opportunities (<.03 lux). The laboratory was maintained at target temperature between 21-23˚C.

*8.3.4 Nutritional Ingredients*

In phase 1 of the project, the combinations and concentrations of nutritional ingredients were tested for inducing physiological predictors of sleepiness (Table 8.1; Halson et al., 2020). The most optimal and least optimal concentrations and combinations of ingredients are described in Tables 8.2 and 8.3 respectively (Halson et al., 2020).

**Table 8.1** Nutritional ingredients included in Stage 1 of the project (Halson et al., 2020).

| Ingredient | Doses | Manufacturer |
|---|---|---|
| Tart Cherry Juice | 0ml, 50ml, 100ml | Cherry Active Australia |
| High GI carbohydrate | 0g, 25g, 50g | PolyJoule, Australia |
| α-lactalbumin | 0g, 20g, 40g | Davisco Foods, USA |
| Adenosine-5-monophosphate | 0mcg, 26.5mcg, 53mcg | Sigma, USA |
| Valerian | 0mg, 750mg, 1500mg | Martin Baeur Group, Germany |
| Theanine | 0mg, 500mg, 1000mg | Sun Theanine, Japan |

Notes: ml = millilitres; g = grams; mg = milligrams; mcg = micrograms.

**Table 8.2** Optimal concentration and combination ingredients.

| Ingredient | Dose |
|---|---|
| α-lactalbumin | 40g |
| Theanine | 655mg |
| Adenosine-5-monophosphate | 53mg/100ml |
| Valerian | 600mg |

Notes: ml = millilitres; g = grams; mg = milligrams.

**Table 8.3** Least optimal concentration and combination ingredients.

| Ingredient | Dose |
|---|---|
| Tart Cherry Juice | 35ml |
| High GI carbohydrate | 45g |
| α-lactalbumin | 8g |
| Theanine | 100mg |
| Adenosine-5-monophosphate | 4.5mg/100ml |
| Valerian | 500mg |

Notes: ml = millilitres; g = grams; mg = milligrams.

*8.3.5 Measures*

*Polysomnography (PSG)*

Sleep was measured using gold-standard PSG. A standard montage of electrodes was attached to the face and scalp of participants, including three electroencephalographs (i.e., C4-M1, F4-M1, O2-M1), two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram. PSG data were recorded directly to data acquisition, storage, and analysis systems (Grael, Compumedics; Victoria, Australia). PSG records were manually scored in 30-s epochs by an experienced registered polysomnographic technician in compliance with standard criteria (Iber, Ancoli-Israel, Chesson, and Quan, 2007). The sleep variables calculated from each PSG recording included:

- Sleep onset latency (min): the duration of time from lights out to the first epoch of any stage of sleep;

- Total sleep time (min): the sum of minutes spent in any stage of sleep (Stages 1, 2, 3, REM);

- Sleep architecture: time in minutes spent in Stages 1, 2, 3, REM;

- Wake after sleep onset (min): the duration of time spent awake between sleep onset and lights on;

- Sleep efficiency (%): total sleep time divided by time in bed, multiplied by 100.

*Subjective scales*

- Subjective sleep quality: 7-point scale, where 1 = "extremely poor", 2 = "very poor", 3 = "poor", 4 = "average", 5 = "good", 6 = "very good" and 7 = "extremely good".

- Subjective sleep quantity and subjective sleep onset latency: participants were asked "how much sleep do you think you got?" and "how long did it take you to fall asleep?".

- Subjective sleepiness: Karolinska Sleepiness Scale (Akerstedt and Gillberg, 1990), where 1 = "extremely alert", 2 = "very alert", 3 = "alert", 4 = "fairly alert", 5 = "neither alert nor sleepy", 6 = "some signs of sleepiness", 7 = "sleepy, but no effort to keep alert", 8 = "sleepy, some effort to keep alert", 9 = "very sleepy, great effort to keep alert, fighting sleep".

- Subjective alertness: Participants rated their current level of alertness by placing a vertical mark on a non-numeric, 100-mm line between two opposing statements (not at all alert – very alert), providing a score between 0 and 100.

- Self-perceived capacity: Participants rated their ability to perform as fast as possible by placing a vertical mark on a non-numeric, 100-mm line between two opposing statements (not fast at all – very fast), providing a score out of 100. Participants rated their ability to perform as accurately as possible by placing a vertical mark on a non-numeric, 100-mm line between two opposing statements (not accurately at all – very accurately), providing a score between 0 and 100.

*Spatial configuration task*

The spatial configuration task utilises the visual search paradigm to measure selective attention (Horowitz et al., 2003). The task is conducted on a desktop computer. Participants are required to identify whether a target (i.e., the number 2) is present amongst distractors (i.e., the number 5). Participants respond by pressing the "yes" key if the target is present on the screen or by pressing the "no" key if the target is not present. The dependent variables were the number of correct responses, mean reaction time and the number of errors made.

*Psychomotor vigilance task*

The psychomotor vigilance task is a widely used reaction time task to measure neurobehavioral performance. The task was conducted using a portable, hand-held unit that contains a display and a response button (PVT-192, Ambulatory Monitoring Inc., USA). Participants were instructed to respond to each visual stimulus as quickly as possible. The dependent variables were the number of lapses (i.e., response times >500 ms), number of false starts (response before stimuli) and mean reaction time (ms). \

*Postural sway test*

Postural sway was assessed using an Accusway computerised force platform (AMTI, Watertown, MA) in conjunction with Swaywin software (AMTI, Watertown, MA). The force platform measures the three-dimensional forces (Fx, Fy, Fz) and the three-dimensional moments (Mx, My, Mz) involved in balance. These provide centre of pressure (COP) coordinates, which allow postural sway to be calculated in square centimetres. Participants performed two postural balance tasks each for 30 seconds – standing still with both feet on the platform with eyes open and standing still with both feet on the platform with eyes closed.
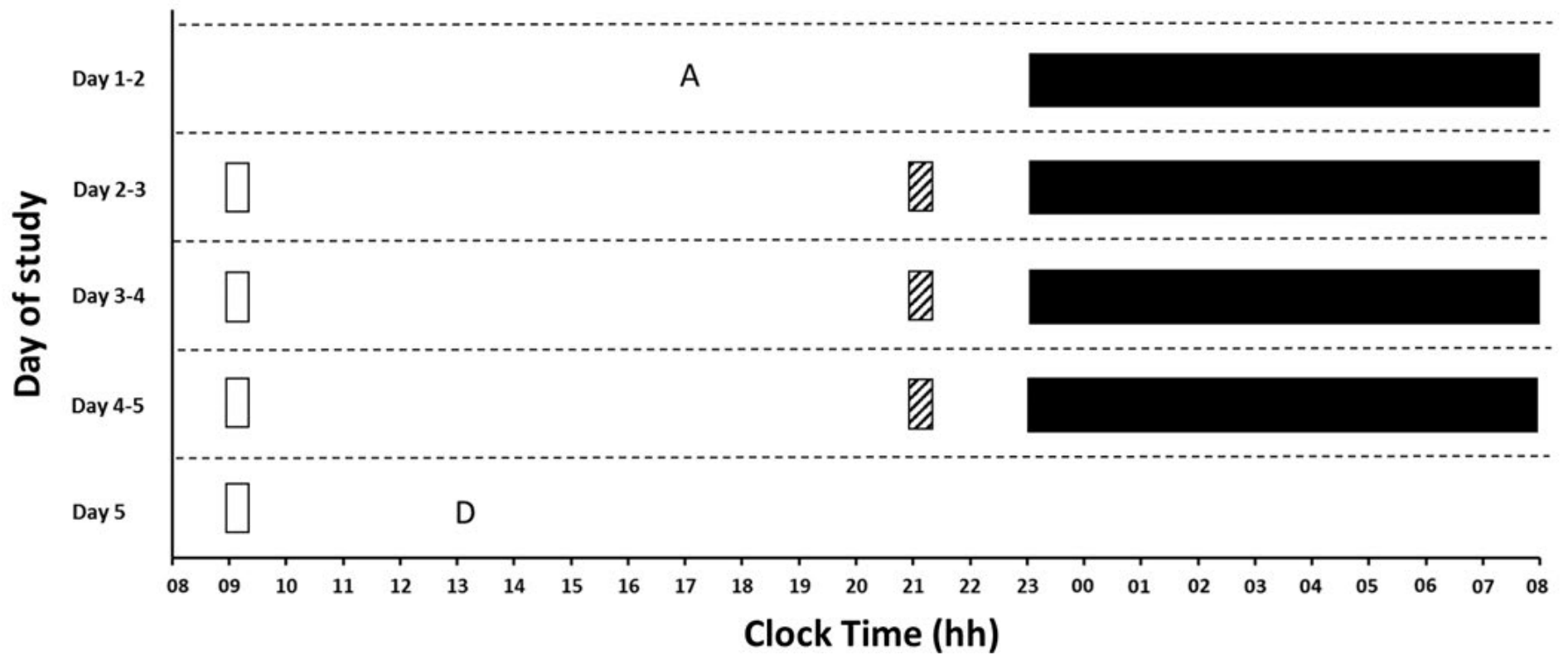
**Figure 8.1** Illustration of the study design. "A" = participant arrival, white bars = test battery, lined bars = experimental conditions (i.e. placebo, least optimal drink, optimal drink), black bars = sleep opportunity, "D" = participant departure.

*8.3.6 Procedures*

Three separate groups of six participants completed the study over 4 consecutive nights. The arrival night was used as an adaptation night to familiarise participants with the equipment used to monitor sleep. The next three nights were experimental nights on which participants received one of three interventions in a randomised, counterbalanced order: 1) placebo, 2) least optimal drink, and 3) optimal drink. Participants were given a 300-ml bottle containing each drink at 21:00h and were required to consume it within 5 min. Participant were not permitted to consume any water from 21:00h onwards. While in the laboratory, participants were provided with four meals (breakfast, lunch, snack, and dinner) consisting of a similar macronutrient profile to a Western diet (26% protein, 20% fat, 54% carbohydrate). Water was provided *ad libitum*. Subjective sleepiness was assessed half hourly between 20:00h and 22:00h. In all conditions, electrodes were applied in the 90 minutes preceding lights out – with participant remaining in bed with the lights turned off between 23:00h and 08:00h.

A standardised test battery was performed at 09:00h on days 2-5. These test batteries consisted of seven tasks performed in the following order: subjective sleepiness; subjective alertness; self-perceived capacity; spatial configuration task, psychomotor vigilance task; and postural sway test. To simulate minimum daily activity (e.g., walking to public transport), participants were taken on two 10-min walks at normal walking pace (14:30h and 16:30h on days 2-5). During free time, participants performed sedentary activities such as reading, studying, or watching movies.

*8.3.7 Data analysis*

Two separate General Linear Mixed Models were fitted using the R package lme4 (R Core Team, 2016). The first model assessed the impact of condition (fixed effect) on all sleep variables. A random intercept for participants was included to account for intraindividual dependencies and interindividual heterogeneity. The second model assessed the impact of condition (i.e., placebo,

least optimal, optimal) on test battery variables. Again, a random intercept for participants was included to account for intraindividual dependencies and interindividual heterogeneity. An $\alpha$-level of 0.05 was used to indicate statistical significance. Data are presented as mean $\pm$ standard deviation.

## 8.4 Results

There was a main effect of condition on sleep onset latency ($F(2,34)=4.29; p=0.02$). Sleep onset latency was significantly lower in the optimal condition compared to the placebo and least optimal condition (Figure 8.2). There were no significant differences between conditions for any of the other sleep variables or test battery variables. Figure 8.3 illustrates that participants who had the longest sleep onset latency (i.e., 23–140 min) on the placebo night experienced a greater reduction in sleep onset latency following consumption of the optimal drink compared to participants who had average (7–15 min) or short (2–5 min) sleep onset latencies on the placebo night.

**Table 8.4** Objective and subjective sleep variables as a function of condition.

| Sleep Variables | Condition | | |
| --- | --- | --- | --- |
| | Placebo | Least optimal | Most Optimal |
| TST (min) | 507.2 ± 60.5 | 503.3 ± 40.8 | 519.5 ± 42.2 |
| WASO (min) | 43.2 ± 47.3 | 40.6 ± 26.8 | 40.6 ± 36.1 |
| SE (%) | 88.9 ± 10.6 | 88.3 ± 7.1 | 91.1 ± 7.4 |
| SOL (min) | 19.6 ± 32.0 | 26.1 ± 37.4 | 9.9 ± 12.3[a] |
| REM latency (min) | 90.3 ± 69.7 | 75.7 ± 25.6 | 82.7 ± 33.3 |
| Stage 3 latency (min) | 14.8 ± 5.9 | 14.9 ± 6.8 | 18.5 ± 13.3 |
| Stage 1 (min) | 30.8 ± 11.6 | 29.1 ± 15.2 | 32.7 ± 12.8 |
| Stage 2 (min) | 204.8 ± 42.5 | 201.9 ± 43.8 | 223.7 ± 44.9 |
| Stage 3 (min) | 139.7 ± 49.5 | 143.2 ± 52.2 | 135.4 ± 39.7 |
| REM sleep (min) | 131.9 ± 26.8 | 129.2 ± 208 | 127.7 ± 29.3 |
| Arousals - total (count) | 97.3 ± 25.8 | 98.5 ± 29.1 | 106.1 ± 32.9 |
| Arousals - REM sleep (count) | 25.0 ± 11.1 | 24.2 ± 7.7 | 23.0 ± 13.2 |
| Arousals - NREM sleep (count) | 72.3 ± 28.7 | 74.3 ± 32.6 | 83.1 ± 36.1 |
| Awakenings (count) | 24.7 ± 6.6 | 26.1 ± 7.9 | 26.2 ± 8.3 |
| Stage Shifts (count) | 173.4 ± 25.7 | 178.2 ± 46.5 | 188.2 ± 38.0 |
| KSS 2000h (units) | 4.3 ± 1.2 | 4.6 ± 0.9 | 4.8 ± 0.9 |
| KSS 2030h (units) | 4.8 ± 1.2 | 4.9 ± 1.0 | 4.9 ± 1.0 |
| KSS 2100h (units) | 5.3 ± 1.4 | 5.0 ± 1.1 | 5.3 ± 1.1 |
| KSS 2130h (units) | 5.9 ± 1.1 | 5.3 ± 1.2 | 5.6 ± 1.0 |
| KSS 2200h (units) | 6.1 ± 1.1 | 5.8 ± 1.1 | 6 ± 1.1 |
| Subjective sleep quality (units) | 4.9 ± 1.1 | 4.6 ± 1.0 | 4.7 ± 0.9 |
| Subjective sleep quantity (h) | 8.0 ± 1.1 | 7.9 ± 1.4 | 8.1 ± 1.0 |
| Subjective SOL (min) | 18.8 ± 14.0 | 22.9 ± 17.8 | 15.8 ± 9.7 |

Data are mean ± standard deviation. [a] indicates a significant difference compared to placebo and least optimal conditions (p <0.05). TST; total sleep time, WASO; wake after sleep onset, SOL; sleep onset latency, REM; rapid eye movement, NREM; non-rapid eye movement KSS; Karolinska Sleepiness Scale.

**Table 8.5** Test battery results as a function of condition.

| Variables | Condition | | |
|---|---|---|---|
| | **Placebo** | **Least optimal** | **Most optimal** |
| *Subjective variables* | | | |
| KSS (units) | 4.3 ± 1.3 | 4.6 ± 1.3 | 4.3 ± 1.1 |
| VAS alertness (units) | 60.7 ± 20.3 | 57.4 ± 18.6 | 63.1 ± 18.0 |
| VAS speed (units) | 65.9 ± 17.4 | 63.7 ± 14.9 | 65.0 ± 16.1 |
| VAS accuracy (units) | 66.3 ± 19.4 | 66.6 ± 17.7 | 67.4 ± 16.9 |
| *PVT* | | | |
| Reaction time (ms) | 260.5 ± 41.0 | 268.8 ± 49.6 | 262.1 ± 39.2 |
| Lapses (count) | 1.6 ± 2.8 | 1.6 ± 3.1 | 1.4 ± 2.3 |
| False starts (count) | 0.8 ±1.3 | 0.8 ± 1.0 | 0.5 ± 1.2 |
| *Spatial Configuration Task* | | | |
| Correct responses (%) | 97.8 ± 2.4 | 97.5 ± 2.8 | 97.5 ± 2.2 |
| Reaction time (ms) | 1.2 ± 1.2 | 1.3 ± 1.4 | 1.3 ± 1.1 |
| Errors (count) | 3.5 ± 1.2 | 3.4 ± 1.1 | 3.8 ± 1.6 |
| *Balance task* | | | |
| Postural sway ($cm^2$) | 0.4 ± 0.2 | 0.4 ± 0.2 | 0.4 ± 0.4 |

Data are mean ± standard deviation. KSS; Karolinska Sleepiness Scale, VAS; visual analogue scale,

PVT; psychomotor vigilance task.
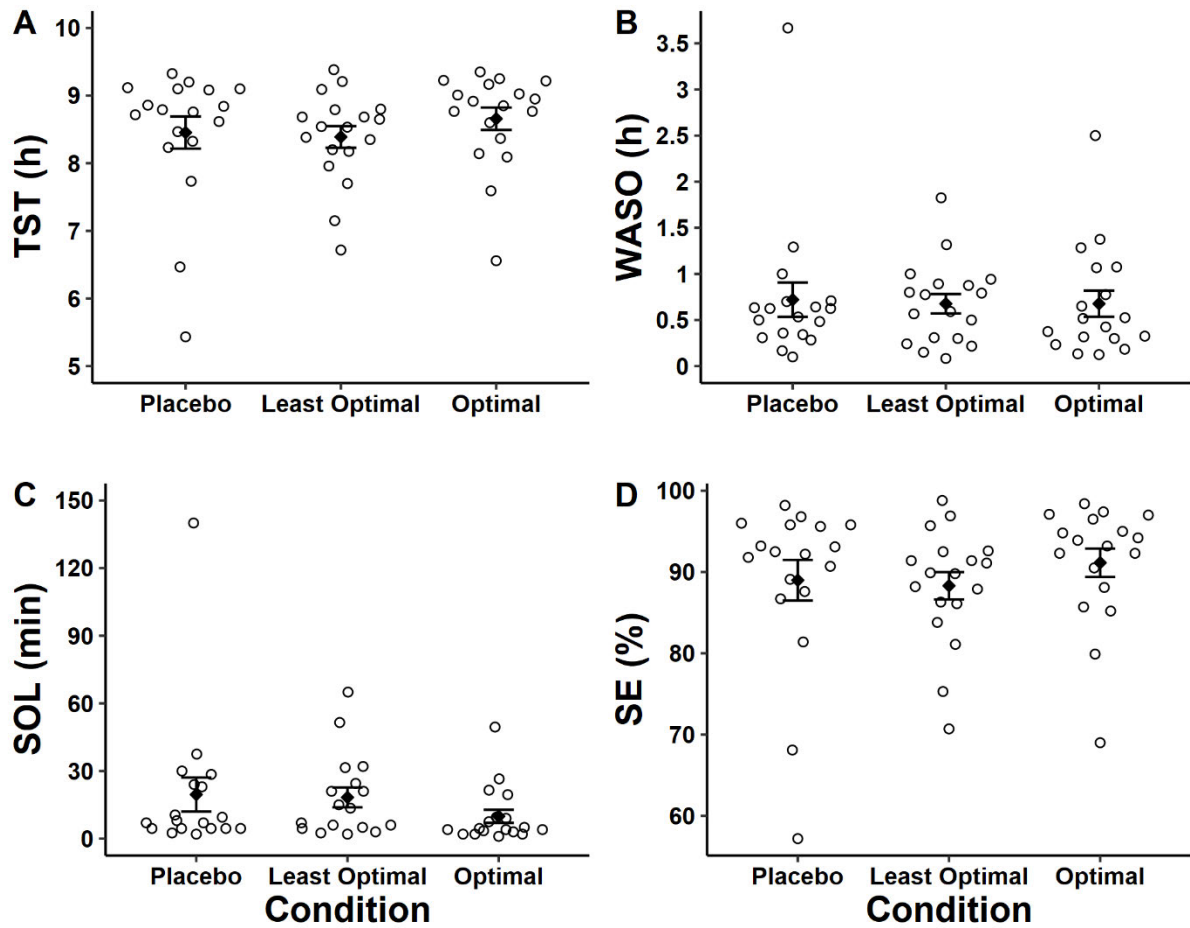
**Figure 8.2** Individual data points for sleep variables: total sleep time (A; TST), wake after sleep onset (B; WASO), sleep onset latency (C; SOL) and sleep efficiency (D; SE) as a function of condition. Individual un-filled circles represent a participant; filled circles represent the mean; error bars represent standard deviation.
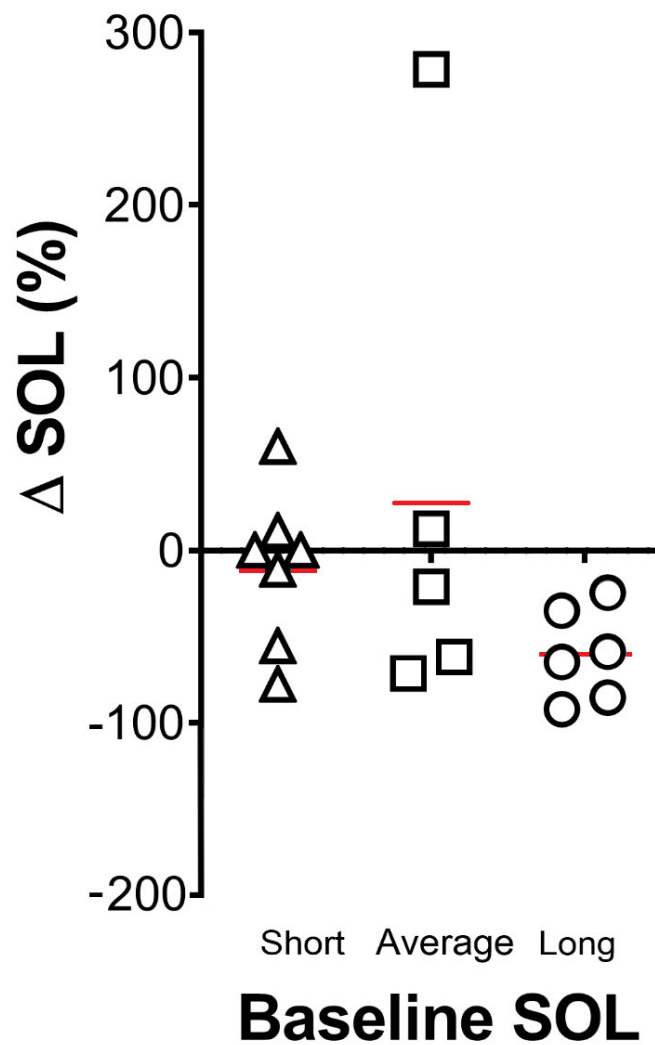
**Figure 8.3** Baseline sleep onset latency (SOL) and the magnitude of the effect of the drink. Short SOL = 2-5 min; Average SOL = 7-15 min; Long SOL = 23-140 min. Red horizontal lines represent group means; triangles represent individual means in the short SOL group; squares represent individual means in the average SOL group; circles represent individual means in the long SOL group.

**8.5 Discussion**

The aim of this study was to examine the impact of a formulated sleep drink on the sleep of healthy young individuals. The main finding was that sleep onset latency was significantly lower in the optimal drink condition compared with the placebo and least optimal drink conditions. Following consumption of the optimal drink, participants fell asleep in under 10 minutes.

The potential sleep promoting effects of the included ingredients have previously been examined in isolation. The mechanism by which such dietary supplements may enhance sleep is by increasing availability of tryptophan – a dietary pre-cursor of serotonin within the brain. The optimal concentration and dosages of the ingredients included in this study were modelled to optimise levels of blood tryptophan. Ingestion of ingredients in the current study likely increased uptake of L-tryptophan to the brain– therefore increasing melatonin release (Halson, 2014). A previous study has shown that consumption of tryptophan-rich protein resulted in less wake during sleep and increased subjective sleep quality (Hudson et al., 2005). In addition to tryptophan-rich supplements, α-lactalbumin and tart cherry juice have been shown to result in favourable sleep outcomes (Markus, Jonkman et al., 2005; Howatson, Bell et al., 2012). The present study is the first to examine the combined effect of these ingredients on sleep.

The findings of this study suggest that the optimal sleep drink may be used to promote the onset of sleep but will not have an impact on sleep architecture (i.e., sleep staging). On average, participants that exhibited a longer sleep onset latency following the placebo drink tended to show the greatest reduction in sleep onset latency following consumption of the optimal drink (Figure 8.3). Therefore, the optimal drink may be more effective for individuals experiencing difficulty initiating sleep. Importantly, the optimal sleep drink does not appear to have adverse effects on next-day cognitive performance. This could be considered the most practically applicable outcome in the context of elite sport. Pharmaceutical sleep medications may have an addictive effect and result in daytime sleepiness and/or next day fatigue (Waterhouse, Reilly et al., 2004). In contrast, the sleep drink examined in the present study did not result in any changes to next-day cognitive ability. Therefore, athletes seeking a

non-pharmaceutical means for reducing sleep onset latency may utilise this combination of ingredients, without potentially compromising next day cognitive performance.

A number of boundary conditions should be considered when interpreting results of this study. Firstly, the participants were healthy young males with no sleep disturbances (Stewart, 2021). It is unclear how the formulated doses of ingredients may impact the sleep of individuals with sleep disorders or other health conditions. Future research should examine the effects of this optimised sleep drink in females, poor sleepers, or in individuals who have difficulty initiating sleep onset.

# References

Akerstedt, T., and Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. The *International journal of neuroscience*, 52(1-2), 29-37.

Doherty, R., Madigan, S., Warrington, G., and Ellis, J. (2019). Sleep and Nutrition Interactions: Implications for Athletes. *Nutrients*, 11(4), 822.

Halson, S.L. (2014). Sleep in elite athletes and nutritional interventions to enhance sleep. *Sports medicine*, 44(1), 13-23.

Halson, S. L., Shaw, G., Versey, N., Miller, D. J., Sargent, C., Roach, G. D., Nyman, L., Carter, J. M., and Baar, K. (2020). Optimisation and Validation of a Nutritional Intervention to Enhance Sleep Quality and Quantity. *Nutrients*, 12(9), 2579.

Horowitz, T. S., Cade, B. E., Wolfe, J. M., and Czeisler, C. A. (2003). Searching night and day: a dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychological science*, 14(6), 549–557.

Howatson, G., Bell, P. G., Tallent, J., Middleton, B., McHugh, M. P., and Ellis, J. (2012). Effect of tart cherry juice (Prunus cerasus) on melatonin levels and enhanced sleep quality. *European journal of nutrition*, 51(8), 909–916.

Hudson, C., Hudson, S. P., Hecht, T., and MacKenzie, J. (2005). Protein source tryptophan versus pharmaceutical grade tryptophan as an efficacious treatment for chronic insomnia. *Nutritional neuroscience*, 8(2), 121–127.

Juliff L.E., Halson S.L., Peiffer J.J. Understanding sleep disturbance in athletes prior to important competitions. *Journal of science and medicine in sport*. 2015 Jan;18(1):13-18.

Kim, J., Lee, S. L., Kang, I., Song, Y. A., Ma, J., Hong, Y. S., Park, S., Moon, S. I., Kim, S., Jeong, S., and Kim, J. E. (2018). Natural Products from Single Plants as Sleep Aids: A Systematic Review. *Journal of medicinal food.* 21(5), 433–444.

Leeder, J., Glaister, M., Pizzoferro, K., Dawson, J., and Pedlar, C. (2012). Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *Journal of sports sciences*, 30(6), 541–545.

Markus, C. R., Jonkman, L. M., Lammers, J. H., Deutz, N. E., Messer, M. H., and Rigtering, N. (2005). Evening intake of alpha-lactalbumin increases plasma tryptophan availability and improves morning alertness and brain measures of attention. *The American journal of clinical nutrition*, 81(5), 1026–1033.

Miller, D. J., Sargent, C., Roach, G. D., Scanlan, A. T., Vincent, G. E., & Lastella, M. (2020). Moderate-intensity exercise performed in the evening does not impair sleep in healthy males. *European journal of sport science*, 20(1), 80–89.

Sargent C, Halson S, Roach G.D. (2014). Sleep or swim? Early-morning training severely restricts the amount of sleep obtained by elite swimmers. *European journal of sport science*. 14(Suppl 1:S310-315).

Schneider, N., Mutungi, G., and Cubero, J. (2018). Diet and nutrients in the modulation of infant sleep: A review of the literature. *Nutritional neuroscience*, 21(3), 151–161.

St-Onge, M. P., Mikic, A., and Pietrolungo, C. E. (2016). Effects of Diet on Sleep Quality. *Advances in nutrition*, 7(5), 938–949.

Waterhouse, J., Reilly, T., and Edwards, B. (2004). The stress of travel. *Journal of sports sciences*, 22(10), 946-966.

# 9. General Discussion

The studies conducted as part of this thesis aimed to examine the influence of factors pertaining to training, competition, and travel on the amount and quality of sleep obtained by athletes. Taken together, the findings have important implications for the sleep of athletes. Indeed, a valid objective measurement of sleep is needed to examine any reciprocal relationship between sleep and factors that may impact sleep. By showing that commercially available devices can provide accurate sleep/wake detection, the findings from Chapters 3 and 4 provide valuable insights into any context in which sleep is being estimated outside of the laboratory. For example, the finding that sleep/wake auto-detection is comparable to manual manipulation provides a safeguard for participant non-compliance for providing sleep/wake times. This problem was encountered in Chapter 7, where low compliance for actigraphy-based measurement of sleep resulted in the exclusion of data. It is reasonable to suggest that these data may not have been excluded if a wearable device capable of providing automatic detection of sleep was utilised. To provide practical applications to athletes and coaches, each Chapter in this dissertation address several factors that may influence or be influenced by sleep.

## 9.1 Summary of studies

The studies conducted in Chapters 3 and 4 examined the effectiveness of a commercially available wearable device (i.e., the WHOOP strap) to estimate sleep. Data comparisons were conducted against PSG, the gold standard of sleep measurement. The results from Chapter 3 indicate that the WHOOP strap is a viable alternative for measuring 2-stage sleep (i.e., sleep or wake) when sleep times were manually adjusted into the WHOOP smart phone application. In addition, the WHOOP strap was found to be comparable to other sleep wearables in estimating 4-stage sleep (i.e., wake, light sleep, SWS, REM). In Chapter 4, the same gold-standard comparisons were made between the WHOOP strap and PSG. However, data were also collected using the WHOOP strap's sleep auto detection function as well as research grade actigraphy. The results indicated that the WHOOP strap, when automatically detecting sleep, has a similar sensitivity and specificity to actigraphy for 2-stage categorisation of sleep and can be used as a practical alternative to PSG for 2-stage categorisation of sleep and 4-stage categorisation of sleep.

The findings presented in Chapters 3 and 4 have meaningful practical implications for athletes looking to monitor their sleep. Research grade actigraphy devices require certain expertise to utilise and are usually implemented over extended periods without immediate feedback (e.g., 2 weeks). In comparison, wearable devices such as the WHOOP strap provide easily accessible data within minutes of a sleep period ending. Furthermore, the findings of Chapter 4 show that the WHOOP strap performs well when automatically detecting sleep. This means that acquiring an accurate measurement of 2-stage sleep does not rely upon the user (e.g., athlete) to manually enter the times in which they go to bed and wake up – alleviating the risk of error and/or non-compliance. In addition, Chapter 4 highlights the way in which the WHOOP strap can be utilised depending on the variable of interest. If an athlete is having difficulty falling asleep, utilising the WHOOP strap under manual adjustment functionality will provide a timestamp in which to begin the sleep opportunity, thus providing a more accurate measure of sleep onset latency. If the variable of interest is total sleep time, automatic detection allows for comparable accuracy and less data input compared to manual adjustment.

The study in Chapter 7 examined the effectiveness of a light exposure schedule for enhancing circadian adaptation following transmeridian travel. A group of travelling athletes were assigned to a control group that received no recommendations, or an intervention group who were given detailed instructions on when to seek/avoid light to enhance entrainment to the destination time zone (i.e., alleviate jet lag). The findings of this study suggest that the intervention was ineffective in enhancing the rate at which the athletes entrained to the destination time zone. There were no statistical differences between groups. Another important finding of this study is that it took 7-days for young male athletes to adapt to the +8.5-h destination time zone. Given that no records of compliance were collected for this study, it is also possible that the athletes did not adhere to the recommended light schedules. It is also possible that participants in the control group were exposed to incidental light during critical times of the day – despite being advised not to. In isolation of the intervention, this study is the first to collect a biomarker of circadian phase in the field with athletes. Future investigations aiming to optimise circadian adaptation schedules can utilise the methods used in this study to collect urinary aMT6s.

The study in Chapter 6 aimed to validate a single question to assess subjective jet lag in athletes. Commonly utilised subjective measures of jet lag are primarily based upon subjective perceptions of jet lag and not circadian biomarkers. This study was the first to compare a subjective measure of jet lag against a circadian biomarker (i.e., aMT6s). Daily urine samples from twenty-seven young athletes were collected and assayed for the concentration of the urinary metabolite of melatonin – aMT6s. On each day, participants completed a single question jet lag scale – providing a subjective measure of their experience of jet lag. There was a moderate correlation between the single question jet lag scale such that subjective jet lag decreased as urinary aMT6s increased. This finding is consistent with the entrainment of the circadian system to destination time zones following travel (i.e., increasing melatonin concentration during nocturnal sleep). A single question jet lag scale may be used by athletes or coaches as an estimation of jet lag.

The study in Chapter 7 examined the impact of moderate-intensity aerobic exercise and moderate intensity resistance exercise performed in the evening on the sleep of healthy young males. The findings indicated that evening exercise did not impact the sleep of participants. Core body temperature increased in both exercise conditions, with aerobic exercise resulting in significantly higher core body temperature than moderate intensity exercise. However, in the 90-minute period between exercise and bedtime, core body temperature returned to pre-exercise levels. It is unclear whether sleep may have been impacted if core body temperature remained elevated when participants began attempting to sleep. The findings of this study suggest that young healthy males may perform moderate-intensity exercise in the evening without disrupting sleep – if there is a 90-minute period between exercise cessation and bedtime. Future research should examine the impact of higher intensity exercise on core body temperature and ascertain whether exercise-induced elevated core body temperature in the 30-60 min prior to sleep onset increased or reduces sleep duration.

The study in Chapter 8 examined the impact of a formulated sleep drink on the sleep of healthy young males. Compared to a placebo drink, and the least optimal drink, the optimal sleep drink significantly reduced sleep onset latency. On average, participants that presented with shorter sleep

onset latency tended to show the greatest reduction in sleep onset latency in the optimal drink condition. Importantly, there was no effect of condition on next day cognitive performance. Therefore, athletes seeking a non-pharmaceutical means for reducing sleep onset latency may utilise the combination of ingredients used in this study, without potentially compromising next-day performance.

## 9.2 Limitations

### 9.2.1 Study sample

Apart from Chapters 3 and 4, the sample demographic that were examined in this dissertation was predominantly healthy young males. Furthermore, due to recruitment constraints, some of the studies were not conducted in a professional athlete population. This approach may restrict the generalisability of the findings across different populations. For sleep, there are age-related differences in sleep duration, sleep architecture, and timing of sleep onset and offset (Ohayon et al., 2004). Humans tend to acquire less sleep, have more wake in sleep, obtain less REM and SWS as they age (Figure 9.1; Ohayon et al., 2004). Age and sex differences are also present for measures of the human body clock (e.g., chronotype; Fischer et al., 2017). Adolescents and younger individuals tend to have a later chronotype (i.e., go to bed later) and therefore have later circadian rhythms (e.g., core body temperature). Therefore, it is possible that the studies with experimental components (i.e., Chapters 5, 6, 7 and 8) may yield differing outcomes across different age cohorts.
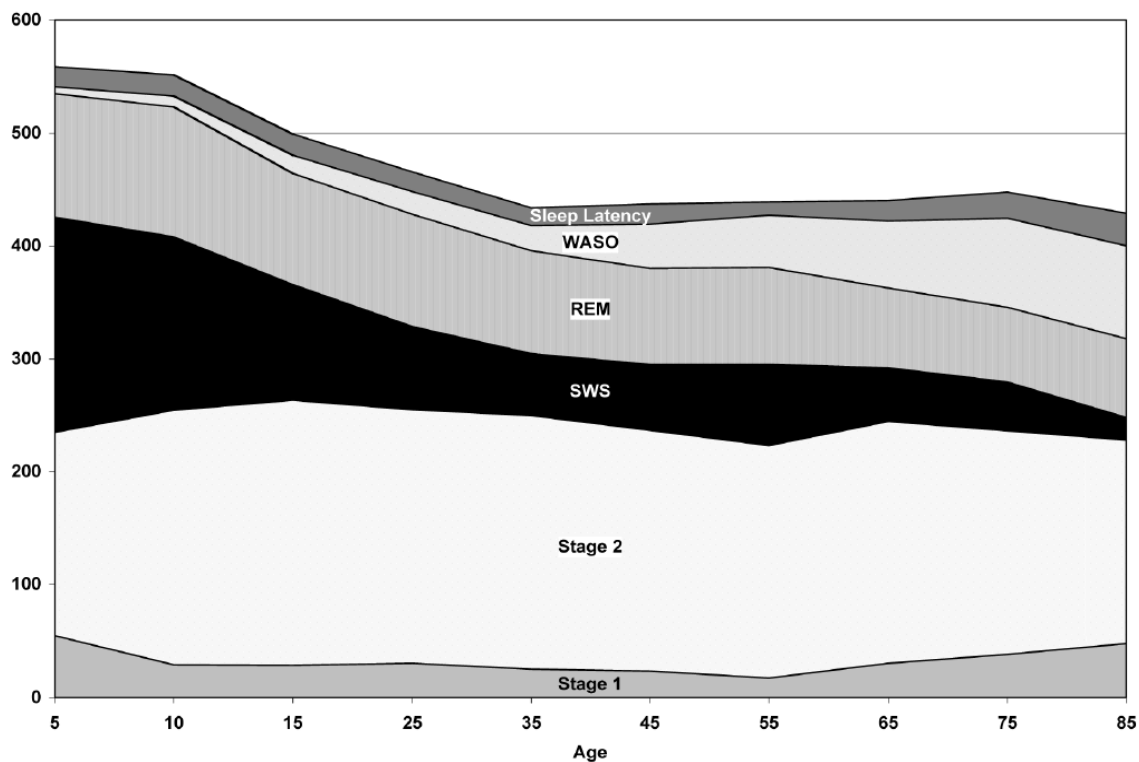
**Figure 9.1** Age-related sleep trends (Ohayon et al., 2004).

Regarding sex, it is important to consider the potential impact of the menstrual cycle on sleep related outcomes. In typical ovulatory menstrual cycles, there are cyclical changes in hormones, as well as core body temperature (Baker and Driver, 2007). Specifically, core body temperature rises by approximately 0.4°C during the 15–16-day luteal phase of menstruation (de Mouzon, Testart, Lefevre, Pouly, and Frydman, 1984; Halbrecht, 1945; Marshall, 1963). Previous research has shown that the increase in core body temperature associated with the luteal phase of menstruation typically reduces the proportion of REM sleep obtained (Baker, Driver, Paiker, Rogers, and Mitchell, 2002; Baker, Driver, Rogers, Paiker, and Mitchell, 1999; Baker et al., 2001; Driver, Dijk, Werth, Biedermann, and Borbély, 1996; Lee, McEnany, and Zaffke, 2000; Parry et al., 1999). Based on these sex differences, the findings from this dissertation may not be generalisable to females. For example, Chapter 5 showed that core body temperature among healthy young males returned to baseline levels within 90 minutes from evening exercise cessation. It is possible that core body temperature remains elevated after evening exercise (i.e., longer than 90 minutes) in females who are in the luteal phase of menstruation. Such sex considerations are not unique to this dissertation, with a high percentage of research conducted among

male cohorts. Consequently, future research should focus on examining measurement tools, and factors impacting sleep in female cohorts.

### 9.2.3 Experimental setting

A major strength of this dissertation is the use of gold-standard PSG in the laboratory-based studies. To validate the wearable device in Chapters 3 and 4, it was essential to assess sleep in a laboratory environment using PSG. For Chapters 6 and 8, a laboratory environment was utilised to implement randomised, repeated measures experiments. PSG is the only means of obtaining a valid measure of sleep onset latency and sleep staging – both of which were major outcome variables for Chapters 6 and 8. In addition to this, the laboratory environment was essential to ensure that the dependent variable was only affected by the experimental manipulation (i.e., the independent variable) rather than other factors (e.g., uncontrolled sleeping environment). This was particularly important for Chapter 6 where conducting the exercise sessions in a controlled environment ensured that participants were being exposed to the same intensity, duration, and timing of exercise 6).

While there are often logistical difficulties conducting field-based studies, it was essential to collect "real-world" data for Chapters 5 and 7. The greatest logistical hurdle for these studies was ensuring the athletes were compliant with the data collection and intervention protocols. As indicated by the data reflecting circadian adaptation, the collection and analysis of urine was a major strength of both studies. For Chapter 7, the primary logistical challenge was ensuring compliance to the intervention protocol. With limited access to the athletes, paper copies of the light exposure schedule were used, along with daily electronic reminders to seek/avoid light. Unfortunately, objective sleep estimations and reliable measures of protocol compliance were not collected. These are both limitations of Chapter 7 that should be accounted for in future studies by utilising automatic measurements of sleep and light exposure.

## 9.3 Concluding remarks

The volume of literature regarding the sleep of athletes has increased significantly in recent years (Lastella, Memon et al., 2020). The general consensus among researchers is that athletes do not obtain sufficient sleep (Leeder, Glaister et al., 2012; Sargent, Halson et al., 2014). However, many factors that impact sleep, and the countermeasures associated with them, have not been examined extensively. The findings from this dissertation have provided valuable insights regarding how sleep could be measured and managed using wearable technology.

Three central themes emerged from the results presented in this dissertation: (1) validation; (2) practical methodology; and (3) sleep onset latency. Chapters 3, 4, and 5 address the validity of approaches to measure sleep and jet lag in athletes. Such approaches are of value given that PSG and some estimations of the human body clock are often impractical to implement in the field. The outcome of the two validation studies examining the WHOOP strap to estimate sleep indicate that athletes and practitioners can utilise the WHOOP strap to estimate the sleep when PSG is impractical. The validation of a single question jet lag scale is the first to be compared to a circadian biomarker (aMT6s) for estimating jet lag in athletes. The outcome of this study may be used by travelling athletes and practitioners to estimate jet lag following transmeridian travel.

The second theme – practical methodology – can be applied to Chapters 3-7. Chapters 3 and 4 highlight the different ways in which wearables, in this case the WHOOP strap, are able to collect sleep data (i.e., manual adjustment and auto-detection). The distinction between the methodologies used to estimate sleep are important – both from an interpretation and a practical application standpoint. Firstly, if an athlete or coach are interested in estimating sleep with minimal self-reported inputs required, they can do so knowing that sleep auto-detection provides accurate data. However, it is important to consider the main outcome variable that an athlete or coach is investigating. For example, if the aim of the athlete is to estimate total sleep time, then auto-detection or manual adjustment, depending on their preference, may be used. However, if the aim is to examine sleep onset latency, then it is best to utilise the manual adjustment method as it provides the wearable with a timestamp in which to start assessing the time taken to fall asleep. This distinction, which has been neglected in previous literature, is an important

outcome of this dissertation. Chapter 5 contributes to the theme of practical methodology given the ease of use of implementing a single question to assess jet lag following transmeridian travel. In comparison to other subjective jet lag questionnaires, the question utilised in this Chapter can be implemented quickly, without significant burden to athletes or coaches. The practical methodologies related to Chapter 6 and 7 relate more to future research than immediate practical applications in the field. Chapter 6 was the first study to investigate the impact of different evening exercise modalities on sleep, highlighting the complexity of the relationship between exercise and sleep. There are many combinations of exercise timing, intensity, duration, and modality that may impact sleep. To provide meaningful recommendations for evening exercise across different populations, future research should utilise similar repeated measures methodology to make direct comparisons, and control exercise and sleep-related factors across different populations. Chapter 7 highlighted that the collection of nocturnal urine to measure concentrations of aMT6s is an option for estimating circadian adaption in athletes. The experimental manipulations in this study did not have a significant effect on adaptation; however future studies can utilise the data collection methods to investigate strategies in a range of athletes following different time zone changes. An additional finding that can be of use to athletes or practitioners managing jet lag is that all the athletes had adapted to the destination time zone (+8.5 hours) 7 days following arrival. Therefore, in absence of implementing a proven adaptation enhancement protocol, planning of training or other activities following similar time zone changes (e.g., direction of travel) may be conducted under the assumption that athletes will adapt within this timeframe.

Chapters 6 and 8 both address the final theme of this dissertation – difficulty falling asleep (i.e., sleep onset latency). Chapter 6 examined sleep onset latency after exercise, a stimulus that has been discouraged in close proximity to attempting sleep (Zarcone, 1994). The findings suggested that the evening aerobic and resistance exercise did not differ to a control, no exercise condition. An important finding from this Chapter was related to thermoregulation – in that core body temperature returned to baseline resting levels prior to bedtime. This is an important finding for athletes who participate in evening training and/or competition. While additional data are needed, ensuring that core body

temperature has returned to baseline levels prior to attempting sleep could be a practice utilised avoid increased sleep onset latency following evening training or competition. It is however, important to consider the additional factors that may impact the sleep onset latency of athletes prior to, during and after competition (i.e., performance rumination, caffeine consumption, media commitments). Contrary to Chapter 6, Chapter 8 examined the effect a stimulus (i.e., a formulated sleep drink) to reduce sleep onset latency. The findings of this study suggest that the optimal sleep drink used in this study impacts those who need it most – subjects with longer sleep onset latency. Importantly, there were no reductions in next-day cognitive performance following the consumption of the optimal sleep drink. Athletes and coaches seeking a non-pharmacological means of reducing sleep onset latency may use this concentration of ingredients as a viable intervention. With these findings in mind, the practical applications of this dissertation can now loop back around to the use of wearables to measure sleep onset latency. Athletes and coaches, informed on how to implement such devices, can examine the effect of potential interventions, such as a sleep drink, on wearable-derived sleep onset latency.

Given that the research areas addressed in this dissertation are in their infancy, there is scope to utilise similar methods to expand our knowledge regarding the sleep of athletes. For measurement of sleep in the field, it is important to continue the validation of developing technology to estimate sleep. As hardware and proprietary algorithms develop, the efficacy of sleep staging via wearable technology may improve. Therefore, research investigating proper implementation and interpretations of sleep wearables is imperative for informing athletes and coaches. One of the biggest challenge for researchers working in the field is the gap between chronobiologically-sound interventions, and successful implementation of the interventions in target cohorts. The use of developing wearable technology (e.g., WHOOP) in conjunction with innovative methodologies to collect actionable data in the field (e.g., urinary melatonin as a circadian biomarker) will allow for considerable growth in this research area. This dissertation can be considered a springboard on which such research can build to actionable outcomes amongst the athlete population.

**9.4 References**

Baker, F. C., and Driver, H. S. (2007). Circadian rhythms, sleep, and the menstrual cycle. *Sleep medicine*, 8(6), 613–622.

Baker FC, Driver HS, Paiker J, Rogers GG, Mitchell D. (1985). Acetaminophen does not affect 24-h body temperature or sleep in the luteal phase of the menstrual cycle. *Journal of applied physiology*, 92(4):1684-91.

Baker, F. C., Driver, H. S., Rogers, G. G., Paiker, J., and Mitchell, D. (1999). High nocturnal body temperatures and disturbed sleep in women with primary dysmenorrhea. *The American journal of physiology*, 277(6).

Baker, F. C., Waner, J. I., Vieira, E. F., Taylor, S. R., Driver, H. S., and Mitchell, D. (2001). Sleep and 24 hour body temperatures: a comparison in young men, naturally cycling women and women taking hormonal contraceptives. *The journal of physiology*, 530(Pt 3), 565–574.

de Mouzon, J., Testart, J., Lefevre, B., Pouly, J. L., and Frydman, R. (1984). Time relationships between basal body temperature and ovulation or plasma progestins. *Fertility and sterility*, 41(2), 254–259.

Driver, H. S., Dijk, D. J., Werth, E., Biedermann, K., and Borbély, A. A. (1996). Sleep and the sleep electroencephalogram across the menstrual cycle in young healthy women. *The journal of clinical endocrinology and metabolism*, 81(2), 728–735.

Fischer, D., Lombardi, D. A., Marucci-Wellman, H., and Roenneberg, T. (2017). Chronotypes in the US - Influence of age and sex. *PloS one*, 12(6), e0178782.

Halbrecht, I. (1945). Ovarian function and body temperature. *Lancet*, 2(6380), 668.

Lastella, M., Memon, A.R., and Vincent, G.E. (2020). Global Research Output on Sleep Research in Athletes from 1966 to 2019: A Bibliometric Analysis. *Clocks and sleep*, 2(2), 99-119.

Lee, K. A., McEnany, G., and Zaffke, M. E. (2000). REM sleep and mood state in childbearing women: sleepy or weepy?. *Sleep*, 23(7), 877–885.

Leeder, J., Glaister, M., Pizzoferro, K., Dawson, J., and Pedlar, C. (2012). Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *Journal of sports sciences*, 30(6), 541–545.

Marshall J. (1963). Thermal Changes in the Normal Menstrual Cycle. *British medical journal*, 1(5323), 102–104.

Ohayon, M. M., Carskadon, M. A., Guilleminault, C., and Vitiello, M. V. (2004). Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*, 27(7),

Parry, B. L., Mostofi, N., LeVeau, B., Nahum, H. C., Golshan, S., Laughlin, G. A., and Gillin, J. C. (1999). Sleep EEG studies during early and late partial sleep deprivation in premenstrual dysphoric disorder and normal control subjects. *Psychiatry research*, 85(2), 127–143.

Sargent C, Halson S, Roach G.D. (2014). Sleep or swim? Early-morning training severely restricts the amount of sleep obtained by elite swimmers. *European journal of sport science*. 14(Suppl 1:S310-315).

Zarcone, V. P. (1994). Sleep hygiene *Principles and Practice of Sleep Medicine* (Vol. 2, pp. 542-546). Philadelphia: W.B. Saunders Co.

# Appendices

## Appendix A: Study information sheets

Appleton Institute for Behavioural Science

Central Queensland University

**The impact of international travel on sleep, body clock, and performance in elite athletes**

You are invited to participate in a research study collecting data on the impact of international travel and jet-lag on the sleep and performance of elite athletes. Before agreeing to participate in the study, it is important that you read and understand the following explanation of the study and procedures. This sheet describes the purpose, procedures, and benefits. If you choose to participate it is important to understand that you have the right to withdraw from the study at any time.

Data is being collected from you in order to:

- learn more about the impact of international travel on the **sleep** of elite athletes;
- learn more about the impact of international travel on the **body clock** of elite athletes;
- learn more about **strategies** that might be useful for reducing the symptoms of jet-lag in elite athletes following international travel.

**How will we collect information about sleep, body clock, and performance?**

*Sleep.* We will collect information about your sleep on two occasions. We will monitor your sleep on two nights before you travel overseas and then for seven nights after you return home from overseas. We will use a tiny device called an 'activity monitor' (same size and appearance

as a wrist watch) to monitor your sleep.  You will need to <u>wear your activity monitor at all times</u> during the study, except in situations where the monitor may get wet (e.g. showering or swimming) or damaged.  The activity monitor measures movement of the body and the information is stored in the monitor to be downloaded using special software at a later date.  This software will determine the quantity and quality of your sleep.  In addition to the activity monitor, we will also ask you to keep a **sleep diary**.  In this diary, you will record important details about your sleep, such as bed times and get up times.  <u>You will need to record these details for every sleep period (i.e. main sleeps and daytime naps).</u>

*Body clock.*  We will collect information on the timing of your body clock on two occasions.  We will assess the timing of your body clock on two nights before you travel overseas and then on seven nights after you return home from overseas.  On each night, you will be asked to collect samples of your urine between 9pm and 9am.  We will measure the amount of melatonin in your urine.  Melatonin is a hormone produced in the body than can provide information about the timing of your body clock.

*Performance.*  We will measure your performance on two occasions.  We will assess your performance on two days before you travel overseas and on seven days after you return home from overseas.  Each day, you will be asked to complete a short reaction time task using a computerised hand-held device.

*Jet-Lag.*  For seven days after you have returned home from international travel, we will ask you to ask to rate your symptoms of jet-lag.  We will also ask you to rate your levels of sleepiness and alertness.

**Who is involved in running this study?**

This study is being conducted by Central Queensland University. Associate Professor Charli Sargent and Dr Matthew Driller are two of the researchers who will be conducting the study. If you are interested in participating or have any further questions, you should feel free to contact Charli at charli.sargent@cqu.edu.au or (08) 8378 4516 or Matthew at mdriller@waikato.ac.nz

**What happens to the results?**

All data from this study will be stored on computers at the Appleton Institute for Behavioural Science with USB devices as backups in lockable filing cabinets. All data will be stored for 5 years. Summarised results will be stored as hard copies at the Appleton Institute for Behavioural Science and may be used in journal publications and at conferences. All records containing personal information will remain STRICTLY CONFIDIENTAL and no information, which could lead to the identification of any individual, will be released.

**What will I get out of the study?**

You may not directly benefit from participating in this study. However, upon completion of the study you will receive individual feedback about your own sleep and performance in the form of a short report.

**Voluntary participation - What happens if I say no?**

Before deciding whether or not to take part in this study, you may wish to discuss the matter with a relative or friend. You should feel free to do this. It is important that you understand that your participation in this trial is voluntary. If you do not wish to take part you are under no obligation to do so. If you decide to take part but later change your mind, you are free to withdraw from the study at any stage.

**Concerns / Complaints**

Please contact the CQUniversity Office of Research should there be any concerns about the nature and/or conduct of this research project (tel: 0749 23 2603, email: ethics@cqu.edu.au; Mailing address: Building 361, CQUniversity, Rockhampton, QLD 4702).

**Exercise before bed: How and when to train.**

**INFORMATION SHEET**

You have been invited to participate in our research study. Before agreeing to participate, it is important that you read and understand the following explanation of the study and the procedures. This form describes the purpose, procedures, benefits, risks and discomforts associated with the research study. If you choose to participate, you have the right to withdraw from the study at any time and this will not disadvantage you in any way.

The Project Team

The Principal Investigators are Dr Michele Lastella (CQUniversity), Associate Professor Charli Sargent (CQUniversity), Dr Aaron Scanlon (CQUniveristy), Dr Grace Vincent (CQUniversity), Ms Georgia Romyn (CQUniversity) and Mr Dean Miller (CQUniversity).

<span style="background-color:#8CC63F">**Aims and purpose of the research project**</span>

Traditional sleep recommendations generally advise against exercise close to bedtime. However, there is some evidence to suggest that people who exercise before bed report better sleep compared with people who exercise during the day. The aim of this project is to examine whether exercise before bed affects sleep.

<span style="background-color:#8CC63F">**Who CAN participate in this research project?**</span>

Healthy, non-smoking, well-trained males between 18 to 35 years old, taking no sleep medication and who have a regular sleep pattern are eligible to participate in this study. Before commencing the study, you will be asked to complete questionnaires related to your general health, physical activity and sleep habits. You will also be required to complete a 3-night sleep/wake diary and wear a wrist activity monitor (sleep watch). These measures will allow us to assess your regular sleep patterns.

<span style="background-color:#8CC63F">**Who CANNOT participate in this research project?**</span>

As we are interested in sleep and performance variables, we must exclude those who are unable to complete all testing. Participants who are injured or ill will not be able to participate in the study.

Individuals who have undertaken overseas travel in the 3 months prior to testing, those who suffer from, or have been diagnosed with, a sleep disorder such as sleep apnea, insomnia or narcolepsy, those who undertake shift work, and those who take any form of sleep medication or supplements known to affect sleep (such as melatonin) will also be excluded from the study.

<span style="background-color:#8CC63F">**How will my consent be obtained?**</span>

All interested participants will be provided with an information sheet, a general health questionnaire, Profile of Mood State questionnaire, Morningness-Eveningness questionnaire and the Pittsburgh Sleep Quality Index. The general health questionnaire will be used to determine your eligibility for the study based on the study's inclusion/exclusion criteria. The Pittsburgh Sleep Quality Index will be used to determine your sleep habits. After this, you will be asked to keep a sleep diary and wear a wrist activity

monitor to assess your sleep patterns. Those who continue to express interest in participating in the study and who meet the inclusion/exclusion criteria will be asked to attend the sleep laboratory for a tour and to sign an informed consent form.

**What is required of you to participate in this research project?**

The study will be carried out at the Appleton Institute's sleep laboratory. We will be collecting data with participants in groups of six, so you will carry out the study with five other people. Each of you will have your own private bedroom, living room and bathroom.

If you are eligible and agree to participate in the study, you will be required to attend the sleep laboratory for 6 nights (5 days). On each night, you will complete a different exercise condition. The conditions are:

**Condition A:** You will not complete any exercise before bed.

**Condition B**: You will complete an aerobic exercise session 90 min prior to bed.

**Condition C:** You will complete a resistance training session 90 min prior to bed.

During the study, we will monitor your sleep, your core body temperature, your physical performance, your cognitive function and your training load.

- Sleep. Your sleep will be monitored each night using polysomnography. This process involves the placement of small leads on your head and face using tape and paste. The leads are connected to a small box next to your bed and are long enough such that you are able to sleep in your normal position when in bed. These leads measure brain activity, eye movements, and muscle tone and provide important information about your sleep (i.e., whether you are in light sleep, deep sleep, or dreaming sleep).
- Sleep diary. You will be required to fill out a sleep diary each morning. This diary will include rating your feelings of sleepiness, fatigue, sleep quality, and muscle soreness on different scales. The diary will be used to measure your perception of your sleep and your post-exercise recovery.
- Physical performance. You will be required to perform an exercise screening test when you come for familiarisation prior to the study. This session will provide baseline data for both aerobic and resistance exercise.
  During the study you will be required to complete aerobic and resistance training tasks. The aerobic exercise block will consist of a 10-minute low-intensity warm up, followed by two 30-min rides at high intensity (70-80% of maximal heart rate). This will be followed by a warm-down. The resistance training block will consist of a standardised warm-up, followed by a 60-minute session involving weighted push ups, bicep curl, triceps press, shoulder raise, shoulder press and inverted row.

- <u>Heart rate.</u>  During all exercise sessions, you will be asked to wear a chest strap and transmitter that will measure your heart rate.
- <u>Body temperature.</u>  During all exercise sessions and while you sleep, we will be monitoring your body temperature. Body temperature will be measured using a tiny capsule (which contains a temperature sensor) that you will be asked to swallow (about the size of a multi-vitamin). The capsule transmits a signal to a small monitor pack that will be worn around the waist.  You will have one capsule traveling through your digestive system each time you stay in the laboratory. The capsules take approximately two days to pass through your system (although this may vary between individuals) and you will be required to swallow one capsule at each visit (three over the course of the study.

When you are not completing performance tasks you may read, watch TV/DVDs, draw, listen to music etc. You will not be able to perform any strenuous activity outside of the planned exercise sessions

The room in which you will sleep is similar to a hotel room and it will be dark and quiet during the night. You will not share the room with anyone else and your room has its own bathroom. The room has a video camera so that research staff monitoring you overnight can see what's happening in the room when the lights are out. It also has an audio system, so they can talk to you and hear you from their monitoring area outside the room. Research staff may come into the room to detach the wires if you need to get up during the night.

All of your meals will be prepared for you during the study and served at the same time each day. You will not be able to bring food with you nor eat food outside of the specified times. In addition to meals, snacks will be provided at regular intervals each day. Prior to the start of the study, you will be given a menu to review. You can use this menu to choose meals and food items that you like and to specify any dietary requirements that you have. All of your meals (except snacks) will be served in a dining room that you will share with the other five participants in your group. It is important to remember that throughout the study, and in the 24h prior to the first night, you will be asked to refrain from consuming any alcohol, caffeine or medication.

## What do I need to bring?

You will be staying in the sleep laboratory for one night on three separate occasions, so you will need to bring enough clothes and toiletries to cover this time. We will provide you with all of your linen

(sheets, blankets, towels, pillows etc.) but you may bring your own pillow from home if you wish. You will have your own bedroom, living area, and shower and toilet facilities for the duration of the study. Because there will be periods of spare time throughout the study, please feel free to bring along your own music, books or movies.

## What are the risks, inconvenience or discomfort that could reasonably be expected to be experienced during the study?

Each night during the study, we will monitor your sleep. You may experience some minor skin irritation from the small leads that we will attach to your face. The leads will be taped to your skin next to your eyes, below your chin, in the middle of your forehead and on your collarbone. If you experience any irritation, we will alternate the position of the leads.

There is a low risk of sustaining an injury during the exercise sessions. This is a risk associated with any physical activity. All efforts will be made to reduce the risk of injury by monitoring and maintaining any equipment, monitoring and maintaining the exercise space to ensure it is safe and free from obstacles. A standard warm up will also be completed prior to each exercise session to reduce the likelihood of injury.

The capsule that we are asking you to swallow (to record your body temperature) transmits data via low power radio frequency. This produces a small level of electromagnetic radiation, but it will be at level of exposure that is considered safe. For example, mobile phones also produce electromagnetic radiation, reaching peak levels of ~2000 milliwatts. The level produced by the ingestible capsule is approximately 1000 times lower than the level produced by a mobile phone. It is important, however, to notify Research Staff if you have been exposed to radiation in other research projects, or as a part of investigation (X-Rays) or treatment (Radiotherapy) in the past year.

In rare instances, you may experience some gastrointestinal discomfort following ingestion of the temperature capsule, or it may become lodged in the intestines. We will be monitoring you closely at all times during the study, and it is important for you to let Research Staff know if you are experiencing any untoward symptoms so that the appropriate action and treatment can be initiated. You may experience some difficulty when swallowing the capsule, which could result in a potential choking

hazard. At all times, the capsules will be administered to you according to the manufacturer's recommendation (i.e. the capsule will always be swallowed with water or other suitable liquid) and you will be instructed not to chew the capsule before swallowing.

If you decide to participate in the study, you will spend three nights in the laboratory. Because of this, you may experience feelings of isolation, anxiety, mood changes, etc. You will be able to interact frequently with the other participants and researchers and contact your family at allocated times. In addition, a clinical psychologist, who is independent of this study, will be available to talk to you if required.

What are the benefits to you?

While you will not directly benefit from the study, you will be compensated for the inconvenience and loss of time associated with your participation. If you complete the study, you will receive a $250 honorarium. Upon request, you will also receive individual feedback about your sleep and your exercise performance as well as the final research report.

How will my privacy and confidentiality be maintained?

All data from this study will be stored on computers or portable hard drives at the Appleton Institute. These computers will be password-protected and will only be accessed by members of the research team. Data that has been collected in hard copy (e.g., questionnaires) will be stored in lockable filing cabinets at the Appleton Institute. All information collected as part of the study will be destroyed at the end of seven years. Summarised results will be stored as hard copies at the Appleton Institute, and may be used in theses, journal publications and conferences. All records containing personal information will remain STRICTLY CONFIDENTIAL. You will be asked to keep the identity of your fellow participants confidential. This is in accordance with the CQUniversity Human Research Ethics Committee.

## What will happen to my information?

Research papers arising from the study will be submitted for publication in scientific journals and presented at conferences. No publications arising from this work will enable any participant to be identified. No information that will lead to the identification of any individual will be released and only pooled data will be reported. No case studies will be reported to protect your privacy. Should any of your data yield concerning or abnormal results we will notify you and refer you to a relevant professional.

## Right to Withdraw

Before deciding whether or not to take part in this trial, you may wish to discuss the matter with a relative, friend or your local doctor. You should feel free to do this. It is important that you understand that your participation in this trial is voluntary. If you do not wish to take part you are under no obligation to do so. If you decide to take part but later change your mind, you are free to withdraw from the project at any stage without explanation, and without prejudice from any member of the research team.

Any questions regarding this project may be directed to

Dr Michele Lastella, Postdoctoral Research Fellow;

E: m.lastella@cqu.edu.au

Associate Professor Charli Sargent, Research Fellow,

E: charli.sargent@cqu.edu.au

## Any concerns or complaints may be directed to

*Please contact CQUniversity's Office of Research (Tel: 07 4923 2603; E-mail: ethics@cqu.edu.au;*

*Mailing address: Building 32, CQUniversity, Rockhampton, QLD, 4702) should there be any concerns*

*about the nature and/or conduct of this research project.*

**Project Title:   Development and assessment of the effectiveness of a Sleep Drink to enhance**

**sleep quality and quantity.**

**Investigator (s)**     Dr. Shona Halson          AIS Performance Recovery

Greg Shaw                   AIS Sports Nutrition

Prof Louise Burke         AIS Sports Nutrition

Prof Keith Baar         UC Davis, California, USA

Dr James Carter       Gatorade Sport Science Institute

Dr Michele Lastella         Appleton Institute, CQU

Dr Charli Sargent           Appleton Institute, CQU

Prof Greg Roach         Appleton Institute, CQU

**Contact Person**     Shona Halson

**Background**

Elite athletes are keenly interested in finding non-pharmacological ("drug-free") means of enhancing sleep. Research completed at the AIS has identified that some athletes have difficulties falling asleep and staying asleep during both training and competition. A number of dietary factors or supplements are claimed to decrease the time taken to fall asleep and to increase quality of sleep. While many of these products have been investigated singly, no studies have examined them in combinations to determine their interactive effects on sleep.  Our goal is to find both the optimal dose and optimal timing

of intake of the ideal sleep product. To study this, we will use the gold standard of sleep measurement - polysomnography.

**Aim of the Study**

There are two aims of this study: 1) to determine the optimal quantities of various dietary factors or supplements on sleep to produce the ideal Sleep Drink (PART A) and 2) to assess the effectiveness of this optimal Sleep Drink on sleep quality and quantity using actigraphy and polysomnography (PART B) by comparing it to both a version of the drink identified as least optimal and a placebo drink.

**What Does Your Participation Involve?**

***You may volunteer to be involved in Part A or Part B of the study, or both.***

PART A- Testing the Sleep Drink during the day

You will be asked to attend the AIS Physiology Department Laboratory on one occasion only. For the 24 hours prior to this, you will be provided with a standardised diet and asked to observe a standardised bed and wake time.

On arrival in the lab, a cannula will be inserted in a vein in your arm to allow us to take a series of blood samples over next ~ 3 hours of recovery. You will then receive one of the dietary supplement treatments that we are trialling for our sleep product. This may be one of the doses of the following products:

- tart cherry juice
- Polyjoule, a high Glycemic Index (GI) carbohydrate powder
- an egg protein powder
- the herbal product Valerian root
- an amino acid called L-Theanine (found in green tea)
- a supplement called Adenosine-5-monophosphate.

All products will be pure and are considered completely safe to consume. However, it is recommended you seek additional advice if you have any questions or concerns regarding any of the products. Athletes

governed by the WADA code should seek advice from the ASADA telephone advice service (13 000 ASADA (27232)) and 'Check your substances' online tool available at www.asada.gov.au if they concerns.

At the time of consuming your supplement treatment, we will not tell you which one you have received. Once the study is completed, we will let you know what it was.

Blood samples will be taken at 0, 30, 60, 90, 120 and 180 min following the ingestion of the nutritional supplement. Each blood sample will be 8ml.

During the trial, your brain activity will be measured via electroencephalography. This technique monitors brainwaves to produce a visual map of the type and location of rhythms in the brain. TO undertake the QEEG measurements, you will need to wear a special cap fitted with 19-electrodes that record brain activity (see side picture). The measurements will require you to sit quietly for 5 minutes with your eyes open, followed by eyes closed. You will also be asked to complete 2 questionnaires about sleepiness.

PART B- Testing of Sleep Drinks during the night.

Based on the results of Part A, two sleep drinks will be formulated, containing either 1) the optimal dosages of ingredients that aid sleep and 2) the least optimal dosages of ingredients. This study will assess the effectiveness of the drinks, based on studying your sleep responses on three separate nights at the Appleton Institute (Sleep Laboratory), Central Queensland University with a technique called polysomnography. This technique is considered the gold standard of sleep measurement and places electrodes at designated sites of the body (head, face and limbs) to measure your sleep stages throughout the night. You will also wear a Sleep Watch (actigraph) to collect other information. You will be asked to undertake three sleep trials: On one occasion you will be provided with the optimal sleep drink, one occasion the least optimal drink and on another occasion you will be provided with a drink that looks and tastes similar, but has no effects (placebo). The drink will be given at a certain time prior to sleep which will be determined in Part A. You will be asked to attend the Sleep Laboratory on one additional occasion prior to the trials to complete a familiarization night (no data is collected, it is simply to familiarize yourself with the overnight monitoring).

At the time of each sleep trial, you will not be told which drink you are receiving. However, at the completion of the study we will reveal the order of your trials.

Before going to sleep and upon waking we will ask you a series of questionnaires to assess sleepiness and alertness. In the morning you will be asked to complete some additional tasks that assess memory and attention (using a computer-based task). Finally, you will be asked to complete a balance test which consists of 2 X 30 second periods standing on a force platform. One 30 second period will be eyes open and the other eyes closed. The tests that will be completed in the morning will take approximately 20 minutes. A summary of these tests is included below:

*Cognitive Function, Self-perceived Capacity, and Postural Sway.*

Tests of cognitive function in the 20-min battery:

- Cognitive throughput will be assessed using the Addition/Subtraction Task.
- Working memory will be assessed using the Digit Symbol Substitution Test.
- Short-term memory will be assessed using the Probed Recall Memory Test.
- Sustained attention will be assessed using the Psychomotor Vigilance Task.

Tests of self-perceived capacity in the 20-min battery:

- Subjective sleepiness will be assessed using the 9-point Karolinska Sleepiness Scale.
- Subjective fatigue will be assessed using the 7-point Samn-Perelli Fatigue Checklist.
- Subjective mood will be assessed using the Profile of Mood States.
- Self-ratings of alertness and performance capacity will be made using Visual Analogue Scales.
- The intensity of experienced arousal will be measured by the Pre-Sleep Arousal Scale.
- The intensity of neutral cognitive activity will be assessed by the Neutral Cognitive Activity Scale.
- Ratings of mental arousal/alertness, physical tension, and sleepiness, will be measured by the Arousal Level as Present State (ALPS) rating scale.

You will be provided with a taxi voucher to return to your home.

**Voluntary Participation and Withdrawal from the Study**

Your participation in this study is entirely voluntary. You may withdraw at any time without discrimination or prejudice. All information is treated as confidential and no names or other details that might identify you will be used in any publication arising from the research. Please make sure that you ask any questions you may have, and that all your questions have been answered to your satisfaction before you agree to participate.

**Benefits of the Study**

The results of the study will:

- Enhance the research teams knowledge on sleep in athletes and provide direction as to whether nutritional interventions should be used for athletes to enhance sleep

If you have any questions about this project please feel free to contact me, Dr Shona Halson, on 0422224491. Thank you for your assistance with this research project.

**If you have any concerns with respect to the conduct of this study, you may contact the Secretary of the AIS Ethics Committee (Mrs Helene Rushby) 02 6214 1577**

**Appendix B: Study consent forms**

# CONSENT FORM

## Exercise Before Bed Study

**I consent to participation in this research project and agree that:**

1. I am aged 18 years or older;

2. An Information Sheet has been provided to me that I have read and understood. Any questions I have had about the project have been answered to my satisfaction by the Information Sheet and any further verbal explanation;

3. I understand that I have the right to withdraw from the project at any time without penalty;

4. I freely consent to donating blood and saliva samples for the above-named project;

5. I understand the procedure for the donation of blood and saliva samples. Any concerns or doubts regarding the procedure have been clarified to my satisfaction;

6. I understand the statement concerning compensation for taking part in the study, which is contained in the Information Sheet;

7. I understand that to preserve anonymity and maintain confidentiality of participants, no personally identifiable information will be used in any publication(s);

8. I understand the research findings will be included in the researchers' publication(s) on the project, and these publications may include articles written for conferences, journals as well as other methods of dissemination;

9. I acknowledge that the results of the sleep study are unknown and that I may experience a range of symptoms including excessive sleepiness and fatigue, and/or irritability. This list of symptoms is not exhaustive;

10.     I release and indemnify the University, its employees, students and agents against liability in respect of all claims, costs and expenses and for all loss, damage, injury or death to persons or property caused or contributed by me in connection with my failure to follow the after-study instructions;

11.     I have discussed the risks of the study with, and I have had the opportunity to seek advice from someone independent from the study such as a relative, friend, doctor or lawyer.

**Participant name** (please print): _____.

**Signature of Participant**:_____**Date**:_/_____ /_____

| I wish to have a plain English statement of results sent to me | **Yes** | **No** |
|---|---|---|

*If yes, please provide* E-mail address:_____.

I have explained the study to the participant and consider that he/she understands what is involved:

**Researcher:** _____.

**Signature of Researcher**:_____**Date**:_/_____ /_____

**Participant Consent Form**

Appleton Institute for Behavioural Science

Central Queensland University

**Project Title:**

The impact of international travel on sleep, body clock, and performance in elite athletes

**Principal Researchers' names**:

Associate Professor Charli Sargent, Dr Matthew Driller, Professor Greg Roach

I have read the information sheet, and the nature and the purpose of the research project has been explained to me.  I understand and agree to take part.

I understand that I may not directly benefit from taking part in the study.

I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.

I understand that I can withdraw from the study at any stage and that this will not affect my status now or in the future.

**Name of Subject** ..................................................................................................

**Signed** ..................................................................................................

**Dated** ..................................................................................................

I have explained the study to subject and consider that he/she understands what is involved.

**Appleton Institute Representative**

**Researcher** ..................................................................................................

**Signed** ..................................................................................................

**Dated** ............

**Appendix C: Self-report sleep diary**

# Sleep Diary

**Name:**

**Date:**

**Monitor ID:**

Dr Michele Lastella

Appleton Institute for Behavioural Science

Central Queensland University

Email: m.lastella@cqu.edu.au


Associate Professor Charli Sargent


Appleton Institute for Behavioural Science

Central Queensland University

Email: charli.sargent@cqu.edu.au


Mr Dean Miller


Appleton Institute of Behavioural Science

Central Queensland University

Email: d.j.miller@cqu.edu.au


## Personal Details

Age: _____ yr          DOB: _____

Height: _____ cm          Weight: _____ kg

## Sleep History

How many hours of sleep do you need to feel rested?                    _____ hours

How satisfied are you with the amount of sleep you get?

       *Very dissatisfied*   1   2   3   4   5   6   7   8   9   10   *Very Satisfied*

Overall, how would you rate the quality of your sleep?

    ☐ very poor   ☐ poor   ☐ fair   ☐ good   ☐ very good   ☐ excellent

Thank you for participating in this study. The lessons we learn about your sleep patterns during training and competition could help to devise strategies to maximise your performance. To enable us to monitor your sleep, you will need to:

(i)     wear an activity monitor

(ii)    keep a sleep diary

(iii)   keep a training/game diary

# What is an Activity Monitor?

An Activity Monitor is a small device worn like a wristwatch that continuously records body movement.

It is a device that can be used to provide information about the amount and quality of your sleep.

- It is best if you wear your Activity Monitor **at all times**.
- The Activity Monitor is water resistant **but not waterproof** – so please remove it when showering or swimming. The Activity Monitor can be worn in the rain.
- It is important that you always wear the Activity Monitor on the **same wrist**.
- When training or competing, do not wear the activity monitor if there is a chance that it could be damaged or cause injury to yourself or others.
- Please take care of your Activity Monitor – **replacement value = $3,000**

# Keeping a Sleep Diary

In this booklet, you will find your sleep diary. The purpose of the diary is to record the times when you are attempting to sleep. This information will be used in conjunction with data from the activity monitor to determine when you fell asleep and woke up.

**Instructions**

- Complete a single line of the sleep diary for every sleep period (main sleeps and naps).
- Please complete the diary straight after every sleep period to aid accuracy. This will have a big influence on the quality of the data that we collect.
- Date - the date that you go to bed
- Sleep Location – the city where the sleep occurs
- Bed Time - the time that you start attempting to sleep. Don't include time spent reading, watching TV, etc

- Get Up Time - the time that you stop attempting to sleep.  Don't include time spent reading, watching TV, etc.
- Sleep Quality - the quality of your sleep compared to a 'normal' sleep period - see the bottom of your sleep diary for scale.

## What will your data look like?

**This figure shows an example of the type of data that you will collect:**

- The figure represents 7 days of data.
- Each line represents a day of data, from midnight to midnight.
- The **red** horizontal bars represent training times we will get these from your diary.
- The **blue** horizontal bars represent bed times we will get these from your diary.
- The **black** vertical bars represent the level of movement we will get these from your activity monitor.
- By combining the information from your sleep diary and activity monitor, we will use special software to determine (i) what time you went to sleep, (ii) what time you woke up, (iii) how much sleep you got, and (iv) how good or bad your sleep was.
- At the end of the sleep audit, you will get a report about your sleep/wake patterns that will include a figure like the one above.
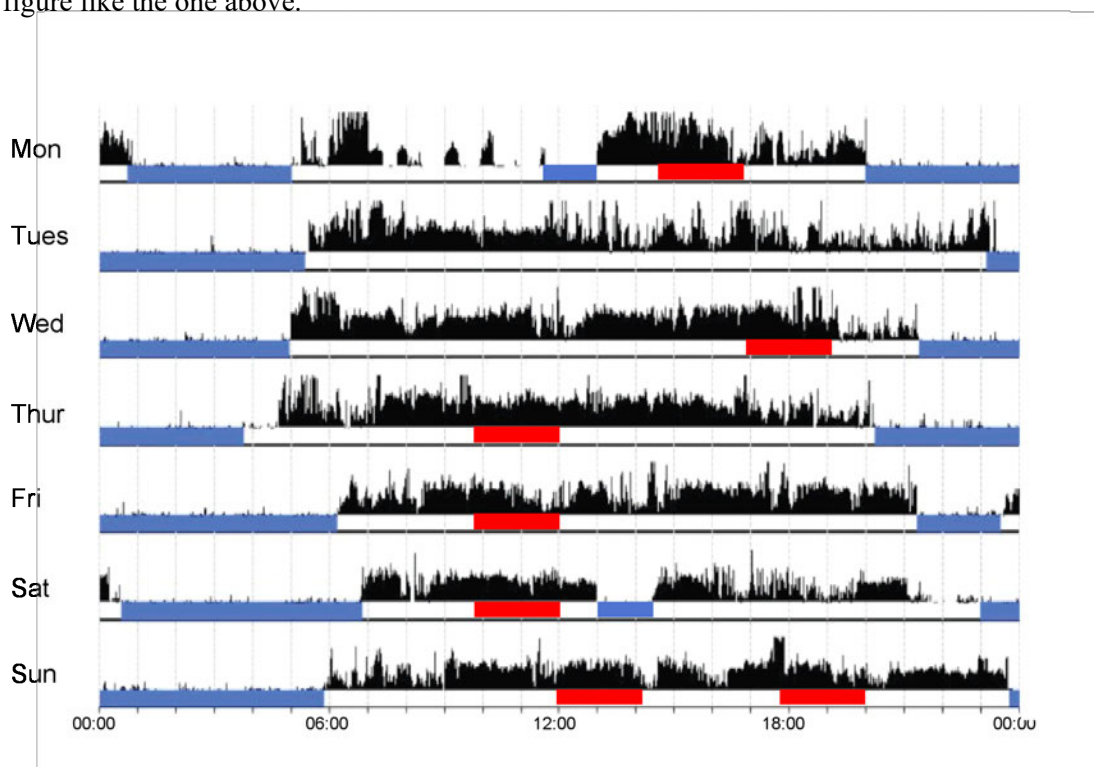
Name: _____

Study ID: _____

# Sleep Diary

| | BEFORE SLEEP | | | | AFTER SLEEP | | |
|---|---|---|---|---|---|---|---|
| **Date** day/mth | **Did you watch TV in bed before sleep?** If yes, for how long? | **Bed Time** (Local Time) hh:mm | **Pre-sleep Fatigue Level** please circle | | **Get up Time** (Local Time) hh:mm | **Post-sleep Fatigue Level** please circle | **Sleep Quality** please circle |
| e.g. 17/7 | | 10:00 am/pm | 1 2 3 4 ⑤ 6 7 | | 07:30 am/pm | ① 2 3 4 5 6 7 | 1 ② 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |
| | | am/pm | 1 2 3 4 5 6 7 | | am/pm | 1 2 3 4 5 6 7 | 1 2 3 4 5 |

**Instructions**

Please record all time periods in local time.

Complete a single line on the sleep diary for every sleep period (i.e. main sleeps and naps) that you have in bed or elsewhere.

Wear the activity monitor at all times except when bathing/swimming – it is water resistant but not waterproof.

Please take care of the Activity Monitor – replacement = Cost $3000.

Always wear the Activity Monitor on the same wrist.

**Pre/Post Sleep Fatigue Scale**

1 = fully alert, wide awake

2 = very lively, responsive, but not at peak

3 = okay, somewhat fresh

4 = a little tired, less than fresh

5 = moderately tired, let down

6 = extremely tired, very difficult to concentrate

7 = completely exhausted, unable to function effectively

**Sleep Quality**

1 = very poor

2 = poor

3 = average

4 = good

5 = very good

**Appendix D: Peer-reviewed publications associated with dissertation**

# Moderate-intensity exercise performed in the evening does not impair sleep in healthy males

D. J. Miller, C. Sargent, G. D. Roach, A. T. Scanlan, G. E. Vincent & M. Lastella

# A validation study of the WHOOP strap against polysomnography to assess sleep

Dean J. Miller , Michele Lastella , Aaron T. Scanlan , Clint Bellenger , Shona L. Halson , Gregory D. Roach & Charli Sargent

📅 Published online: 26 Jul 2020.

✏ Submit your article to this journal ↗

🔍 View related articles ↗

⬛ View Crossmark data ↗

*Article*

# Optimisation and Validation of a Nutritional Intervention to Enhance Sleep Quality and Quantity

Shona L. Halson [1,*], Gregory Shaw [2], Nathan Versey [3], Dean J. Miller [4], Charli Sargent [4], Gregory D. Roach [4], Lara Nyman [5], James M. Carter [5] and Keith Baar [6]

[1] School of Behavioural and Health Sciences, Australian Catholic University, Banyo QLD 4014, Australia
[2] High Performance Unit, Swimming Australia, Brisbane QLD 4519, Australia; greg.shaw@swimming.org.au
[3] Rowing Australia, Yarralumla ACT 2600, Australia; nathan.versey@rowing.ausportnet.com
[4] Appleton Institute for Behavioural Science, Central Queensland University, Wayville 5034, Australia; dean.miller@cqu.edu.au (D.J.M.); charli.sargent@cqu.edu.au (C.S.); greg.roach@cqu.edu.au (G.D.R.)
[5] Gatorade Sport Science Institute, PepsiCo Global Research and Development-Life Sciences, Purchase, NY 10577, USA; lara.nyman@pepsico.com (L.N.); james.carter@pepsico.com (J.M.C.)
[6] Departments of Neurobiology, Physiology & Behavior and Department of Physiology and Membrane Biology, University of California Davis School of Medicine, Davis, CA 95817, USA; kbaar@ucdavis.edu
* Correspondence: shona.halson@acu.edu.au

check for updates

**Abstract:** Background: Disturbed sleep may negatively influence physical health, cognitive performance, metabolism, and general wellbeing. Nutritional interventions represent a potential non-pharmacological means to increase sleep quality and quantity. Objective: (1) Identify an optimal suite of nutritional ingredients and (2) validate the effects of this suite utilising polysomnography, and cognitive and balance tests. Methods: The optimal and least optimal combinations of six ingredients were identified utilising 55 male participants and a Box–Behnken predictive model. To validate the model, 18 healthy, male, normal sleepers underwent three trials in a randomised, counterbalanced design: (1) optimal drink, (2) least optimal drink, or (3) placebo were provided before bed in a double-blinded manner. Polysomnography was utilised to measure sleep architecture. Cognitive performance, postural sway, and subjective sleep quality, were assessed 30 min after waking. Results: The optimal drink resulted in a significantly shorter sleep onset latency (9.9 ± 12.3 min) when compared to both the least optimal drink (26.1 ± 37.4 min) and the placebo drink (19.6 ± 32.0 min). No other measures of sleep, cognitive performance, postural sway, and subjective sleep quality were different between trials. Conclusion: A combination of ingredients, optimised to enhance sleep, significantly reduced sleep onset latency. No detrimental effects on sleep architecture, subjective sleep quality or next day performance were observed.

**Keywords:** nutrition; polysomnography; sleep onset latency

## 1. Introduction

Sleep has important biological functions in a myriad of physiological processes including learning, memory, and cognition [1,2]. Restricting sleep to less than 6 h per night for four or more consecutive nights has been shown to impair cognitive performance and mood [3], disturb glucose metabolism [4], appetite regulation [5], and immune function [6]. Common prescription medications for insomnia may result in adverse side effects, including tolerance and addiction, rebound insomnia upon cessation, and reduced emotional and cognitive function with long-term use [7]. Furthermore, pharmaceutical sleep interventions may result in next-day psychomotor impairment [8]. Therefore, novel interventions, as alternatives to pharmacological interventions, are needed.

A number of neurotransmitters are associated with the sleep–wake cycle, and include: serotonin (5-hydroxytryptamine; 5-HT), gamma-aminobutyric acid (GABA), orexin, melatonin, galanin, noradrenaline, and histamine [9]. Dietary interventions that act upon these neurotransmitters in the brain and their role in changing sleep quality and quantity have become of interest to those looking to improve sleep without the use of pharmaceutical intervention.

Dietary precursors can influence the rate of synthesis and function of a small number of neurotransmitters, including serotonin [10]. Synthesis of serotonin (5-HT) may influence sleep and is dependent on the availability of its precursor, the amino acid L-tryptophan (Trp), in the brain. Trp is transported across the blood–brain barrier by a system that shares other transporters including several large neutral amino acids (LNAA). Thus, the ratio of Trp/LNAA in the blood is crucial to the transport of Trp into the brain. Since Trp is the least abundant amino acid in food proteins, ingestion of many protein-rich foods generally decreases the uptake of Trp into the brain, due to the relative rise and preferential transport of LNAAs into the brain [10]. However, carbohydrate (CHO) intake increases brain Trp as a net result of the action of insulin. Insulin stimulates the uptake of LNAA into skeletal muscle (increasing the ratio of Trp:LNAA in the blood), but also reduces blood concentrations of free-fatty acids, which results in increased binding of Trp (reducing blood levels of free Trp) [11]. Additionally, ingestion of specific protein fractions, high in Trp, have been show to enhance Trp availability and improve sleep related measures [12,13].

Dietary supplements that have some evidence of positively influencing sleep include: CHO, $\alpha$-lactalbumin, tart cherry juice, valerian, L-theanine, and nucleotides [14–18]. Intake of these factors represents a potential intervention for enhancing natural sleep. A number of recent reviews have highlighted the potential for a number of these ingredients to influence sleep and jetlag [16,19]. However, each of these reviews make mention of the need for additional randomised controlled trials, using objective measures of sleep to enhance the quality of research in the area. Since there are multiple mechanisms, both known and unknown, by which nutrition may affect sleep, we hypothesised that a combination of nutrients would be more effective than any of the single nutrients studied to date. Furthermore, since these agents have only been tested in isolation, we hypothesised that there would be interactions between different sleep promoting pathways. Therefore, is was necessary to perform an experiment that would allow us to simultaneously determine the effect of all six nutrients at one time, regardless of mechanism and bias.

The aim of this study was to (1) identify the optimal combination and dosage of this suite of nutritional ingredients and (2) validate the optimal ingredient mix identified using gold standard sleep monitoring (polysomnography) by determining the effects of supplementation on sleep quantity and quality as well as subjective sleep, balance, and cognitive function in healthy adults.

## 2. Methods—Optimisation of Nutritional Intervention

Fifty-five healthy males (Age: 27.4 ± 6.2 year; Weight 83.2 ± 11.5 kg; Height; 181.2 ± 6.7 cm) attended the laboratory on two occasions (one screening plus baseline session and one testing session). Participants were provided with a standardised diet controlling for caffeine, total, carbohydrate (40%), and protein (30%) for the 24 h prior to visit 2. To determine the optimal concentration and combination of the identified nutritional ingredients, a six-factor Box–Behnken model was produced using Design-Expert software (Stat-Ease, Inc., Minneapolis, MN, USA). The Box–Behnken design (a subset of design of experiment methodology) is an incomplete factorial design that uses three levels of each factor being tested and analyses the resulting outcome data for effects, via quadratic response surface plots. According to the Box–Behnken design, each subject was provided with one intervention, making for 55 independent trials (48 intervention trials, plus 7 centre point trials). The 7 centre point trials consisted of participants being provided the mid-range dose for each ingredient. The seven participants who received the central combination of nutrients were used to determine the biological variability of the measures, and provide the statistical basis for the optimal selection. As each participant completed one trial only, the subjects were allocated a supplement in the order they were recruited.

The study was approved by the Australian Institute of Sport Ethics Committee (Approval number: 20131003) and all participants completed written informed consent (ClinicalTrials.gov Identifier: NCT03288077).

### 2.1. Dietary Supplement

The dietary supplement was given at time point 0 and blood samples taken at 0 (pre-drink ingestion), 30-, 60-, 90-, 120-, 150-, and 180-min post ingestion of the supplement. The drink provided to the participants was a combination of the below ingredients:

1. Tart cherry juice (0 ml, 50 mL, 100 mL) (Cherry Active Australia, Sydney, Australia)
2. High GI CHO (0 g, 25 g, 50 g) (PolyJoule, Nutricia, Sydney, Australia)
3. α-lactalbumin (0 g, 20 g, 40 g) (Davisco Foods, Le Sueur, MN, USA)
4. Adenosine-5-monophosphate (5-AMP)-(0 mcg, 26.5 mcg, 53 mcg) (Sigma, Ronkonkoma, NY, USA)
5. Valerian (0 mg, 750 mg, 1500 mg) (Martin Baeur Group, Vestenbergsgreuth, Germany)
6. Theanine (0 mg, 500 mg, 1000 mg) (Suntheanine, Osaka, Japan)

### 2.2. Assessment of Effectiveness

Free tryptophan was assessed using ultra-performance liquid chromatography tandem mass spectrometry (UPLC-MS/MS) at each time point (baseline and every 30 min until 180 min). A three-choice vigilance task (3CVT) lasting 5 min was completed at each time point (baseline and every 30 min until 180 min). The 3CVT requires participants to discriminate one primary target (presented 70% of the time) from two secondary non-target geometric shapes that are randomly interspersed over the 5-min test period. Each 3CVT consists of a single stimulus type (target or non-target) presented for 0.2 s, with fixed timing between stimuli. Participants were instructed to respond as quickly as possible to each stimulus presentation by selecting the left arrow to indicate target stimuli, and the right arrow to indicate non-target stimuli. Performance measurements include reaction time and accuracy (i.e., percentage of correct responses). Participants were administered a "practice round" before beginning the 3CVT to ensure that the participants fully understood the requirements of the task.

### 2.3. Model Optimisation

The Box–Behnken model was optimised for 3CVT and blood tryptophan levels, since these measures are believed to be quantitative physiological predictors of sleepiness. Following data collection, the 3CVT and tryptophan measures were fed back into the Design-Expert software and a model of treatment versus response was produced. This model was then used to predict the optimal and least optimal combination and concentration of each of the ingredients. The model was validated in a separate cohort of participants using sleep monitoring (see below).

## 3. Methods—Validation

### 3.1. Design and Procedures

Eighteen healthy males participated in the validation component of the study (Age: 25.1 ± 6.2 year; Weight 71.6 ± 11.3 kg; Height; 176 ± 8.3 cm). None of the participants had a history of a sleep disorder and none were taking medication during the time of testing. Participants were all non-smokers. None of the participants had undertaken transmeridian travel or shift work in the month prior to participation in the study. Over the previous month, participants' self-reported bedtime was 23:12 ± 00:54 h; self-reported get-up time was 08:12 ± 01:18 h; self-reported sleep onset latency was 15.0 ± 10.2 min and self-reported total sleep time was 8.1 ± 1.1 h. Participants were excluded if their usual self-reported bedtime was later than 01:00 h or their usual self-reported get-up time was later than 10:00 h. On average, over the past month, participants consumed 1.0 ± 0.9 caffeinated beverages

per day and consumed 1.6 ± 1.1 units of alcohol per day (each unit equivalent to 14 g of pure alcohol). There were no participant dropouts throughout the study.

A double-blind, placebo-controlled crossover experimental design was employed. Participants attended the Appleton Institute Sleep Laboratory and completed overnight polysomnography on four consecutive nights. The first night served as an adaptation night and was used to familiarise participants with the equipment for monitoring sleep. The next three nights were experimental nights on which participants received one of three interventions in a randomised, counterbalanced order: (1) optimal drink, (2) least optimal drink, or (3) placebo. On each night, participants consumed the drink at 21:00 h and were instructed to consume each drink within 5 min. Electrodes for monitoring sleep were then applied to participants. Participants were given a 9.5-h sleep opportunity from 22:30 h to 08:00 h. Participants rated their subjective sleepiness level every thirty minutes from 20:00 h until 22:30 h. Participants were not permitted to consume any water after 21:00 h. At 08:30 h the following morning (i.e., 30 min after waking), participants rated their subjective sleepiness, sleep quality, sleep duration, sleep latency, and completed a gastrointestinal symptom scale. At 09:00 h, participants completed a test battery to assess subjective alertness, self-perceived capacity, cognitive performance, and postural sway as outlined below. The test battery was always conducted in the same order and took approximately 30 min to complete.

### 3.2. Living Conditions

Participants were housed in a purpose-built accommodation suite at Central Queensland University's Appleton Institute. The suite, configured like a serviced apartment, can accommodate six participants concurrently, each with a private bedroom, lounge room, and bathroom. During scheduled daytime wake episodes, participants were free to engage in activities such as reading, watching movies, and listening to music. Participants were not permitted to sleep outside of scheduled time in bed. Researchers monitored participants for compliance in person and via CCTV. At the end of the protocol, participants were provided with transport to their home to minimise any risk of driving if sleep was at all disturbed.

### 3.3. Meals

All meals provided to participants were calorie-controlled and the participants consumed the same approximate number of calories at each meal. The participants were served breakfast, lunch, and dinner at 09:30 h, 12:30 h, and 19:00 h, respectively. An afternoon snack was also provided at 16:00 h. On average, participants consumed 10,061 ± 1112 kJ per day containing 51.8 ± 6.2% carbohydrate, 21.7 ± 6.5% protein, and 23.8 ± 6.7% fat. Participants did not have access to any other food or beverages (other than water) outside of the designated meal/snack times and were not permitted to consume caffeine or alcohol at any time during the protocol.

### 3.4. Sleep Assessment

Sleep was recorded using polysomnography equipment (Grael; Compumedics, Melbourne, VIC, Australia) with a standard montage of electrodes. Electrodes were applied in the 60 min prior to lights out and included three electroencephalograms (C4-M1, F4-M1, O2-M1), two electrooculograms (left/right outer canthus), and a submental electromyogram. All sleep records were blinded and manually scored in 30-s epochs by the same technician according to established criteria [20]. Stages of sleep were identified as non-rapid eye movement sleep (stages N1, N2, N3) and rapid eye movement sleep (R), with N1 considered the lightest phase of sleep and N3 considered the deepest phase. The following dependent variables were calculated from each sleep recording: total sleep time (min), the time spent in any stage of sleep (i.e., N1, N2, N3, R) during time in bed; time spent in stages N1, N2, N3 and R sleep (min); sleep onset latency (min), the time between lights-out to the first epoch of any stage of sleep (i.e., N1, N2, N3, R); wake after sleep onset (min), the time spent in bed awake minus sleep onset latency; sleep efficiency (%), total sleep time divided by time in bed x 100; arousals, (count);

arousals in NREM (count); arousals in REM (count); awakenings (count); stage shifts (count); stage R onset latency (min); and stage N3 onset latency.

### 3.5. Subjective Sleepiness

Subjective sleepiness was assessed using the Karolinska sleepiness scale (KSS) [21]. The KSS is a 9-point scale where 1 = "extremely alert", and 9 = "very sleepy, great effort to keep awake, fighting sleep". Participants were instructed to circle the number on the scale that corresponded to their current level of sleepiness.

### 3.6. Subjective Sleep Quality, Sleep Duration, Sleep Latency

Sleep quality was assessed using a 7-point scale, where 1 = "extremely poor", 2 = "very poor", 3 = "poor", 4 = "average", 5 = "good", 6 = "very good", and 7 = "extremely good". Participants were verbally asked "how much sleep do you think you got?" and "how long did it take you to fall asleep?".

### 3.7. Gastrointestinal Symptom Scale

The presence of gastrointestinal symptoms was assessed using a 16-item questionnaire. Participants used a 10-point Likert scale to rate whether they had experienced a gastrointestinal symptom (for e.g., reflux, heartburn, bloating, nausea, etc.) since bedtime the previous night. Responses ranged from 1, "no problem at all" to 10, "the worst it has ever been".

### 3.8. Subjective Alertness and Self-Perceived Capacity

Subjective alertness was assessed using a visual analogue scale. Participants rated their current level of alertness by placing a vertical mark on a non-numeric, 100-mm line that was anchored with "not at all" at one end and "completely" at the other end. Self-perceived capacity was assessed using two separate visual analogue scales. Participants rated their ability to perform as fast as possible by placing a vertical mark on a non-numeric, 100-mm line that was anchored with "not fast at all" at one end and "very fast" at the other end. Participants rated their ability to perform as accurately as possible by placing a vertical mark on a non-numeric, 100-mm line that was anchored with "not accurately at all" at one end and "very accurately" at the other end.

### 3.9. Cognitive Performance

Sustained attention was assessed using the psychomotor vigilance task (PVT-192; Ambulatory Monitoring Inc., New York, NY, USA). The PVT is a hand-held device with an upper surface that contains a four-digit LED display and two push-button response keys. Participants attended to the LED display for the duration of the test (10 min) and pressed the appropriate response key with the thumb of their dominant hand as quickly as possible after the appearance of a visual stimulus (presented at a variable interval of 2–10 s). If the correct response key was pressed, the LED display exhibited the participant's response time, in milliseconds, for 500 ms. If the wrong response key was pressed, an error message was displayed (ERR). If a response was made prior to the stimulus being presented, a false start message was displayed (FS). The dependent measures derived from the PVT included response time (ms); the number of lapses (i.e., response latency exceeding 500 ms) and the number of errors (i.e., false starts and incorrect button pushes). For all analyses, anticipated responses (i.e., those with response time less than 100 ms) were excluded. Cognitive throughput was assessed using a computer-based visual spatial-configuration search task [22,23]. The task is self-paced and consists of 52 trials in which participants are required to search for a target (i.e., the number 5) amongst distractors (i.e., the number 2). Visual searches were performed for set sizes of 10, 20, 30, or 40 distractor stimuli. Set size was equally distributed across the task. The dependent variables obtained from the task were the number of errors and the time taken to complete the task.

### 3.10. Postural Sway

Postural sway was assessed using an Accusway computerised force platform (AMTI, Watertown, MA, USA) in conjunction with Swaywin software (AMTI, Watertown, MA, USA). The force platform measures the three-dimensional forces (Fx, Fy, Fz) and the three-dimensional moments (Mx, My, Mz) involved in balance. These provide centre of pressure (COP) coordinates, which allow postural sway to be calculated. Participants performed two postural balance tasks each for 30 s; standing still with eyes open and standing still with eyes closed. The dependent variable obtained from the task was the area of the 95% confidence ellipse enclosing the COP.

### 3.11. Statistical Analysis

All data were analysed with a general linear mixed model using the R package lme4 (R Core Team). A random intercept for 'subjects' was included to account for intraindividual dependencies and interindividual heterogeneity. All models were estimated using restricted maximum likelihood. Data points with a value that was greater than 2 standard deviations from the mean were removed. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity but showed indications of heavy tails against the normal distribution. This was accommodated by obtained bootstrapped confidence intervals. *p*-values were obtained using Type II Wald F tests with Kenward-Roger degrees of freedom as implemented in the R package. Results are reported as mean estimates and 95% confidence intervals

### 4.1. Model Generation

Total tryptophan was dependent only on the amount of valerian and α-lactalbumin within the drink (Figure 1). The more of each of these nutrients, the greater the serum tryptophan levels.
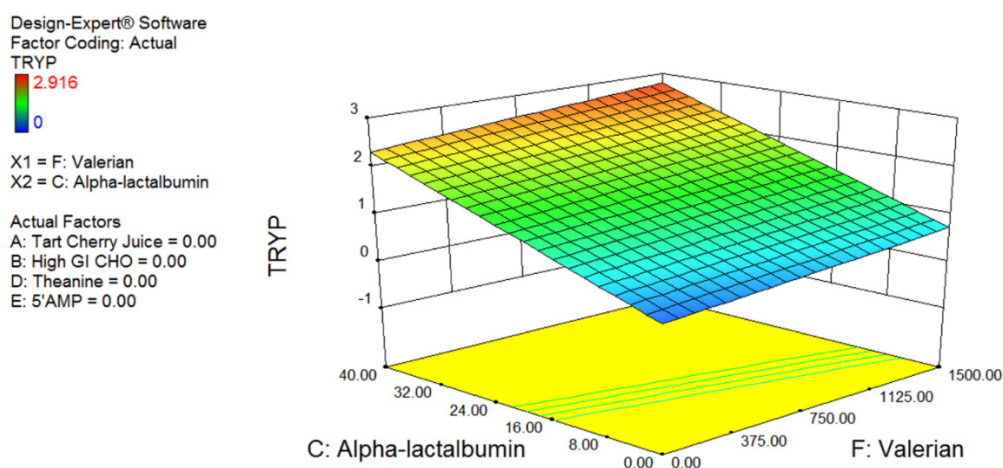


**Figure 1.** Quadratic response surface plot highlighting the effect of α-lactalbumin and valerian on blood tryptophan levels. The response surface represents all 55 measures. No other components affected this response surface.

A lower 3CVT score is an indicator of greater sleepiness. Contrary to tryptophan, 3CVT scores showed the most sensitivity to 5'AMP, theanine, and α-lactalbumin. For 3CVT, the most significant effect was from 5'AMP which showed a direct relationship between increasing levels of 5'AMP and decreasing 3CVT. As with tryptophan, increasing α-lactalbumin had a positive relationship with sleepiness (decreasing 3CVT); however, unlike tryptophan the effects of valerian were to decrease sleepiness (Figure 2).
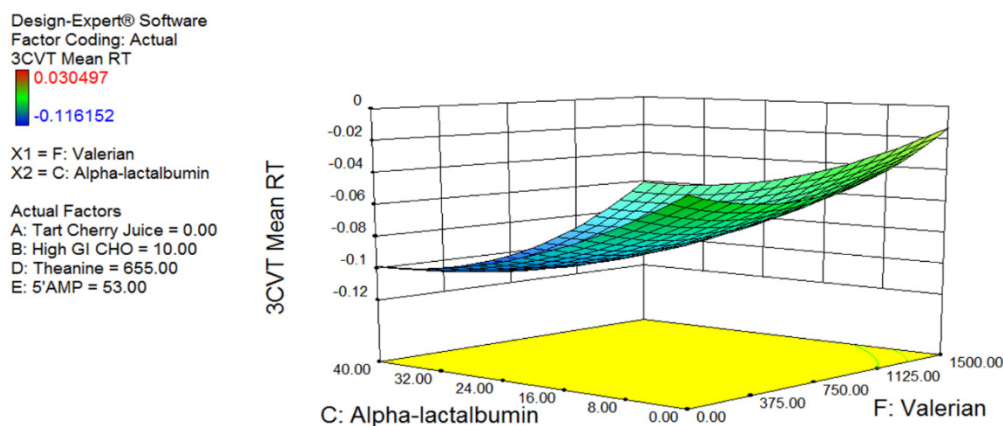
**Figure 2.** Quadratic response surface plot highlighting the effect of α-lactalbumin and valerian on three-choice vigilance task (3CVT) scores. A decrease in 3CVT score indicates greater sleepiness. The response surface represents all 55 measures. The 5′AMP linearly shifted the response down and is shown at its highest level. Theanine and high GI CHO had variable effects and are shown at their optimal.

When the Box–Behnken design was optimised for both tryptophan and 3CVT, a combination of 10 g high GI carbohydrate, 40 g α-lactalbumin, 655 mg theanine, 53 mcg 5′AMP, and 600 mg of valerian was predicted to be the best. This combination of ingredients was predicted to increase tryptophan to 2.25 μg/mL and decrease 3CVT score by 0.104 s. For the sake of validating the model in the subsequent sleep study, the worst combination and concentration of nutrients were predicted to be: 35 mL tart cherry, 45 g high GI carbohydrate, 8 g α-lactalbumin, 1000 mg theanine, 4.5 mcg 5′AMP, and 500 mg of valerian. These combinations are predicted to increase tryptophan 0.48 μg/mL and decrease 3CVT score by 0.001 s.

*4.2. Sleep Study Validation*

Objective and subjective sleep variables following drink consumption are reported in Table 1. Sleep onset latency was significantly lower in the optimal trial compared to both the placebo and least optimal trials ($p = 0.02$) (Figure 3D). None of the other variables (i.e., total sleep time, wake after sleep onset, sleep efficiency), were different between trials (Figure 3). Similarly, there was no change in the duration of time spent in sleep stages 1–3 or REM sleep (Table 1). Visual inspection of Figure 4 suggests that individuals with longer sleep onset latency respond in a more positive manner, relative to those with shorter sleep onset latency (Figure 4). However, this was not statistically significant as a consequence of the lack of power to address this question.

**Table 1.** Objective and subjective sleep variables and test battery results for least optimal, placebo, and optimal nutritional combinations (mean ± standard deviation (sd)).

| | Least Optimal (Mean ± SD) | Placebo (Mean ± SD) | Optimal (Mean ± SD) |
|---|---|---|---|
| **Sleep Variables** | | | |
| Total Sleep Time (min) | 503.3 ± 40.8 | 507.2 ± 60.5 | 519.5 ± 42.2 |
| Wake after sleep onset (min) | 40.6 ± 26.8 | 43.2 ± 47.3 | 40.6 ± 36.1 |
| Sleep Efficiency (%) | 88.3 ± 7.1 | 88.9 ± 10.6 | 91.1 ± 7.4 |
| Sleep onset latency (min) | 26.1 ± 37.4 | 19.6 ± 32.0 | 9.9 ± 12.3 * |
| REM Latency (min) | 75.7 ± 25.6 | 90.3 ± 69.7 | 82.7 ± 33.3 |
| Stage 3 Latency (min) | 14.9 ± 6.8 | 14.8 ± 5.9 | 18.5 ± 13.3 |
| Stage 1 (min) | 29.1 ± 15.2 | 30.8 ± 11.6 | 32.7 ± 12.8 |
| Stage 2 (min) | 201.9 ± 43.8 | 204.8 ± 42.5 | 223.7 ± 44.9 |
| Stage 3 (min) | 143.2 ± 52.2 | 139.7 ± 49.5 | 135.4 ± 39.7 |
| REM (min) | 129.2 ± 208 | 131.9 ± 26.8 | 127.7 ± 29.3 |
| Arousals—total (count) | 98.5 ± 29.1 | 97.3 ± 25.8 | 106.1 ± 32.9 |
| Arousals—REM (count) | 24.2 ± 7.7 | 25.0 ± 11.1 | 23.0 ± 13.2 |
| Arousals—NREM (count) | 74.3 ± 32.6 | 72.3 ± 28.7 | 83.1 ± 36.1 |
| Awakenings (count) | 26.1 ± 7.9 | 24.7 ± 6.6 | 26.2 ± 8.3 |
| Stage Shifts (count) | 178.2 ± 46.5 | 173.4 ± 25.7 | 188.2 ± 38.0 |
| KSS 2000 h (AU) | 4.6 ± 0.9 | 4.3 ± 1.2 | 4.8 ± 0.9 |
| KSS 2030 h (AU) | 4.9 ± 1.0 | 4.8 ± 1.2 | 4.9 ± 1.0 |
| KSS 2100 h (AU) | 5.0 ± 1.1 | 5.3 ± 1.4 | 5.3 ± 1.1 |
| KSS 2130 h (AU) | 5.3 ± 1.2 | 5.9 ± 1.1 | 5.6 ± 1.0 |
| KSS 2200 h (AU) | 5.8 ± 1.1 | 6.1 ± 1.1 | 6.0 ± 1.1 |
| Subjective Sleep Quality (AU) | 4.6 ± 1.0 | 4.9 ± 1.1 | 4.7 ± 0.9 |
| Subjective Sleep Quantity (h) | 7.9 ± 1.4 | 8.0 ± 1.1 | 8.1 ± 1.0 |
| Subjective SOL (min) | 22.9 ± 17.8 | 18.8 ± 14.0 | 15.8 ± 9.7 |
| **Test Battery** | | | |
| Mean Reaction Time (ms) | 268.8 ± 49.6 | 260.5 ± 41.0 | 262.1 ± 39.2 |
| Lapses (count) | 1.6 ± 3.1 | 1.6 ± 2.8 | 1.4 ± 2.3 |
| False Starts (count) | 0.8 ± 1.0 | 0.8 ± 1.3 | 0.5 ± 1.2 |
| KSS 0900 h (AU) | 4.6 ± 1.3 | 4.3 ± 1.3 | 4.3 ± 1.1 |
| VAS Alertness 0900 h (AU) | 57.4 ± 18.6 | 60.7 ± 20.3 | 63.1 ± 18.0 |
| VAS Speed 0900 h (AU) | 63.7 ± 14.9 | 65.9 ± 17.4 | 65.0 ± 16.1 |
| VAS Accuracy 0900 h (AU) | 66.6 ± 17.7 | 66.3 ± 19.4 | 67.4 ± 16.9 |
| Postural Sway—Area 95 ($cm^2$) | 0.4 ± 0.2 | 0.4 ± 0.2 | 0.4 ± 0.4 |

* indicates significantly different from least optimal and placebo. SOL = sleep onset latency; REM = rapid eye movement; NREM = non rapid eye movement; KSS = Karolinska sleepiness scale; AU = arbitrary units; VAS = visual analogue scale.
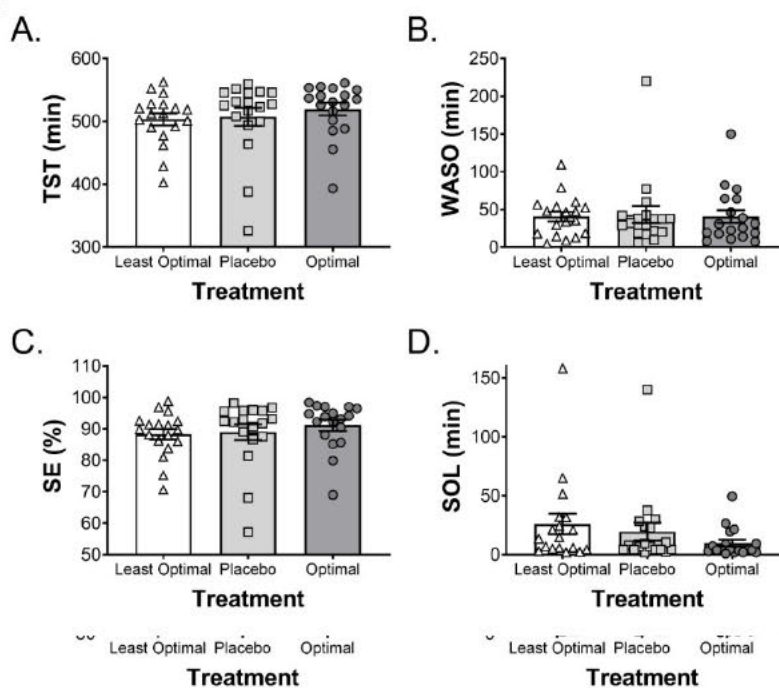
**Figure 3.** Individual data points for (A) total sleep time (TST), (B) wake after sleep onset (WASO), (C) sleep efficiency (SE), and (D) sleep onset latency (SOL). Bars and error bars are means and SD, respectively. Each point represents an individual.
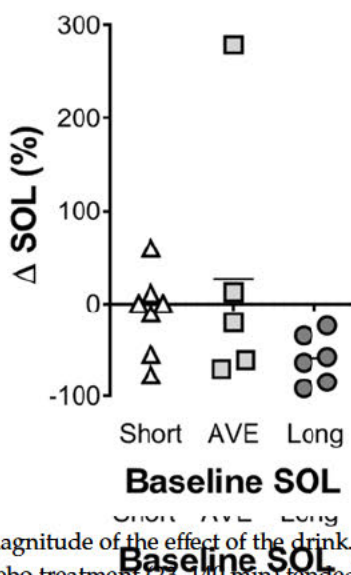


**Figure 4.** Baseline SOL and the magnitude of the effect of the drink. Individuals with the longest sleep onset latency (SOL) with the placebo treatment (23–140 min) tended to decrease sleep onset more than those with average (AVE; 7–15 min) or short (2–5 min) SOL when given the optimal drink. Lines are means and each dot is an individual.

### 4.3. Gastrointestinal Symptoms, Self-Perceived Capacity, and Performance

There was no difference in next-day gastrointestinal symptoms, subjective alertness, self-perceived capacity, cognitive performance or postural sway between trials (Table 1).

## 5. Discussion

This study describes the use and validation of a predictive model to identify the optimal combination and dose of a suite of ingredients that have been independently suggested to improve sleep. Furthermore, this model was tested using polysomnography to determine the effects of the intervention on sleep quality and quantity. Overall, there was a 49% reduction in sleep onset latency following the consumption of the optimal drink compared to a placebo control drink. Interestingly,

combining the same ingredients in the predicted least optimal manner resulted in a 33% increase in sleep onset latency when compared to the placebo control drink. These data suggest that the Box–Behnken model optimised for blood tryptophan and 3CVT was effective at predicting changes in sleep onset latency in human subjects. Importantly, the observed changes in sleep onset latency occurred despite all participants being considered normal or good sleepers. Furthermore, there were no adverse effects on sleep architecture and the optimal drink did not result in cognitive or balance impairments in the following morning. Of note, the improvements in sleep onset latency with the optimised intervention in the current study are similar to the improvements observed with pharmaceutical interventions [24].

Both tryptophan and high glycaemic index (GI) carbohydrate consumption prior to sleep are thought to influence neurotransmitters that are involved in the sleep–wake cycle [15]. Tryptophan shares blood brain barrier transporters with several large neutral amino acids, and an increase in the ratio of tryptophan to LNAAs may increase serotonin synthesis. Intake of a high GI carbohydrate may increase the Trp:LNAA in the brain via the insulin-stimulated [15] uptake of LNAA by muscle. Therefore, both $\alpha$-lactalbumin (a milk protein containing high levels of tryptophan) and glucose supplementation prior to sleep have the potential to influence sleep.

Yajima and co-workers [25] investigated the effects of a high CHO, compared with a high fat, meal prior to sleep and reported a decrease in slow-wave sleep (Stage 3 or deep sleep) in the first half of the night in the high CHO trial. When low and high GI CHO-rich meals (consumed 4 h before bedtime) and a high GI meal (consumed 1 h prior) were compared, the high GI meal consumed 4 h before bedtime reduced sleep onset latency [26]. In a second study by the same authors [27], a very low CHO diet was compared to an energy-matched control diet, consumed 4 h prior to sleep. The very low CHO diet increased slow wave sleep and all stages of NREM sleep, whereas the control diet decreased REM sleep.

The potential sleep promoting ingredients investigated in the current study have been examined individually (i.e., not in combination with other ingredients) in a small number of studies. Research investigating the effects of tryptophan-rich protein on sleep indicates a reduction in time awake during the night, increased sleep efficiency, and increased subjective sleep quality [12]. When comparing $\alpha$-lactalbumin to placebo (casein) prior to sleep, improved alertness the following morning has also been demonstrated [13]. The potential role for tryptophan in sleep is also supported by tryptophan depletion studies, which demonstrate an increase in sleep fragmentation [28].

Improved sleep quality and quantity has also been demonstrated with tart cherry juice consumption for 7 days [29]. Tart cherries have high concentrations of melatonin and other phenolic compounds which may have antioxidant and anti-inflammatory properties [16] that could potentially positively influence sleep. It should be noted that tart cherry juice may also contain high concentrations of carbohydrate, estimated to be approximately 30 g per 250 mL, which as mentioned above, may influence insulin concentrations.

Valerian (*Valeriana officinalis*) is one of the most commonly consumed 'natural remedies' for treating insomnia, and acts to inhibit sympathetic nervous system activity via the neurotransmitter GABA [30]. In a recent systematic review of various single plants to aid sleep, the authors concluded that due to different methods to measure sleep in the available literature (17 human trials) as well as varied doses used (225–1060 mg), studies demonstrated conflicting results. An earlier meta-analysis suggested that valerian may be effective for improving subjective sleep and is considered safe; however, there is insufficient objective data to determine its effectiveness.

L-Theanine is an amino acid found in tea which is reported to have anti-anxiolytic effects and induces changes in serotonergic and dopaminergic transmission [31]. Ingestion of 200 mg of L-Theanine has been shown to reduce wake after sleep onset, increase sleep efficiency, and have minimal daytime residual effects [31].

Nucleotides (purine adenosine 5'monophosphate (5'AMP), guanosine 5'monophosphate (5'GMP), and uridine 5'monophosphate (5'UMP) have been suggested to have a sleep-inducing role [14]. Specifically, the 5'AMP mechanism of action is thought to be via stimulation of the release of GABA [14].

The role of nucleotides in enhancing sleep has previously been investigated in infants. This is due to the higher concentrations of nucleotides in breast milk at night, demonstrating a circadian rhythm that was hypothesised to improve sleep. The minimal research in this area suggests that sleep was improved in infants who received a nucleotide-rich formula [32].

As a way to investigate all of the above-mentioned sleep promoting nutrients in one study, the current work tested all possible combinations of these six ingredients at three levels using an incomplete factorial design. This experiment would have required 729 different trials with groups of 10 or more participants per trial using traditional single factor methodology. Using a Box–Behnken incomplete factorial design, a model of the interactions between ingredients, and the necessity of each ingredient, was determined using only 48 trials. This model was then used to predict the optimal combination of ingredients based on both the maximal amount of tryptophan in the blood and the decrease in 3CVT measured following ingestion. A second combination of these same ingredients was established that was predicted to minimise blood tryptophan and increase 3CVT. This drink was predicted to negatively affect sleep and was used to validate the predictive capacity of the model. These combinations were then compared with a placebo control in the subsequent sleep studies.

The optimal combination of ingredients identified was a limited amount of glucose (10 g); the highest tested level of $\alpha$-lactalbumin (40 g); no tart cherry juice (possibly due to the time of ingestion; 1 h prior to sleep); 600 mg of valerian (approximately mid-range based on previous research); a modest dose of theanine (655 mg); and the highest tested level of 5′AMP (53 mcg). The levels of $\alpha$-lactalbumin and 5′AMP were highest in the optimal and lowest in the least optimal formulation suggesting that these ingredients have the strongest effect on the model. This is the first study of its kind to investigate this combination of ingredients, and there is limited research investigating the effects of more than one ingredient on sleep. Therefore, it is not possible to definitively describe the synergistic or potentially additive interaction of the ingredients. In the validation arm of the study, the optimal drink decreased sleep onset latency by 49% compared to the placebo control drink, whereas the least optimal combination of the same ingredients increased sleep onset latency by 33% compared to the placebo control drink. These results validate the predictive model and highlight the importance of understanding the precise combination and concentration of active ingredients within a supplement. Whether the effective concentration of the ingredients varies with body size is an important consideration that needs to be addressed in further research. Although not statistically significant, those participants who displayed the poorest sleep onset latency tended to show the greatest reduction in sleep onset latency with the optimal drink (Figure 4). This suggests that the optimal drink may be more efficacious for poor sleepers, or in individuals experiencing acute issues that impair sleep onset. However, both the optimal and least optimal drink had no effect on any other measure of sleep, which were all in the expected range for normal sleepers. Therefore, further research exploring the effects of this optimised sleep drink in poor sleepers, or in individuals who have difficulty initiating sleep onset, is warranted. Finally, an examination of the possible influence of this nutritional intervention on overnight muscle protein synthesis would be of interest. Previous work has demonstrated that protein ingestion prior to sleep can increase overnight muscle protein synthesis [33,34]. Furthermore, ingestion of protein prior to sleep was shown to be digested and absorbed throughout the night, muscle protein synthesis increased without disturbing sleep in both young and older healthy males [35,36]. If a nutritional intervention can both enhance sleep and improve overnight muscle protein synthesis, it may assist the recovery and performance of athletes.

In summary, an optimised combination of several nutritional ingredients, which had each demonstrated an influence on sleep independently, was determined. This intervention, when assessed utilising the gold standard of sleep monitoring in a controlled laboratory setting, significantly reduced sleep onset latency. Importantly, unlike pharmacological treatments, no detrimental effect on sleep architecture, subjective sleep measures, or next day performance were observed with the optimised nutritional supplement.

## References

1. Cirelli, C.; Tononi, G. Is sleep essential? *PLoS Biol.* **2008**, *6*, e216. [CrossRef]
2. Dattilo, M.; Antunes, H.K.M.; Medeiros, A.; Neto, M.M.; Souza, H.S.; Tufik, S.; de Mello, M.T. Sleep and muscle recovery: Endocrinological and molecular basis for a new and promising hypothesis. *Med. Hypotheses* **2011**, *77*, 220–222. [CrossRef] [PubMed]
3. Belenky, G.; Wesensten, N.J.; Thorne, D.R.; Thomas, M.L.; Sing, H.C.; Redmond, D.P.; Russo, M.B.; Balkin, T.J. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *J. Sleep Res.* **2003**, *12*, 1–12. [CrossRef] [PubMed]
4. Spiegel, K.; Leproult, R.; Van Cauter, E. Impact of sleep debt on metabolic and endocrine function. *Lancet* **1999**, *354*, 1435–1439. [CrossRef]
5. Spiegel, K.; Tasali, E.; Penev, P.; Van Cauter, E. Brief communication: Sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. *Ann. Intern. Med.* **2004**, *141*, 846–850. [CrossRef] [PubMed]
6. Krueger, J.M.; Majde, J.A.; Rector, D.M. Cytokines in immune function and sleep regulation. *Handb. Clin Neurol.* **2011**, *98*, 229–240.
7. Frase, L.; Nissen, C.; Riemann, D.; Spiegelhalder, K. Making sleep easier: Pharmacological interventions for insomnia. *Expert Opin. Pharmacother.* **2018**, *19*, 1465–1473. [CrossRef]
8. Brandt, J.; Leong, C. Benzodiazepines and Z-Drugs: An Updated Review of Major Adverse Outcomes Reported on in Epidemiologic Research. *Drugs R D* **2017**, *17*, 493–507. [CrossRef]
9. Saper, C.B.; Scammell, T.E.; Lu, J. Hypothalamic regulation of sleep and circadian rhythms. *Nature* **2005**, *437*, 1257–1263. [CrossRef]
10. Silber, B.Y.; Schmitt, J.A. Effects of tryptophan loading on human cognition, mood, and sleep. *Neurosci Biobehav. Rev.* **2010**, *34*, 387–407. [CrossRef]
11. Fernstrom, J.D.; Wurtman, R.J. Brain serotonin content: Physiological dependence on plasma tryptophan levels. *Science* **1971**, *173*, 149–152. [CrossRef] [PubMed]
12. Hudson, C.; Hudson, S.P.; Hecht, T.; MacKenzie, J. Protein source tryptophan versus pharmaceutical grade tryptophan as an efficacious treatment for chronic insomnia. *Nutr. Neurosci.* **2005**, *8*, 121–127. [CrossRef] [PubMed]
13. Markus, C.R.; Jonkman, L.M.; Lammers, J.H.; Deutz, N.E.; Messer, M.H.; Rigtering, N. Evening intake of alpha-lactalbumin increases plasma tryptophan availability and improves morning alertness and brain measures of attention. *Am. J. Clin. Nutr.* **2005**, *81*, 1026–1033. [CrossRef] [PubMed]
14. Schneider, N.; Mutungi, G.; Cubero, J. Diet and nutrients in the modulation of infant sleep: A review of the literature. *Nutr. Neurosci.* **2018**, *21*, 151–161. [CrossRef] [PubMed]
15. Halson, S.L. Sleep in elite athletes and nutritional interventions to enhance sleep. *Sports Med.* **2014**, *44* (Suppl. 1), S13–S23. [CrossRef]
16. Doherty, R.; Madigan, S.; Warrington, G.; Ellis, J. Sleep and Nutrition Interactions: Implications for Athletes. *Nutrients* **2019**, *11*, 822. [CrossRef]
17. St-Onge, M.P.; Mikic, A.; Pietrolungo, C.E. Effects of Diet on Sleep Quality. *Adv. Nutr.* **2016**, *7*, 938–949. [CrossRef]
18. Kim, J.; Lee, S.L.; Kang, I.; Song, Y.A.; Ma, J.; Hong, Y.S.; Park, S.; Moon, S.I.; Kim, S.; Jeong, S.; et al. Natural Products from Single Plants as Sleep Aids: A Systematic Review. *J. Med. Food* **2018**, *21*, 433–444. [CrossRef]

19. Janse van Rensburg, D.C.C.; Jansen van Rensburg, A.; Fowler, P.; Fullagar, H.; Stevens, D.; Halson, S.; Bender, A.; Vincent, G.; Claassen-Smithers, A.; Dunican, I.; et al. How to manage travel fatigue and jet lag in athletes? A systematic review of interventions. *Br. J. Sports Med.* **2020**, *54*. [CrossRef]

20. Iber, C.; Ancoli-Israel, S.; Chesson, A.L., Jr.; Quan, S.F. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1st ed.; American Academy of Sleep Medicine: Westchester, IL, USA, 2007.

21. Akerstedt, T.; Gillberg, M. Subjective and objective sleepiness in the active individual. *Int. J. Neurosci.* **1990**, *52*, 29–37. [CrossRef]

22. Burke, T.M.; Scheer, F.; Ronda, J.M.; Czeisler, C.A.; Wright, K.P., Jr. Sleep inertia, sleep homeostatic and circadian influences on higher-order cognitive functions. *J. Sleep Res.* **2015**, *24*, 364–371. [CrossRef] [PubMed]

23. Horowitz, T.S.; Cade, B.E.; Wolfe, J.M.; Czeisler, C.A. Searching night and day: A dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychol. Sci.* **2003**, *14*, 549–557. [CrossRef] [PubMed]

24. Wilt, T.J.; MacDonald, R.; Brasure, M.; Olson, C.M.; Carlyle, M.; Fuchs, E.; Khawaja, I.S.; Diem, S.; Koffel, E.; Ouellette, J.; et al. Pharmacologic Treatment of Insomnia Disorder: An Evidence Report for a Clinical Practice Guideline by the American College of Physicians. *Ann. Intern. Med.* **2016**, *165*, 103–112. [CrossRef] [PubMed]

25. Yajima, K.; Seya, T.; Iwayama, K.; Hibi, M.; Hari, S.; Nakashima, Y.; Ogata, H.; Omi, N.; Satoh, M.; Tokuyama, K. Effects of nutrient composition of dinner on sleep architecture and energy metabolism during sleep. *J. Nutr. Sci. Vitaminol.* **2014**, *60*, 114–121. [CrossRef] [PubMed]

26. Afaghi, A.; O'Connor, H.; Chow, C.M. High-glycemic-index carbohydrate meals shorten sleep onset. *Am. J. Clin. Nutr.* **2007**, *85*, 426–430. [CrossRef]

27. Afaghi, A.; O'Connor, H.; Chow, C.M. Acute effects of the very low carbohydrate diet on sleep indices. *Nutr. Neurosci.* **2008**, *11*, 146–154. [CrossRef]

28. Arnulf, I.; Quintin, P.; Alvarez, J.C.; Vigil, L.; Touitou, Y.; Lèbre, A.S.; Bellenger, A.; Varoquaux, O.; Derenne, J.P.; Allilaire, J.F.; et al. Mid-morning tryptophan depletion delays REM sleep onset in healthy subjects. *Neuropsychopharmacology* **2002**, *27*, 843–851. [CrossRef]

29. Howatson, G.; Bell, P.G.; Tallent, J.; Middleton, B.; McHugh, M.P.; Ellis, J. Effect of tart cherry juice (Prunus cerasus) on melatonin levels and enhanced sleep quality. *Eur. J. Nutr.* **2012**, *51*, 909–916. [CrossRef]

30. Fernandez-San-Martin, M.I.; Masa-Font, R.; Palacios-Soler, L.; Sancho-Gomez, P.; Calbo-Caldentey, C.; Flores-Mateo, G. Effectiveness of Valerian on insomnia: A meta-analysis of randomized placebo-controlled trials. *Sleep Med.* **2010**, *11*, 505–511. [CrossRef]

31. Rao, T.P.; Ozeki, M.; Juneja, L.R. In Search of a Safe Natural Sleep Aid. *J. Am. Coll. Nutr.* **2015**, *34*, 436–447. [CrossRef]

32. Cubero, J.; Chanclon, B.; Sanchez, S.; Rivero, M.; Rodriguez, A.B.; Barriga, C. Improving the quality of infant sleep through the inclusion at supper of cereals enriched with tryptophan, adenosine-5′-phosphate, and uridine-5′-phosphate. *Nutr. Neurosci.* **2009**, *12*, 272–280. [CrossRef] [PubMed]

33. Snijders, T.; Trommelen, J.; Kouw, I.W.K.; Holwerda, A.M.; Verdijk, L.B.; van Loon, L.J.C. The Impact of Pre-sleep Protein Ingestion on the Skeletal Muscle Adaptive Response to Exercise in Humans: An Update. *Front. Nutr.* **2019**, *6*, 17. [CrossRef] [PubMed]

34. Trommelen, J.; Holwerda, A.M.; Kouw, I.W.; Langer, H.; Halson, S.; Rollo, I.; Verdijk, L.B.; van Loon, L.J.C. Resistance Exercise Augments Postprandial Overnight Muscle Protein Synthesis Rates. *Med. Sci. Sports Exerc.* **2016**, *48*, 2517–2525. [CrossRef] [PubMed]

35. Trommelen, J.; Kouw, I.W.K.; Holwerda, A.M.; Snijders, T.; Halson, S.L.; Rollo, I.; Verdijk, L.B.; van Loon, L.J.C. Presleep dietary protein-derived amino acids are incorporated in myofibrillar protein during postexercise overnight recovery. *Am. J. Physiol. Endocrinol. Metab.* **2018**, *314*, E457–E467. [CrossRef]

36. Kouw, I.W.; Holwerda, A.M.; Trommelen, J.; Kramer, I.F.; Bastiaanse, J.; Halson, S.L.; Wodzig, W.K.; Verdijk, L.B.; van Loon, L.J. Protein Ingestion before Sleep Increases Overnight Muscle Protein Synthesis Rates in Healthy Older Men: A Randomized Controlled Trial. *J. Nutr.* **2017**, *147*, 2252–2261. [CrossRef]

*Article*

# A Validation Study of a Commercial Wearable Device to Automatically Detect and Estimate Sleep

Dean J. Miller [1,*], Gregory D. Roach [1], Michele Lastella [1], Aaron T. Scanlan [2], Clint R. Bellenger [3,4], Shona L. Halson [5] and Charli Sargent [1,*]

1   The Appleton Institute for Behavioural Science, CQUniversity, Adelaide, SA 5043, Australia; greg.roach@cqu.edu.au (G.D.R.); m.lastella@cqu.edu.au (M.L.)
2   Human Exercise and Training Laboratory, CQUniversity, Rockhampton, QLD 4700, Australia; a.scanlan@cqu.edu.au
3   Alliance for Research in Exercise, Nutrition and Activity (ARENA), Allied Health and Human Performance, University of South Australia, Adelaide, SA 5001, Australia; clint.bellenger@unisa.edu.au
4   South Australian Sports Institute, Adelaide, SA 5001, Australia
5   School of Behavioural and Health Sciences, Australian Catholic University, Brisbane, QLD 4014, Australia; shona.halson@acu.edu.au
*   Correspondence: d.j.miller@cqu.edu.au (D.J.M.); charli.sargent@cqu.edu.au (C.S.); Tel.: +61-8378-4523 (D.J.M.)

**Abstract:** The aims of this study were to: (1) compare actigraphy (ACTICAL) and a commercially available sleep wearable (i.e., WHOOP) under two functionalities (i.e., sleep auto-detection (WHOOP-AUTO) and manual adjustment of sleep (WHOOP-MANUAL)) for two-stage categorisation of sleep (sleep or wake) against polysomnography, and; (2) compare WHOOP-AUTO and WHOOP-MANUAL for four-stage categorisation of sleep (wake, light sleep, slow wave sleep (SWS), or rapid eye movement sleep (REM)) against polysomnography. Six healthy adults (male: $n = 3$; female: $n = 3$; age: $23.0 \pm 2.2$ yr) participated in the nine-night protocol. Fifty-four sleeps assessed by ACTICAL, WHOOP-AUTO and WHOOP-MANUAL were compared to polysomnography using difference testing, Bland–Altman comparisons, and 30-s epoch-by-epoch comparisons. Compared to polysomnography, ACTICAL overestimated total sleep time (37.6 min) and underestimated wake ($-37.6$ min); WHOOP-AUTO underestimated SWS ($-15.5$ min); and WHOOP-MANUAL underestimated wake ($-16.7$ min). For ACTICAL, sensitivity for sleep, specificity for wake and overall agreement were 98%, 60% and 89%, respectively. For WHOOP-AUTO, sensitivity for sleep, wake, and agreement for two-stage and four-stage categorisation of sleep were 90%, 60%, 86% and 63%, respectively. For WHOOP-MANUAL, sensitivity for sleep, wake, and agreement for two-stage and four-stage categorisation of sleep were 97%, 45%, 90% and 62%, respectively. WHOOP-AUTO and WHOOP-MANUAL have a similar sensitivity and specificity to actigraphy for two-stage categorisation of sleep and can be used as a practical alternative to polysomnography for two-stage categorisation of sleep and four-stage categorisation of sleep.

**Keywords:** consumer sleep technology; wearables; PSG; sleep staging; sleep monitoring; sleep quality

## 1. Introduction

Polysomnography (PSG) is the gold standard method of objectively assessing sleep [1]. However, PSG is expensive, time consuming and impractical in some field settings [2]. The most accepted alternative to PSG is research-grade actigraphy [3]. Actigraphy uses algorithms based on the association of movement and wakefulness, allowing for an objective measurement of sleep and wake [4]. However, acquiring accurate actigraphy data can be cumbersome and requires certain resources and expertise (e.g., proprietary software, reliance of self-reported sleep times, retrospective data extraction) [3,5]. Modern commercial wearable technology provides a user-friendly, accessible alternative to PSG and actigraphy that provides easily accessible aggregated sleep data [3,5]. Unlike actigraphy,

which relies solely on accelerometer-based movement detection to measure sleep and wake, sleep wearable technology utilises accelerometers and heart rate tracking technology (photoplethysmography) to provide two-stage categorisation and four-stage categorisation of sleep [3]. Photoplethysmography provides a convenient and accurate indication of autonomic nervous system status by measuring heart rate variability [6].

While providing a convenient alternative to PSG for measuring sleep, the use of actigraphy relies on self-report data from the user (i.e., bedtime, get up time) for retrospective manual adjustment of bed and wake times by the researcher [3]. In comparison, most commercial sleep wearables estimate sleep under two functionalities: auto-detection (i.e., automatic detection of sleep onset and sleep offset) or manual adjustment (i.e., manual input of bed and wake times after a sleep period). The distinction between auto-detected and manually adjusted sleep records is important as they are two different methods of measuring sleep. However, most validation studies for sleep wearables have either analysed manually adjusted data, [5,7,8] or have not reported their methods of data acquisition (i.e., auto-detection versus manual) [9–11]. For manually adjusted data validations, the adjustments are performed by researchers in a controlled laboratory setting [5,12]. Therefore, the accuracy of sleep wearables in situations where manual adjustment of sleep times is performed by the user may vary. In this context, actigraphy and sleep wearables utilising a manual adjustment function are subject to compliance of wearing the device and accurately reporting bed and wake times. Compliance for self-report measures is a common obstacle in acquiring accurate data in clinical settings [13] and elite sport [14]. Therefore, sleep wearables that are capable of accurately auto-detecting sleep, and therefore eliminating the non-compliance of users reporting bed and wake times, would provide an attractive alternative for measuring sleep in non-laboratory settings.

The WHOOP strap is a sleep wearable capable of estimating sleep [5]. When using manually adjusted sleep records (i.e., WHOOP-MANUAL), the WHOOP strap has been validated as an alternative for two-stage categorisation of sleep (i.e., sleep and wake) and four-stage categorisation of sleep (i.e., wake, light sleep, slow wave sleep (SWS), rapid eye movement sleep (REM)) when PSG is impractical [5]. However, the ability of the WHOOP strap to automatically detect (i.e., WHOOP-AUTO) and categorise two-stage sleep and four-stage sleep has not been examined. Therefore, the two aims of this study were to: (1) compare the ability of WHOOP-AUTO, WHOOP-MANUAL and research grade actigraphy (ACTICAL) for two-stage categorisation of sleep against PSG, and; (2) compare the ability of WHOOP-AUTO and WHOOP-MANUAL for four-stage categorisation of sleep against PSG.

## 2. Methods

### 2.1. Participants

Six healthy, young adults (male: $n$ = 3; female: $n$ = 3; age: 23.0 $\pm$ 2.2 yr; height: 170.5 $\pm$ 7.2 cm; weight: 65.8 $\pm$ 3.6 kg) participated in this study. Participants were excluded if they reported any existing medical conditions or sleep disorders or had a recent history of shift work and/or transmeridian travel. The study was approved by the Central Queensland University Human Research Ethics Committee.

### 2.2. Laboratory Setting

The study was conducted in a purpose-built accommodation suite at the Appleton Institute for Behavioural Science, Central Queensland University, Adelaide, Australia. The suite is sound-attenuated, free from external environmental cues and simultaneously houses six participants with private bedrooms and bathrooms.

### 2.3. Design

Data were collected as part of a larger experimental study. Participants lived in a sleep laboratory for ten consecutive nights/days and were given sleep opportunities of varying durations. Participants were given 9 h sleep opportunities on nights 1 (23:00–08:00 h)

and 2 (03:00–12:00 h); 7 h sleep opportunities on days 3–8 (14:30–21:30 h) and day 9 (08:30–15:30 h; Figure 1). Participants completed simulated work shifts on days 3–9 and performed sedentary tasks during free time (i.e., reading, watching movies; Figure 1). Data
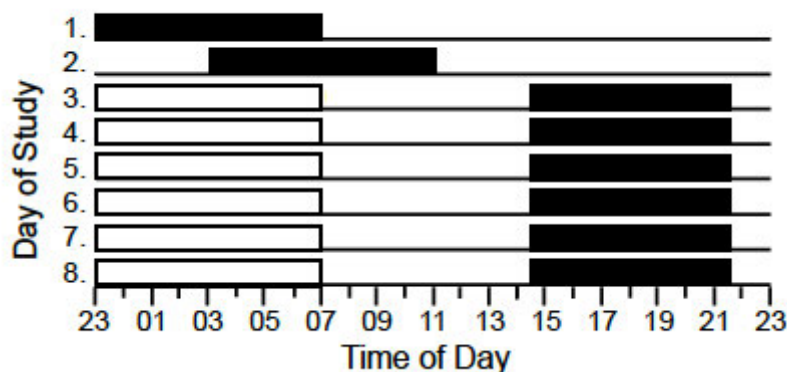


**Figure 1.** Illustration of the study design. Black horizontal bars indicate sleep opportunities. White horizontal bars indicate simulated work periods. *Y*-axis: "Day of Study". *X*-axis: "Time of Day".

*2.4. Measures and Procedure*

Sleep was measured using PSG. A standard montage of electrodes was attached to the face and scalp of participants (i.e., C4-M1, F4-M1, O2-M1), including two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram [15]. PSG data were recorded directly to data acquisition, storage, and analysis systems (Grael, Compumedics; Melbourne, Victoria, Australia). PSG records were manually scored in 30-s epochs by an experienced registered polysomnographic technician in compliance with standard criteria [16]. The commercially available wearable device used in this study was the WHOOP strap (Generation 2.0, CB Rank, Greater Boston, New England). The research grade activity monitor used in this study was the Actical Z-series (ACTICAL; Mini-Mitter Philips Respironics, Inc., Bend, OR, USA) [17]. Participants wore the WHOOP strap and ACTICAL on their non-dominant wrist, with the WHOOP strap placed 1 cm above the wrist bone—proximal to the ACTICAL. Prior to the study, clock time was manually synchronised on all devices (i.e., laboratory computers, ACTICAL, mobile devices running the WHOOP iOS application).

Data for automatically detected sleep (WHOOP-AUTO) and for manually adjusted sleep (WHOOP-MANUAL) were provided by the manufacturer for comparison to PSG. WHOOP-AUTO data were provided first to ensure that the manufacturer was blind to sleep times. Once WHOOP-AUTO data were received, the start and end times of each sleep opportunity were manually entered by a researcher into the WHOOP iOS application and the manually adjusted data (WHOOP-MANUAL) were then provided by the manufacturer. Epoch-by-epoch ACTICAL data were obtained using accompanying software (30-s epochs, medium sensitivity threshold; Actiware version 3.4; Mini-Mitter Philips Respironics, Inc.) [17].

The following sleep variables were collected during the study:

- Total sleep time (TST): the sum of minutes spent in any stage of sleep (N1, N2, N3, REM).
- Wake: the sum of minutes spent awake during the sleep opportunity.
- Light sleep: the sum of minutes spent in stage N1 or N2 sleep.
- Slow wave sleep (SWS): the sum of minutes spent in stage N3 sleep.
- Rapid eye movement sleep (REM): the sum of minutes spent in stage REM.
- Sleep onset latency (SOL): the duration of time from lights out to the first epoch of any stage of sleep.

PSG and WHOOP-MANUAL provided records of all the above variables. WHOOP-AUTO provided records of all the above variables, except for sleep onset latency. ACTICAL provided records of total sleep time and wake only.

To ensure that the WHOOP-AUTO, WHOOP-MANUAL and ACTICAL data were properly aligned to PSG data for each sleep record, agreement was calculated for offset adjustments of $\pm 3$ min in 30-s increments [18]. In all cases, agreement was not substantially improved by applying an offset, so all subsequent analyses were based on unadjusted data.

## 3. Data Analysis

Differences in TST, wake, light sleep, SWS and REM between PSG, WHOOP-AUTO and WHOOP-MANUAL were tested using separate General Linear Mixed Models (R package lme4; R Core Team, 2016). Differences in TST and wake time between PSG and ACTICAL were analysed using separate General Linear Mixed Models (R package lme4; R Core Team, 2016). A random intercept for participants was included in each model to account for intraindividual dependencies and interindividual heterogeneity.

Agreement between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL was tested using the Bland–Altman limits of agreement method for repeated measurements [19]. For each sleep variable, the difference between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL (i.e., bias) and the 95% limits of agreement (i.e., bias $\pm 1.96{*}SD$) were plotted. Each plot was examined for heteroscedasticity and proportional bias using the Breusch–Pagan test and least ordinary squares regression, respectively. If proportional bias or heteroscedasticity was present, the bias and 95% limits of agreement were adjusted accordingly [20].

To conduct epoch-by-epoch comparisons for two-stage categorisation of sleep, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL data were arranged in 30-s epochs and aligned with the corresponding PSG record. The following measures were then calculated for WHOOP-AUTO, WHOOP-MANUAL and ACTICAL:

- Sensitivity: the percentage of PSG-determined sleep epochs correctly identified as sleep by each method;
- Specificity: the percentage of PSG-determined wake epochs correctly identified as wake by each method;
- Agreement: the percentage of PSG-determined sleep and wake epochs correctly identified as sleep or wake by each method.

To conduct epoch-by-epoch comparisons for four-stage categorisation of sleep, WHOOP-AUTO and WHOOP-MANUAL data were arranged in 30-s epochs and aligned with the corresponding PSG record. The following measures were then calculated for WHOOP-AUTO and WHOOP-MANUAL:

- Sensitivity for wake: the percentage of PSG-determined wake epochs correctly identified as wake by each method;
- Sensitivity for light sleep: the percentage of PSG-determined N1 and N2 epochs correctly identified as light sleep by each method;
- Sensitivity for SWS: the percentage of PSG-determined N3 epochs correctly identified as SWS by each method;
- Sensitivity for REM: the percentage of PSG-determined REM epochs correctly identified as REM by each method;
- Agreement: the percentage of PSG-determined N1, N2, REM, and wake epochs correctly identified as light sleep, deep sleep, REM, or wake by each method.

Cohen's kappa (κ) was calculated to evaluate agreement between PSG and WHOOP-AUTO, WHOOP-MANUAL and ACTICAL beyond what could be expected by chance [21]. Agreement was interpreted against recommended guidelines as: *slight* agreement = 0–0.20; *fair* agreement = 0.21–0.40; *moderate* agreement = 0.41–0.60; *substantial* agreement = 0.61–0.80; *almost perfect* agreement = 0.81–0.99; and *perfect* agreement = 1 [22]. Intraclass correlation co-efficients were calculated to assess the reliability of WHOOP-AUTO, WHOOP-MANUAL and ACTICAL for two- and four-stage categorisation of sleep [23]. Intraclass correlation coefficients were interpreted against recommended guidelines as: "*poor*" = <0.40; "*fair*" = 0.40–0.59; "*good*" = 0.60–0.74; and "*excellent*" = 0.75–1.00 [24].

Aggregated data were collated from previous studies to compare WHOOP-AUTO, WHOOP-MANUAL and ACTICAL, respectively, against previous validations of sleep wearables [5,7,9–12,25–31].

## 4. Result

Data acquired using WHOOP-AUTO (*n* = 54), WHOOP-MANUAL (*n* = 54) and ACTICAL (*n* = 54) were included in the analyses for comparison to PSG. No data were lost and WHOOP-AUTO correctly identified 100% of the 54 sleep opportunities.

For two-stage categorisation of sleep, there was no significant difference between WHOOP-AUTO and PSG for TST or wake time (Table 1). Epoch-by-epoch data showed high sensitivity and moderate specificity for WHOOP-AUTO against PSG (Table 2). Cohen's kappa coefficient indicated *moderate* agreement (κ = 0.44) between the WHOOP-AUTO and PSG for two-stage categorisation of sleep [22]. Intraclass coefficient correlation indicated *fair* reliability (0.45) between WHOOP-AUTO and PSG for two-stage categorisation of sleep.

**Table 1.** Comparison of sleep variables determined by PSG, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL.

| Variable (min) | PSG vs. WHOOP-AUTO | | | | PSG vs. WHOOP-MANUAL | | | PSG vs. ACTICAL | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSG | Bias | AE | F | Bias | AE | F | Bias | AE | F |
| TST | 392.8 (60.7) | −17.8 (61.1) | 40.0 | 1.7 | 16.7 (35.6) | 25.4 | 2.4 | 37.6 * (85.6) | 38.1 | 12.2 |
| Wake | 53.9 (45.7) | 17.8 (61.1) | 40.0 | 2.8 | −16.7 * (35.6) | 25.4 | 6.3 | −37.6 * (85.6) | 38.1 | 35.1 |
| Light | 197.1 (50.8) | −8.9 * (55.9) | 43.8 | 0.8 | 13.9 (59.9) | 47.0 | 2.0 | | | |
| SWS | 101.4 (21.6) | −15.5 ** (30.1) | 24.7 | 13.1 | −6.1 (25.4) | 20.7 | 2.8 | | | |
| REM | 94.3 (28.9) | 6.5 (39.5) | 33.0 | 0.9 | 8.8 (42.0) | 33.0 | 1.9 | | | |
| SOL | 5.3 (5.9) | | | | −0.2 (4.8) | 2.8 | 0.01 | | | |

Notes: PSG; polysomnography, AE; absolute error (minutes), F; F-statistic, TST; total sleep time, Wake; wake time; Light; light sleep; SWS; slow wave sleep, REM; rapid eye movement sleep; SOL; sleep onset latency. Negative bias indicates an underestimation of the sleep variable by WHOOP-AUTO, WHOOP-MANUAL and ACTICAL when compared to PSG. * indicates significant difference to PSG with $p < 0.05$; ** indicates significant difference to PSG with $p < 0.001$. Data are mean (*SD*).

For two-stage categorisation of sleep, there was no significant difference between WHOOP-MANUAL and PSG for TST, but WHOOP-MANUAL significantly underestimated wake time compared to PSG (Figure 2; Table 1). Epoch-by-epoch data showed high sensitivity, but low specificity compared to PSG (Table 2). Cohen's kappa coefficient indicated *moderate* agreement (κ = 0.48) between the WHOOP-MANUAL and PSG for two-stage categorisation of sleep. Intraclass coefficient correlation indicated *fair* reliability (0.48) between WHOOP-MANUAL and PSG for two-stage categorisation of sleep.

**Table 2.** Epoch-by-epoch concordance statistics for WHOOP-AUTO (2-stage and 4-stage categorisation of sleep), WHOOP-MANUAL (2-stage and 4-stage categorisation of sleep) and ACTICAL (2-stage categorisation of sleep) against PSG.

| Measure | Value (%) |
|---|---|
| **2-stage comparison** | |
| WHOOP-AUTO | |
| Sensitivity for sleep | 90 |
| Specificity for wake | 60 |
| Overall agreement | 86 |
| WHOOP-MANUAL | |
| Sensitivity for sleep | 97 |
| Specificity for wake | 45 |
| Overall agreement | 90 |
| ACTICAL | |
| Sensitivity for sleep | 98 |
| Specificity for wake | 60 |
| Overall agreement | 89 |
| **4-stage comparison** | |
| WHOOP-AUTO | |
| Sensitivity for wake | 60 |
| Sensitivity for light sleep | 61 |
| Sensitivity for SWS | 63 |
| Sensitivity for REM | 66 |
| Overall agreement | 63 |
| WHOOP-MANUAL | |
| Sensitivity for wake | 45 |
| Sensitivity for light sleep | 67 |
| Sensitivity for SWS | 61 |
| Sensitivity for REM | 66 |
| Overall agreement | 62 |

Notes: SWS; slow wave sleep, REM; rapid eye movement sleep.

For four-stage categorisation of sleep, there was no significant difference between WHOOP-AUTO and PSG for TST, wake time, light sleep or REM. WHOOP-AUTO significantly underestimated SWS and overestimated sleep onset latency (Figure 3; Table 1). There was moderate overall agreement for four-stage categorisation of sleep between PSG and WHOOP-AUTO and moderate sensitivity for wake, light sleep, SWS and REM (Table 2). Cohen's kappa coefficient indicated moderate agreement ($\kappa = 0.47$) between WHOOP-AUTO and PSG for four-stage categorisation of sleep [22]. Intraclass coefficient correlation indicated *fair* reliability (0.48) between WHOOP-AUTO and PSG for four-stage categorisation of sleep.

A four-stage error matrix comparing WHOOP-AUTO and PSG is presented in Table 3. When WHOOP-AUTO misclassifies wake, it classifies it as light sleep. When WHOOP-AUTO misclassifies light sleep, it classifies it as either wake or REM. When WHOOP-AUTO misclassifies SWS, it classifies it as light sleep. When WHOOP-AUTO misclassifies REM, it classifies it as light sleep. Bland–Altman plots comparing WHOOP-AUTO to PSG for each sleep variable are depicted in Figure 2. Proportional bias (i.e., whether the differences between a device and PSG change as a function of duration) and heteroscedasticity (i.e., whether variance changes as a function of duration) were present for TST, wake time, and sleep onset latency, but not for light sleep, SWS or REM.
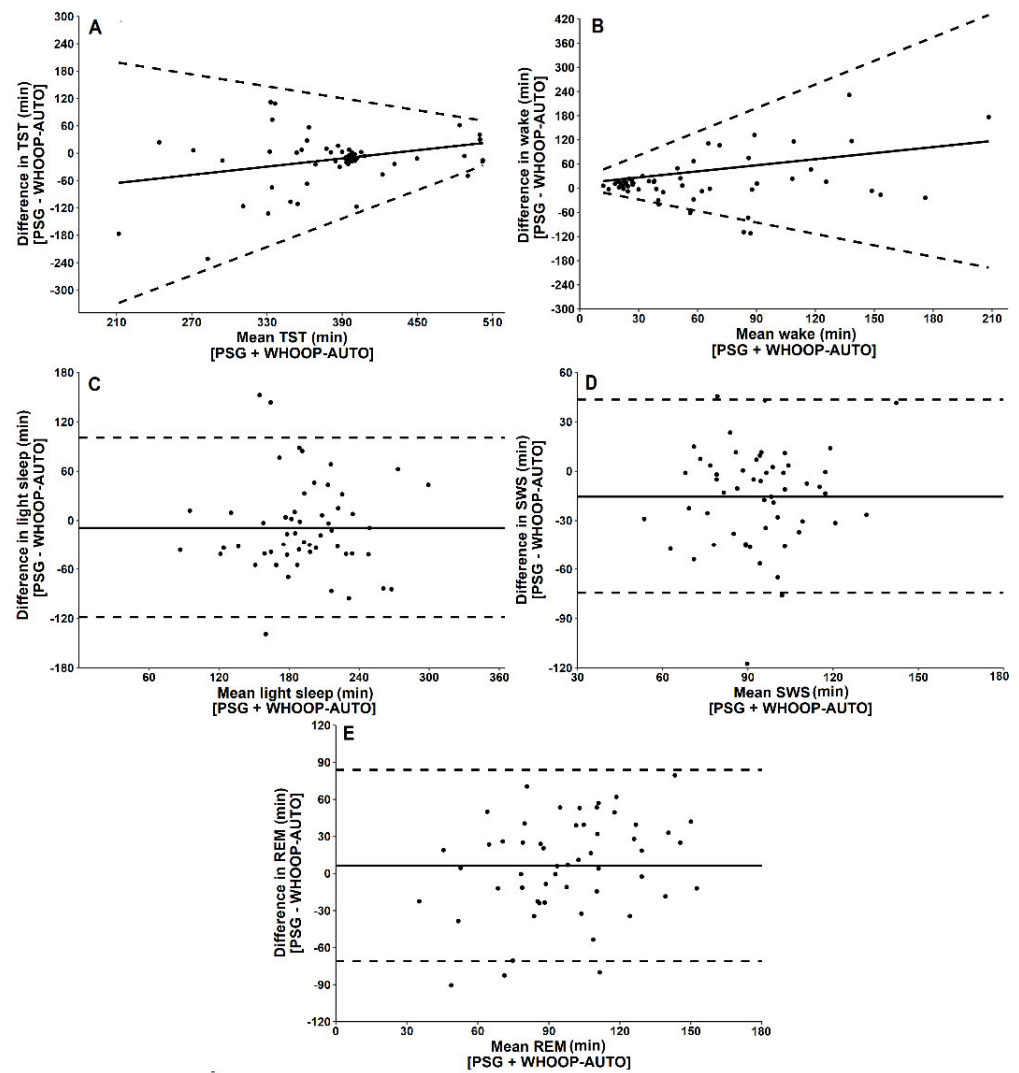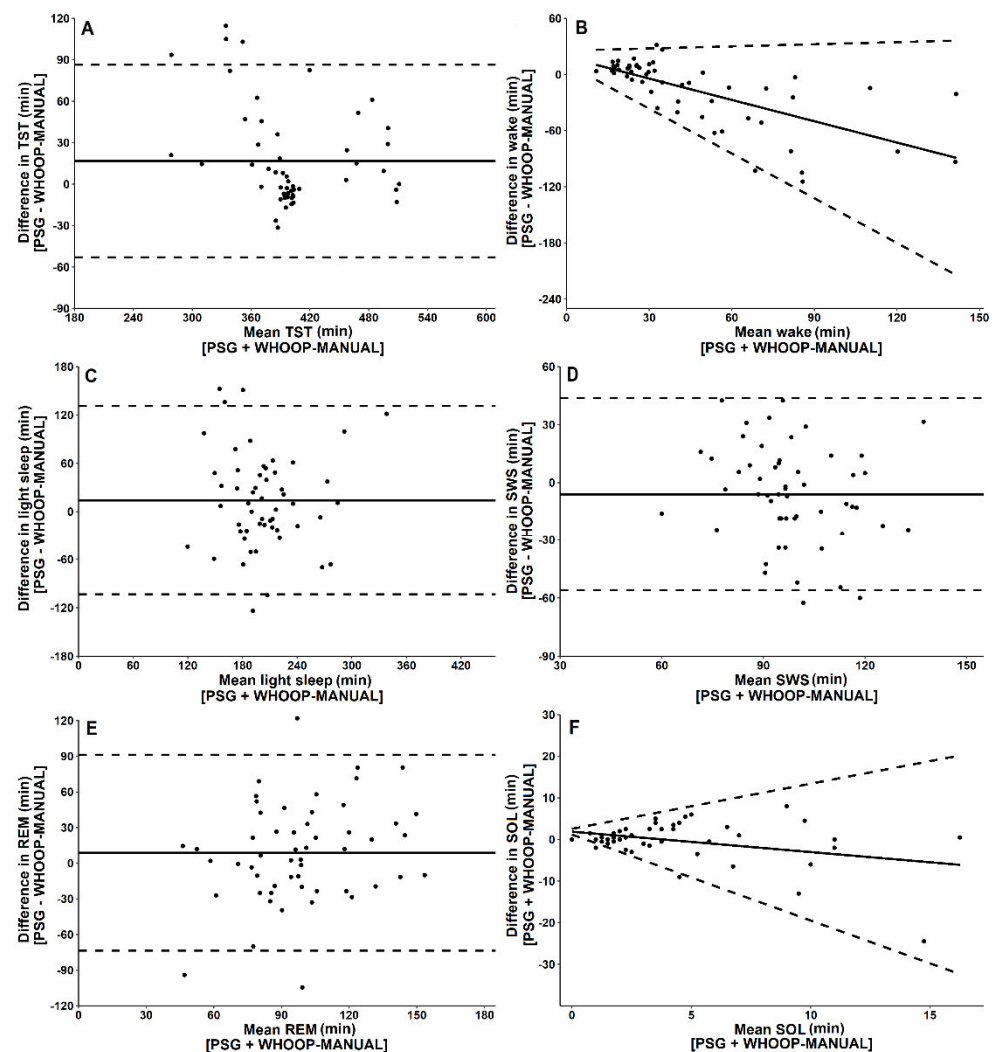
**Figure 2.** Bland–Altman plots for WHOOP-AUTO and PSG-derived measures of (**A**) total sleep time (TST), (**B**) wake time, (**C**) light sleep, (**D**) slow wave sleep (SWS), (**E**) rapid eye movement sleep (REM). Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from WHOOP-AUTO and PSG. The y-axes represent the difference between the values, such that positive values indicate that WHOOP-AUTO overestimates relative to PSG and negative values indicate that WHOOP-AUTO underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement ($\pm$1.96 standard deviations) [20].

**Table 3.** Four-stage error matrix for WHOOP-AUTO and PSG.

| | | WHOOP-AUTO | | | |
|---|---|---|---|---|---|
| | **Stage** | **Wake** | Light sleep | SWS | REM |
| **PSG** | Wake | 60% | 26% | 1% | 12% |
| | Light sleep | 14% | 61% | 10% | 15% |
| | SWS | 6% | 28% | 64% | 2% |
| | REM | 6% | 27% | 1% | 66% |

Notes: This matrix presents the percentage of each sleep stage that WHOOP-AUTO has correctly or incorrectly classified compared to PSG. Shaded cells indicate correctly classified sleep. SWS; slow wave sleep, REM; rapid eye movement sleep.

**Figure 3.** Bland–Altman plots for WHOOP-MANUAL and PSG-derived measures of (**A**) total sleep time (TST), (**B**) wake time, (**C**) light sleep, (**D**) slow wave sleep (SWS), (**E**) rapid eye movement sleep (REM) and (**F**) sleep onset latency (SOL). Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from WHOOP-MANUAL and PSG. The y-axes represent the difference between the values, such that positive values indicate WHOOP-MANUAL overestimates relative to PSG and negative values indicate WHOOP-MANUAL underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement (±1.96 standard deviations) [20].

ACTICAL significantly overestimated TST and underestimated wake when compared to PSG (Figure 4; Table 1). For two-stage categorisation of sleep, ACTICAL had high sensitivity (i.e., ability to detect sleep) and moderate specificity (i.e., ability to detect wake; Table 2). Cohen's kappa coefficient for two-stage categorisation of sleep ($\kappa = 0.23$) indicated *fair* agreement between ACTICAL and PSG. Intraclass coefficient correlation indicated *poor* reliability (0.26) between ACTICAL and PSG for two-stage categorisation of sleep. Bland–Altman plots comparing ACTICAL to PSG for each sleep variable are depicted in Figure 4.
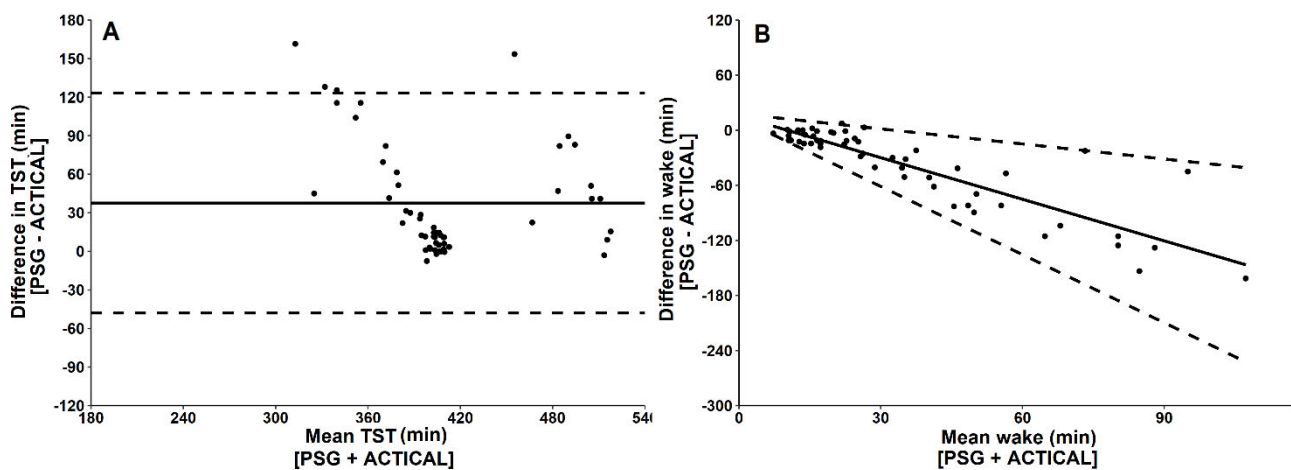
**Figure 4.** Bland–Altman plots for ACTICAL and PSG-derived measures of (**A**) total sleep time (TST) and (**B**) wake time. Data points represent one sleep opportunity. The x-axes represent the mean of the values obtained from ACTICAL and PSG. The y-axes represent the difference between the values, such that positive values indicate that ACTICAL overestimates relative to PSG and negative values indicate that ACTICAL underestimates relative to PSG. Solid horizontal lines indicate the mean bias from PSG, and broken lines indicate the 95% limits of agreement ($\pm$1.96 standard deviations) [20].

For four-stage categorisation of sleep, there was no significant difference between WHOOP-MANUAL and PSG for TST, light sleep, SWS, REM or sleep onset latency. WHOOP-MANUAL significantly underestimated wake time compared to PSG (Table 1). There was *moderate* overall agreement for four-stage categorisation of sleep, *moderate* sensitivity for light sleep, SWS and REM, and *low* sensitivity for wake time between PSG and WHOOP-MANUAL (Table 2). Cohen's kappa coefficient indicated *moderate* agreement ($\kappa = 0.49$) between WHOOP-MANUAL and PSG for four-stage categorisation of sleep [22]. Intraclass coefficient correlation indicated *fair* reliability (0.47) between WHOOP-MANUAL and PSG for two-stage categorisation of sleep.

A four-stage error matrix comparing the WHOOP-MANUAL and PSG is presented in Table 4. When WHOOP-MANUAL misclassifies wake, it classifies it as light sleep. When WHOOP-MANUAL misclassifies light sleep, it classifies it as REM. When WHOOP-MANUAL misclassifies SWS, it classifies it as light sleep. When WHOOP-MANUAL misclassifies REM, it classifies it as light sleep. Bland–Altman plots comparing WHOOP-MANUAL to PSG for each sleep variable are depicted in Figure 3. Proportional bias and heteroscedasticity were present for wake time and sleep onset latency, but not for TST, light sleep, SWS or REM.

**Table 4.** Four-stage error matrix for WHOOP-MANUAL and PSG.

| | Stage | WHOOP-MANUAL | | | |
| --- | --- | --- | --- | --- | --- |
| | | Wake | Light sleep | SWS | REM |
| PSG | Wake | 45% | 37% | 1% | 18% |
| | Light sleep | 7% | 67% | 11% | 15% |
| | SWS | 1% | 38% | 61% | 1% |
| | REM | 1% | 31% | 2% | 66% |

Notes: This matrix presents the percentage of each sleep stage that the WHOOP-MANUAL has correctly or incorrectly classified compared to PSG. Shaded cells indicate correctly classified sleep. SWS; slow wave sleep, REM; rapid eye movement sleep.

## 5. Discussion

The two aims of this study were to: (1) compare WHOOP-AUTO, WHOOP-MANUAL and research grade actigraphy (ACTICAL) for two-stage categorisation of sleep against PSG, and; (2) compare WHOOP-AUTO and WHOOP-MANUAL for four-stage categorisation of sleep against PSG.

## 5.1. Two-Stage Categorisation of Sleep

Actigraphy is commonly utilised as an objective measure of sleep and wake by practitioners [3,30,32]. However, the process of acquiring sleep data using actigraphy requires certain expertise and is usually a retrospective analysis of an extended data collection period—rather than the immediate day-by-day data that are provided by modern sleep wearables. The accuracy of actigraphy and sleep wearables utilising a manual adjustment function is subject to the compliance of the user wearing the device and accurately reporting bed and wake times. In this context, it is important to compare the performance of actigraphy (i.e., ACTICAL) to modern sleep wearables that can automatically detect sleep and provide easily accessible data (e.g., the WHOOP strap).

Regarding the two-stage detection of sleep, WHOOP-AUTO, WHOOP-MANUAL and ACTICAL had high sensitivity for sleep (97, 90 and 98%, respectively), but WHOOP-MANUAL had lower specificity for wake (45%) than WHOOP-AUTO (60%) and ACTICAL (60%). Chance-corrected agreement was *fair* for ACTICAL (κ = 0.23) and *moderate* for WHOOP-AUTO (κ = 0.44) and WHOOP-MANUAL (κ = 0.48). Intraclass correlation coefficients showed that WHOOP-AUTO (0.45) and WHOOP-MANUAL (0.48) had *fair* reliability for two-stage classification of sleep, compared to *poor* reliability for ACTICAL (0.26). Comparisons of reliability based on intraclass correlations should be made across devices within the same study as there is no clear threshold at which a device can be considered "valid" [3]. It should be noted that a previous validation study conducted in the same laboratory found WHOOP-MANUAL to have a 51% specificity for wake when compared to PSG [5]. These findings support a previous validation of WHOOP-MANUAL two-stage categorisation of sleep [5] and provide novel support for WHOOP-AUTO as a practical alternative for two-stage categorisation of sleep in the absence of PSG.

## 5.2. Four-Stage Categorisation of Sleep

For four-stage categorisation of sleep, WHOOP-AUTO and WHOOP-MANUAL had similar overall agreement (63% and 62%, respectively) and sensitivity to light sleep (61% and 67%, respectively), SWS (63% and 61%, respectively) and REM (66% and 66%, respectively). Chance-corrected agreement for four-stage categorisation of sleep was *moderate* for WHOOP-AUTO (κ = 0.47) and WHOOP-MANUAL (κ = 0.49). As a reference point, the chance-corrected agreement between expert sleep scorers independently scoring a common set of PSG records was substantial rather than perfect (κ = 0.78) [33]. These results support a previous validation of WHOOP-MANUAL to measure four-stage and provide the first validation of WHOOP-AUTO as a practical alternative for four-stage categorisation of sleep in the absence of PSG. The main disparity between WHOOP-AUTO and WHOOP-MANUAL for four-stage categorisation of sleep compared to PSG was that WHOOP-AUTO exhibited 16% higher sensitivity for wake compared to WHOOP-MANUAL. However, WHOOP-MANUAL can provide an accurate measure of onset latency (Table 1). Depending on the variable of interest, practitioners seeking to utilise the WHOOP strap to measure sleep can selectively utilise WHOOP-AUTO or WHOOP-MANUAL functions. For example, in situations where the WHOOP strap is utilised for two-stage or four-stage categorisation of sleep for sleep opportunities between 7 and 9 h, WHOOP-AUTO appears to be the more practical, better performing function. However, given that WHOOP-MANUAL utilises a reference point for when an individual begins to attempt sleep, it should be used in situations where sleep onset latency is the variable of interest.

The difference for estimating wake between WHOOP-AUTO and WHOOP-MANUAL in this study highlights the need for future validation research to report the ability of sleep wearables to measure a range of sleep measures under auto-detection and manual function. Previous validation studies for consumer sleep wearables do not explicitly report whether data were acquired using the automatic detection of sleep or manual entering of sleep times [7–11], thus limiting practitioners' ability to best utilise sleep wearables to measure specific sleep variables. Overall, the findings of this study suggest that WHOOP-AUTO

and WHOOP-MANUAL may be used as a practical alternative for two-stage categorisation of sleep and four-stage categorisation of sleep when PSG is not available.

### 5.3. Comparison to Other Sleep Wearables

Due to an increase in consumer devices providing measures of sleep, it is important to conduct cross-device comparisons. Ideally, within-study comparisons like in the present study should be made to provide meaningful comparison. However, interpretations of cross-study comparisons can be made with consideration to differences in study methodologies (i.e., sleep opportunity, sample, sleep environment). A comparison of the performance of WHOOP-AUTO, WHOOP-MANUAL and ACTICAL, respectively, to previous sleep wearable validations can be seen in Figure 5 [5,7,9–12,25–31].
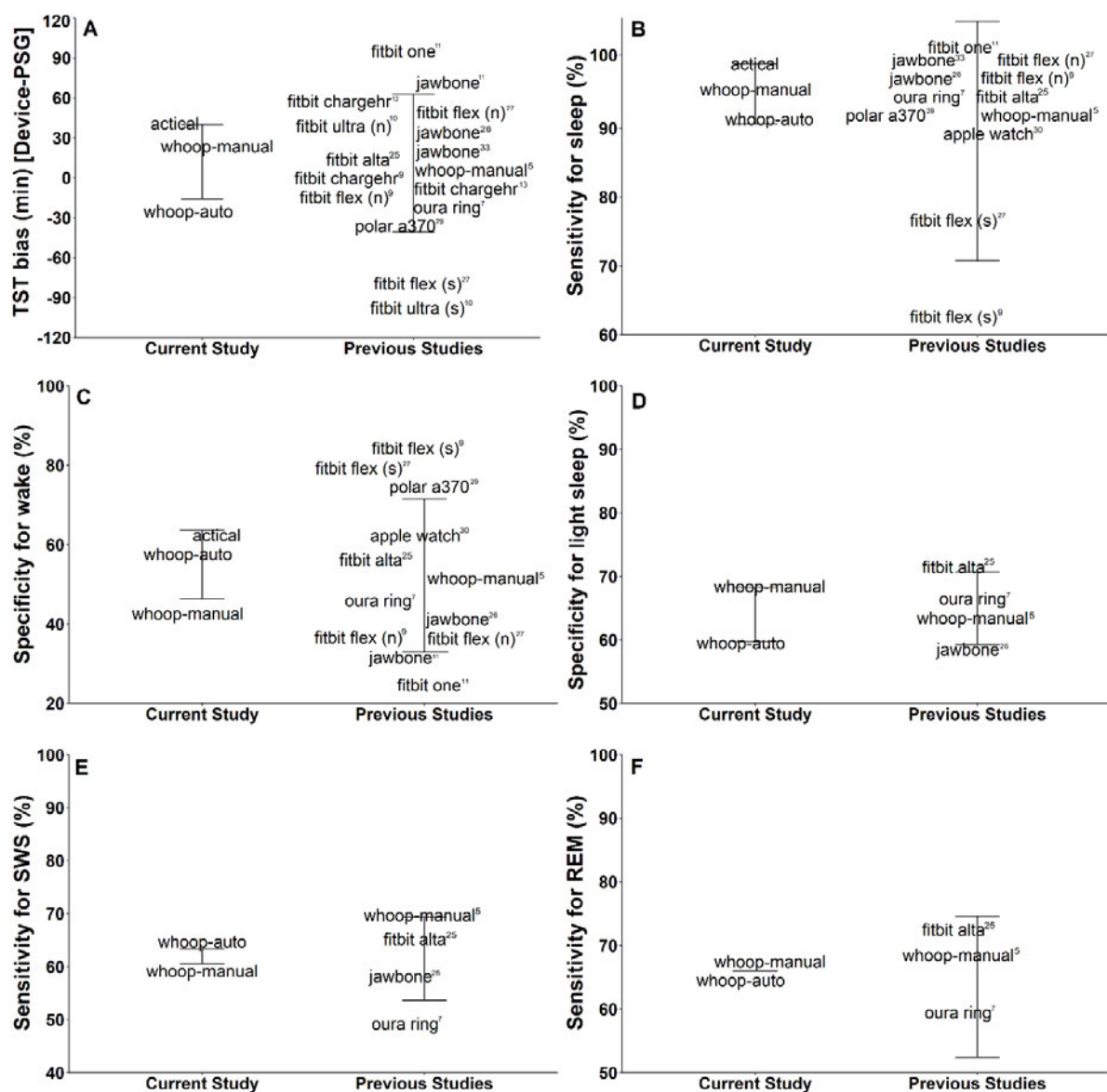


**Figure 5.** Performance of ACTICAL, WHOOP-AUTO, WHOOP-MANUAL and other sleep wearables for (**A**) total sleep time bias (TST), (**B**) sensitivity for sleep, (**C**) specificity for wake, (**D**) sensitivity for light sleep, (**E**) sensitivity for slow wave sleep (SWS) and (**F**) sensitivity for rapid eye movement sleep (REM). Error bars represent standard deviation. Fitbit Flex (N); Fitbit Flex with normal sensitivity, Fitbit Flex (S); Fitbit Flex with high sensitivity. Superscript numbers represent respective validation studies [5,7,9–12,25–31].

The WHOOP strap, in both automatic and manual functions, fell within the standard deviation for TST bias, sensitivity for sleep, specificity for wake, sensitivity for light sleep, sensitivity for SWS, and sensitivity for REM compared to previous validations (Figure 5). Previous validations of sleep wearables have shown that there is an apparent "trade-off" between sensitivity and specificity [3], such that higher sensitivity may result in decreased specificity, and vice versa. For example, a validation study conducted with the Fitbit One had high sensitivity but had low specificity compared to PSG (Figure 5) [10]. Compared to the WHOOP strap, other sleep wearables have shown higher sensitivity to individual sleep stages (Figure 5). However, both WHOOP-AUTO and WHOOP-MANUAL appear to be consistent across all four sleep stages and do not seem to exhibit a large "trade-off" between sensitivities for all sleep stages.

According to the methodologies of previous studies, WHOOP-AUTO provides the only epoch-by-epoch comparison to PSG using sleep auto-detection [7,9–12,25–27]. From a practical perspective, WHOOP-AUTO provides a measure of sleep comparable to manually adjusted data and eliminates the risk of non-compliance for entering bed times. Overall, the findings of this validation study suggest that the WHOOP strap, under both automatic and manual detection of sleep, performs well in comparison to other commercially available sleep wearables.

*5.4. Boundary Conditions and Future Research*

This validation study was conducted on the WHOOP-AUTO and WHOOP-MANUAL functions of the WHOOP strap. The validation of other WHOOP metrics (i.e., heart rate, heart rate variability) was outside of the scope of this project. The algorithms used by WHOOP to score sleep are proprietary, and epoch-by-epoch data are not accessible through the WHOOP smart phone application. Findings should also be interpreted within the boundary conditions of the sleep environment (laboratory), time in bed opportunities (7–9 h) and sample (healthy young adults). Future investigations should validate the WHOOP strap and other sleep wearables with reference to all available functionalities (i.e., auto-detection and manual adjustment) and across a wider range of conditions (sleep opportunities of different lengths, disturbed sleep periods, unhealthy and/or older populations).

## References

1. Kushida, C.A.; Littner, M.R.; Morgenthaler, T.; Alessi, C.A.; Bailey, D.; Coleman, J.; Friedman, L.; Hirshkowitz, M.; Kapen, S.; Kramer, M.; et al. Practice parameters for the indications for polysomnography and related procedures: An update for 2005. *Sleep* **2005**, *28*, 499–521. [CrossRef] [PubMed]
2. Yi, H.; Shin, K.; Shin, C. 'Development of the Sleep Quality Scale': Erratum. *J. Sleep Res.* **2006**, *15*, 465. [CrossRef] [PubMed]
3. Zambotti, M.; Cellini, N.; Goldstone, A.; Colrain, I.M.; Baker, F.C. Wearable Sleep Technology in Clinical and Research Settings. *Med. Sci. Sports Exerc.* **2019**, *51*, 1538–1557. [CrossRef] [PubMed]
4. Pollak, C.P.; Tryon, W.W.; Nagaraja, H.; Dzwonczyk, R. How accurately does wrist actigraphy identify the states of sleep and wakefulness? *Sleep* **2001**, *24*, 957–965. [CrossRef]
5. Miller, D.J.; Lastella, M.; Scanlan, A.T.; Bellenger, C.; Halson, S.L.; Roach, G.D.; Sargent, C. A validation study of the WHOOP strap against polysomnography to assess sleep. *J. Sports Sci.* **2020**, *38*, 2631–2636. [CrossRef] [PubMed]
6. Stein, P.K.; Pu, Y. Heart rate variability, sleep and sleep disorders. *Sleep Med. Rev.* **2012**, *16*, 47–66. [CrossRef]
7. De Zambotti, M.; Rosas, L.; Colrain, I.M.; Baker, F.C. The Sleep of the Ring: Comparison of the OURA Sleep Tracker against Polysomnography. *Behav. Sleep Med.* **2019**, *17*, 124–136. [CrossRef]
8. De Zambotti, M.; Baker, F.C.; Willoughby, A.R.; Godino, J.G.; Wing, D.; Patrick, K.; Colrain, I.M. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiol. Behav.* **2016**, *158*, 143–149. [CrossRef]
9. Kang, S.G.; Kang, J.M.; Ko, K.P.; Park, S.C.; Mariani, S.; Weng, J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *J. Psychosom. Res.* **2017**, *97*, 38–44. [CrossRef]
10. Meltzer, L.J.; Hiruma, L.S.; Avis, K.; Montgomery-Downs, H.; Valentin, J. Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. *Sleep* **2015**, *38*, 1323–1330. [CrossRef]
11. Maskevich, S.; Jumabhoy, R.; Dao, P.D.M.; Stout, J.C.; Drummond, S.P.A. Pilot Validation of Ambulatory Activity Monitors for Sleep Measurement in Huntington's Disease Gene Carriers. *J. Huntingt. Dis.* **2017**, *6*, 249–253. [CrossRef]
12. Sargent, C.; Lastella, M.; Romyn, G.; Versey, N.; Miller, D.J.; Roach, G.D. How well does a commercially available wearable device measure sleep in young athletes? *Chronobiol. Int.* **2018**, *35*, 754–758. [CrossRef] [PubMed]
13. Stone, A.A. Patient non-compliance with paper diaries. *BMJ* **2002**, *324*, 1193–1194. [CrossRef] [PubMed]
14. Saw, A.E.; Main, L.C.; Gastin, P.B. Monitoring athletes through self-report: Factors influencing implementation. *J. Sports Sci.* **2015**, *14*, 137–146.
15. Miller, D.J.; Sargent, C.; Roach, G.D.; Scanlan, A.T.; Vincent, G.E.; Lastella, M. Moderate-intensity exercise performed in the evening does not impair sleep in healthy males. *Eur. J. Sport Sci.* **2019**, *20*, 80–89. [CrossRef]
16. Iber, C.; Ancoli-Israel, S.; Chesson, A.; Quan, S. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specifications*; American Academy of Sleep Medicine: Westchester, IL, USA, 2007.
17. Kosmadopoulos, A.; Sargent, C.; Darwent, D.; Zhou, X.; Roach, G.D. Alternatives to polysomnography (PSG): A validation of wrist actigraphy and a partial-PSG system. *Behav. Res. Methods* **2014**, *46*, 1032–1041. [CrossRef]
18. Sargent, C.; Lastella, M.; Halson, S.L.; Roach, G.D. The validity of activity monitors for measuring sleep in elite athletes. *J. Sci. Med. Sport* **2016**, *19*, 848–853. [CrossRef]
19. Bland, J.M.; Altman, D.G. Agreement between methods of measurement with multiple observations per individual. *J. Biopharm. Stat.* **2007**, *17*, 571–582. [CrossRef] [PubMed]
20. Ludbrook, J. Confidence in Altman-Bland plots: A critical review of the method of differences. *Clin. Exp. Pharmacol. Physiol.* **2010**, *37*, 143–149. [CrossRef] [PubMed]
21. Sim, J.; Wright, C.C. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Phys. Ther.* **2005**, *85*, 257–268. [CrossRef]
22. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [CrossRef] [PubMed]
23. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
24. Cicchetti, D.V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* **1994**, *6*, 284. [CrossRef]
25. Cook, J.D.; Eftekari, S.C.; Dallmann, E.; Sippy, M.; Plante, D.T. Ability of the Fitbit Alta HR to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: A comparison against polysomnography. *J. Sleep Res.* **2019**, *28*, e12789. [CrossRef]
26. Cook, J.D.; Prairie, M.L.; Plante, D.T. Ability of the Multisensory Jawbone UP3 to Quantify and Classify Sleep in Patients with Suspected Central Disorders of Hypersomnolence: A Comparison Against Polysomnography and Actigraphy. *J. Clin. Sleep Med.* **2018**, *14*, 841–848. [CrossRef]
27. Cook, J.D.; Prairie, M.L.; Plante, D.T. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *J. Affect Disord.* **2017**, *217*, 299–305. [CrossRef]
28. De Zambotti, M.; Goldstone, A.; Claudatos, S.; Colrain, I.M.; Baker, F.C. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. *Chronobiol. Int.* **2018**, *35*, 465–476. [CrossRef]
29. Pesonen, A.K.; Kuula, L. The Validity of a New Consumer-Targeted Wrist Device in Sleep Measurement: An Overnight Comparison against Polysomnography in Children and Adolescents. *J. Clin. Sleep Med.* **2018**, *14*, 585–591. [CrossRef]

30. Walch, O.; Huang, Y.; Forger, D.; Goldstein, C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep* **2019**, *42*, zsz180. [CrossRef]

31. De Zambotti, M.; Baker, F.C.; Colrain, I.M. Validation of Sleep-Tracking Technology Compared with Polysomnography in Adolescents. *Sleep* **2015**, *38*, 1461–1468. [CrossRef]

32. Halson, S.L. Sleep Monitoring in Athletes: Motivation, Methods, Miscalculations and Why it Matters. *Sports Med.* **2019**, *49*, 1487–1497. [CrossRef] [PubMed]

33. Magalang, U.J.; Chen, N.H.; Cistulli, P.A.; Fedson, A.C.; Gíslason, T.; Hillman, D.; Penzel, T.; Tamisier, R.; Tufik, S.; Phillips, G.; et al. Agreement in the scoring of respiratory events and sleep among international sleep centers. *Sleep* **2013**, *36*, 591–596. [CrossRef] [PubMed]